An Investigation of the Prevalence, Diversity, Evolution, Physiology, and Community Dynamics of Dissimilatory Phosphite Oxidizing Microorganisms

by

Sophia Dagmar Ewens

A dissertation submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor John D. Coates, Chair
Professor Arash Komeili
Professor Douglas S. Clark

Spring 2022

Abstract

An Investigation of the Prevalence, Diversity, Evolution, Physiology, and Community Dynamics of Dissimilatory Phosphite Oxidizing Microorganisms

by

Sophia Dagmar Ewens

Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor John D. Coates, Chair

Phosphite is the most energetically favorable chemotrophic electron donor known, with a half-cell redox potential ($E^{o\prime}$) of −650 mV for the HPO_4^{2-}/HPO_3^{2-} couple. Dissimilatory phosphite oxidizing microorganisms (DPOM) harness the free energy from phosphite oxidation to support cellular growth and are universally capable of CO_2 fixation. The prevalence of this unique metabolism remained largely uncharacterized since its discovery in 2000, as only two species had been identified in three discrete locations prior to the writing of this dissertation. A false notion of rarity has consequently limited our understanding of the diversity and environmental distribution of DPOM. However, phosphite has been detected in several environments at concentrations that suggest a contemporary P redox cycle that might sustain a greater diversity of DPOM than is currently recognized. This dissertation significantly expands the known diversity of DPOM and uses enrichments and genome-resolved metagenomics to describe their prevalence, diversity, evolutionary history, and metabolic potential. Physiological studies further describe the biochemistry and community dynamics of cultures that perform dissimilatory phosphite oxidation (DPO).

Chapter 1 of this dissertation reviews our current understanding of phosphite biogeochemistry and DPOM. I begin with a historical review of phosphorus research to demonstrate the significant role phosphorus plays in both geochemical and anthropogenic contexts. Phosphorus redox chemistry is then reviewed to support a subsequent discussion on the geochemistry and distribution of reduced phosphorus species, with particular emphasis on the DPO substrate of phosphite. Establishing our understanding of phosphite distribution aids later discussions on the potential impact of DPOM on the global phosphorus cycle, for which I also review our current understanding of biological phosphite utilization, including assimilatory phosphite oxidation (APO) and DPO. This chapter ends with a review of DPOM diversity and biochemistry, which briefly outlines the status of DPOM research that had been published prior to the work presented in this dissertation. **Chapter 1** is intended to prime the reader for subsequent chapters, whose separate introductions explore the pertinent aspects of DPOM in more detail.

Upon initiating this thesis work, the only DPOM known were *Desulfotignum phosphitoxidans* FiPS-3 and *Phosphitivorax anaerolimi* Phox-21, of which only FiPS-3 had been isolated. **Chapter 2** describes the establishment and cultivation of phosphite-enriched microbial communities. I focus on the methodology that successfully identified 21 additional DPOMs, whose

metagenomic data supports later chapters' analyses of DPOM prevalence, diversity, evolution, and metabolic potential. I also describe the cultivation and characterization of a highly enriched phosphite oxidizing culture (HEPO), which is used for the physiological studies of subsequent chapters.

In **Chapter 3,** the genomes of those DPOM that were enriched in Chapter 2 are used to identify DPO marker genes, whose sequences are used to build profile hidden Markov models (pHMMs). Select marker genes are compared with over 17,000 publicly available metagenomes, and it is found that DPO metabolism exists globally in diverse anoxic environments, including wastewaters, sediments, and subsurface aquifers. By assigning taxonomy to metagenome-assembled genomes of DPOM, we find that DPOMs are also phylogenetically diverse, spanning six classes of bacteria, including the *Negativicutes*, *Desulfotomaculia*, *Synergistia*, *Syntrophia*, *Desulfobacteria*, and *Desulfomonilia_A*. By comparing DPOM taxonomy to the phylogeny of DPO marker genes, we conclude that DPO metabolism likely has ancient evolutionary origins predating the split of monoderm and diderm bacteria.

Chapter 4 discusses the physiology and metabolism of DPOM. The metagenomes of the DPOM enriched in Chapter 2 are analyzed using DRAM, a novel software that profiles the metabolic potential of microbial genomes. In this chapter, these DRAM predictions are used to refine and expand upon current models for DPOM energy conservation. I pay particular attention to the way in which DPOM appear to utilize CO_2 as both an electron acceptor for energy conservation and as a carbon source for biomass generation. Physiological studies using the HEPO culture established in Chapter 2 are leveraged to test hypotheses generated from our DRAM analyses, with primary emphasis on the identification of the CO_2-reduction end-product.

Throughout this dissertation, physiological experiments are designed to understand the activity of the DPOM that exist within the HEPO culture. When interpreting results, it is therefore important to remember that the HEPO culture is a mixed microbial community. **Chapter 5** aims to characterize and understand the potential interactions that may be occurring between DPOM and the non-DPOM community members. I consider the viability of alternative symbiotic relationships and pursue the hypothesis that DPO activity is dependent on nutrient-exchange with the non-DPOM community. I explore the potential role of alternative community members, with particular attention given to methanogens and those community members with high relative abundance. I then discuss the potential role of certain nutrients that may be exchanged, including corrinoids, coenzyme M (CoM), and S-adenosyl L-methionine (SAM).

Finally, **Chapter 6** summarizes the main conclusions of this dissertation and highlights the future research that will be needed to resolve the many questions about DPOM that remain unanswered.

Dedication

This work is dedicated to *Gunter Stein* and *Nancy Stein*, for their generous support of my academic pursuits;

to my parents, *Tom Ewens* and *Ruth Kreichauf*, whose sage wisdom has consistently guided me through the most wonderful and most difficult moments of my life;

and to my grandmother, *Joanne Hanson*, who passed during the writing of this thesis and whose strength will forever be an inspiration.

Table of Contents

List of Abbreviations

2HADH: D-2-hydroxyacid dehydrogenase protein family
ANI: average nucleotide identity
APO: assimilatory phosphite oxidation
BES: 2-bromoethanesulfonate
bp: base-pairs
CoM: Coenzyme M
DPO: dissimilatory phosphite oxidation
DPOM: dissimilatory phosphite oxidizing microorganism
E°': standard reduction potential for electro biochemical half reactions measured at 25 °C, 1 atmosphere and a pH of 7 in aqueous solution.
E2B: ethyl-2-butynoate
ETC: electron transport chain
ΔG°': Gibbs free energy at 25 °C, 1 atmosphere and a pH of 7 in aqueous solution.
F81: *Phosphitivorax anaerolimi* F81
FiPS-3: *Desulfotignum phosphitoxidans* FiPS-3
GC-MS: gas-chromatography mass-spectrometry
GHRB: glyoxylate/hydroxypyruvate reductase B protein family
GTDB-Tk: Genome Taxonomy Database Toolkit
Gya: billion years ago
HEPO: highly enriched phosphite-oxidizing (culture)
HGT: horizontal gene transfer
HPLC: high-performance liquid chromatography
IC: ion chromatography
IMG/M: Joint Genome Institute Integrated Microbial Genomes and Metagenomes database
LC-MS: liquid-chromatography mass-spectrometry
MAG: metagenome-assembled genome
MPN: most probable number
MWTF: municipal wastewater treatment facilities
NCBI: National Center for Biotechnology Information
NMR: nuclear magnetic resonance spectroscopy
OD_{600}: optical density at 600nm
OTU: operational taxonomic unit
P: phosphorus
pHMM: profile hidden Markov model
Phox-21: *Phosphitivorax anaerolimi* Phox-21
PtxD: phosphite dehydrogenase
qPCR: quantitative polymerase chain reaction
rpS8/rpS3: ribosomal protein S8 / ribosomal protein S3
Rnf: sodium translocating ferrodoxin:NADH oxidoreductase
RT: retention time
SAM: S-adenosyl L-methionine
SCG: single copy gene
THF: tetrahydrofolate
TMA: trimethylamine
WLP: Wood-Ljungdahl pathway

Acknowledgements

The completion of this dissertation was dependent on the generous support of the community that surrounded me. My deepest gratitude goes to my advisor, *John D. Coates*, without whom this work would not have been possible, due to his dependable financial support, intellectual guidance, and unwavering faith in my success. I would also like to give my sincerest gratitude to all the members of the Coates Lab, who collectively ushered me through my graduate degree: *Anna Engelbrektson* and *Misha Mehta* for acquainting me with the lab and teaching me the fundamentals of anaerobic microbiology; *Yi Liu* for keeping the lab running smoothly and for her crucial technical and logistical support; *Tyler Barnum* for his invaluable mentorship in computational biology; *Israel Figueroa* for sharing his expertise on the study of dissimilatory phosphite oxidizing organisms; *Hans Carlson* for his constant excitement, intellectual guidance, and logistical support for this project and its experiments; *Anchal Mehra, Victor Reyes-Umana, Kyle Metcalfe,* and *David Meier* for their companionship, collaboration, and guidance throughout different stages of my doctoral degree; and to my incredible undergraduate students *Alexa Gomberg, Mia Silverberg, Emily Barnhart-Ross,* and *Rebecca Salcedo* for all of their extraordinary help in the lab. I would also like to thank all of my collaborators for their essential contributions to this work: *Mikayla A. Borton* and *Kelly Wrighton*, for their help analyzing the metabolic potential of our metagenomes; *Cameron Thrash* and *John Taylor* for offering their insight on evolutionary analyses; *Kenny Mok* for his patient mentorship and training in corrinoid extraction and analyses; *Shana McDevitt* for her substantial guidance through the metagenomic sequencing process; *Jeffrey Pelton* for his help running and analyzing NMR samples; *Zhongrui Zhou* for her help with GC-MS analyses; *Anthony Lavarone* for his help with LC-MS analyses; and *Judy Walker*, *Aloke Vaid*, *Bob Wandro*, *Robert Huffstutler*, *Jan Guy, Pete Dallabetta*, *Nimisha Patel*, and *Jimmie Truesdell* for helping me collect samples from each of their wastewater facilities. Thank you to the *UC Berkeley Plant and Microbial Biology Department*, for providing the platform and education for my graduate studies. Thank you to the *Energy & Biosciences Institute, Shell Global,* and the *U.S. Department of Energy Genomic Sciences Program* for their financial support. Finally, thank you most warmly to my parents, *Ruth Kreichauf* and *Tom Ewens*, for preparing me to tackle this academic challenge and supporting me every step of the way.

Chapter 1
Introduction

Parts of this chapter have been drawn from the following publication:

Ewens, S. D., Gomberg, A. F. S., Barnum, T. P., Borton, M. A., Carlson, H. K., Wrighton, K. C., Coates, J. D. (2021). The diversity and evolution of microbial dissimilatory phosphite oxidation. PNAS.

Abstract

This introductory chapter aims to orient the reader by reviewing the current understanding of phosphite biogeochemistry and microbial phosphite oxidation. I begin with a perspective on esoteric metabolisms, followed by a historical review of phosphorus research, which demonstrates the significant role phosphorus plays in both geochemical and anthropogenic contexts. Phosphorus redox chemistry is then reviewed to support a subsequent discussion on the geochemistry and distribution of reduced phosphorus species, with particular emphasis on phosphite, which is the substrate for dissimilatory phosphite oxidation. Microbial interactions with phosphite are subsequently reviewed, including a brief introduction on the differences between assimilatory phosphite oxidation and dissimilatory phosphite oxidation. Dissimilatory phosphite oxidation is the subject of this dissertation, and the final section of this chapter briefly outlines the status of research on dissimilatory phosphite oxidizing microbes. The introductory material of subsequent chapters reviews pertinent information in more detail. I conclude with a summary of the contributions of this dissertation, and how they advance our understanding of the esoteric metabolism of dissimilatory phosphite oxidation.

Revisiting Earth's Esoteric Biogeochemical Cycles

At the intersection of microbiology and geochemistry lie the interdisciplinary subjects of biogeochemistry and geomicrobiology, whose origins can be traced back to the mid-1800's[1]. Sergei Winogradsky christened these interdisciplinary fields with his early research on the role of sulfur oxidation in elemental sulfur precipitation[2], the role of iron oxidation in both ferric oxide deposits and bicarbonate geochemistry[3], and the influence of nitrifying bacteria on nitrogen availability in agricultural practices[4]. In concert with other scientists of this era, Winogradsky laid the intellectual groundwork for microbiologists to identify novel metabolisms and characterize their interactions with global geochemical cycles[5].

Microorganisms interact with their environment by utilizing chemical species that fulfill the four requirements for microbial life: i) an energy source, ii) an electron donor, ii) an electron acceptor, and iv) nutritional sources. Their acquisition of these substrates can be classified as assimilatory or dissimilatory. Assimilatory pathways are used to fulfill the nutritional requirements of the cell for biomass accumulation. Dissimilatory pathways are used to supply energy to the cell via redox reactions between an electron donor and an electron acceptor. Dissimilatory pathways are of particular importance to biogeochemical cycles, as metabolized substrates involve redox reactions that change the oxidation state of the environment (e.g., iron oxidation to form rust), and end-products are immediately exchanged with the environment, rather than integrated into biomass[6,7].

Extensive work in major nutrient cycles (i.e., carbon, nitrogen, and sulfur) has exemplified the diversity of mechanisms that microorganisms can use to transform environmental compounds, thereby influencing the surrounding geochemistry in ways that are both geologically and anthropogenically relevant. Carbon is perhaps the most widespread and influential cycle in Earth's biogeochemical system[8]. Although most of the Earth's carbon is stored in the lithosphere and is broadly inaccessible to life on Earth, a substantial amount of global carbon is stored in living terrestrial biomass[9]. Consequently, the linchpin between geochemical carbon deposits and life on Earth is the process of biological fixation of inorganic carbon to biomass. Most carbon fixation pathways are exclusively used for carbon *assimilation*, but the Wood-Ljungdahl pathway (WLP) is also used for energy conservation via *dissimilatory* carbon dioxide (CO_2) reduction[10]. Using the WLP, methanogens reduce CO_2 to methane, a greenhouse gas with 28-34 times the global warming potential of CO_2 in a 100 year timespan[11]. The biogeochemistry of dissimilatory CO_2 reduction to methane consequently amplifies the effects of anthropogenic climate change[12]. Like carbon, nitrogen is highly redox-active, and bioavailable compounds include oxidized nitrate (NO_3^-) and nitrite (NO_2^-), as well as dinitrogen (N_2) and reduced ammonium (NH_4^+)[13]. Multiple biotic processes have been described that interface with these nitrogen species, including nitrification, denitrification, ammonia oxidation, and nitrogen fixation, which collectively influence agricultural nitrogen availability, eutrophication, and marine microbial community structure[13]. Meanwhile, sulfur undergoes biotransformations that include sulfate (SO_4^{2-}) reduction, sulfide (H_2S) oxidation, and sulfur (S^0) disproportionation[14,15], all of which are also important mechanisms for biological energy conservation and have simultaneous geochemical influences like hydrogen-sulfide corrosion, metal-sulfide precipitation, and sulfur precipitation[16]. The complex global cycles of carbon, nitrogen, and sulfur are well-studied and heavily influenced by microbial activity, and because these elements have broad ranges of oxidation states, their cycles are particularly influenced by microbial redox reactions (Table 1.1).

The diversity of microbial energy metabolisms extends beyond the most well-studied global nutrient cycles, as there exist numerous metabolisms involving substrates that are comparatively rare in the environment. Studies continue to unveil esoteric dissimilatory microbial metabolisms that utilize elements spanning the periodic table, including non-metals[17,18], metalloids[19], transition metals[20], actinides[20], and rare earth elements[21]. The identification of these esoteric metabolisms inherently suggests that those substrates are supporting microbiology on Earth, which is also reciprocally influencing the host-environment. With the identification of increasingly diverse microbial metabolisms, it is crucial to characterize their biochemistry and their influence on elemental cycles. Such studies will better describe the true prevalence and diversity of those metabolisms which may still be considered 'niche', allowing for wholistic models of global nutrient cycles. This dissertation focuses on the characterization of dissimilatory phosphite oxidation (DPO), a microbial metabolism that redox cycles phosphorus.

History of Phosphorus

The element of phosphorus (P) has a tragicomic place in human history that began with its discovery in 1669 by the German alchemist, Hennig Brandt[22]. In an attempt to identify the philosopher's stone, Brandt distilled an excess of urine in the presence of sand and coal to yield a chemiluminescent, waxy substance that he named 'phosphorus' based on Greek ontology describing its phosphorescence. Phosphorus quickly gained the attention of nobles for its mystical glow and was used in fallacious medicine throughout the 17th and 18th centuries. A century passed before Antoine Lavoisier recognized P as the 15th element on the periodic table[22]. By way of his distillation process, Brandt had identified white P, which is one of several elemental P allotropes, each with distinctive properties[23]. White P is the most reactive allotrope, and its flammability was harnessed in the mid-1800's for the production of matches[22]. White P is also highly toxic, and it has since been replaced in matches with the non-toxic, red P allotrope. Unfortunately, the combined reactivity and toxicity of white P continues to be harnessed for chemical warfare[24].

Despite its many allotropes, elemental P is highly reactive and rarely found in nature[23]. Given this rarity, the continuously growing demand for P since its discovery in urine has led to a persistent search for increasingly concentrated and accessible sources. In the mid-1700's, bones were found to contain high concentrations of calcium phosphate, and bone ash served as the primary P source until the mid-1800's. Around this time, scientists had identified P as one of the essential elements in fertilizer which could support agricultural food production, pushing the importance of P into the realms of both military and food security. This precipitated the acquisition of guano (high-phosphorus seabird excrement) from uninhabited oceanic islands, and most recently, the mining of P from mineral-phosphate deposits[22].

It is now abundantly recognized that P is required and irreplaceable for all life on Earth, as it is biochemically required for information transfer (e.g., DNA), energy transfer (e.g., ATP), cell structure (e.g., phospholipids), and metabolic regulation (e.g., protein phosphorylation). Yet P is often the limiting nutrient for the outgrowth of life due to its low abundance and low bioavailability[25,26]. Consequently, P serves as a key metric to determine biological productivity, especially in marine systems. Given the finite and limited quantities of P on earth, the unsustainable linearization of P utilization in modern agriculture has generated concern[22], and scientists estimate the global population will only have ~200 years before this necessary and non-

renewable resource is depleted[27]. As such, models of the P cycle typically focus on the bulk content and transport of P and its conversion between mineral-bound, organic, and ionic states[28].

Phosphorus is one of the six most common elements in biological systems, which include carbon (C), nitrogen (N), and sulfur (S), as well as hydrogen (H) and oxygen (O). Yet, unlike any of the other six major life elements, geochemical models of the P cycle largely disregard the effects of microbial redox transformations. This is in contrast to recent findings, which show that P occupies several redox states that are both environmentally relevant and biologically accessible[29,30]. Given the timely importance of depleting P reserves[31], it is therefore essential that the global P cycle is described with increasing resolution and to greater completion, which necessitates the incorporation of microbial redox transformations.

The Global Phosphorous Cycle

Current models of the global P cycle are generally defined by four major phenomena: 1) tectonic uplift, which supplies terrestrial systems with buried, sedimentary P; 2) erosion and weathering of rocks, which exposes dissolved and particulate P to local water systems; 3) transport of P to the ocean via rivers; and 4) oceanic sedimentation of P associated with organic and mineral matter, which is then returned to terrestrial systems via tectonic uplift[28]. In the context of these models, "P" encompasses mineral P, inorganic P, and organic P, all of which are generally assumed to contain oxidized P in the +5-valence state (P^{+5}). This assumption is made because P on Earth predominates in its oxidized form (P^{+5}) and because the environmental impact of P redox chemistry remains poorly defined. However, P is known to occupy multiple redox states that are detected in the environment and involved in biological activity[29]. As such, environmental redox cycling of P is considered a critical frontier in need of further study and consideration[28,32].

Phosphite (HPO_3^{2-}) is the substrate of focus for this dissertation. It is a reduced phosphorus (P) compound in which the P atom occupies the +3 valence state[33]. Phosphite is 1000x more soluble than phosphate[34], and it can account for over 30% of the total dissolved inorganic phosphorus in diverse environments[29,35]. Phosphite is kinetically stable on the order of years, with its longevity inversely proportional to the number of oxidizing radicals in solution[34]. However, phosphite is thermodynamically reactive, as represented by its highly negative reduction potential when oxidized to phosphate ($E^{o'}$ = -650 mV for HPO_4^{2-}/HPO_3^{2-})[33]. Emerging evidence suggests that phosphite has played a critical role in the development of early life and continues to be an important substrate for microorganisms on neoteric earth. Yet its neogenesis and interaction with the global P cycle remains ambiguous.

Phosphite on Early Earth

Evidence suggests that meteorite impacts deposited substantial quantities of phosphite on early Earth. It is estimated that 10^{15} to 10^{18} kg of reduced P was delivered during the heavy bombardment period (3.8 – 3.9 billion years ago [Gya]), and as much as 10% of the total P on Earth's surface could be derived from meteorite impacts[34]. Schreibersite is a significant component of meteorites and can abiotically corrode over geological timescales (1-10^4 years) in the presence of water, with phosphite representing >50% of the total soluble P produced[34]. Meteoric impacts are therefore likely to have deposited a substantial quantity of phosphite on the early Earth. Additional phosphite is predicted to have been derived from lightning discharges reducing

mineral-bound phosphate[36]. Under the reducing conditions of Archean Earth, extraterrestrial and electrochemically deposited phosphite is predicted to have had a half-life of 10^8 – 10^{10} years and could have accumulated in the early ocean to concentrations as high as 10^{-5} to 10^{-2} moles per liter[34,37]. These findings support the early hypothesis by Gulick (1955) that reduced phosphorus compounds may have played a significant role in the emergence of life on Earth[38], which is estimated to have occurred ~3.5 - 4 Gya[39]. This hypothesis was inspired by "the phosphate problem", which notes that phosphate is a poor substrate for the phosphorylation of prebiotic organic molecules due to its low solubility and poor reactivity. In contrast, reduced P species, such as phosphite and hypophosphite, are significantly more soluble and reactive towards organic carbon and nitrogen and could have facilitated the emergence of phosphorylated biomolecules. In support of this, schreibersite corrosion has been linked to the synthesis of phosphorylated biomolecules from phosphite radicals[40].

Phosphite on Neoteric Earth

Despite the likely prevalence of reduced P on Archean Earth, the Great Oxygenation Event (~2.5 Gya) is assumed to have rendered phosphite negligible in neoteric environments due to its abiotic oxidation to phosphate on geological timescales[34]. However, phosphite continues to be detected in diverse environments at concentrations ranging from 0.1 to 1.3 μM, accounting for 1 to 33% of total dissolved P in these systems[35,41,42]. Most environments in which phosphite has been detected are reducing, including river and lake sediments, swamps, and geothermal pools, but up to 1 μM phosphite has also been observed in some surface waters[34]. This suggests contemporary neogenesis, since ancient phosphite in oxygen rich surface waters would have been oxidized on geological timescales.

The primary mode of phosphite neogenesis has not been well established, but several alternative mechanisms have been proposed. Some phosphite is likely due to anthropogenic activity. Phosphite can be a byproduct of industrial phosphonate production (compounds with C-P bonds), which are used as herbicides, detergents, and chelating agents[30,43,44]. Additionally, phosphite itself is used as a reducing agent in electroplating processes[45] and as a fungicide in agriculture[46]. Phosphite can therefore be a component of industrial waste and agricultural runoff, and has been detected in wastewater treatment influent[43,47]. Furthermore, higher concentrations of phosphite are observed in heavily polluted lakes compared to less impacted areas[41]. In pristine systems, geothermal and hydrothermal activity may generate phosphite through metal phosphide corrosion and iron-mediated phosphate reduction[48,49]. Phosphite might also be produced by the partial oxidation of phosphine (O_2 and/or UVB), which is a ubiquitous gas at low concentrations and has been measured at higher concentrations in highly reduced sediments, organically enriched wastes (e.g., animal manure), and animal flatus[50,51]. However, the neogenesis of phosphine is equally enigmatic, providing little opportunity to integrate phosphite neogenesis into the larger P redox cycle. Alternatively, phosphite may be derived from biological phosphonate degradation or anomalous phosphate reduction[33,52–54]. However, none of the hypotheses for biological phosphite generation have been experimentally validated, and the enigma of contemporary environmental phosphite (bio)synthesis is yet to be resolved.

Phosphite Oxidation in Biology

Phosphorus is an essential nutrient for all life on Earth, but the only form of P that is immediately accessible for cellular intake and incorporation is soluble orthophosphate (HPO_4^{2-}) in which P is in the +5-valence state (P^{+5}). While the majority of P on Earth is found as orthophosphate, this is predominantly trapped in mineral deposits that are biologically unavailable, resulting in P limitation for most ecosystems[55]. In response to P limitation, organisms have evolved mechanisms that expand their capacity for P acquisition when soluble, inorganic orthophosphate is unavailable. Organisms will degrade mineral phosphates and organophosphates to release molecularly bound organophosphate[56]. An increasing number of organisms have also been shown to utilize reduced P species[30]. Phosphonates are a form of organic P in which the P is in the +3-valence state and is directly bound to an organic, carbon backbone via a C-P bond. Phosphorus from phosphonates is bioavailable via oxidation following enzymatic cleavage[53]. The bioavailability of phosphonates represents one of the first examples of reduced P impacting the global P cycle at a significant scale, as their degradation has been shown to generate oceanic methane emissions[57,58]. Only ~20 years ago did scientists discover that phosphonates can account for up to 25% of dissolved organic P in marine environments and are sourced by both industrial and biosynthetic processes[30,59,60].

Microbes can also utilize *inorganic* reduced P species such as phosphite and hypophosphite to accommodate P needs[29,30,61]. Promiscuous C-P lyases that typically act on phosphonates, and bacterial alkaline phosphatases that typically act on organophosphates, can each convert phosphite to bioavailable organophosphate[62,63]. However, the phosphite dehydrogenase (PtxD) is the only known dedicated phosphite dehydrogenase, which exclusively couples phosphite oxidation to NAD^+ reduction[64]. The utilization of phosphite for nutritional P is termed assimilatory phosphite oxidation (APO), and this capacity is predicted to be widespread across a variety of environments and phylogenetic lineages: roughly 1.5% of genomes in the Integrated Microbial Genomes and Metagenomes (IMG/M) database contain *ptxD* homologs[29], and up to 67% of cultured bacteria have been found to utilize phosphite as a P source[65].

Dissimilatory Phosphite Oxidation

Microorganisms can alternatively harness the extremely low redox potential of phosphite oxidation ($E^{o'}$ = -650 mV for HPO_4^{2-}/HPO_3^{2-}), using phosphite as the sole electron donor and energy source while excreting biogenic phosphate from cells[30]. Phosphite represents the most energetically favorable chemotrophic microbial electron donor described, yet only two DPO microbes (DPOM) had been cultured prior to the writing of this dissertation, and only one of these representatives was successfully isolated.

Dissimilatory phosphite oxidation was first identified in *Desulfotignum phosphitoxidans* strain FiPS-3, an autotrophic homoacetogenic facultative sulfate-reducing bacterium, isolated from Venetian brackish sediments[66]. DPO in FiPS-3 is attributed to the *ptx-ptd* gene cluster (*ptxDE-ptdCFGHI*), which FiPS-3 likely acquired through horizontal gene transfer (HGT)[67–69]. The most closely related cultured isolate to FiPS-3 is incapable of DPO although the organisms share 99% 16S rRNA gene identity[18]. The second known DPOM, Ca. *Phosphitivorax anaerolimi* Phox-21, was enriched from wastewater collected in Oakland, California, and recently, another *Phosphitivorax* strain with the *ptx-ptd* genes (Ca. *P. anaerolimi* F81) was identified in Danish wastewater[70,71]. Phox-21 grows chemolithoautotrophically with phosphite and CO_2 as the sole

electron donor and electron acceptor, respectively, and it is the first naturally occurring species proposed to fix CO_2 via the reductive glycine pathway[70,72,73]. The reductive glycine pathway has since been confirmed to naturally fix CO_2 in wild-type *Desulfovibrio desulfuricans*[74]. Phox-21 harbors all *ptx-ptd* genes, but unlike FiPS-3, lacks *ptdG* (a putative transcriptional regulator) and shows no evidence of horizontal acquisition of the *ptx-ptd* cluster[67]. The latest DPOM to be identified is *Phosphitispora fastidiosa*: a spore-forming bacterium belonging to the *Firmicutes* phylum that was isolated during the writing of this dissertation from anaerobic sewage sludge in Germany[75]. *P. fastidiosa* grows exclusively with phosphite as the electron donor and CO_2 as the electron acceptor, and its genome has not yet been analyzed for the presence or absence of the *ptx-ptd* genes. Understanding the evolutionary history of DPO metabolism is consequently limited by the existence of only two genomically characterized DPOM (i.e., Phox-21 and FiPS-3) whose *ptx-ptd* clusters exhibit deviating patterns of composition and inheritance.

Scarce representation also limits our understanding of the genes, organisms, and environments that support DPO. It is difficult to predict the range of DPO taxa because *D. phosphitoxidans* FiPS-3, Ca. *P. anaerolimi,* and *P. fastidiosa* are members of distinct taxonomic classes (*Desulfobacteria, Desulfomonilia_A,* and *Desulfotomaculia*), and their closest relatives are either uncultured or unable to catalyze DPO[29]. The environmental context of DPO remains ambiguous since active DPOM have only been identified in three distinct locations globally[18,70,75]. Furthermore, the *ptx-ptd* cluster has unresolved genetic diversity. *D. phosphitoxidans* FiPS-3 and Ca. *P. anaerolimi* species have *ptx-ptd* clusters with alternative synteny and gene composition, and the PtxD proteins from FiPS-3 and Phox-21 share only 55% amino acid sequence similarity[70]. The *ptx-ptd* genes of *P. fastidiosa* have not yet been identified or characterized.

In this dissertation, I present my findings on the prevalence, diversity, evolution, physiology, and community dynamics of DPOM. In **Chapter 2,** selective DPO enrichments are established using inoculum from the digester sludge of wastewater facilities around the San Francisco Bay area. Metagenomic sequencing is performed to establish the datasets for downstream computational analyses, and a hyper-enriched DPO culture (HEPO) is established for downstream physiological analyses. In **Chapter 3**, phylogenetic analyses of the DPO enrichment metagenomes identify 21 DPOM spanning three disparate phyla. Marker genes applied to public metagenomes confirms this diversity and shows that DPOM exist globally in diverse anoxic environments. This newfound diversity of DPOM allows for evolutionary analyses that suggest DPO is an ancient metabolism that evolved ~3.2 Gya. In **Chapter 4**, comparative genomics reveals conservation of energy generation and carbon utilization pathways among DPOM genomes, despite their taxonomic diversity. Physiological studies of the HEPO culture further test various metagenome-derived hypotheses, resulting in the identification of an unknown carbon compound that accumulates temporally with phosphite oxidation, providing insight into the CO_2-reduction pathway utilized by *Phosphitivorax* species. In **Chapter 5,** community dynamics of the HEPO culture are explored to describe potential relationships that DPOM may have within their microbial environment. In these community-based studies, it is found that corrinoids are likely exchanged with the DPOM, who appear to require these corrinoids for DPO activity. **Chapter 6** then serves to summarize the findings of this dissertation, which collectively show that DPO, which until recently was considered a niche microbial metabolism, is widespread across diverse environments and bacterial taxa, and likely represents a vestige of ancient microbial life with important implications in primary production.

Figures

Table 1.1: Redox potentials of various electron donors, electron acceptors, and common cofactors involved in microbial electron transfers at pH 7. For a more complete collation of reduction potentials, the reader is referred to Bard, et al.[76].

Element/Category	Equation (ox/red)	Reduction potential (mV)	Reference
Phosphorus (P)	$H_2PO_2^-/P_4$	-1220	[33]
	HPO_3^{2-}/PH_3	-760	[33]
	$HPO_3^{2-}/H_2PO_2^-$	-740	[33]
	$H_2PO_2^{-/}PH_3$	-660	[33]
→	**$H_2PO_4^-/HPO_3^{2-}$**	**-650**	[33]
	P_4/PH_3	-480	[33]
Hydrogen (H)	H^+/H_2	-414	[77]
Sulfur (S)	SO_4^{2-}/HSO_3^-	-516	[77]
	$S_2O_3^{2-}/HS^- + HSO_3^-$	-402	[77]
	S^0/HS^-	-270	[77]
	$HSO_3^-/S_3O_6^{2-}$	-173	[77]
	HSO_3^-/HS^-	-116	[77]
	$S_3O_6^{2-}/S_2O_3^{2-} + HSO_3^-$	+225	[77]
Carbon (C)	CO_2/formate	-432	[77]
	CO_2/acetate	-290	[77]
	CO_2/CH_4	-244	[77]
	acetaldehyde/ethanol	-197	[77]
	pyruvate$^-$/lactate$^-$	-190	[77]
	dihydroxyacetone phosphate/glycerol-phosphate	-190	[77]
	oxaloacetate^{2-}/malate^{2-}	-173	[77]
	glycine/acetate$^-$ + NH_4^+	-10	[77]
	fumarate/succinate	+33	[77]
Transition metals	Mo $(E'_{01})^{a,b}$	-265	[78]
	Mo^{5+}/Mo^{4+}	-18	[78]
	Mo^{6+}/Mo^{5+}	+110	[78]
	Fe^{3+}/Fe^{2+}	+772	[77]
Nitrogen (N)	NO_2^-/NO	+350	[77]
	NO_3^-/NO_2^-	+433	[77]
	NO/N_2O	+1175	[77]
	N_2O/N_2	+1355	[77]

Table 1.1: *continued.*

Element/Category	Equation (ox/red)	Reduction potential (mV)	Reference
Iodine (I)	IO_3^-/I^-	+720[d]	[79]
Oxygen (O)	O_2/H_2O	+818	[77]
Chlorine (Cl)	ClO_3^-/ClO_2^-	+744	[80]
	ClO_4^-/ClO_3^-	+813	[80]
	ClO^-/Cl^-	+1261	[80]
	ClO_2^-/ClO^-	+1280	[80]
Cofactors[c]	F_{430}: Ni^{2+}/Ni^+	-650	[81]
	cob(II)alamin/cob(I)alamin; *base-off*	-572	[82]
	cob(II)alamin/cob(I)alamin	-526	[82]
	Ferredoxin ox/red (E'_{01})	-398	[77]
	F_{420} ox/red	-380	[83]
	flavodoxin ox/red (E'_{01})	-371	[77]
	NAD/NADH	-320	[77]
	cytochrome c3 ox/red	-290	[77]
	$FAD/FADH_2$	-220	[77]
	$FMN/FMNH_2$	-190	[77]
	flavodoxin ox/red (E'_{02})	-115	[77]
	menaquinone ox/red	-74	[77]
	$APS/AMP + HSO_3^-$	-60	[77]
	rubredoxin ox/red	-57	[77]
	acrylyl-CoA/propionyl-CoA	-15	[77]
	2-demethylvitamin K_2 ox/red	+25	[77]
	ubiquinone ox/red	+113	[77]
	cob(III)alamin/cob(II)alamin	+273	[82]
	Cu^{2+}/C^0	+341	[84]

[a] *E'_{01} signifies a single electron transfer*
[b] *The referenced study did not assign the precise oxidation states of the Mo atom.*
[c] *The reduction potential of cofactors is highly variable, depending on their coordination state in the associated protein.*
[d] *Reduction potential reported at pH8.*

Chapter 2

Selective Enrichment and Preliminary Characterization of DPOM

Parts of this chapter have been drawn from the following publication:

Ewens, S. D., Gomberg, A. F. S., Barnum, T. P., Borton, M. A., Carlson, H. K., Wrighton, K. C., Coates, J. D. (2021). The diversity and evolution of microbial dissimilatory phosphite oxidation. PNAS.

Abstract

Dissimilatory phosphite oxidation (DPO) is a chemolithotrophic microbial metabolism in which phosphite is used as an electron donor to support cellular growth. *Desulfotignum phosphitoxidans* FiPS-3 was the first bacterium shown to perform DPO, and it was enriched and isolated in the year 2000 from Venetian canal sediment. The second known DPOM, *Phosphitivorax anaerolimi* Phox-21, was identified in 2017 from the metagenome of an active DPO enrichment whose inoculum was from a San Francisco Bay area wastewater facility, and another *P. anaerolimi* strain (F81) with the *px-ptd* genes was recently identified in Danish wastewater. The latest DPOM to be identified is *Phosphitispora fastidiosa,* which was isolated in 2021 from German anaerobic sewage sludge during the writing of this dissertation. *Phosphitispora fastidiosa* and FiPS-3 remain the only isolated dissimilatory phosphite oxidizing microorganisms (DPOM) to date. To survey the prevalence and diversity of DPOM, selective enrichments were inoculated with wastewater sludge from six different wastewater facilities around the San Francisco Bay area. Ion chromatography was used to monitor metabolic activity and identify active DPO enrichments, and metagenomic sequencing of active enrichments allowed for the identification of 21 DPOM. The metagenomic data from DPOM enrichments that is introduced in this chapter also serves as a central dataset for later chapters. This chapter also presents various attempts to isolate DPOM from select enrichments. While isolation was never achieved, these efforts resulted in a hyper-enriched phosphite oxidizing (HEPO) culture. In this chapter, basic physiological characterization of HEPO is used to outline several hypotheses that are tested in future chapters, where HEPO serves as the primary model-system for physiological studies.

Introduction

Phosphite (HPO_3^{2-}) is a highly soluble, reduced phosphorus (P) compound that can account for over 30% of the total dissolved inorganic phosphorus in diverse environments[29,35]. Evidence suggests that meteorite impacts deposited substantial quantities of phosphite on early Earth, but its abiotic oxidation to phosphate after the Great Oxidation Event (~2.5 billion years ago [Gya]) is assumed to have rendered phosphite negligible in neoteric environments[34]. Surprisingly, phosphite has been detected in diverse, extant, reducing environments, and up to 1 μM has been observed in some surface waters, suggesting contemporary neogenesis[34,35]. Geothermal and hydrothermal systems may generate phosphite through metal phosphide corrosion and iron-mediated phosphate reduction[48,49]. Alternatively, phosphite may be partially derived from biological phosphonate degradation or anomalous phosphate reduction[52,53]. Some phosphite accumulation is likely attributable to anthropogenic activity as well because comparatively higher concentrations of phosphite have been identified in contaminated environments and industrial wastewaters[29,35,41].

Municipal wastewater treatment facilities (MWTF) have shown elevated concentrations of phosphite as high as 34 ug/L, representing 2.27% of total soluble P[43]. Yu et al. observed that phosphite concentrations were seasonal and decreased as wastewater treatment progressed, suggesting that phosphite may be exogenously sourced. Alternatively, phosphite may be generated during the wastewater treatment process via mechanisms that are similar to those proposed above, including microbial phosphonate degradation, microbial phosphate reduction, or chemical activity that is stimulated by the relatively high P concentrations found in municipal wastewater[85]. P removal is a crucial step in wastewater treatment to prevent downstream environmental eutrophication, and phosphite poses a particular challenge in this process because it is highly soluble (1000x phosphate) and kinetically stable[86]. In an effort to capture phosphite in the P-removal process, researchers are developing novel methods to expedite phosphite oxidation for subsequent phosphate precipitation and removal[86].

The HPO_4^{2-}/HPO_3^{2-} redox couple has an extremely low redox potential ($E^{o'}$ = -650 mV) which can be harnessed by microorganisms performing DPO. When used for DPO, phosphite represents the most energetically favorable chemotrophic microbial electron donor described[33]. Yet the highly energetic metabolism of DPO is still considered niche, as only three species have been identified to date[18,70,75]. As reviewed in **Chapter 1**, *Desulfotignum phosphitoxidans* strain FiPS-3 was the first DPOM to be identified, and it is one of only two DPOM isolates. FiPS-3 is an autotrophic homoacetogenic facultative sulfate-reducing bacterium that was isolated from Venetian brackish sediments and couples phosphite oxidation to the reduction of carbon dioxide (CO_2), sulfate (SO_4^{2-}) and nitrate (NO_3^-)[66,67]. DPO in FiPS-3 is attributed to the *ptx-ptd* gene cluster (*ptxDE-ptdCFGHI*), which FiPS-3 likely acquired through horizontal gene transfer (HGT)[67–69]. The taxonomy of FiPS-3 was not predictive of novel DPOM, as its most closely related cultured isolate was incapable of DPO even though the organisms share 99% 16S rRNA gene identity[18]. The second known DPOM, Ca. *Phosphitivorax anaerolimi* Phox-21, was enriched from MWTF sludge collected in Oakland, California[70]. A *Phosphitivorax* strain (Ca. *P. anaerolimi* F81) was also recently identified through metagenomic analyses of a Danish wastewater microbial community; however the capacity for F81 to perform DPO has not been biochemically validated[71]. Phox-21 grows chemolithoautotrophically with phosphite and carbon dioxide (CO_2) as the sole electron donor and acceptor, respectively. No alternative donors or acceptors have been identified to support the growth of Phox-21. Like FiPS-3, Phox-21 harbors all *ptx-ptd* genes, but unlike FiPS-3, it lacks *ptdG* (a putative transcriptional regulator) and shows no evidence of horizontal

acquisition of the *ptx-ptd* cluster[70]. The shared genes between Phox-21 and FiPS-3 (*ptxDE* and *ptdCFHI*) therefore serve as the most promising markers for genomic identification of DPOM.

Microbial cultures are pivotal to the characterization of esoteric metabolisms that have scarce representation, such as DPO[87]. Reviews by Rosenberg et al. and Lewis et al. offer thorough analyses of the classic and forthcoming strategies that may improve the cultivation of environmental microbes[88,89], and several of these strategies were used in this chapter to enrich for diverse DPOM. Traditional selective enrichments sourced from environmental samples, a technique first established during the era of Winogradsky, served as the model for my methodology[5]. Below is a set of considerations and historical observations that influenced the experimental design of my DPO enrichments.

Inoculum: When choosing an inoculum, the Bass Becking hypothesis (everything is everywhere, but the environment selects) suggests that any inoculum source eventually provides one with an active enrichment of their metabolism of interest. However, the choice of inoculum is important when one seeks to increase the likelihood of identifying an environmentally representative microorganism performing a particular metabolism. In the case of DPOM, wastewater digester sludge serves as a promising initial candidate due to the previous identification of Phox-21 and F81 in MWTF, as well as the inherent microbial diversity and elevated concentrations of phosphite previously observed[43,86,90].

Culture Conditions: Microbial growth is dependent on i) an energy source, ii) an electron donor, iii) an electron acceptor, and iv) a nutrient source. Since DPOM is a metabolism defined by the use of phosphite as an electron donor and energy source, amending culture media with phosphite as the only electron donor and P source can selectively support the preferential growth of DPOM. DPOM have been shown to couple DPO to the reduction of carbon dioxide, sulfate, and nitrate [18,67,70], suggesting at least one of these electron acceptors must be present as well. Finally, one must also account for environmental and nutritional factors that support growth, including but not limited to temperature, pH, dissolved oxygen content, conductivity, redox state, pressure, salinity, and nutritional content. The nutritional content is the most challenging to parameterize, as environmental concentrations of certain nutrients are complex, transient, and may include undetectable quantities of essential substrates. Generalists without auxotrophies are among the easiest to enrich, as a defined media with an appropriate electron donor and acceptor adequately supports growth. On the other hand, specialists may require media made by sterile-filtering liquid from the host environment and amending with supplemental nutrients, donors, and acceptors. Working with filtered environmental media is challenging because of the large quantities needed for experimentation and the potential spatiotemporal variability of essential substrates at the source. When additional media is needed, experimental replication is complicated by changing environmental conditions that are hard to characterize and laborious to describe. Soil or sediment samples whose aqueous content is only a fraction of the sample volume further exacerbate this challenge. A compromise can be made by designing defined media that is supplemented with an undefined nutrient mix. This was the strategy used by Figueroa et al. to optimize growth of Phox-21: rumen fluid was added to defined growth media and increased the rate of DPO six-fold[70]. As is often the conundrum for undefined media, it is still unclear which component of rumen fluid stimulates DPO activity.

Monitoring: Following the curation of an active enrichment culture, microbial growth must be measured to reliably identify active enrichments. Choosing which method to use for monitoring ultimately depends on a compromise between throughput, accuracy, and precision. Methods such

as Most Probable Number (MPN) analysis, cell counting, or dry cell weight measurements are used commonly, each with unique benefits and drawbacks. Optical density measurements are commonly used to assess turbidity using light at a wavelength equivalent to the size of a bacterial cell. However, optical density is non-selective, can measure non-viable cells and abiotic particles, does not properly account for the effect of cell morphology, and is further complicated in environmental enrichments by opaque sedimentary solutes in the inoculum. Modern approaches such as quantitative PCR (qPCR) circumvent many of these challenges and can be used to monitor the growth proportional to the quantitative amplification of a marker gene from extracted DNA[91]. Redundant primers for a phylogenetic marker gene (i.e., 16S rRNA gene) are typically used to measure community composition over time, but this is inadequate for identifying novel organisms whose phylogeny has not yet been defined. Primers can instead be designed to monitor the amplification of a physiological marker gene (e.g., *ptxD*) whose abundance increases with its host. For metabolisms whose diversity has not yet been characterized, redundant primers capable of amplifying all unknown variants of a physiological marker gene would likely result in false negatives. While not directly indicative of cell mass accumulation, consumption or production of a substrate directly involved in the growth metabolism of an organism can complement growth data. For instance, ions such as phosphite and phosphate can be measured using ion chromatography and their changing ratio is representative of DPO activity[70]. While ion chromatography is sensitive and accurate when analyzing active enrichments, throughput is on the order of hours or days. Colorimetric assays are an alternative that utilize the reactivity of substrates of interest for high-throughput detection. The Malachite Green colorimetric assay detects inorganic phosphate[92] and could theoretically be used to detect phosphate accumulation from DPO activity, but phosphate precipitation and non-specific biological phosphate consumption deteriorate the correlation between phosphate accumulation and DPO activity. Combining several monitoring techniques enables one to describe potential energy flows and metabolic processes within a complex microbial community.

Genomics: While frequently staged against each other for their alternative utility, metagenomics and microbial culturing can be used synchronously to better understand the prevalence and diversity of novel metabolisms[93]. When enrichments fail to yield isolates, genome-resolved metagenomics is a powerful alternative for phylogenetically and metabolically characterizing a microbe of interest that remains embedded in a microbial community. This strategy was used to identify and characterize Phox-21from within its microbial community[70]. Active DPO enrichments were temporally sequenced to identify which organisms positively responded to the presence of phosphite, and the ultimate identification of Phox-21 was based on the presence of the *ptx-ptd* gene cluster[70]. It is important to note that the genes used to identify Phox-21 as a DPOM had initially been identified using proteomics of a pure culture (FiPS-3), emphasizing the utility of isolation and biochemistry, which allow for the attribution of biological activity to definitive gene sets, especially those that are novel[69]. The gold standard for the biochemical characterization of novel genes is via genetic manipulation (i.e., gene knockouts and heterologous expression) and protein biochemistry, which have been limited in the studies of DPO metabolism.

Materials and Methods

Enrichment Conditions and Sampling

Enrichment inocula were obtained from six wastewater treatment facilities in the San Francisco Bay area of California (Table 2.1). Serum bottles (150 mL volume) (Bellco, Vineland, NJ, USA) containing basal media (45 mL) were each inoculated with sludge (5 mL) and incubated at 37 ºC. Anoxic medium was prepared by boiling under N_2/CO_2 (80:20, v/v) to remove dissolved O_2, and dispensed under N_2/CO_2 (80:20, v/v) into anaerobic pressure tubes or serum bottles. These were capped with thick butyl rubber stoppers and sterilized by autoclaving (15 min at 121 ºC). The basal medium was composed of (per 1 L of deionized water): 5 g $NaHCO_3$, 12 g HEPES buffer, 1 g NH_4Cl, 0.5 g KCl, 1.5 g $MgCl_2$, 0.15 g $CaCl_2$ ($2H_2O$), 0.5 g L-cysteine HCl and 10 mL each of vitamins and trace minerals[94]. Saline medium additionally contained 20 g/L NaCl. Salt solutions of Na_2HPO_3, Na_2SO_4, and $NaNO_3$ were added from sterile anoxic stocks as needed. Rumen fluid (Bar Diamond Inc, Parma, ID, USA) was prepared by degassing (30 minutes with N_2) and autoclaving (121 ºC for 30 min) and was added to the basal media as required. Heat killed controls were autoclaved at 121 ºC for 1 h. Samples for DNA extraction were pelleted by 30 min centrifugation at 10,000 rcf and stored at -80 ºC. Samples for ion determination were filtered and stored at 4 ºC prior to ion chromatography (IC) using the method described previously[70]. Cell growth was measured as optical density at 600nm (OD_{600}) using a Genesys^{TM} 20 Visible Spectrophotometer (Thermo Scientific).

Cultivation and Physiology of HEPO

Agar plugs were prepared by serially diluting active cultures into the anaerobic basal bicarbonate buffered media with 10 mM phosphite, 1% rumen fluid, and the same electron acceptor (i.e., CO_2 only or CO_2+SO_4^{2-}) that was available in the source-culture. For each serial dilution, liquid culture (7 mL) was injected into anaerobic tubes containing autoclaved 3.3% agar (3 mL) that was maintained in a molten state in a water bath (60 ºC). Following injection, tubes were rapidly mixed and inserted into an ice slurry to solidify. Agar plugs were inverted and incubated at 37 ºC until colonies were observed. Colonies were picked in aseptic conditions using a sterilized glass Pasteur pipette with a cotton filter and rubber bulb under a nitrogen stream and transferred to anaerobic basal bicarbonate buffered media with 10 mM phosphite and variable electron acceptors. Colony transfers were incubated at 37 ºC and monitored for growth.

Culture purity was determined via 16S rRNA gene sequencing, microscopy, and whole genome sequencing. PCR with the standard 27F and 1492R primers was used to amplify the 16S rRNA gene from DNA, which was extracted using the DNeasy PowerLyzer Microbial Kit (Qiagen). Amplified 16S rRNA gene products were Sanger sequenced at the UC Berkeley DNA Sequencing Facility. Sequence purity was estimated by visualizing the electropherogram in AliView[95]. Cells were visualized with a Zeiss Axioskop 2-plus phase contrast microscope equipped with a QImaging Retiga-2000R digital camera. DNA extract from the highly enriched phosphite oxidizing (HEPO) culture was sequenced with an Illumina MiSeq (150 bp paired-end reads) at the UC Berkeley Vincent J. Coates Genomics Sequencing Laboratory. Reads were trimmed and filtered using Sickle v1.33 (quality threshold value of 20)[96]. Reads were combined and co-assembled using MEGAHIT v1.1.2[97] using the meta-sensitive preset. Reads were mapped to assembled contigs using BWA-MEM v0.7.17[98] with default parameters. Open reading frames

were predicted from selected genomes using Prodigal v2.6.3[99] and the profile hidden Markov model (pHMM) for rpS3 was obtained from Wu et al. and applied to the metagenome[100].

Physiological studies were conducted in anaerobic glass tubes (Bellco, Vineland, NJ, USA) with 9 mL of the basal bicarbonate media and 1 mL of the HEPO culture. Phosphite ($Na_2HPO_3 \bullet 5$ H2O), phosphate (Na_3PO_4), ammonium ($NH_4MgPO_4 \bullet 6H_2O$), and sodium chloride (NaCl) (Sigma Aldrich) were added from sterile, anaerobic stocks to the basal medium as needed. Optical density and phosphite oxidation were monitored as described above.

Metagenomic Assembly, Binning, and Annotation of Diverse DPO Cultures

Eleven DPO communities were sequenced for metagenomic analysis. Communities were grown in triplicate cultures amended with 5% rumen fluid with or without 10 mM phosphite (Fig. 1.1). DNA was extracted from the no-phosphite triplicates in stationary phase (-Ps), and the 10 mM phosphite triplicates in exponential phase (+Pe) and stationary phase (+Ps) (Fig. 2.1). Community R1 failed to reach stationary phase and was only represented by samples -Ps and +Pe. Communities LM1, R3, SL1, and SL3 failed to reproduce activity and were instead sampled from two previously active enrichments (E1 and E2) (Fig. 2.1). DNA was extracted using the DNeasy PowerLyzer Microbial Kit (Qiagen) and sequenced with an Illumina HiSeq 4000 (150 bp paired-end reads) at the UC Berkeley Vincent J. Coates Genomics Sequencing Laboratory. Reads were trimmed and filtered using Sickle v1.33 (quality threshold value of 20)[96]. Gene-level taxonomy was assigned using Centrifuge v1.0.1[101]. Reads for each of the 11 communities were combined and co-assembled using MEGAHIT v1.1.2[97] using the meta-sensitive preset. Reads were mapped to assembled contigs using BWA-MEM v0.7.17[98] with default parameters. Contigs over 1000 bp in length from each combined assembly were binned into individual genomes using Anvi'o v5.4.0[102]. Communities with < 30,000 contigs (LM3, M1, R1, SM1, SM3, SV1, SV3) were binned manually using patterns of hierarchical clustering, sequencing coverage, GC content, and gene-level taxonomic assignments. Communities with > 30,000 contigs (LM1, R3, SL1, SL3) were binned automatically using CONCOCT then manually refined with the Anvi'o graphical interface[103]. Quality of metagenome-assembled genomes (MAGs) was measured from lineage-specific, conserved, single-copy marker genes using the CheckM v1.0.18 lineage workflow[104]. The resulting 11 co-assemblies consisted of 1900 Mbp, 1.99 million contigs, and 574 draft genomes (Table 2.2). Only draft genomes of medium quality or greater (>50% completion; <10% redundant)[105] were subjected to further study, resulting in 239 metagenome-assembled genomes (MAGs) that represent 60% (647 Mbp) of the binned contigs (Table 2.2). Open reading frames were predicted from selected genomes using Prodigal v2.6.3[99] and contigs of interest were functionally annotated with Prokka v1.14.6[106]. Annotated proteins from all MAGs were searched for known DPO genes (*ptxDE-ptdCFGHI*) with my custom pHMMs (see **Chapter 3**). MAGs were operationally considered capable of DPO if they included at least one gene from the *ptx-ptd* gene cluster. The *ptx-ptd* genes that were absent from MAGs were searched for in all remaining contigs of the respective community.

Data Deposition

All metagenomic reads, assemblies, and curated metagenome-assembled genomes (MAGs; quality metrics >50% complete and <10% redundant) are available through the NCBI BioProject accession PRJNA655520.

Results

Selective Enrichment

DPOM were hypothesized to be cultivatable from wastewater sludge because phosphite can represent up to 2.27% of total dissolved wastewater phosphorus[43] and because both strains of Ca. *P. anaerolimi* were identified in wastewater digester sludge[70,71]. Accordingly, sludge from six San Francisco Bay area facilities was used to inoculate 30 enrichment cultures (Table 2.1). All cultures were grown in bicarbonate-buffered basal medium amended with 10 mM phosphite and multivariate exogenous electron acceptors (CO_2-only, CO_2+SO_4^{2-}, or CO_2+NO_3^-) (Table 2.1). The variable electron acceptors were chosen based on what FiPS-3 and Phox-21 were observed to utilize, and enrichments were exposed to these alternative electron acceptors to capture a broader range of metabolic diversity. Rumen fluid (5% by volume) was also added to all enrichments as a complex nutrient source to stimulate DPOM growth because DPO activity in the Phox-21 enrichment was significantly increased in the presence of rumen fluid[70].

Phosphite oxidation was observed in 26 of 30 enrichments and across all six wastewater facilities, where enrichments were considered active if 50% of the initial phosphite had been consumed (Fig. 2.2, Table 2.1). When stationary phase enrichments were re-spiked with phosphite, DPO activity resumed. No phosphite oxidation occurred in autoclaved controls (Fig. 2.2A). Based on prior experience[70], the high percentage of active DPO enrichments was unpredicted, indicating a greater prevalence of DPOM than previously assumed.

Preferential Growth Conditions

Salinity

Inocula were introduced to defined media of either 0 or 2% salinity. Active enrichments grown in saline media supported higher average phosphite oxidation rates (0.70 ± 0.17 mM HPO_3^{2-}/day) than those in non-saline media (0.44 ± 0.10 mM HPO_3^{2-}/day) (Fig. 2.2B). While this could indicate a metabolic preference for saline conditions, DPO was still possible in non-saline media. Similarly, FiPS-3 was shown to grow optimally in 2% salinity while Phox-21 was grown in 0% salinity. Consequently, it is unlikely that saline conditions are a universal necessity for DPO metabolism, and salinity preference may instead be indicative of the inoculum's environmental conditioning prior to enrichment. Typical raw municipal sewage has high concentrations of total dissolved solids (200-3000 mg/L)[107], contributing to high solute:solvent ratios that may contribute to higher salinity preferences. Furthermore, municipal sewage near coastal areas, such as the San Francisco Bay area, are frequently more saline due to the input of seawater into the watershed[108].

Electron Acceptor

DPO was also impacted by the amended electron-acceptor. Active enrichments with only CO_2 supplied as an electron acceptor supported the highest average phosphite oxidation rate (0.64 ± 0.17 mM HPO_3^{2-}/day for CO_2 versus 0.56 ± 0.10 and 0.50 ± 0.20 mM HPO_3^{2-}/day for NO_3^- and SO_4^{2-}, respectively). CO_2 also supported DPO from all six sample sites. Despite the availability of nitrate and sulfate, neither electron acceptor was definitively coupled to phosphite oxidation (Fig. 2.2B & C). While all amended cultures consumed nitrate, it was metabolized before phosphite oxidation was complete, suggesting its utilization was independent of DPO. In fact, when compared to other cultures with the same inoculum, nitrate delayed or even excluded DPO (Fig. 2.2C). Meanwhile, although sulfate was consistently consumed at the expected ratio if reduction to sulfide was coupled to phosphite oxidation (1 mol sulfate per 4 mols phosphite), the timing of sulfate consumption was variable and frequently offset from DPO (Fig. 2.2C). This suggests that sulfate reducers may be utilizing a reduced carbon substrate that is generated from DPO carbon metabolism. Consistent with this, both characterized DPOM either grow preferentially (FiPS-3) or exclusively (Phox-21) by autotrophy and utilize CO_2 as an electron acceptor. In the case of FiPS-3, the reduced carbon end-product is acetate[18], which is readily utilized by sulfate reducers. These results consequently support a DPOM preference for CO_2 and indicate that alternative electron acceptors (such as nitrate) may inhibit DPO activity (Fig. 2.2).

DPOM Identification

To identify and characterize DPOM, active CO_2-only enrichments were transferred to triplicate bottles that were monitored for activity and growth. Cells were temporally collected for DNA extractions, which were short read sequenced, assembled, and binned into metagenome-assembled genomes (MAGs) using a computational pipeline that was developed in-house. To identify candidate DPOM, I searched all MAGs for each of the seven *ptx-ptd* genes[67,69] using custom-built pHMMs (See Materials and Methods from **Chapter 3**). In total, 21 genomes had at least one gene from the *ptx-ptd* cluster (DPO MAGs), and of these 21 DPO MAGs, 19 were of high quality (>90% complete; <5% redundant) (Table 2.3)[105]. I found that DPO MAGs were specifically enriched in all phosphite-amended communities (compared to no-phosphite controls) (Fig. 2.3A). DPO MAGs were also the dominant community member in all but one enrichment (SL1) (Fig. 2.3B). Furthermore, every sequenced community had at least one DPO MAG (Fig. 2.3B). Collectively these results confirmed that 1) the phosphite-amended enrichments selectively promoted the growth of DPOM, 2) DPO activity in phosphite amended enrichments was dependent on the *ptx-ptd* genes, and 3) *ptx-ptd* genes serve as effective probes for DPOM. The metagenomic sequencing and identification of DPOM also supplied substantial metagenomic data which supports the phylogenetic, metabolic, and evolutionary analyses of later chapters.

Hyper-Enrichment of DPOM

The only DPOMs that have been isolated to date are *Desulfotignum phosphitoxidans* FiPS-3 and *Phosphitispora fastidiosa*[18,75]. The majority of DPOM that have been identified, either previously (*Phosphitivorax anaerolimi*) or in this study, are taxonomically diverse and remain within complex microbial communities. A novel DPOM isolate would consequently contribute substantially to the biochemical characterization of DPO metabolism. To this end, a subset of

active cultures was subjected to various isolation attempts. Eight of the most active enrichments (SM1, SM2, SM3, SM4, LM3, LM4, SV3 and VR3) were serially diluted and inoculated into agar plugs that contained the same media composition as their respective enrichment culture (Table 2.1), including 10mM phosphite, 1% rumen fluid, and 1% agar. Colonies were observed after approximately 6 weeks of incubation (Fig. 2.4). Despite the growth of colonies, it was not possible to successfully culture any subsequently picked colony in liquid media amended with phosphite (Fig. 2.5). The failure to culture DPOM from colonies motivated the hypothesis that colonies were not DPOM but instead were fermentative organisms growing on carbon substrates from the supplemental rumen fluid.

Rumen fluid was previously shown to stimulate DPO activity[70] and likely supported the high success-rate of my enrichment series (Fig. 2.2B). However, the chemical composition of rumen fluid is complex[109] and could support the growth of diverse microorganisms. It was hypothesized that the removal of rumen fluid might increase the selectivity of the growth media and assist in the isolation of DPOM. By transferring a subset of active enrichments into media without rumen fluid, three enrichments were identified (SM3, SM4, and SV3) as capable of robust DPO activity without rumen fluid (Fig. 2.6). Each of these enrichments was serially diluted to extinction in fresh media without rumen fluid. Sanger sequencing of the 16S rRNA gene from the extracted DNA of transferred cultures showed that the only sequence showing no signal of contamination in the sequencing chromatogram was from the highest SV3 dilution, and a BLAST alignment of the NCBI database identified its closest relatives to be *Phosphitivorax anaerolimi* Phox-21 (98.72% sequence identity) and other members of the GW-28 clade[70]. Phase contrast microscopy showed that SV3 cultures were dominated by non-motile rods that formed aggregates which were frequently associated with an unidentified matrix material, which complicated purity diagnoses (Fig. 2.7). However, inoculation of the serially diluted SV3 culture into rich media subsequently showed that contaminant organisms were still present, as determined via Sanger sequencing. Short-read (meta)genome sequencing (March, 2019) confirmed the presence of at least 16 organisms using the rpS3 single copy gene (SCG) marker (Table 2.4). Subsequent isolations were attempted using multiple iterations of agar plugs and dilution series, but none were successful in isolating DPOM from the SV3 enrichment.

This highly enriched phosphite-oxidizing (HEPO) culture continues to be maintained by regular batch feedings with phosphite and batch media transfers. Revival from cold storage of HEPO is unreliable, suggesting it is sensitive to cold treatment. Similarly, the Phox-21 enrichment has been inconsistently revived from freezer stocks (unpublished observations). Freezer stocks of the HEPO culture exist, but their viability is yet to be verified. The batch fed HEPO culture is consequently the source of all subsequent physiological experiments in this chapter and later chapters.

Physiological Characterization of HEPO

Phosphite

Phosphite is the sole electron donor required for DPO growth in the HEPO community. To test the range of phosphite in which HEPO can grow by DPO, phosphite was added to growth medium at concentrations between 0 and 50mM (Fig. 2.8). Growth yield increased with phosphite concentrations up to 10mM, after which additional phosphite did not support further net growth (Fig. 2.8A & E). This suggests that phosphite concentrations below 10mM are limiting, and

concentrations above 10mM are in excess for standard HEPO growth conditions. In phosphite-limiting conditions (< 10 mM), 0.018 ± 0.006 OD_{600} units were gained per 1mM phosphite. When compared to *Escherichia coli,* a single unit of OD_{600} is equivalent to 0.3 – 0.5 g/L of cells[110]. Using these values, we can estimate that each 1 mole of phosphite (which should generate 1 mol of ATP; see **Chapter 4**) generates 5.4 – 9 g of cells. This is energetically similar to the 10.3 g of cells per mole of ATP that are generated when *E. coli* ferment glucose anaerobically[111].

When phosphite is present in excess (> 10 mM), net cell yield no longer correlates with increasing phosphite concentrations and instead reaches a threshold of 0.156 ± 0.008 OD_{600} units (Fig. 2.8E). Excess phosphite (> 10 mM up to 30 mM) also maintains cells in stationary phase, suspending the onset of death phase that is otherwise observed in phosphite-limiting conditions (< 10 mM) (Fig. 2.8A). This suggests that the growth media is either limiting or inhibiting cell replication, but that DPO metabolic activity may continue at a basal level (Fig. 2.8C). Concentrations of phosphite that are greater than 40 mM appear to inhibit HEPO growth by DPO (Fig. 2.8A & E).

The growth rate of HEPO mirrored the observations for net cell yield, where growth rate increased with phosphite up to 10mM, after which the growth rate remained steady (Fig. 2.8B). However, the rate of phosphite oxidation increased linearly with phosphite concentrations, even in excess phosphite where growth rate was limited (Fig. 2.8D). This supports the hypothesis that growth limitations are due to cellular growth needs and not due to the DPO protein machinery.

It was hypothesized that the growth limitations of HEPO in excess phosphite could be due to either: 1) phosphate accumulation and inhibition, 2) sodium ion toxicity from excess Na_2HPO_3, 3) the accumulation of carbon fixation end-products, or 4) nutrient limitation. The following sections present the results for experiments that test hypotheses 1 and hypothesis 2. Hypothesis 3 is explored in **Chapter 4**, and hypothesis 4 is explored in **Chapter 5**.

Hypothesis 1: Phosphate accumulation inhibits HEPO growth.

Hypothesis 1 suggests that growth limitations in excess phosphite may be due to the accumulation of phosphate from DPO activity. To test the response of HEPO to accumulated phosphate, phosphate was added to growth medium at concentrations between 0 and 50 mM (Fig. 2.9), which represents the range of phosphate that would stoichiometrically accumulate in the growth media with excess phosphite (Fig. 2.8). Phosphate concentrations up to 50 mM did not affect HEPO growth nor DPO activity, as the net growth yield (Fig. 2.9A) and net phosphite oxidized (Fig. 2.9B) remained relatively consistent despite increasing concentrations of phosphate. This suggests that growth limitation in excess phosphite is not due to phosphate inhibition.

Hypothesis 1b: Phosphate accumulation precipitates ammonium.

To supplement hypothesis 1, I predicted that phosphate accumulation from DPO activity might result in the precipitation and removal of ammonium as struvite ($NH_4MgPO_4 \bullet 6H_2O$), as was observed for FiPS-3[18]. The oxidation of 10 mM phosphite in standard growth media results in the accumulation of 10 mM phosphate (Fig. 2.2A), which could maximally precipitate 10 mM of ammonium from the media. The media used for HEPO contains 18.7 mM of ammonium chloride. As such, the dissimilatory oxidation of excess phosphite (> 18.7 mM) might entirely remove

ammonium from the growth media, resulting in ammonium limitation for HEPO. To test this, the HEPO culture was grown in excess phosphite with increasing amounts of ammonium (Fig. 2.10). As was previously shown, 30 mM phosphite was in excess, as 30 mM phosphite could not be oxidized to completion (Fig. 2.10B) and 30 mM phosphite did not increase growth yields (0.289 ± 0.002 OD_{600} units) as compared to 20 mM phosphite controls (0.278 ± 0.022) (Fig. 2.10A). The addition of increasing concentrations of ammonium did not increase net growth yields of the HEPO culture (0.3 ± 0.01 OD_{600} units) as compared to the control (0.289 ± 0.002 OD_{600} units) (Fig. 2.10A). Increasing concentrations of ammonium also did not increase net phosphite oxidation (71 ± 20 % initial phosphite) as compared to the control (78 ± 20 % initial phosphite) (Fig. 2.10B). This was consistent with the observations of the HEPO culture's activity in excess phosphate (Fig. 2.9), reinforcing the suggestion that neither phosphate accumulation nor ammonium depletion are limiting growth of the HEPO culture by DPO when excess phosphite is present.

Hypothesis 2: HEPO growth is inhibited by increasing salinity.

Growth and phosphite oxidation by the HEPO community were temporally monitored across low (0-3%) and high (3-10%) salinities (Fig. 2.11). DPO growth rate increased linearly (GR= 0.0031(S) + 0.0006) with salinity until 1.5% (0.005 ± 0.002 OD_{600}/day) (Fig. 2.11E), where 'GR' represents growth rate, and 'S' represents percent salinity. When salinity was increased to 2%, the growth rate increased significantly (0.017 ± 0.001 OD_{600}/day), deviating from the linear trend observed between 0 and 1.5% (Fig. 2.11E). This increased growth rate remained consistent (0.016 ± 0.003 OD_{600}/day) as salinity was increased beyond 2% (Fig. 2.11E). Similarly, net cell yields increased with salinity between 0% (0.004 ± 0.005 OD_{600} units) and 2% (0.153 ± 0.033 OD_{600} units), but cell yields at higher salinities were not correlated with salinity (Fig. 2.11G), suggesting ~2% salinity represents a threshold for maximal cell replication in these growth conditions (Fig. 2.11A & B). Salinities above 3% delayed HEPO growth and activity, but after an extended lag phase, phosphite was completely oxidized with cell yields (0.156 ± 0.02 OD_{600} units) equivalent to the 2% threshold (0.153 ± 0.033 OD_{600} units) (Fig. 2.11B & D), suggesting HEPO is capable of acclimating to high salinity after an adjustment period. Net phosphite consumption was linearly correlated with net cell yield (Fig. 2.11G), yet cells in 0% salinity were capable of oxidizing 3.9 mM of phosphite despite their relative lack of growth. Meanwhile, the rate of phosphite oxidation during exponential growth (0.88 ± 0.187 mM phosphite consumed) remained consistent between 0-3% salinities, despite differences in growth rate and absolute cell yield (Fig. 2.11F). The cumulative effects of salinity on growth rate and net cell yields suggest that DPOM may require a threshold concentration of NaCl (2% = 342 mM) for cell replication, but that DPO metabolism likely acts independently of NaCl since the rate of phosphite oxidation remains consistent across the range of salinities tested. Furthermore, it is unlikely that elevated sodium concentrations from the addition of excess phosphite (Na_2HPO_3) are the reason for growth limitations beyond 10 mM phosphite, because a maximum of 50 mM phosphite was added to HEPO, which would equate to ~ 392mM Na^+ in the growth medium (Fig. 2.8). Here, HEPO grew optimally in up to 3% salinity with 10mM phosphite, which is equivalent to 523mM Na^+ (Fig. 2.11).

Temperature

In addition to characterizing the nature of growth limitation in the HEPO culture, growth and phosphite oxidation by the HEPO culture were temporally monitored at incubation temperatures of 30, 37, 40, 50, and 60 °C (Fig. 2.12). The optimal temperature for growth of the HEPO culture was at 37 °C, with a growth rate of 0.013 ± 0.004 OD_{600}/day and phosphite oxidation rate of 0.79 ± 0.1 mM/day (Fig. 2.12). The growth and phosphite oxidation rates at 40 °C were within standard deviation of 37 °C. At the temperatures tested below 37 °C and above 40 °C, growth and phosphite oxidation activity were reduced nearly 5-fold, leaving a narrow temperature range for HEPO cultivation. The HEPO culture originated from wastewater digester sludge maintained at 37 °C, which likely explains optimized activity at this temperature (Table 2.1; see *SV3*).

Discussion

Cultivation-based investigations coupled to high-resolution metagenomics were used to enrich and identify DPOM. This enrichment strategy increased the database of known DPOM ten-fold (from 2 to 23 genomes), with an additional DPOM (*P. fastidiosa*) identified during the writing of this dissertation. The resulting metagenomic data is crucial for the downstream analyses presented in **Chapters 3, 4, and 5.** Furthermore, this successful enrichment campaign demonstrates that DPOM are more prevalent than previously recognized. This appears to be particularly true for wastewater treatment facilities, as all facilities sampled in this study harbored at least one DPOM. The previously identified strains of *Phosphitivorax* were also identified in wastewater digesters[70,71]. Wastewater may select for DPOM, but it is unclear whether their prevalence is due to endogenous phosphite production, phosphite influx, or the growth of DPOM using a non-DPO metabolism. Given the necessity for phosphorus removal in wastewater systems, DPOM may have implications on the conversion of highly-soluble phosphite (1000x more soluble than phosphate) to phosphate, whose removal is dependent on low-solubility and precipitation[112].

The biogeochemical parameters of wastewater digester sludge may also provide guidance for future sampling campaigns, in which DPOM might be most prevalent in mesophilic, anoxic, microbially rich environments. However, it remains unclear as to which geochemical parameters support DPOM, and efforts to identify pristine ecosystems that support DPOM would be informed by comprehensive biogeochemical analyses that describe P speciation while monitoring DPOM-specific biomarkers and DPO activity. In **Chapter 3**, pHMMs from this study are used as such biomarkers to probe genomic data for the *ptx-ptd* genes of environmental metagenomic data. In the future, these pHMMs may also be useful to predict the metabolic potential for DPO of new environmental samples prior to their cultivation. Meanwhile, continuously improving measurement techniques will continue to support the identification of reduced P species in diverse environments[42,47,113–115].

The high success rate of DPOM cultivation in this study may have been due to the addition of rumen fluid across all primary enrichments. Figueroa et al. first identified the stimulatory effects of rumen fluid on the growth of Phox-21[70]. Rumen fluid contains a complex mixture of various cofactors and nutrients, and it remains unknown which substrate(s) stimulate DPOM. The identification of this substrate would provide substantial insight into the metabolic requirements of DPOM. In this chapter, I identified three DPO communities that could grow via phosphite oxidation without rumen fluid supplementation. The DPOM in these cultures either utilized

alternative metabolic machinery which negated the requirement for rumen fluid, or alternatively established a symbiotic relationship with community members that could provide the substrate biogenically. I had hypothesized that the removal of rumen fluid from the culture media would ease isolation, but subsequent attempts remained unsuccessful. This could be explained by a dependence on a symbiotic partner, a concept that is further explored in **Chapter 5**.

While isolation of DPOM remained unsuccessful, a hyper-enriched phosphite oxidizing (HEPO) culture with robust growth in the absence of rumen fluid was established. The cultivation of HEPO has allowed for physiology experiments in which activity can be largely attributed to DPOM, and the HEPO culture serves as the model system for physiology experiments in **Chapters 4 and 5**. In this chapter, the HEPO culture is optimized for growth on 10 mM phosphite, and concentrations of phosphite above 10 mM are in excess, as net cell yields fail to increase with the addition of > 10 mM phosphite. This observed growth limitation was hypothesized to be due phosphate accumulation, sodium ion toxicity (from Na_2HPO_3), the accumulation of carbon fixation end-products, or nutrient limitation. However, data provided here demonstrates that phosphate accumulation does not inhibit HEPO growth or activity, and that HEPO is capable of withstanding sodium concentrations far higher than what would have been introduced to from the range of tested phosphite concentrations. As such, nutrient limitation or the accumulation of carbon fixation end-products are the next most plausible explanations for growth limitation. It is possible that NaCl represents a limiting nutrient, as cell yields from HEPO grown on 10 mM phosphite are maximized between 1.5 and 2% NaCl. Therefore, increasing concentrations of phosphite may require additional NaCl to support cell replication. However, the mechanism of this growth benefit is unclear, and it appears to be a requirement that is distinct from the DPO metabolic machinery, since the rate of DPO and net phosphite consumption in HEPO are not correlated with salinity. Future studies will be needed to understand the mechanistic role of NaCl, and to determine if it is the sole nutrient that is limiting DPOM growth in excess phosphite.

In the HEPO culture, phosphite is the sole electron donor and CO_2 is the sole exogenous electron acceptor. CO_2 must therefore be reduced by the HEPO community, as electrons from phosphite oxidation must be transferred to an electron acceptor to conserve energy. CO_2 may even be the preferred electron acceptor for most DPOM, as alternative electron acceptors (i.e., SO_4^{2-}, NO_3^-) were not used and were potentially inhibitory for DPO activity across enrichments. If CO_2 is used as a terminal electron acceptor, we would expect a reduced carbon end-product to be secreted by DPOM. In FiPS-3, acetate is the end-product of CO_2 reduction via the Wood-Ljungdahl pathway[18]. However, Phox-21 does not have all the genes for a complete WLP and is instead predicted to fix CO_2 via the recently validated reductive glycine pathway[70,74]. Yet in the Phox-21 model for energy conservation, the reductive glycine pathway is attributed to carbon fixation for biomass accumulation and the end-product of dissimilatory CO_2 reduction remains to be identified. Likewise, the identity of the CO_2-reduction product of the HEPO culture remains unknown, and studies in **Chapter 4** are dedicated to its identification. The identification of this reduced product will be crucial to 1) calculate the free energy gained by DPOM in the HEPO culture via phosphite oxidation; 2) test the hypothesized role of DPO inhibition by a reduced carbon product; and 3) describe the community structure of the HEPO culture, since the CO_2 fixed by DPOM would provide primary carbon substrates for the remaining microbial community members. Identification of the fixed carbon product will be an important step towards biochemically describing the mechanism of energy conservation in a DPOM other than FiPS-3.

Figures

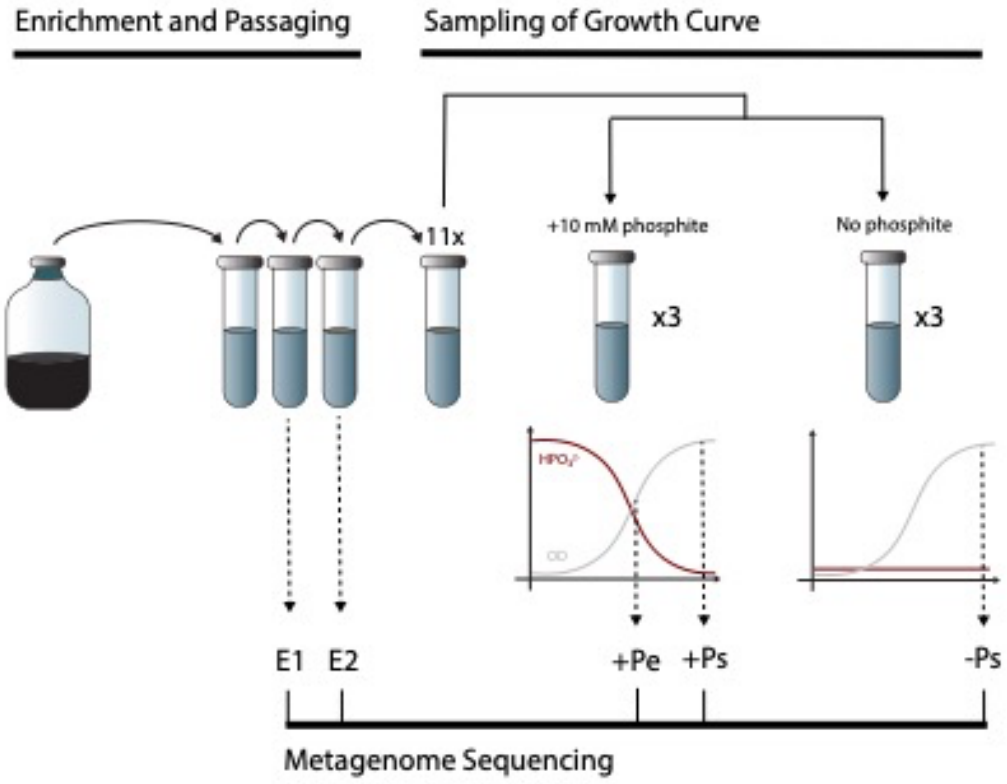


Figure 2.1: Metagenome sequencing workflow. *Enrichment and Passaging:* Inocula from active enrichments were transferred to fresh media upon reaching stationary phase. Mature cultures were then stored at 4 °C. *Sampling of Growth Curve:* Eleven communities that were active with only CO_2 as an electron acceptor were selected for metagenome sequencing. Selected communities were used to inoculate triplicate cultures that were amended with 5% rumen fluid and either 10 mM phosphite or no phosphite. The phosphite concentration of cultures was monitored with IC, and cell growth was monitored by measuring optical density (OD_{600}). *Metagenome Sequencing:* DNA samples were collected from 10 mM phosphite-cultures at exponential phase (+Pe) and early stationary phase (+Ps), and from no phosphite-cultures at early stationary phase (-Ps). The DNA from triplicates was pooled and submitted for short-read metagenome sequencing, so that each community was represented by three DNA samples (-Ps, +Pe, and +Ps). Community R1 did not reach stationary phase, leaving only samples for -Ps and +Pe. Four communities failed to show DPO activity during sampling of the growth curve (LM1, R3, SL1, SL3). To characterize these communities, DNA samples were obtained from the most recent, active enrichments from previous passages (E1 and E2).

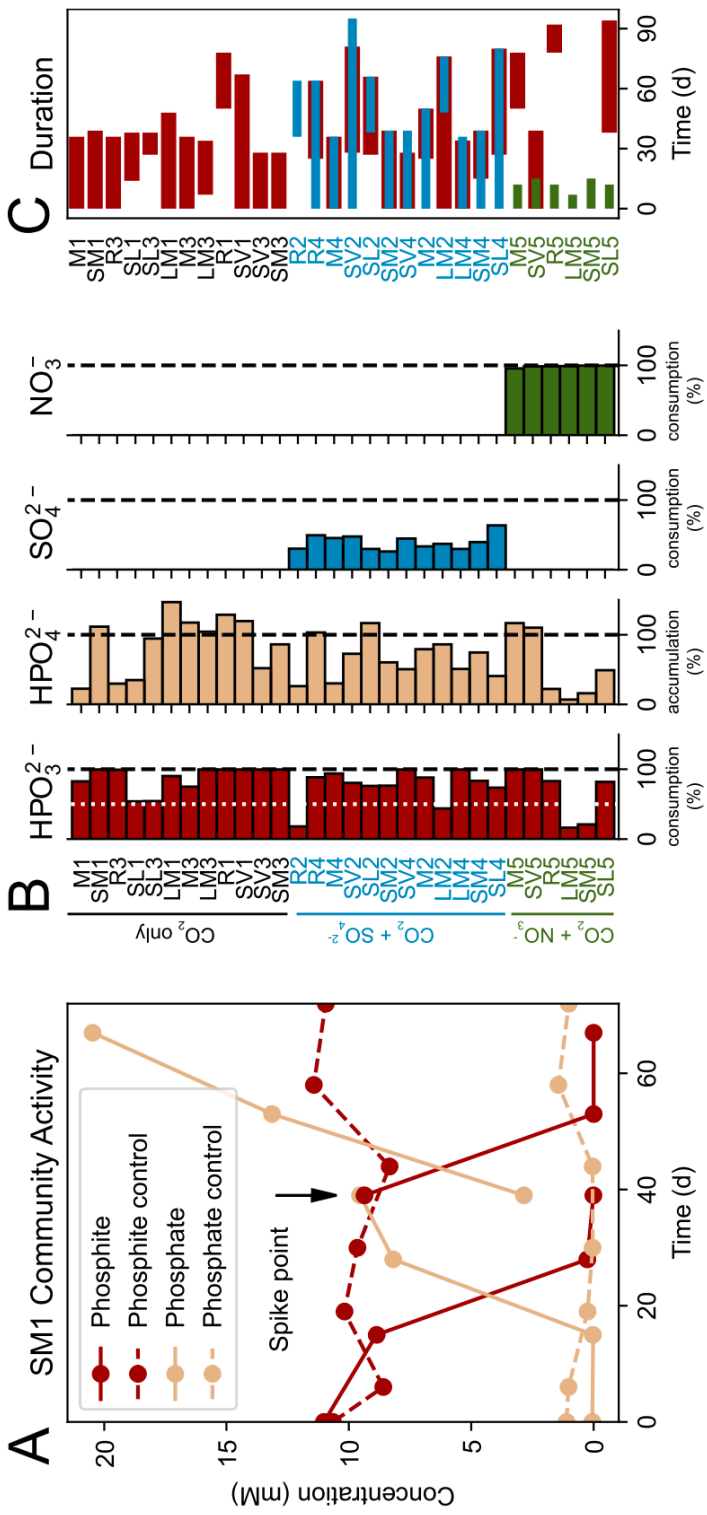


Figure 2.2: DPO enrichment activity. (A) Representative phosphite oxidation by the SM1 community. Temporal ion concentrations are shown for live (solid lines) or autoclaved (dashed lines) inoculum. Enrichments were amended with 10 mM phosphite at the spike point. (B) Percent change of measured ions for each enrichment community. Each row represents one community; each column displays the percent-accumulation or consumption of each titled ion. Row labels are colored according to the added electron acceptor (black, CO_2 only; blue, CO_2+SO_4^{2-}; green, CO_2+NO_3^-). A white dotted line denotes 50% consumption of HPO_3^{2-}. All percentages were calculated from concentration values *prior* to the first spike point. (C) Duration of ion depletion. Horizontal bars show the time frame for metabolic activity of each measured ion. Colors correspond with panel B (red, HPO_3^{2-}; blue, SO_4^{2-}; green, NO_3^-).

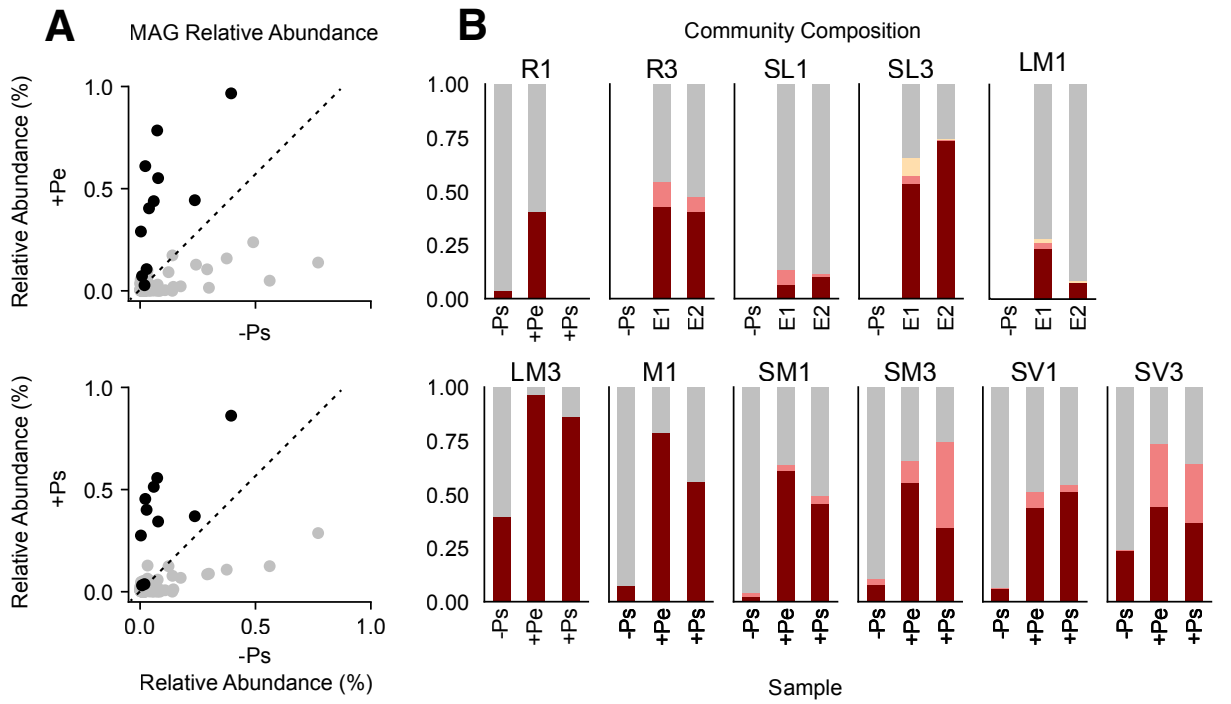


Figure 2.3: Relative abundance of DPO MAGs. (A) Relative abundance of MAGs across samples. Each point represents one MAG. Color represents the presence (black) or absence (grey) of any *ptx-ptd* genes. Top panel compares samples from phosphite-amended exponential phase (+Pe) to no-phosphite (-Ps) controls. Bottom panel compares samples from phosphite-amended stationary phase (+Ps) to no-phosphite (-Ps) controls. (B) Relative abundance of MAGs across time. Each subplot represents one community, while each stacked bar represents the community composition of one sample. Colors indicate the dominant (maroon), second dominant (pink) and third dominant (yellow) DPO members, and all remaining community members (grey). Relative abundance was calculated by dividing the mean coverage of a single MAG by the sum of mean coverages for all MAGs in the respective sample.

Figure 2.4: Colony formation in agar plugs. DPOM enrichment cultures formed colonies when incubated in agar media at 37 °C. This representative image shows that colonies were typically small, white, and disk-shaped.

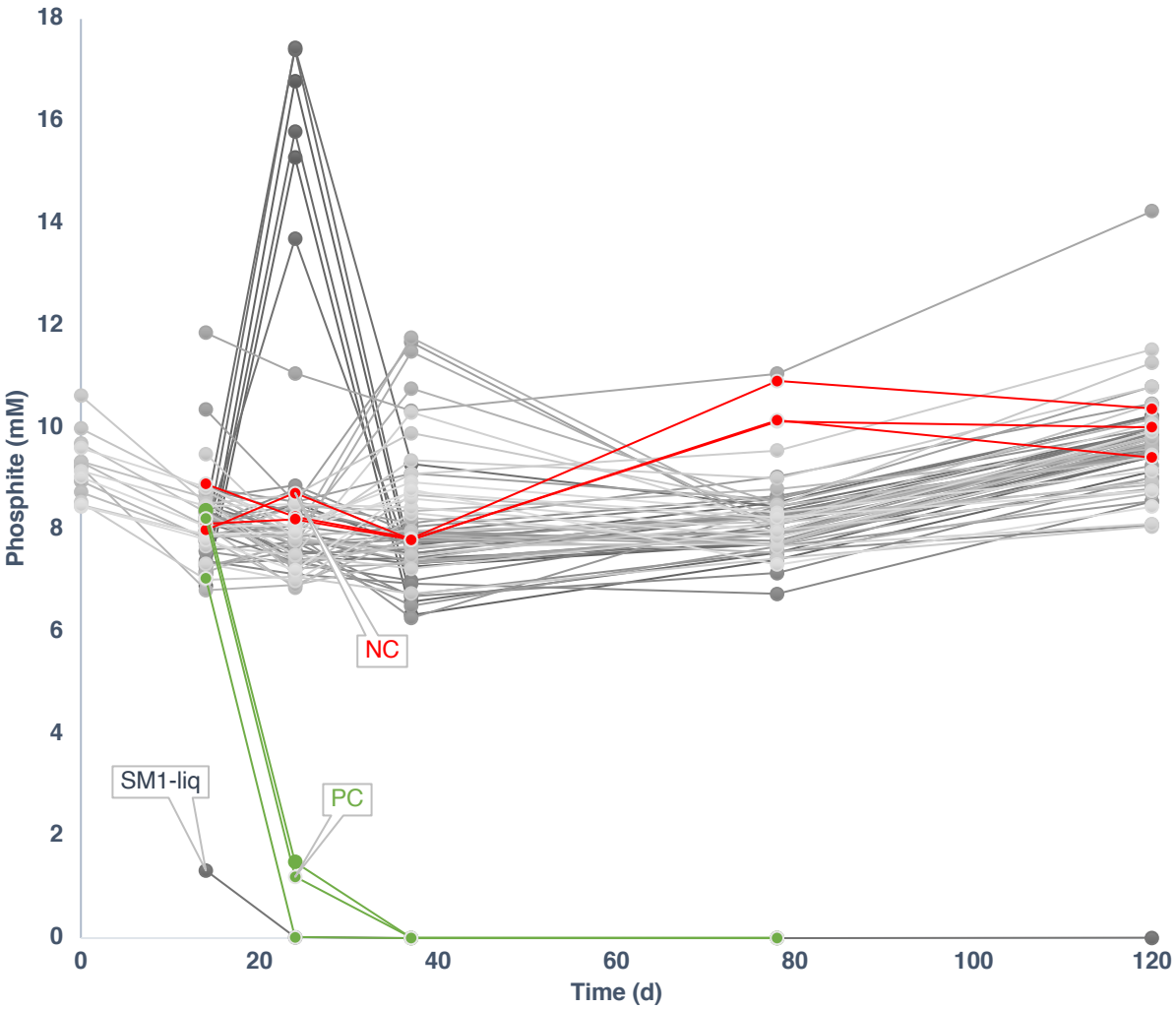


Figure 2.5: DPOM activity of colonies transferred to liquid media. Temporal phosphite concentrations are plotted for liquid media that was inoculated with 60 colonies picked from agar plugs (grey lines). A triplicate negative control (NC) shows temporal phosphite concentrations for a sterile inoculation of liquid media (red line). A triplicate positive control (PC) shows temporal phosphite consumption in liquid media that was inoculated with an active enrichment culture. SM1-liq shows temporal phosphite consumption in liquid media that was inoculated with the original inoculum for a subset (n=8) of the agar plugs.

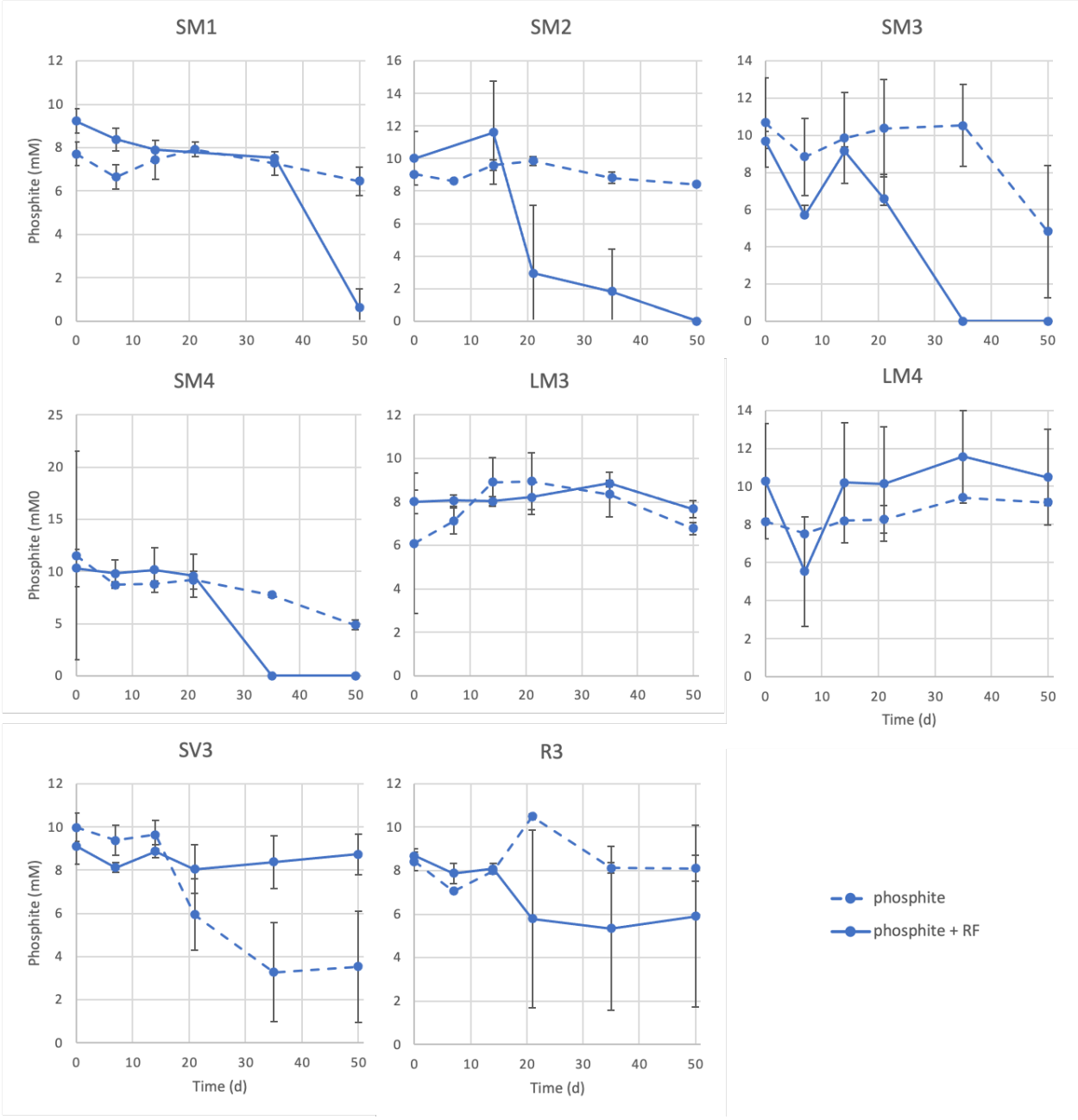


Figure 2.6: Necessity of rumen fluid across DPO enrichments. Temporal phosphite oxidation of DPO enrichments grown with 10 mM phosphite in the presence of 1% rumen fluid (solid line) and in the absence of rumen fluid (dashed line).

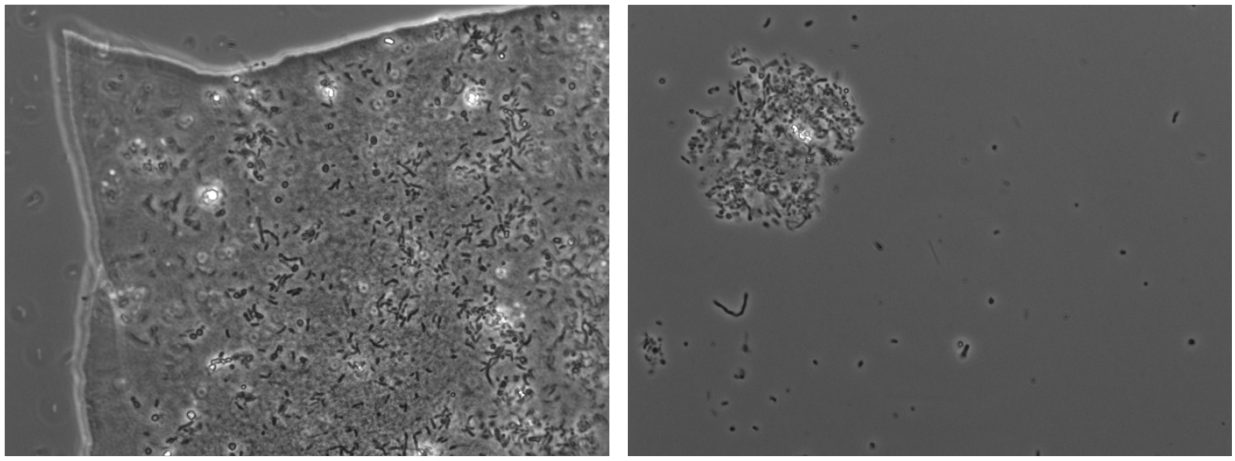


Figure 2.7: Phase-contrast microscopy of SV3 cultures. Wet-mounts of serially diluted SV3 cultures were visualized at 100x via phase-contrast microscopy.

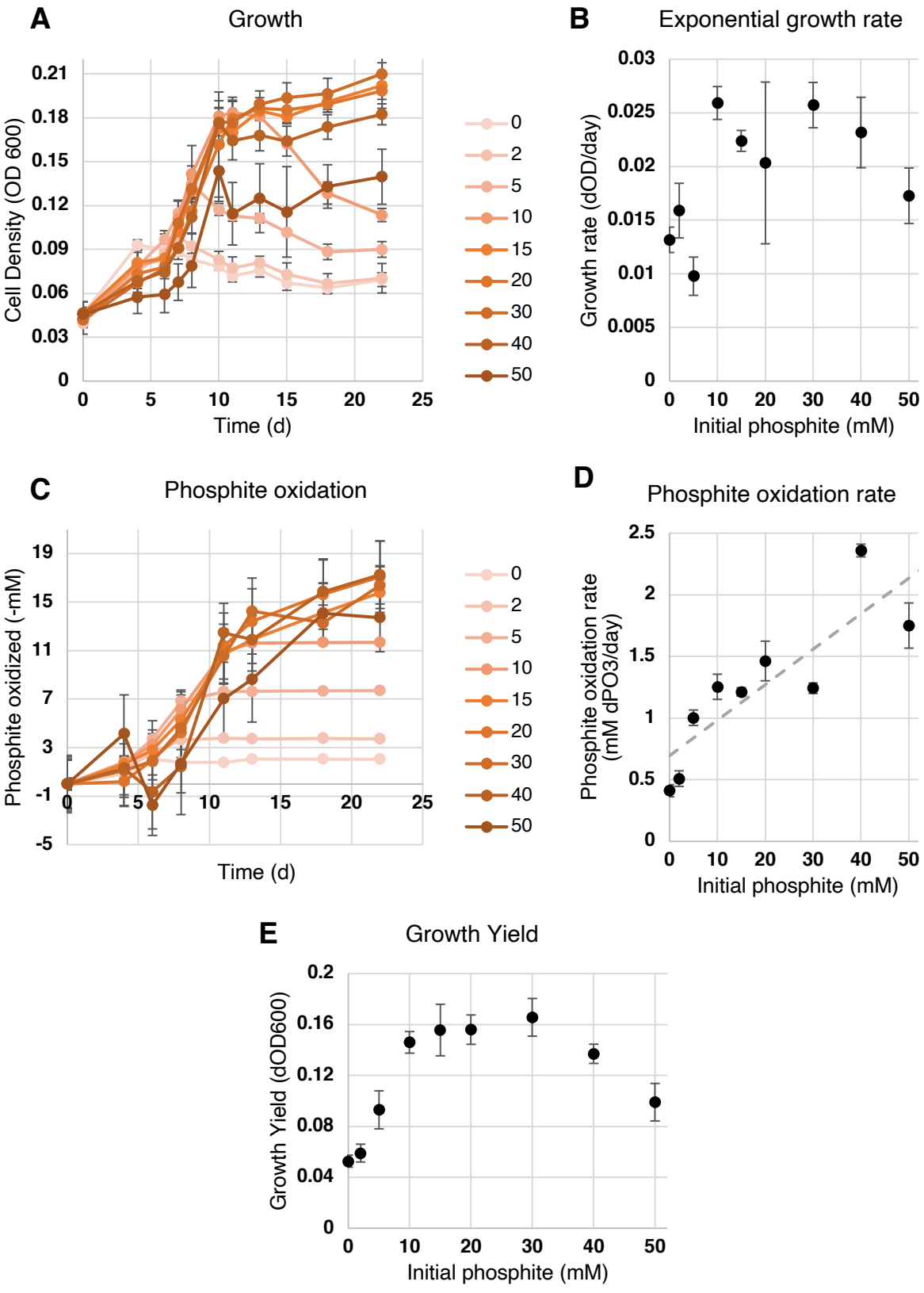

A
Growth
Cell Density (OD 600)
Time (d)
0
2
5
10
15
20
30
40
50
B
Exponential growth rate
Growth rate (dOD/day)
Initial phosphite (mM)
C
Phosphite oxidation
Phosphite oxidized (-mM)
Time (d)
D
Phosphite oxidation rate
Phosphite oxidation rate (mM dPO3/day)
Initial phosphite (mM)
E
Growth Yield
Growth Yield (dOD600)
Initial phosphite (mM)

Figure 2.8: HEPO growth in variable concentrations of electron donor (phosphite). HEPO was inoculated into triplicate tubes with media containing 0 to 50 mM phosphite and monitored for (A) growth and (B) phosphite utilization. The (C) growth rate and (D) phosphite oxidation rate are plotted, with linear best fits shown as a grey dashed line.

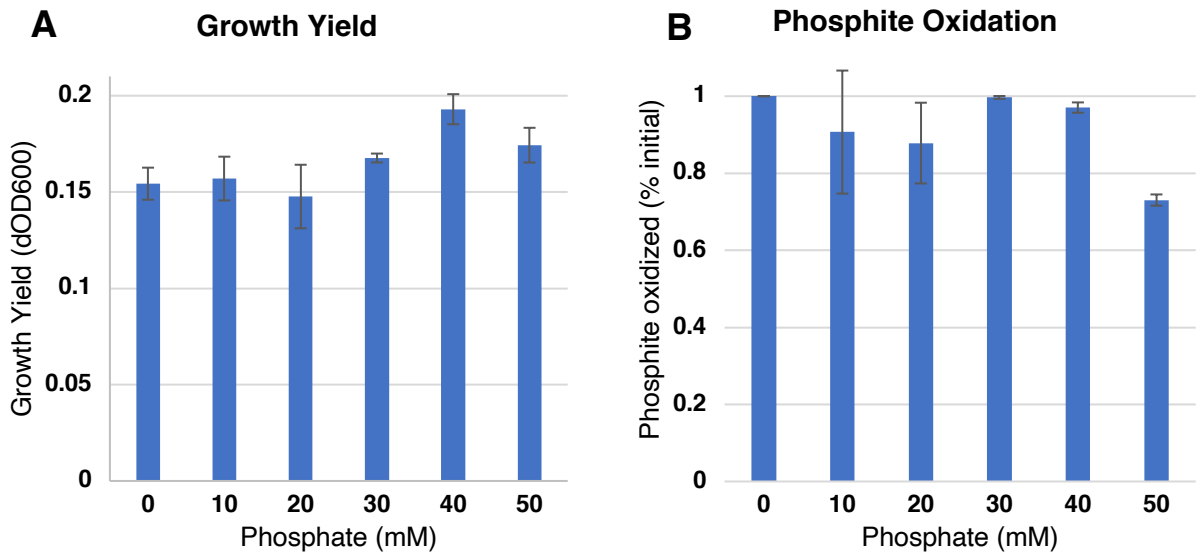


Figure 2.9: HEPO growth with variable concentrations of exogenous phosphate. The HEPO culture was inoculated into triplicate tubes with media containing10 mM of electron donor (phosphite) and 0 to 50 mM of phosphate and monitored for (A) net growth and (B) net phosphite utilization.

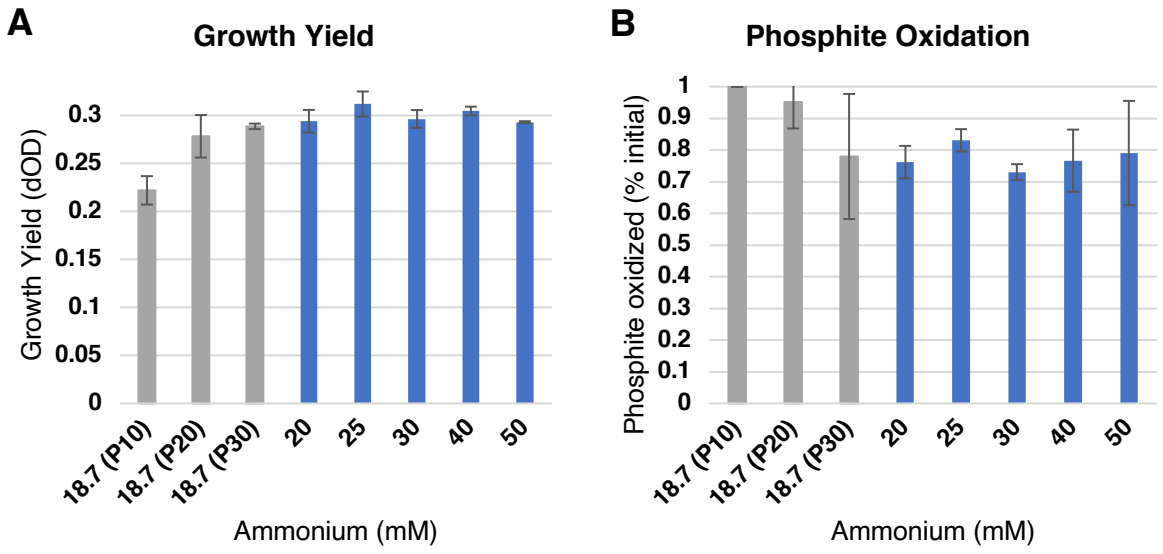
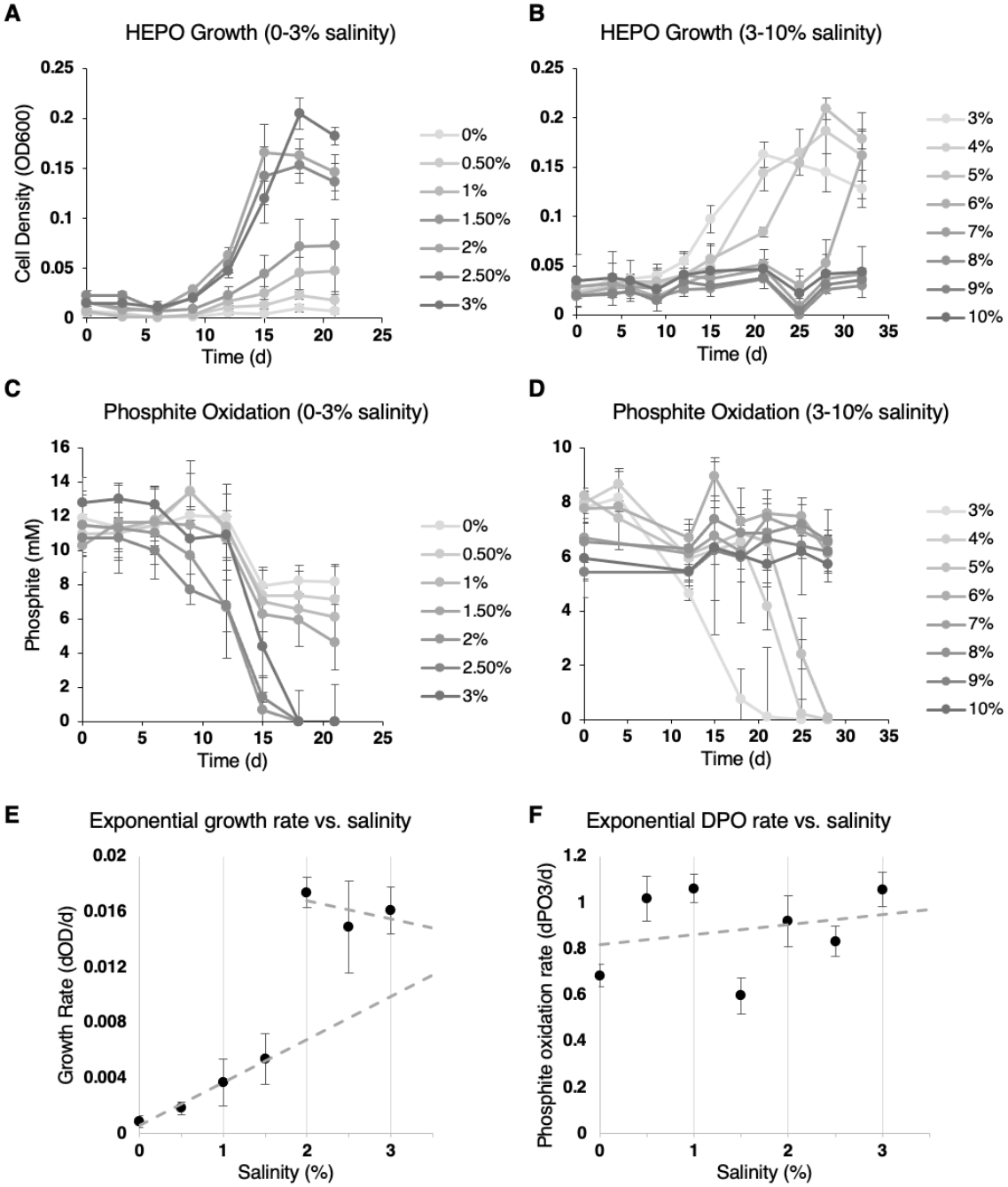


Figure 2.10: HEPO growth with variable concentrations of ammonium. The HEPO culture was inoculated into triplicate tubes with excess phosphite and variable concentrations of ammonium. (A) Net growth yield and (B) net phosphite oxidation are plotted for each condition. Grey bars signify conditions in which the HEPO culture was grown in basal bicarbonate media with standard ammonium concentration (18.7 mM) and phosphite concentrations of 10 mM (P10), 20 mM (P20), or 30 mM (P30). Blue bars signify conditions in which the HEPO culture was growth in basal bicarbonate media with excess phosphite (30mM) and variable ammonium concentrations ranging from 20 mM to 50 mM.


A
HEPO Growth (0-3% salinity)
Cell Density (OD600)
Time (d)
0%
0.50%
1%
1.50%
2%
2.50%
3%
B
HEPO Growth (3-10% salinity)
Time (d)
3%
4%
5%
6%
7%
8%
9%
10%
C
Phosphite Oxidation (0-3% salinity)
Phosphite (mM)
Time (d)
0%
0.50%
1%
1.50%
2%
2.50%
3%
D
Phosphite Oxidation (3-10% salinity)
Time (d)
3%
4%
5%
6%
7%
8%
9%
10%
E
Exponential growth rate vs. salinity
Growth Rate (dOD/d)
Salinity (%)
F
Exponential DPO rate vs. salinity
Phosphite oxidation rate (dPO3/d)
Salinity (%)

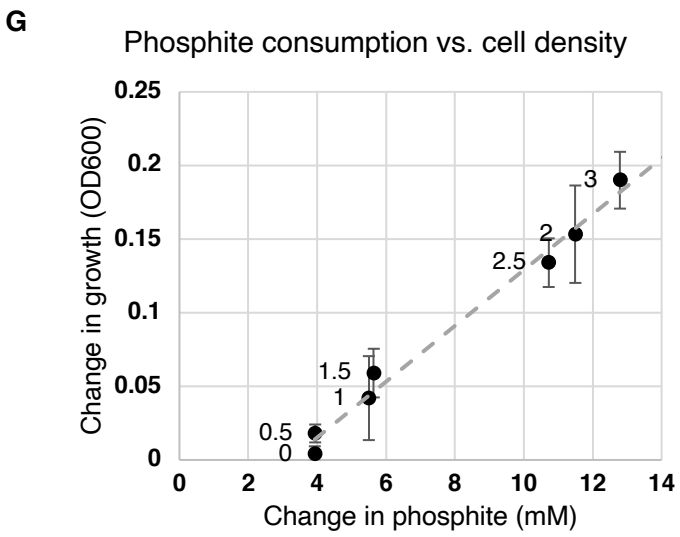


Figure 2.11: HEPO growth in variable salinity. HEPO was inoculated into triplicate tubes with media containing 10 mM phosphite and variable salinities that ranged from (A & C) 0-3% or (B & D) 3-10%. For each experimental range of salinities, (A & B) growth and (C & D) phosphite oxidation were temporally monitored. For the 0-3% experimental range, (E) the exponential growth rates, (F) the exponential phosphite oxidation rates, and (G) the change in phosphite vs. cell density were plotted for each salinity value. Salinity is listed next to each point in plot G. The grey dotted lines represent the linear best fit for subsets of the data.

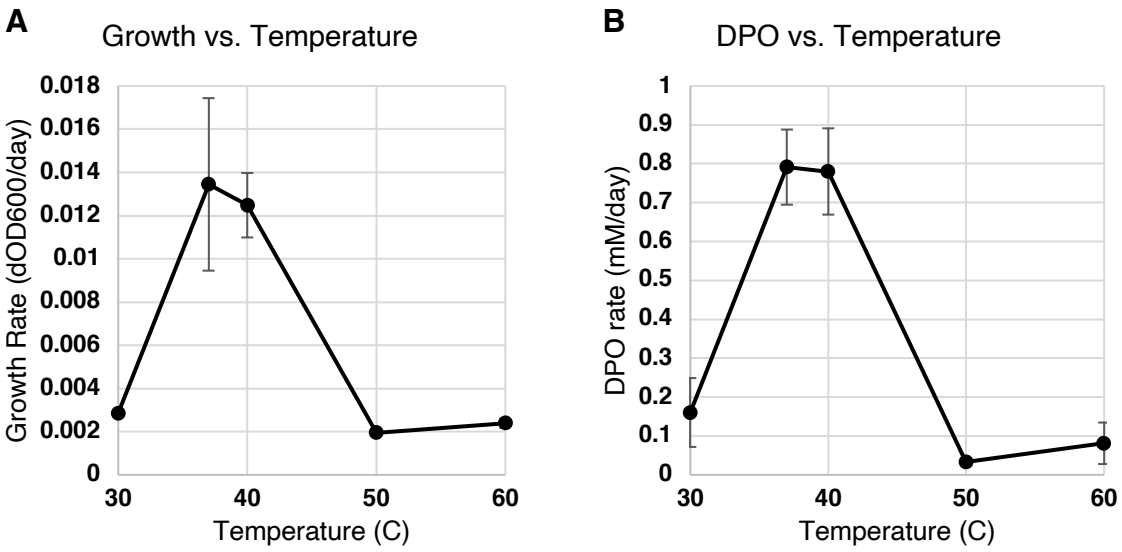


Figure 2.12: Temperature dependence of the HEPO culture. HEPO was inoculated into triplicate tubes with media containing 10mM phosphite and incubated at temperatures ranging from 30 °C to 60 °C. (A) Growth rate and (B) phosphite oxidation rate are shown for each incubation temperature.

Table 2.1: **Summary of conditions for each enrichment community.** 'Community name' lists the abbreviations used throughout this dissertation for each enrichment, which represent the inoculum source (letters) and media composition (numbers). Subsequent columns describe the conditions for each enrichment culture. CO2, carbon dioxide; SO4, sulfate; NO3, nitrate.

Community Name	Inoculum Source	Electron Acceptor	Salinity (%)	Phosphite (mM)	Rumen Fluid (%)	Temperature (C)	pH
M1	Sewerage Agency of Southern Marin, Digester 1	CO_2 only	0	10	5	37	7
M2	Sewerage Agency of Southern Marin, Digester 1	CO2 + SO4	0	10	5	37	7
M3	Sewerage Agency of Southern Marin, Digester 1	CO2 only	2	10	5	37	7
M4	Sewerage Agency of Southern Marin, Digester 1	CO2 + SO4	2	10	5	37	7
M5	Sewerage Agency of Southern Marin, Digester 1	CO2 + NO3	0	10	5	37	7
R1	Veolia Water North America, Richmond, Digester 2	CO2 only	0	10	5	37	7
R2	Veolia Water North America, Richmond, Digester 2	CO2 + SO4	0	10	5	37	7
R3	Veolia Water North America, Richmond, Digester 2	CO2 only	2	10	5	37	7
R4	Veolia Water North America, Richmond, Digester 2	CO2 + SO4	2	10	5	37	7
R5	Veolia Water North America, Richmond, Digester 2	CO2 + NO3	0	10	5	37	7
SL1	San Leandro Water Treatment, Digester 1	CO2 only	0	10	5	37	7
SL2	San Leandro Water Treatment, Digester 1	CO2 + SO4	0	10	5	37	7
SL3	San Leandro Water Treatment, Digester 1	CO2 only	2	10	5	37	7
SL4	San Leandro Water Treatment, Digester 1	CO2 + SO4	2	10	5	37	7
SL5	San Leandro Water Treatment, Digester 1	CO2 + NO3	0	10	5	37	7
LM1	City of Livermore Water Resources Department, Digester 2	CO2 only	0	10	5	37	7
LM2	City of Livermore Water Resources Department, Digester 2	CO2 + SO4	0	10	5	37	7
LM3	City of Livermore Water Resources Department, Digester 2	CO2 only	2	10	5	37	7
LM4	City of Livermore Water Resources Department, Digester 2	CO2 + SO4	2	10	5	37	7
LM5	City of Livermore Water Resources Department, Digester 2	CO2 + NO3	0	10	5	37	7
SM1	San Mateo Wastewater Treatment Plant, Digester 1	CO2 only	0	10	5	37	7
SM2	San Mateo Wastewater Treatment Plant, Digester 1	CO2 + SO4	0	10	5	37	7
SM3	San Mateo Wastewater Treatment Plant, Digester 1	CO2 only	2	10	5	37	7
SM4	San Mateo Wastewater Treatment Plant, Digester 1	CO2 + SO4	2	10	5	37	7
SM5	San Mateo Wastewater Treatment Plant, Digester 1	CO2 + NO3	0	10	5	37	7
SV1	Silicon Valley Clean Water, Digester 3	CO2 only	0	10	5	37	7
SV2	Silicon Valley Clean Water, Digester 3	CO2 + SO4	0	10	5	37	7
SV3→ HEPO	Silicon Valley Clean Water, Digester 3	CO2 only	2	10	5	37	7
SV4	Silicon Valley Clean Water, Digester 3	CO2 + SO4	2	10	5	37	7
SV5	Silicon Valley Clean Water, Digester 3	CO2 + NO3	0	10	5	37	7

Table 2.2: Assembly statistics for each metagenome-sequenced community.

Each community was assembled using MEGAHIT, and the resulting contigs were structurally annotated with Prodigal, and functionally annotated within Anvi'o using pHMMs from multiple published SCG collections. Rows 2-9 list standard statistics for each metagenome, where NXX is defined as the length N of the contig at which XX% of all bases in the total sequence have been assembled into contigs of length <N. Rows 10-13 identify the number of genes from each of the SCG collections. Rows 14-17 predict the number of archaea, bacteria, and eukarya in each community based on SCGs. Rows 18-21 summarize binning statistics, where the Binned Length lists the number of base pairs that were included in the binning step. bp, base pair; Kbp, kilobase pair; Mbp, megabase pair.

Community	SV3	SV1	SM3	SM1	SL3	SL1	R3	R1	M1	LM3	LM1
Total Length, Mbp	145.37811	138.66598	99.922768	99.890685	254.20631	285.59402	246.1601	109.94996	126.71662	127.16279	266.35047
Num Contigs	111861	109363	73271	67017	350097	345204	293887	96858	111337	104341	322958
Num Genes	216310	208541	145624	142348	500364	519433	449060	173420	198804	192155	487230
Longest Contig, Kbp	660.683	486.241	563.73	439.752	613.691	575.371	646.91	391.207	580.093	436.323	297.852
Shortest Contig, bp	200	200	200	200	200	200	200	200	200	200	200
N50, bp	4992	5363	1472	1817	58269	38157	36145	7512	5647	3983	43658
N75, bp	30953	30177	16318	13827	172507	149951	128012	31991	34580	30393	144062
N90, bp	70394	69138	44483	38501	270999	256881	218207	64680	73893	67552	241146
Rinke_et_al	4922	5535	3403	3732	9999	10593	9114	4514	4705	4036	10110
Campbell_et_al	6060	6223	4174	4175	11906	13013	10852	5193	5385	5005	12279
Ribosomal_RNAs	54	57	43	41	90	105	72	49	53	45	94
BUSCO_83_Protista	268	303	194	202	455	445	420	233	239	247	471
archaea (Rinke_et_al)	5	8	3	7	10	9	9	6	6	4	11
bacteria (Campbell_et_al)	41	46	28	25	66	76	83	31	34	33	88
eukarya (BUSCO_83_Protista)	0	0	0	0	0	0	0	0	0	0	0
predicted community size	46	54	31	32	76	85	92	37	40	37	99
Binned Length, Mbp	87.784583	96.233072	70.381788	74.58603	106.68482	144.74504	126.31662	72.582582	82.381951	84.642024	134.90931
Number of Bins	38	41	34	43	71	78	72	39	45	36	77
Binned Contigs	15724	21133	11017	13492	34621	40074	39298	19804	18307	17255	45370

Table 2.3: **MAG binning statistics.** The MAGs that contained genes from the *ptx-ptd* gene cluster are highlighted. Each row represents a genome, whose 'Full ID' indicates the enrichment community (e.g., LM1) and bin number (e.g., Bin_11_1) separated by a dash. Lineage specific completeness ('Comp., %'), contamination ('Cont., %'), and strain heterogeneity (Strain het., %) estimates were calculated with CheckM based on the reported 'CheckM marker lineage'. The '# of Contigs', 'N50', and 'GC content' were reported by the Anvi'o interface following binning. N50 is defined as the length N of the contig at which 50% of all bases in the total sequence have been assembled into contigs of length <N. bp, base pair; Mbp, megabase pair.

Full ID	Comp. (%)	Cont. (%)	Strain het. (%)	Genome Size (Mbp)	# of Contigs	N50 (bp)	GC content (%)	CheckM marker lineage
LM1-Bin_11_1	61.29	4.6	9.09	2.158608	1011	2336	60.62264509	c__Deltaproteobacteria (UID3216)
LM1-Bin_12_1	69.24	2.77	25	1.944739	902	2359	50.48358826	p__Firmicutes (UID1022)
LM1-Bin_17	77.7	3.88	33.33	2.273518	881	3061	68.42170168	k__Bacteria (UID2982)
LM1-Bin_18	84.14	9.37	20	3.01549	851	4182	50.31709528	c__Clostridia (UID1085)
LM1-Bin_21	98.9	9.27	80.95	2.474819	796	3911	37.75227393	k__Bacteria (UID2495)
LM1-Bin_22	89.62	5.24	5.56	3.332388	788	4998	45.53633898	k__Bacteria (UID2570)
LM1-Bin_25	75.26	2.63	0	1.414607	714	2113	40.02348578	k__Bacteria (UID209)
LM1-Bin_27	91.02	7.72	53.33	5.289283	705	10457	64.30990343	c__Gammaproteobacteria (UID4274)
LM1-Bin_30	71.87	5.16	20	1.294685	664	2074	47.68995651	k__Bacteria (UID1452)
LM1-Bin_32	81.49	2.58	10	2.76951	615	5771	44.62124257	p__Bacteroidetes (UID2605)
LM1-Bin_35	94.44	1.99	16.67	3.283129	529	8555	50.03429874	k__Bacteria (UID2569)
LM1-Bin_37_1	89.95	4.9	60	2.651873	349	12072	49.75869405	p__Bacteroidetes (UID2605)
LM1-Bin_39	94.83	0.57	100	3.403826	496	11727	54.04479989	k__Bacteria (UID2495)
LM1-Bin_40	94.72	8.13	63.64	1.510766	485	3999	62.64006952	p__Euryarchaeota (UID3)
LM1-Bin_42_D	96.77	5.38	11.11	3.523995	38	176023	53.85024092	c__Deltaproteobacteria (UID3216)
LM1-Bin_42_1	91.51	4.84	25	3.454405	258	19239	57.91170944	c__Deltaproteobacteria (UID3216)
LM1-Bin_46	88.66	3.25	23.08	2.024463	349	8510	48.92732048	p__Bacteroidetes (UID2605)
LM1-Bin_49	97.39	2.09	50	2.354626	265	12508	53.49356158	p__Euryarchaeota (UID49)
LM1-Bin_51	98.31	2.31	20	2.680084	83	99156	61.27851477	k__Bacteria (UID209)
LM1-Bin_8	85.29	5.73	16.67	3.628181	1414	2939	59.69412404	k__Bacteria (UID209)
LM3-Bin_1_merge	98.71	1.97	0	4.193681	176	72590	56.34294899	c__Deltaproteobacteria (UID3216)
LM3-Bin_10_1	83.59	9.1	65	2.881568	897	4196	65.43585969	k__Bacteria (UID2982)
LM3-Bin_11_1	97.43	3.42	0	4.901608	689	10289	37.82648977	p__Firmicutes (UID241)
LM3-Bin_12_1	81.36	0	0	3.329374	228	34671	58.40226448	k__Bacteria (UID209)
LM3-Bin_13_1	94.67	1.67	0	2.434541	35	157242	45.39554881	p__Bacteroidetes (UID2605)
LM3-Bin_13_2	98.31	0	0	4.777911	127	92672	58.17178518	k__Bacteria (UID209)
LM3-Bin_14_1	98.89	0.65	0	2.907346	66	77612	61.25536235	p__Euryarchaeota (UID54)
LM3-Bin_2_1	96.71	0.66	0	2.444589	78	45171	59.57730785	p__Euryarchaeota (UID49)
LM3-Bin_2_2	55.93	0.77	20	2.150062	184	18204	62.42602371	k__Bacteria (UID209)
LM3-Bin_2_3	78.04	0	0	2.254344	167	29185	41.17323815	k__Bacteria (UID209)
LM3-Bin_2_4	86	9.95	0	2.902598	346	11547	54.02520538	c__Deltaproteobacteria (UID3216)
LM3-Bin_3_2	98.12	3.76	25	2.574058	112	46035	50.87519899	k__Bacteria (UID209)
LM3-Bin_4_1	99.07	0.7	0	3.765323	35	190131	35.13527188	o__Clostridiales (UID1120)

Table 2.3: *continued.*

Full ID	Comp. (%)	Cont. (%)	Strain het. (%)	Genome Size (Mbp)	# of Contigs	N50 (bp)	GC content (%)	CheckM marker lineage
LM3-Bin_4_2	95.45	3.64	0	5.691286	66	208604	65.28109526	k__Bacteria (UID1452)
LM3-Bin_4_3	100	0	0	3.198789	50	105252	58.90084454	k__Bacteria (UID209)
LM3-Bin_6_1	61.89	9.43	0	3.492651	1660	2242	62.50764675	k__Bacteria (UID2565)
LM3-Bin_9_1	93.01	2.26	25	2.376246	509	6233	47.87541088	k__Bacteria (UID2495)
M1-Bin_1_1	99.3	0.47	0	2.210591	81	118390	37.60569174	f__Spirochaetaceae (UID2535)
M1-Bin_10	98.92	0.81	0	2.621785	75	68317	36.98220227	k__Bacteria (UID2569)
M1-Bin_11_1	93.79	2.38	0	3.195879	224	22727	43.97281495	p__Bacteroidetes (UID2605)
M1-Bin_12_1	83.56	6.04	4.35	2.187482	428	7206	46.60726695	p__Bacteroidetes (UID2605)
M1-Bin_13_1	52.1	5.29	9.09	2.250692	1381	1656	55.03504441	c__Deltaproteobacteria (UID3216)
M1-Bin_13_2	64.53	2.87	0	1.459621	561	2990	58.92806822	k__Bacteria (UID2495)
M1-Bin_14_1	52.19	1.81	0	2.08263	919	2568	52.07206863	k__Bacteria (UID2495)
M1-Bin_14_2	68.85	3.9	14.29	1.250165	543	2568	62.42428593	p__Euryarchaeota (UID3)
M1-Bin_15_1	93.5	3.57	44.44	2.150794	673	4262	41.53412681	k__Bacteria (UID2570)
M1-Bin_17_1	59	6.1	6.25	3.417036	1802	2001	57.77699236	c__Deltaproteobacteria (UID3216)
M1-Bin_17_2	63.56	0	0	3.123185	404	10866	61.90653328	k__Bacteria (UID209)
M1-Bin_18	99.29	3.07	0	4.786322	70	161165	40.20717869	o__Clostridiales (UID1120)
M1-Bin_19	78.57	1.1	100	1.770685	295	8707	36.87970798	k__Bacteria (UID2495)
M1-Bin_2_1	95.19	0.96	0	2.454638	34	201165	38.55566397	p__Firmicutes (UID240)
M1-Bin_20_1	92.47	2.69	20	2.501646	113	38241	39.28430076	k__Bacteria (UID2569)
M1-Bin_21_2	100	2.31	20	0.979502	9	203902	61.42810687	k__Bacteria (UID209)
M1-Bin_22_1	77.82	1.96	66.67	3.639515	377	15785	41.30513918	p__Euryarchaeota (UID49)
M1-Bin_3_1	74.3	5.73	69.23	1.314881	657	2083	59.83117804	p__Euryarchaeota (UID54)
M1-Bin_5_1	98.85	0	0	2.887409	37	209089	55.44820805	k__Bacteria (UID2495)
M1-Bin_6_1	98.71	1.94	0	3.759152	245	34840	55.7546552	c__Deltaproteobacteria (UID3216)
M1-Bin_8_1	70.63	2.38	40	2.44932	305	14857	48.18161662	k__Bacteria (UID209)
M1-Bin_8_2	100	1.75	0	2.036885	141	25351	38.96006022	k__Bacteria (UID209)
M1-Bin_9_1	88.63	6.42	4.55	3.700867	54	164333	51.20957573	c__Clostridia (UID1085)
M1-Bin_9_2	81.32	6.9	0	2.660125	203	23460	50.15933637	k__Bacteria (UID209)
R1-Bin_1	97.71	2.9	40	3.795281	338	28495	55.02929263	c__Deltaproteobacteria (UID3216)
R1-Bin_13_1	92.47	0.33	0	2.022588	45	92232	42.3650481	p__Euryarchaeota (UID49)
R1-Bin_13_2	70.62	5.43	25	2.362729	263	24320	47.36997332	k__Bacteria (UID2495)
R1-Bin_13_3	68.86	1.83	100	0.944075	223	5569	36.39769853	k__Bacteria (UID2495)
R1-Bin_13_4	84.99	3.45	100	1.849772	187	15651	46.92092572	k__Bacteria (UID2495)
R1-Bin_14_1	62.56	2.52	85.71	2.041396	638	3992	62.33922503	p__Euryarchaeota (UID54)
R1-Bin_14_2	95.53	2.82	55.56	3.060461	428	13017	48.77837111	k__Bacteria (UID209)
R1-Bin_15_1	98.04	0.65	0	2.365947	95	46283	52.74338876	p__Euryarchaeota (UID49)
R1-Bin_15_2	92.95	4.76	0	3.70671	38	189659	49.75381871	c__Clostridia (UID1085)
R1-Bin_15_3	100	0.62	0	2.936037	220	57503	58.30749327	k__Bacteria (UID209)
R1-Bin_15_4	71.23	7.18	100	1.587904	634	2803	59.50882261	p__Euryarchaeota (UID54)
R1-Bin_2_1	95.19	0	0	2.585985	293	16724	38.86250848	p__Firmicutes (UID240)
R1-Bin_3	98.85	0.1	100	3.221312	314	23244	51.17628171	k__Bacteria (UID2495)
R1-Bin_4_1	96.19	2.38	0	3.196695	98	71157	43.79793599	p__Bacteroidetes (UID2605)
R1-Bin_5_1	79.28	4.12	75	1.979919	424	6459	37.29099229	k__Bacteria (UID2569)
R1-Bin_7_1	82.99	0.63	50	2.264287	633	4776	46.03867118	k__Bacteria (UID209)
R3-Bin_12	52.02	2.37	0	1.961442	1169	1712	53.20158768	c__Deltaproteobacteria (UID3216)
R3-Bin_14	72.07	7.22	61.54	1.920665	1030	1916	45.27313498	k__Bacteria (UID2495)
R3-Bin_19	53.11	8.45	62.5	1.887751	861	2443	65.76249888	k__Bacteria (UID2982)
R3-Bin_20	98.31	2.54	50	2.699005	813	4627	35.68373617	k__Bacteria (UID209)

Table 2.3: *continued.*

Full ID	Comp. (%)	Cont. (%)	Strain het. (%)	Genome Size (Mbp)	# of Contigs	N50 (bp)	GC content (%)	CheckM marker lineage
R3-Bin_25	98.04	6.8	58.82	3.92036	628	8787	48.39125276	c__Deltaproteobacteria (UID3216)
R3-Bin_29	82.37	2.71	29.41	1.932821	566	4060	46.49434125	p__Bacteroidetes (UID2605)
R3-Bin_30	96.17	3.32	7.69	3.961919	560	10868	41.50940195	c__Clostridia (UID1085)
R3-Bin_31	63.48	4.94	30.77	1.14935	520	2431	48.09346761	k__Bacteria (UID1452)
R3-Bin_32	97.71	2.29	75	2.557611	500	7568	53.94747617	p__Euryarchaeota (UID54)
R3-Bin_35_1	95.02	1.94	0	3.594833	114	62309	56.43505452	c__Deltaproteobacteria (UID3216)
R3-Bin_35_2	98.06	6.49	69.23	4.083962	160	47853	57.89831045	c__Deltaproteobacteria (UID3216)
R3-Bin_38	96.6	2.75	0	2.277893	379	8229	48.01155232	k__Bacteria (UID2495)
R3-Bin_40	98.32	4.42	64	5.005175	263	49636	63.94598831	c__Gammaproteobacteria (UID4274)
R3-Bin_43	100	1	100	2.057812	252	11806	53.45860262	k__Bacteria (UID209)
R3-Bin_46	98.86	2.15	0	3.662032	46	272288	42.56960281	k__Bacteria (UID2565)
R3-Bin_6	56.32	5.11	0	2.568386	1624	1603	44.74700796	k__Bacteria (UID2495)
R3-Bin_8	67.01	5.46	18.18	2.79072	1406	2103	32.2906856	o__Clostridiales (UID1120)
SL1-Bin_12	67.42	3.84	11.11	1.918784	975	2071	68.1739003	k__Bacteria (UID2982)
SL1-Bin_13	80.88	6.76	21.05	2.765802	913	3430	44.40743574	c__Deltaproteobacteria (UID3216)
SL1-Bin_17	81.3	5.6	21.05	1.991261	820	2766	37.9918103	k__Bacteria (UID2569)
SL1-Bin_19	79.45	1.7	0	1.754667	765	2597	59.35842194	k__Bacteria (UID2495)
SL1-Bin_21	66.1	4.39	25	1.572524	718	2414	56.12924067	k__Bacteria (UID209)
SL1-Bin_22	82.38	1.85	33.33	2.742295	692	5023	44.58111628	p__Bacteroidetes (UID2605)
SL1-Bin_24	61.24	0.2	0	1.220318	670	1875	37.7779565	k__Bacteria (UID2495)
SL1-Bin_25	56.43	1.25	50	1.242409	668	2003	53.97068296	k__Bacteria (UID2495)
SL1-Bin_26	65.2	5.99	20	1.222264	660	1882	48.13174808	k__Bacteria (UID203)
SL1-Bin_27	99.84	8.07	75	3.965802	653	10001	48.89801706	k__Bacteria (UID209)
SL1-Bin_28	78.21	2.81	11.11	1.710803	626	3280	48.87283156	p__Bacteroidetes (UID2605)
SL1-Bin_29	96.51	8.06	26.32	3.749274	524	16529	39.19740993	k__Bacteria (UID2569)
SL1-Bin_31_1	98.31	0	0	2.558347	76	49759	60.85528681	k__Bacteria (UID209)
SL1-Bin_31_2	98.49	2.58	20	3.426896	96	67223	57.16435399	c__Deltaproteobacteria (UID3216)
SL1-Bin_31_3	78.98	3.23	0	2.421291	149	25635	62.56509059	c__Deltaproteobacteria (UID3216)
SL1-Bin_32	77.25	4.07	20	1.044055	441	2669	54.04031196	k__Bacteria (UID1452)
SL1-Bin_33	93.06	4.71	30	2.7551	429	8953	45.77369505	p__Bacteroidetes (UID2605)
SL1-Bin_34_1	94.61	2.45	33.33	2.547125	91	43097	39.0515736	p__Firmicutes (UID239)
SL1-Bin_34_2	98.1	0.95	0	3.28182	70	80605	48.92044085	p__Firmicutes (UID1022)
SL1-Bin_34_3	95.14	1.43	0	1.997561	22	137904	46.45028082	p__Bacteroidetes (UID2605)
SL1-Bin_35	95.97	2.06	50	1.518975	425	5097	62.19359989	p__Euryarchaeota (UID3)
SL1-Bin_37	97.71	5.07	80	2.646916	356	12137	44.86715509	p__Euryarchaeota (UID54)
SL1-Bin_41	94.57	3.38	0	5.998366	105	102031	54.94109688	k__Bacteria (UID2982)
SL1-Bin_42	91.82	4.83	14.29	2.864805	249	29255	48.18234867	k__Bacteria (UID1452)
SL1-Bin_43	90.74	2.78	0	4.822721	145	53092	67.88311178	k__Bacteria (UID1452)
SL1-Bin_44	98.31	0	0	4.532432	77	120604	58.86599017	k__Bacteria (UID209)
SL1-Bin_45	94.09	3.76	0	4.433143	70	141120	68.85232933	k__Bacteria (UID2565)
SL1-Bin_47	99.68	1.11	0	3.818114	100	77209	59.86419482	p__Firmicutes (UID1022)
SL1-Bin_48	97.31	2.69	66.67	3.423526	118	59547	50.08075303	k__Bacteria (UID2569)
SL1-Bin_50	98.04	1.96	33.33	2.427392	133	30954	52.58034116	p__Euryarchaeota (UID49)
SL1-Bin_51	98.85	0	0	2.892242	29	338689	55.70284848	k__Bacteria (UID2495)
SL1-Bin_8	60.66	3.44	33.33	2.206414	1254	1771	47.34628682	p__Firmicutes (UID239)

Table 2.3: *continued.*

Full ID	Comp. (%)	Cont. (%)	Strain het. (%)	Genome Size (Mbp)	# of Contigs	N50 (bp)	GC content (%)	CheckM marker lineage
SL3-Bin_10	85.9	6.55	25	3.327998	1138	3626	60.20756412	k__Bacteria (UID209)
SL3-Bin_13	53.91	7.68	14.29	1.743811	1037	1708	47.03803466	k__Bacteria (UID1452)
SL3-Bin_14	58.48	5.48	7.14	1.747571	968	1851	51.7105471	c__Deltaproteobacteria (UID3216)
SL3-Bin_15	71.67	7.13	21.05	1.701877	865	2064	55.40505591	k__Bacteria (UID1452)
SL3-Bin_16	59.93	3.39	50	1.515235	858	1818	63.18321748	k__Bacteria (UID209)
SL3-Bin_20	55.35	5.28	31.25	1.043992	684	1492	49.68156298	c__Betaproteobacteria (UID3888)
SL3-Bin_21	51.69	1.97	0	1.142478	674	1724	60.79856195	p__Euryarchaeota (UID49)
SL3-Bin_22_1	96.61	0	0	2.263039	36	108035	46.71025407	k__Bacteria (UID209)
SL3-Bin_22_2	98.69	0.65	0	4.287371	58	206214	54.85093667	c__Deltaproteobacteria (UID3216)
SL3-Bin_22_3	94.92	0	0	1.784666	67	43727	56.22595718	k__Bacteria (UID209)
SL3-Bin_22_4	88.14	0.31	50	1.40321	167	8915	52.9312656	k__Bacteria (UID209)
SL3-Bin_23	93.33	8.47	66.67	4.060099	595	13085	37.48284202	c__Clostridia (UID1085)
SL3-Bin_27	81.21	0.81	50	1.736853	466	4752	50.6207	o__Clostridiales (UID1120)
SL3-Bin_28	95.25	1.68	20	3.316081	434	10321	50.43368433	k__Bacteria (UID2569)
SL3-Bin_29	99.84	0.63	0	3.578389	410	18972	49.24532012	k__Bacteria (UID209)
SL3-Bin_30	95.1	1.14	100	2.652413	419	9997	44.12246625	p__Euryarchaeota (UID54)
SL3-Bin_33	97.43	4.24	0	5.18417	325	24641	37.33824601	p__Firmicutes (UID241)
SL3-Bin_38	100	2	0	2.757371	191	142201	41.6717807	k__Bacteria (UID209)
SL3-Bin_39	100	4.61	42.86	2.766768	192	43117	54.88672763	p__Euryarchaeota (UID49)
SL3-Bin_41_1	98.71	1.94	0	3.524099	90	103203	56.61946526	c__Deltaproteobacteria (UID3216)
SL3-Bin_6	62.24	7.01	0	2.893558	1509	2041	69.33398667	k__Bacteria (UID2565)
SL3-Bin_7	57.74	3.18	0	2.542201	1474	1755	63.91021581	k__Bacteria (UID2565)
SL3-Bin_9	60.26	4.1	6.25	2.101167	1185	1855	67.08748236	o__Sphingomonadales (UID3310)
SM1-Bin_1_1	98.85	0	0	2.798546	123	41072	56.17750979	k__Bacteria (UID2495)
SM1-Bin_10_1	96.19	1.43	0	2.099756	65	72851	37.85707752	p__Bacteroidetes (UID2605)
SM1-Bin_11_1	98.84	1.9	0	3.901353	55	153921	52.86124189	p__Firmicutes (UID1022)
SM1-Bin_13_1	96.19	1.43	0	3.177205	90	83691	43.71781551	p__Bacteroidetes (UID2605)
SM1-Bin_14_1	94.67	4.94	57.14	2.850332	325	13451	48.26421158	k__Bacteria (UID209)
SM1-Bin_15_1	93.33	2.5	0	3.827328	70	96450	50.64701355	c__Clostridia (UID1085)
SM1-Bin_16_1	93.18	0.07	100	2.887501	450	10407	51.32072924	k__Bacteria (UID2495)
SM1-Bin_17_1	96.71	1.72	0	3.113281	185	31690	52.59282999	o__Clostridiales (UID1212)
SM1-Bin_17_2	70.69	7.16	8.33	3.858483	1150	4184	59.30941914	k__Bacteria (UID209)
SM1-Bin_18_1	91.66	2.07	0	2.82649	381	9741	59.50737469	p__Firmicutes (UID239)
SM1-Bin_18_2	50.36	5.27	0	1.518287	904	1710	63.28076827	k__Bacteria (UID203)
SM1-Bin_19_1	59.73	0.65	0	1.32062	582	2626	54.27905771	p__Euryarchaeota (UID49)
SM1-Bin_19_3	56.09	2.2	0	1.238136	705	1827	36.34227353	k__Bacteria (UID2495)
SM1-Bin_19_5	56.12	4.39	0	1.512742	841	1826	45.58788741	k__Bacteria (UID209)
SM1-Bin_2_1	76.61	7.29	85.71	1.974006	871	2569	62.36747213	p__Euryarchaeota (UID54)

Table 2.3: *continued.*

Full ID	Comp. (%)	Cont. (%)	Strain het. (%)	Genome Size (Mbp)	# of Contigs	N50 (bp)	GC content (%)	CheckM marker lineage
SM1-Bin_3_1	94.84	1.94	0	3.991185	104	62323	56.96827462	c__Deltaproteobacteria (UID3216)
SM1-Bin_7_1	99.84	0	0	2.986772	107	74177	45.88776911	k__Bacteria (UID209)
SM1-Bin_8_1	99.3	1.4	0	2.323822	146	32149	33.6932519	o__Clostridiales (UID1120)
SM1-Bin_9	99.18	0	0	2.089447	22	198514	42.11005798	p__Euryarchaeota (UID49)
SM3-Bin_1	96.43	1.94	0	3.534795	130	62052	56.12219851	c__Deltaproteobacteria (UID3216)
SM3-Bin_10_1	100	0	0	2.356492	67	67601	55.95936802	p__Euryarchaeota (UID49)
SM3-Bin_12_1	97.27	6.09	0	5.262576	130	61470	49.11520069	k__Bacteria (UID1452)
SM3-Bin_12_2	95.63	1.37	33.33	2.313649	92	41579	49.52986331	k__Bacteria (UID2570)
SM3-Bin_13	83.05	0.77	0	2.433496	458	7688	62.16055793	k__Bacteria (UID209)
SM3-Bin_14_1	89.86	1.84	50	3.053212	313	15657	50.64384112	k__Bacteria (UID2495)
SM3-Bin_15_1	97.73	2.94	50	3.764222	441	14364	54.04114854	k__Bacteria (UID2565)
SM3-Bin_16_1	100	1.14	0	2.914218	55	285730	54.17337005	k__Bacteria (UID2495)
SM3-Bin_16_2	100	2.2	0	2.441037	97	59504	47.35754058	k__Bacteria (UID2495)
SM3-Bin_17	99.18	1.33	0	3.946105	329	19930	40.77032353	p__Euryarchaeota (UID49)
SM3-Bin_2	98.89	0.65	0	2.93925	63	100012	60.83437786	p__Euryarchaeota (UID54)
SM3-Bin_3_merge	97.4	2.9	40	4.316883	99	119502	57.79631379	c__Deltaproteobacteria (UID3216)
SM3-Bin_4_1	98.37	1.4	0	4.186474	38	311033	28.79365348	o__Clostridiales (UID1120)
SM3-Bin_5_1	91.53	6.82	16.67	3.684212	247	31716	49.44445254	k__Bacteria (UID2495)
SM3-Bin_5_2	94.92	0	0	2.453539	172	32250	42.54102958	k__Bacteria (UID209)
SM3-Bin_6_1	67.5	2.85	30.77	1.724461	274	10007	45.94566084	p__Bacteroidetes (UID2605)
SM3-Bin_6_2	96.76	2.68	0	2.99693	58	134390	43.39471037	c__Clostridia (UID1085)
SM3-Bin_7_1	56.14	0.49	0	1.333458	557	2832	33.44790892	k__Bacteria (UID237)
SM3-Bin_9_2	75.01	7.16	0	2.353514	1067	2488	42.34573951	p__Bacteroidetes (UID2605)
SV1-Bin_1_1	98.71	2.15	0	3.518499	297	28378	55.34550046	c__Deltaproteobacteria (UID3216)
SV1-Bin_1_2	97.75	0	0	1.855567	20	208079	34.59916129	k__Bacteria (UID2495)
SV1-Bin_1_3	98.85	0	0	2.824057	23	235531	56.2617631	k__Bacteria (UID2495)
SV1-Bin_2_2	100	0.62	0	2.530233	72	93956	61.57121656	k__Bacteria (UID209)
SV1-Bin_3_1_1_1	87.23	4.63	87.5	2.137557	475	5802	62.55661619	p__Euryarchaeota (UID54)
SV1-Bin_3_2	98.92	0.81	0	2.574201	132	38953	37.07831921	k__Bacteria (UID2569)
SV1-Bin_3_3	96.19	1.43	0	3.207085	105	73634	43.4872117	p__Bacteroidetes (UID2605)
SV1-Bin_3_4	96.77	3.23	0	4.145731	78	115056	56.50555637	c__Deltaproteobacteria (UID3216)
SV1-Bin_4_1	97.63	0	0	2.252545	62	67508	58.39660766	k__Bacteria (UID2495)
SV1-Bin_4_2	97.46	2.62	14.29	3.463143	388	13060	59.36841563	k__Bacteria (UID209)
SV1-Bin_5_1	97.46	1.69	100	1.766432	336	8264	44.89220104	k__Bacteria (UID209)
SV1-Bin_5_2	77.37	4.03	9.09	1.897673	834	2489	37.37931669	f__Spirochaetaceae (UID2535)
SV1-Bin_5_3	50.12	7.61	5.88	2.253173	1190	1921	50.38707567	c__Spirochaetia (UID2496)
SV1-Bin_6_1	94.93	5.11	60	3.218031	457	11821	50.8137886	k__Bacteria (UID2495)
SV1-Bin_6_2	77.41	6.09	100	1.556517	558	3373	36.61926061	k__Bacteria (UID2495)
SV1-Bin_6_3	88.28	2.58	0	2.130091	112	30117	53.22382554	c__Deltaproteobacteria (UID3216)
SV1-Bin_6_4	79.61	4.95	66.67	2.221915	188	18982	48.72570987	k__Bacteria (UID209)
SV1-Bin_6_5	87.4	6.9	100	2.887084	278	20626	50.15862592	k__Bacteria (UID209)
SV1-Bin_7_2	97.75	1.12	0	1.938615	35	219880	35.29505858	k__Bacteria (UID2495)

Table 2.3: *continued.*

Full ID	Comp. (%)	Cont. (%)	Strain het. (%)	Genome Size (Mbp)	# of Contigs	N50 (bp)	GC content (%)	CheckM marker lineage
SV1-Bin_7_4	62.94	1.48	0	1.591791	407	5084	36.92893429	p__Bacteroidetes (UID2605)
SV1-Bin_7_5	93.71	1.92	40	1.90238	217	13033	32.16783499	o__Clostridiales (UID1120)
SV1-Bin_8_1	86.02	4.03	37.5	1.495399	554	3229	62.78786655	p__Euryarchaeota (UID3)
SV1-Bin_8_2	71.47	6.2	11.11	1.90713	614	3949	46.57402328	k__Bacteria (UID2495)
SV1-Bin_9_1	94.61	0	0	1.976387	105	30782	42.58257393	p__Euryarchaeota (UID49)
SV1-Bin_9_2	93.27	2.4	33.33	2.465724	209	16214	37.31108679	p__Firmicutes (UID240)
SV1-Bin_9_3_1	80.8	3.45	100	2.181564	216	14781	44.72151909	k__Bacteria (UID203)
SV1-Bin_9_3_2	70.35	4.37	30	2.247105	473	6059	45.52253446	k__Bacteria (UID2570)
SV1-Bin_9_4	72.04	5.11	0	2.824996	505	7455	40.53019715	k__Bacteria (UID2569)
SV3-Bin_1_1	92.26	1.31	33.33	3.778971	66	128804	57.69175353	c__Deltaproteobacteria (UID3216)
SV3-Bin_10_1	98.31	0	0	4.51741	36	251583	59.06037149	k__Bacteria (UID209)
SV3-Bin_11_1	88.27	0	0	2.051917	93	28670	57.23197773	p__Euryarchaeota (UID49)
SV3-Bin_12_1	99.65	1.4	0	2.274249	31	157247	30.56418754	o__Clostridiales (UID1120)
SV3-Bin_13_1	100	0	0	2.290324	63	63057	41.67698582	k__Bacteria (UID209)
SV3-Bin_14_1	93.46	2.94	0	3.757364	199	32232	37.13319229	c__Clostridia (UID1085)
SV3-Bin_15_1	98.06	4.52	9.09	3.38049	198	28500	48.36630011	c__Deltaproteobacteria (UID3216)
SV3-Bin_16_1	81.7	0.65	0	3.410925	169	29982	42.02271925	p__Euryarchaeota (UID49)
SV3-Bin_16_2	74.85	0.11	50	2.05812	173	17632	41.45339822	c__Clostridia (UID1085)
SV3-Bin_17_1	98.76	1.74	20	5.04755	59	149464	67.3482876	o__Rhodospirillales (UID3754)
SV3-Bin_18_1	93.14	4.03	75	1.754252	415	5385	60.28158238	p__Euryarchaeota (UID54)
SV3-Bin_19_1	76.16	4.79	0	2.527575	706	4692	37.24995885	k__Bacteria (UID2570)
SV3-Bin_20_1_1	90.85	1.31	0	1.864808	236	12576	56.02851032	p__Euryarchaeota (UID54)
SV3-Bin_21_1	63.73	2.91	75	1.671206	572	3260	50.63127102	k__Bacteria (UID2570)
SV3-Bin_21_3_1	54.02	5.08	0	1.835603	746	2854	61.20194493	k__Bacteria (UID209)
SV3-Bin_26_1_1	68.69	1.69	10	2.629117	757	4522	48.79748091	p__Bacteroidetes (UID2605)
SV3-Bin_3_1	90.97	1.31	33.33	3.10179	64	93603	57.64817512	c__Deltaproteobacteria (UID3216)
SV3-Bin_30_1	89.5	0.97	0	2.641899	439	9579	45.20911732	k__Bacteria (UID209)
SV3-Bin_4_1	94.92	0.65	0	2.652084	125	37074	61.77394713	p__Euryarchaeota (UID54)
SV3-Bin_6_1	94.52	6.57	84.62	2.458181	572	6063	29.19126171	p__Bacteroidetes (UID2605)
SV3-Bin_7_1	96.76	3.4	0	3.164657	70	145547	42.87419792	c__Clostridia (UID1085)
SV3-Bin_8_1	100	0	0	2.176944	67	73060	46.68787613	k__Bacteria (UID209)
SV3-Bin_9_1	83.08	4.76	66.67	2.508724	604	5466	65.62735518	k__Bacteria (UID2982)
SV3-Bin_9_2	96.39	1.72	0	2.417266	459	7536	46.14567909	k__Bacteria (UID209)

Table 2.4: Composition of the HEPO culture following serial dilution from the SV3 enrichment in the absence of rumen fluid. All rps3 genes from the HEPO culture were identified via pHMMs and subsequently BLAST searched on the NCBI. The closest relative for each rpS3 gene in the NCBI database is listed along with its accession number and percent identity (ID). The shaded rows are predicted to be two strains of DPOM based on sequence homology and metagenomic analyses of the parental SV3 enrichment.

rps3 ID	Closest relative taxonomy	Accession Number	Percent ID
k141_2512_34	Lascolabacillus massiliensis	WP_045089286.1	100
k141_4090_21	Deltaproteobacteria bacterium	MBP8649273.1	94.91
k141_1073	Candidatus Sumerlaea chitinivorans	AXA35169.1	86.71
k141_13518_1	Aminobacterium colombiense	WP_013048034.1	89.66
k141_14450_10	Methanomicrobium sp. W14	WP_209690087.1	86.01
k141_20507_22	Deltaproteobacteria bacterium	MBP8649273.1	94.44
k141_22067_35	Unclassified Bacterium	MBI1292738.1	83.17
k141_25159_1	Desulfomicrobium apsheronum	WP_092376938.1	99.52
k141_28904_4	Chloroflexi bacterium	QRN82250.1	96.89
k141_29216_11	Methanomicrobiales archaeon	PKL55217.1	94.5
k141_31413_5	Clostridiaceae bacterium	NLW10938.1	94.62
k141_33394_7	Synergistaceae bacterium	NLD96077.1	98.33
k141_33754_44	Mesotoga bacterium	WP_099828912.1	100
k141_35434_80	Candidatus Cloacimonetes bacterium	NLN34878.1	93.93
k141_39717_9	Clostridia bacterium	NLO75971.1	83.78
k141_42761_29	Bacteroidales bacterium	MBP7434540.1	94.51

Chapter 3

Prevalence, Phylogenetic Diversity, and Evolution of DPOM

Parts of this chapter have been drawn from the following publication:

Ewens, S. D., Gomberg, A. F. S., Barnum, T. P., Borton, M. A., Carlson, H. K., Wrighton, K. C., Coates, J. D. (2021). The diversity and evolution of microbial dissimilatory phosphite oxidation. PNAS.

Abstract

The environmental distribution, diversity, and evolution of DPO microorganisms (DPOMs) remained enigmatic for over 20 years following the initial discovery of *Desulfotignum phosphitoxidans* FiPS-3, the first known DPOM. Since the discovery of FiPS-3, only one additional species of DPOM, *Phosphitivorax anaerolimi,* had been identified prior to the writing of this dissertation. FiPS-3 and *P. anaerolimi* were each identified in distant geographical locations and were characterized to be disparate organisms with respect to their taxonomy, genomic potential, and physiology. As such, it was not possible to make sound predictions about the prevalence, diversity, or evolutionary history of DPOM organisms at large. In this chapter, metagenomic sequencing of the phosphite-enriched microbial communities from **Chapter 2** enabled the taxonomic classification of 21 additional DPOMs. These DPOMs span six classes of bacteria, including the *Negativicutes*, *Desulfotomaculia*, *Synergistia*, *Syntrophia*, *Desulfobacteria*, and *Desulfomonilia_A*. By comparing DPO marker genes from the genomes of enriched organisms with over 17,000 publicly available metagenomes, we find that DPO metabolism exists globally in diverse anoxic environments, including wastewaters, sediments, and subsurface aquifers. Phylogenetic analyses of DPO marker genes further indicates that the DPO genes form a highly conserved cluster that likely has ancient origins predating the split of monoderm and diderm bacteria.

Introduction

At the time of writing, scarce representation of DPOM (n=2) limited our understanding of the genes, organisms, and environments that support DPO. The taxonomic diversity of DPOM remained unclear because the two known species of DPOM, *Desulfotignum phosphitoxidans* FiPS-3 and Ca. *Phosphitivorax anaerolimi,* belong to distinct taxonomic classes, and their closest relatives are either uncultured or unable to catalyze DPO[29]. *Desulfotignum phosphitoxidans* FiPS-3 was isolated from brackish sediments of the Italian Venice canals, and it belongs to the *Desulfobacteria* class of the *Deltaproteobacteria*[18]. It's closest cultured relative is *Desulfotignum balticum* SaxT, which shares 99% 16S rRNA gene sequence identity with FiPS-3, yet wild-type SaxT is unable to oxidize phosphite for either nutritive phosphorus (APO) or for energy (DPO)[18]. *Phosphitivorax anaerolimi* Phox-21 was identified in a DPO enrichment originating from a San Francisco Bay area wastewater facility[70]. Phox-21 belongs to the GW-28 candidate order of the *Deltaproteobacteria*, which is occupied by uncultured clones whose ability to perform DPO has not yet been tested[70]. An additional strain of *P. anaerolimi* (strain F81) was later identified in a syntrophic enrichment originating from a Danish wastewater facility. F81 contains marker genes for DPO, but has not yet been biochemically shown to perform DPO[71]. The closest cultured relative of *P. anaerolimi* is *Desulfomonile tiedjei* DCB-1, which shares only 84% 16S rRNA gene sequence identity with Phox-21. The taxonomic and geographical separation of FiPS-3 and *P. anaerolimi* consequently obscure the recognized taxonomic diversity and prevalence of DPOM at large.

FiPS-3 and Phox-21 have been empirically shown to couple DPO to growth, and DPO activity in both organisms has been attributed to the *ptx-ptd* gene cluster[69]. The FiPS-3 gene cluster encodes seven proteins (PtxDE-PtdCFGHI)[69], and the *P. anaerolimi* gene cluster encodes six proteins, as it is missing *ptdG*[70,71]. Based on sequence homology, *ptxE* is predicted to encode a LysR family transcriptional regulator; *ptxD* is predicted to encode a phosphite dehydrogenase that couples phosphite oxidation to NAD^+ reduction; *ptdC* is predicted to encode a phosphite:phosphate antiporter; *ptdF* is predicted to encode a UDP-glucose 4-epimerase; *ptdG* is predicted to encode a UspA-family nucleotide-binding universal stress protein; *ptdH* is predicted to encode a B12-dependent radical S-adenosyl L-methionine (SAM) family protein; and *ptdI* is a hypothetical protein with no homologous proteins to support functional predictions[29,69]. Heterologous expression of the PtxD and PtdFCG from FiPS-3 conferred *Desulfotignum balticum* SaxT the ability to utilize phosphite as a phosphorus source, but did not confer the ability to conserve energy via DPO[69]. PtxD and PtdC were necessary and sufficient for *D. balticum* to assimilate phosphite, suggesting other members of the *ptx-ptd* gene cluster are required for energy conservation[69]. However, the biochemical mechanism by which the *ptx-ptd* gene products conserve energy via DPO remains unknown, and models to describe this mechanism are explored further in **Chapter 4**.

The *ptx-ptd* gene cluster distinguishes DPO from APO metabolism, and it was used to identify Phox-21 from a complex microbial community[70]. While homologues of PtxD and PtxE are used by APO microorganisms to oxidize phosphite for phosphorus, no APO microorganisms have been shown to successfully use phosphite as an electron donor[30]. Yet the sequence diversity of the *ptxD* and *ptxE* within the DPO gene cluster, and their divergence from the APO homologues, has not yet been quantitatively demarcated. The synteny of the *ptx-ptd* gene clusters in FiPS-3 and Phox-21 also features alternative gene presence and orientation, and the absence of *ptdG* in the Phox-21 genome highlights the lack of clarity about which core genes are necessary and sufficient for functional DPO activity. Furthermore, the PtxD proteins from FiPS-3 and Phox-21 share only 55%

amino acid sequence similarity, suggesting broad diversity between the homologous genes of DPO gene clusters[70]. This sequence diversity exacerbates the unclear evolutionary history of DPOM: FiPS-3 is the only member of its taxonomic clade to contain the *ptx-ptd* gene cluster, which is flanked by transposase genes. As such, FiPS-3 likely acquired its *ptx-ptd* genes via horizontal gene transfer[67]. However, no evidence exists for the horizontal acquisition of *ptx-ptd* genes in Phox-21[70]. In this chapter, the diverse DPOM genomes from the enrichments of **Chapter 2** allowed for phylogenetic analyses that summarize the likely evolutionary history of DPO metabolism on Earth.

The availability of sequence data for *ptx-ptd* genes from diverse DPOM further allowed the development of profile hidden Markov models (pHMMs) for each of the DPO genes. pHMMs model biological sequence data and can be used to locate sequences of interest within a dataset[116]. A pHMM search is analogous to the familiar BLAST search with better-defined boundaries for what constitutes a "hit", thereby helping to resolve the predictive ambiguity of homology assignments. As such, characterizing greater sequence diversity of a niche metabolism (i.e., DPO) allows for the refinement of pHMMs, further enabling the confident identification of novel representatives. It was essential to ground the modeled genes via activity assays, phylogenetic relationships, and gene neighborhood analyses, as the quality of a pHMM depends on the quality of sequences incorporated into the model[117]. In this chapter, pHMMs of the *ptx-ptd* genes are used to computationally mine for DPOM in metagenomes and serve as a tool for profiling future environmental metagenomes.

Materials and Methods

Phylogenetic Characterization

The DPO MAGs that were identified in the enrichments from **Chapter 2** were phylogenetically characterized using three alternative databases (Genome Taxonomy Database: GTDB, National Center for Biotechnology Information: NCBI, and SILVA) and their respective marker genes (concatenated gene sets, rpS8, and 16S rRNA gene).

Taxonomy assignments utilizing the GTDB toolkit (GTDB-Tk) placed all high-quality MAGs into protein reference trees using concatenated SCG sets[118]. Protein reference trees for the GTDB are sourced from prokaryotic genomes in the NCBI Assembly database and are continuously updated to provide rank-normalized genome-based taxonomies[119].

A pHMM for the rpS8 was obtained from Wu *et al.* and applied to all DPO MAGs[100,116]. The rpS8 gene has been shown to effectively represent whole-genome average nucleotide identity (ANI) values[120] and was present once in each DPO MAG. Each rpS8 was BLAST searched against the NCBI GenBank database to identify the closest relative, closest isolated relative, and informative representatives for phylogenetic analyses. Identified close relatives corresponded to the multi-gene taxonomy assignments of the GTDB (Table 3.1 and Table 3.2). Sequences were aligned using MUSCLE v3.8.1551[121], and an approximately-maximum-likelihood phylogenetic tree was constructed with 1000 bootstrap resamplings using FastTree v2.1.11[122]. Trees were visualized using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

The 16S rRNA gene for each community was reconstructed from metagenomic reads using default parameters in EMIRGE with 50 iterations[123]. Reconstructed genes were classified using SILVA[124] and mapped back to the 16S rRNA gene fragments of DPO MAGs.

The novelty of each DPO MAG was determined by the rank of closest relatives in the GTDB, NCBI (rpS8), and SILVA (16S rRNA gene) databases (Table 3.4). A DPO MAG was considered novel at the specified rank (e.g., species, genus) based on the following thresholds: (i) GTDB: considered novel if there were no logged relatives for that rank; (ii) NCBI (rpS8): considered a novel species if the closest relative was <98.3% identity; (iii) SILVA (16S rRNA gene) considered a novel species if the closest relative shared <96.7% identity and a novel genus if the closest relative shared <94% identity[120]. The novelty of a DPO MAG was assigned based on the lowest resolved taxonomic rank between all searched databases.

Identification of Metagenomic DPO Proteins.

DPO proteins (PtxD, PtdC, PtdF) were identified from publicly available metagenomes. The largest metagenomes (representing 90% of proteins from each ecosystem category) in the JGI Integrated Microbial Genomes and Metagenomes (IMG/M) database were collected (n=17,888) on August 1, 2018 (Table 3.5). Sequence data from the IMG/M database were produced by the U.S. Department of Energy Joint Genome Institute (**http://www.jgi.doe.gov/**) in collaboration with the user community. The FiPS-3 PtxD, PtdC, and PtdF were searched against all proteins using BLASTP with bit score thresholds of 270, 300, and 250 respectively. Positive hits were aligned using MUSCLE v3.8.1551[121] and constructed into an approximately maximum-likelihood phylogenetic tree using FastTree v2.1.11[122] with 1000 bootstrap resamplings. DPO proteins were defined as i) those that formed a phylogenetically distinct clade with proteins from experimentally validated DPOM, ii) were found on a contig near at least one other putative DPO gene, and iii) were at least 90% the length of their homolog protein in FiPS-3. Protein sequences from the identified *ptx-ptd* gene clusters were used to create profile Hidden Markov Models (pHMMs) for each of the PtxDE-PtdCFGHI proteins using HMMER v3.2.1[116,125]. These pHMMs are publicly available and reported in Ewens et al.[117]. Bit score thresholds for stringent *de novo* identification of DPO proteins were determined by a reciprocal pHMM search on a subset of the IMG/M database[117].

Evolutionary Analyses

To compare the evolutionary relationships between the PtxD, PtdC, and PtdF, members of the DPO clade were dereplicated with CD-HIT v4.8.1[126] by clustering proteins with 100% sequence similarity and selecting the largest contig to represent each cluster in a simplified phylogenetic tree. Tanglegrams comparing PtxD to PtdC and PtdF were generated with Dendroscope v3.7.2[127]. Gene synteny was visualized with SimpleSynteny[128], where genes were identified with BLAST and annotated according to the custom pHMMs.

Results

DPOM Taxonomy

DPOM taxonomy assignments were made using i) reconstructed 16S rRNA gene fragments[123], ii) multigene alignments using the genome taxonomy database (GTDB)[118], and iii) alignment of the ribosomal S8 proteins (rpS8)[120]. Assignments were congruent in each instance and visualized in a phylogenetic tree (Fig. 3.1). Prior to this study, DPOM had been identified as

belonging to only two taxonomic classes of the *Desulfobacterota* phylum. In contrast, DPOM in this study's enrichments span the monoderm-diderm taxonomic boundaries and include three phyla (*Desulfobacterota, Firmicutes,* and *Synergistota*) and six classes (*Negativicutes, Desulfotomaculia*, *Synergistia*, *Syntrophia*, *Desulfobacteria* and *Desulfomonilia_A*) (Fig. 3.1).

Desulfomonilia_A was the most sampled class of DPOM (Fig. 3.1), comprising 13 of the 21 DPO MAGs. The *Desulfomonilia_A* were enriched from all six sample sites and were present in nine communities. Furthermore, they were the most relatively abundant DPO MAG (representing >85%) in each of eight communities, indicating either i) environmental dominance, or ii) a possible advantage under my enrichment conditions (Figs. 3.1 & 2.3). *Desulfomonilia_A* is an uncultured bacterial class that has recently been distinguished from the *Desulfomonilia* (https://gtdb.ecogenomic.org/). Consistent with this, the *Desulfomonilia* are represented by *Desulfomonile tiedjei*, which shares just 49% rpS8 sequence identity to the most closely related DPO MAG of the *Desulfomonilia_A*[129,130] (Fig. 3.1). All DPO MAGs of the *Desulfomonilia_A* class belong to the uncultured order UBA1062 (previously denoted GW-28)[70], which includes Ca. *Phosphitivorax* (Table 3.1). The monophyletic separation of the *Desulfomonilia_A* DPOM supports the hypothesis that Ca. *Phosphitivorax* species are part of a unique order, and possibly a unique class, for which DPO is a common metabolic feature[70].

Beyond the *Desulfomonilia_A* DPOM, eight additional genomic representatives were recovered from four diverse classes (*Negativicutes, Desulfotomaculia, Synergistia,* and *Syntrophia)* (Fig. 3.1). While most of these are minority DPOM in their respective communities, at least three (*Pelotomaculaceae* SL1, Ca. *Smithella* SM1, and Ca. *Smithella* LM1) dominate their DPOM populations (>83%) (Figs. 2.3B & 3.1). The Ca. *Negativicutes* and Ca. *Desulfotomaculia* MAGs represent the first DPO genomes taxonomically assigned to the gram-positive *Firmicutes* phylum, highlighting the broad evolutionary divergence of DPOM[131,132]. During the writing of this thesis, the first gram-positive DPOM (*Phosphitispora fastidiosa*) was isolated, confirming the existence of gram-positive DPOM[75]. *Phosphitispora fastidiosa* is a spore-forming bacterium belonging to the *Firmicutes* phylum and was isolated from anaerobic sewage sludge in Germany[75]. NCBI taxonomic classification assigns *P. fastidiosa* to the *Peptococcaceae* family of the *Clostridia* class[75]. The GTDB assigns this taxonomic branch to the *Pelotomaculaceae* family of the *Desulfotomaculia* class[119]. The rps8 gene of *P. fastidiosa* shares 74.24% sequence identity with its closest DPO MAG relative, *Pelotomaculaceae* SL1 (Fig. 3.1).

Comparing the taxonomic marker genes of DPO MAGs to various genomic databases, it was found that DPOM belong to largely uncultivated and uncharacterized classes of bacteria. The closest cultured relatives of DPO MAGs share 57-95% rpS8 amino acid sequence identity (Fig. 3.1), which surpasses the standard rpS8 species threshold (<98.3%)[120]. Furthermore, multigene classification by GTDB designates these related isolates as belonging to different genera[120]. The 16S rRNA gene is the canonical taxonomic marker for resolving microbial speciation. While not present in all DPO MAGs, 86% (n=18) contained fragments of the 16S rRNA gene, enabling refined taxonomic analyses (Table 3.3). To capture the novelty of enriched DPOM, EMIRGE was used to reconstruct full-length 16S rRNA gene sequences that were subsequently BLAST searched in the SILVA database[123,124]. This approach determined that the DPOM represented 14 new strains, six new species, and one new genus based on standardized relatedness metrics[120]. Proposed names and etymologies are provided (Table 3.4). The novel genus, represented by *Cosmobacter schinkii* SL3 (named in recognition of Bernhard Schink, for his exemplary contributions to microbiology and for his isolation of the first DPOM[18], is the second characterized genus of the

Desulfomonilia_A UBA1062 order, in addition to Ca. *Phosphitivorax*. Consequently, UBA1062 was expanded by this study to include two genera and five species (Fig. 3.1).

Syntrophic Taxa

Using taxonomic classifications to predict physiology is often unreliable, especially for taxonomically novel bacteria with few cultured representatives, such as DPOM. However, it is still worth noting that the closest relatives of all characterized DPO MAGs, regardless of taxonomy, are obligately anaerobic chemotrophs. Furthermore, the *Desulfomonilia_A, Desulfotomaculia*, and *Syntrophia* classes each contain cultured representatives that are canonically dependent on syntrophic associations[71,132–135]. Syntrophic organisms live in coordination with a microbial partner to degrade substrates that are energetically unfavorable, and they are notoriously difficult to grow axenically[136,137].

Pelotomaculaceae was the dominant DPO MAG of the SL1 community and was also enriched in the LM1 community (Fig. 3.1). The *Pelotomaculaceae* are a family of gram-positive, spore-forming, strict anaerobes with five cultured representatives that are all capable of syntrophic propionate oxidation in coculture with methanogens[131]. Three of the cultured *Pelotomaculaceae* are obligate syntrophs, while the closest relative of the enriched *Pelotomaculaceae* DPOM, *P. thermopropionicum* SI, can also grow axenically on a limited number of substrates[131,132].

Syntrophales LM1 is most closely related to the isolates *Syntrophorhabdus aromaticivorans* and *Syntrophus aciditrophicus* (Fig. 3.1). All cultured *Syntrophus* are strict anaerobes that can perform syntrophic fermentations in the presence of methanogens and are also universally capable of crotonate metabolism[135,138].

Three DPO MAGS were part of the *Smithellaceae* family: *Ca. Smithella phosphorovis* LM1, Ca. *Smithella antiquus* SM1, and Ca. *Smithella antiquus* SV1 (Fig. 3.1). Ca. *Smithella antiquus* SM1 was the dominant DPOM in the SM1 community and co-existed with *Sporomusaceae* SM1. Ca. *Smithella antiquus* SV1 co-existed with the dominant DPOM *Phosphitivorax anaerolimi* SV1. Members of the *Smithellaceae* family will syntrophically degrade butyrate, propionate, malate, and fumarate in association with methanogens and can also grow axenically on crotonate. However, propionate oxidation by *Smithellaceae* occurs by an unknown catabolic pathway[133,139].

Thirteen DPO MAGs belonged to the *Desulfomonilia_A* class, which contains no representative isolates. The only previously cultured representatives are Ca. *Phosphitivorax anaerolimi* Phox-21, Ca. *Phosphitivorax anaerolimi* F81, and Ca. *Phosphitivorax butyraticus* R76, which had all been enriched from full-scale anaerobic digesters[70,71]. The genomes of all three Ca. *Phosphitivorax* strains include a complete butyrate beta-oxidation pathway and were annotated with multiple common features of syntrophic bacteria[71]. Hao et al. also observed that Ca. *Phosphitivorax* was the third most abundant genus comprising the known syntrophs across all the wastewater treatment plants of their study, which was designed to selectively enrich for syntrophic bacteria[71].

The phylogenetic relatedness of DPO MAGs to notoriously fastidious syntrophic organisms could explain the difficulty in isolating some DPOM[70]. The potential interactions between DPOM and microbial community members is further explored in **Chapter 5**. Isolation of DPOM is not impossible, especially for those DPOM that are not part of canonically syntrophic

taxa. This has been exemplified by *Desulfotignum phosphitoxidans* FiPS-3 and *Phosphitispora fastidiosa,* which are presently the only two isolates of DPOM[18,75]. However, acknowledging the relatedness of most DPOM to characterized syntrophs may guide future campaigns to isolate novel DPOM.

Taxonomy of DPOM in the SV3 Enrichment Community

The taxonomic identity of DPOM in the SV3 community was of particular interest, as SV3 was the parent enrichment that was used to generate the highly enriched phosphite oxidizing (HEPO) culture that is used for most of the physiology experiments of this dissertation (see **Chapter 2**). A preliminary metagenome sequence of the HEPO culture (**Chapter 2**), which was performed after the removal of rumen fluid, detected two rpS3 genes that were predicted to belong to *Phosphitivorax* DPOM (Fig. 2.4). Consistent with this, short-read sequencing of the SV3 community prior to the removal of rumen fluid identified two species of *Desulfomonilia_A* DPOM*,* based on genome-resolved metagenomics coupled to computational binning of the assembled contigs (Table 2.3). The dominant DPOM in exponential phase (37% relative abundance) was *Phosphitivorax anaerolimi* SV3, which is the same species as Phox-21 (Fig. 3.1). *Phosphitivorax pugnaclava* SV3 was subsidiary to *P. anaerolimi* SV3 but still comprised a significant portion (27.5% relative abundance) of the community in exponential phase (Fig. 2.3). The genomes of *P. anaerolimi* and *P. pugnaclava* were highly similar, with a rpS8 sequence identity of 95%. The high sequence similarity prevented the identification or reconstruction of the 16S rRNA gene from short read data and complicated the separation of the two strains during computational binning of assembled contigs. As such, it is largely unclear which genes belong to which species. However, coverage data helped to identify which genes were shared by the two genomes, as the coverage of shared genes (such as the *ptx-ptd* gene cluster) was approximately equal to the sum of coverage for genes that were not shared (such as the rpS8). While it is unclear which metabolic niche each *Phosphitivorax* species inhabits within the enrichment culture, the temporal relationship between these two *Phosphitivorax* species during the enrichment of HEPO and across a growth curve is explored further in **Chapter 5**.

ptx-ptd Cluster Diversity

DPO activity in FiPS-3 and Phox-21 was attributed to the *ptx-ptd* gene cluster, and only organisms with *ptx-ptd* genes were enriched in this study (Fig. 2.3), positing the *ptx-ptd* gene cluster to be the dominant, or possibly sole, metabolic pathway underlying dissimilatory phosphite oxidation[67,68,70]. To determine the prevalence and diversity of DPOM beyond the enrichments established in this study, the PtxD protein sequence from FiPS-3 was used as a marker gene to query the IMG/M protein sequence space (Table 3.5). I recovered 35 positive hits that were phylogenetically compared to the PtxD from the enriched DPO MAGs of this study, and to the three physiologically verified DPOM species: *Phosphitivorax anaerolimi*, *Desulfotignum phosphitoxidans* FiPS-3, and *Phosphitispora fastidiosa* (Fig. 3.3).

The *ptx-ptd* gene cluster analysis revealed that the DPO-PtxD form a monophyletic clade that includes all physiologically validated DPOM (i.e., FiPS-3, Phox-21, *P. fastidiosa*, and the enriched DPOM) (Fig. 3.3). The DPO-PtxD belongs to the glyoxylate/hydroxypyruvate reductase B (GHRB) protein sub-family of the D-2-hydroxyacid dehydrogenases (2HADH) (Fig. 3.2). The

closest relatives of DPO-PtxD are the sugar dehydrogenases and the PtxD homologs involved in phosphorus assimilation[140]. The DPO PtxD can be distinguished from closely related proteins based on the presence of nearby *ptd* genes (Fig. 3.3). The closest non-DPO homolog (Ga0209611_10199181) of the DPO-PtxD lacks the remaining *ptx-ptd* genes in the inclusion-matrix (Fig. 3.3B), demonstrating the specificity of the custom pHMMs[117].

The DPO PtxD was found exclusively in anoxic environments (Fig. 3.3A). The observed inability of DPOM to occupy oxic environments, despite the thermodynamic favorability of DPO coupled to oxygen respiration ($\Delta G^{o'}$ = -283 kJ mol^{-1} HPO_3^{2-}), suggests that metabolic proteins may be oxygen sensitive. Alternatively, DPO metabolism may be dependent on the biochemical pathways of anaerobes. The phylogenetic diversity of PtxD that was identified in global metagenomic datasets was largely representative of the diversity that was identified in the DPO enrichment series (Fig. 3.3). Consistent with the enrichments, *Desulfomonilia_A* were the most-sampled class of PtxD and were found globally in the anaerobic digester sludge of MWTFs (Fig. 3.3). This global distribution suggests that *Desulfomonilia_A* is specifically selected for in wastewater sludge, while divergent DPOM taxonomy are found in alternative environments (Fig. 3.3). However, future sampling campaigns will be required to correlate DPOM taxonomy with environmental geochemistry.

Evolutionary History

The increased sample size of taxonomically diverse DPOM allowed for computational analyses of DPO evolution. The DPO evolutionary history was ascertained using i) genomic features, ii) comparative taxonomic clustering, and iii) syntenic conservation. Within the DPO-PtxD clade, proteins clustered based on host taxonomy, and the PtxD was distinguishable at the genus level (Fig. 3.3A). The only deviation from this pattern was Ca. *Smithella phosphorovis* LM1 of the *Syntrophia* class, which had a PtxD lineage consistent with Ca. *Phosphitivorax* species of the *Desulfomonilia_A* class (Fig. 3.3A). The *ptx-ptd* cluster from Ca. *S. phosphorovis* LM1 occurred on a single contig (13,378 bp) that hosted an IS91 family transposase. This contig had a sequencing depth (64.7x) three-fold that of the bin's average coverage (19.4x), and the GC content (57.4%) was 3.5% higher than the host genome's mean GC content (53.9%). Together these findings suggest that, like FiPS-3[18], Ca. *S. phosphorovis* LM1 likely acquired its *ptx-ptd* genes through HGT. Consistent with this conclusion, the LM1 community assembly did not include taxonomic marker genes for Ca. *Phosphitivorax* species, and the assembly graph supported the binning results, precluding a different bin-assignment for this contig.

In contrast to Ca. *S. phosphorovis* LM1, most PtxD clustered according to host taxonomy, indicating that most DPOM have likely acquired their PtxD via vertical inheritance (Fig. 3.3A). Similar taxonomic clustering occurred for the PtdC and PtdF, further suggesting that the *ptx-ptd* genes are inherited as a metabolic unit (Figs. 3.4 & 3.5). Tanglegram analyses facilitate a coarse approximation of topological similarity between gene phylogenies, where crossing lines ("tangles") indicate alternative evolutionary histories[141]. Comparisons of the PtxD, PtdC, and PtdF exhibited zero tangling, supporting a linked evolutionary history (Fig. 3.6). Although the phylogenetic trees of individual DPO genes showed alternative branching patterns, this was expected, as genes with functional differences are subject to unique selective pressures.

Synteny provides an alternative metric to gauge the unity of *ptx-ptd* gene evolution because: i) linked genes tend to maintain organization throughout evolutionary history, and ii)

closely related taxa show high genomic stability[142,143]. Synteny graphs show that the individual *ptx* and *ptd* genes were always codirectional in the order *ptxED* and *ptdCF(G)HI,* respectively (Fig. 3.7). However, the directionality between the *ptx* and *ptd* gene clusters was variable, and syntenic variation formed four distinct groups (Groups I-IV) that correlated with host taxonomy (Fig. 3.7). Groups I and IV do not contain *ptdG*, suggesting it is nonessential (Fig. 3.7). While other genes were frequently missing from the *ptx-ptd* cluster, synteny analysis suggests this is due to fragmentation of contigs during the sequencing process (Figs. 3.3 & 3.7). For example, *Synergistaceae* SL3 was identified as a DPOM in the active enrichments, but the pHMM search failed to identify its PtxD (Fig. 3.3). Synteny analysis suggests that this PtxD was truncated downstream of PtxE, which was then confirmed by BLAST alignment (Fig. 3.7).

Searching metagenome databases with additional DPO genes will likely reveal other DPO contigs that were split from the PtxD. This was the case when the IMG/M database was mined for PtdC, which identified five additional contigs with divergent PtxD phylogeny (Figs. 3.4 & 3.6). While these divergent genes may indicate further DPO diversity, their contigs showed non-canonical *ptx-ptd* neighborhoods and are not yet represented by validated DPO cultures. Cultured representatives will be necessary to validate that these divergent genes are capable of conferring DPO activity. For those *ptx-ptd* clusters that do confidently represent DPOM, the predominance of vertical transfer was collectively supported by genomic features, taxonomy, and synteny.

Discussion

DPOM were found to span six classes of three bacterial phyla (*Desulfobacterota, Firmicutes* and *Synergistota*). Yet the taxonomic diversity of DPOM was not monophyletic, as representatives were sparsely distributed across the bacterial tree of life. Such sparse representation across diverse taxa is typically indicative of broad-host-range HGT, but phylogenetic analyses of the *ptx-ptd* gene cluster indicate that vertical transfer is the predominant mechanism of inheritance, since DPO gene evolution mirrored the evolutionary history of host taxonomy. Small variations in synteny further support the correlation between gene order and taxonomy while also suggesting that *ptx-ptd* genes have coevolved as a metabolic unit specialized for DPO metabolism.

Given the diversity of DPOM lineages that inherited the *ptx-ptd* gene cluster, it is tempting to speculate the biological timescale for when DPO metabolism originated. The last common node for all known DPOM suggests that DPOM arose before the divergence of monoderm and diderm bacteria[144]. Mapping the divergence of these clades to geological timescales suggests that DPOM evolved ~3.2 Gya[145], contemporaneously with anoxygenic photosynthesis and ~0.8 Gya after the evolution of methanogenesis[145]. This is consistent with the suggestion that phosphite composed 40-67% of dissolved phosphorus species in Archaean oceans (>3.5 Gya)[32,37]. The half-life of oceanic phosphite under a reducing atmosphere is expected to be 0.1-10 billion years, which would have allowed phosphite to persist on early Earth, possibly supporting a robust chemolithoautotrophic DPO population.

One would expect such an ancient metabolism to be detected more broadly across all bacteria. However, oxygenation of Earth's atmosphere since the great oxidation event (~2.5 Gya) has likely depleted ancient natural phosphite reserves, as oxidizing radicals abiotically oxidize phosphite on geological timescales[34,146]. Phosphite would consequently be too rare for DPO in most contemporary environments, and lack of positive selection would promote widespread gene loss[147]. Yet pockets of phosphite (0.1 – 1.3 mM) exist in diverse contemporary

environments[36,43,148], and assimilatory phosphite oxidizing metabolisms still occur in various habitats on extant Earth[30]. Here, environmental metadata from global metagenomes identified DPOM in multiple anoxic environments that are representative of ancient Earth (i.e., oil reservoirs, deep subsurface aquifers) and in environments that serve as potential examples of contemporary phosphite accumulation (i.e., wastewater sludge, freshwater wetlands). The identification of DPO metabolism across global metagenomes suggests that several environments likely continue to support phosphorus redox cycling.

Microbial phosphite oxidation is likely to be the principal mechanism by which localized phosphite is converted to phosphate at biologically relevant timescales. This is because abiotic phosphite oxidation is a slow process - even under hot (95 °C), aerobic conditions, phosphite remains stable on the order of months[34]. Consequently, in areas of phosphite enrichment, DPO may play a significant role in environmental phosphite oxidation due to the substantially higher rates and quantities of phosphate production compared to assimilatory phosphite oxidation (APO), by which phosphate is exclusively incorporated into biomass[29]. For example, FiPS-3 growing by DPO uses roughly 10 mM of phosphite, producing equimolar amounts of phosphate, whereas APO uses nearly two orders of magnitude less phosphite (0.1 mM) to reach a similar optical density[18]. While the impact of DPO on phosphite depletion and phosphate accumulation in environmental contexts remains uncharacterized, it is known that DPOM enrichment cultures can oxidize up to 500uM/day of phosphite[70,117]. Meanwhile, environmental phosphite concentrations rarely exceed 1mM[35], leading one to speculate that environmental DPO activity could quickly deplete localized phosphite reserves in the absence of substantial phosphite neogenesis.

Characterizing the role of DPO in the phosphorus cycle depends on the geochemical distribution of phosphite, and likewise, the distribution of DPO genes may offer a signal for the accumulation and distribution of phosphite in the environment. While the primary modes of phosphite neogenesis remain unclear, the persistence of DPO metabolism suggests that a modern selective pressure exists to maintain the *ptx-ptd* genes in the gene pool. Ion chromatography has proven successful for laboratory measurements because exogenous phosphite is typically added at millimolar concentrations[18,70,117]. However, this is not representative of environmental conditions, where substrates are typically found at sub-micromolar concentrations[35,70]. Consequently, connecting DPO metabolism to its geochemical context requires advancements in the sensitivity of phosphite detection. Significant headway has been made in this area, as Sadeghi et al. recently developed an IC/MS/MS method for detecting phosphite using an ^{18}O-labeled HPO_3^{-2} internal standard that reaches a detection limit of 0.017ug L^{-1} [115]. Using these advanced detection methods, we may be able to establish the fundamental modes of environmental phosphite synthesis, which could in turn guide geochemical and biological sampling campaigns and contextualize the role of DPO redox chemistry in a broader geochemical framework.

Combining metagenomic signals for DPOM with existing geochemical signals for reduced P could cooperatively identify hot spots for microbial P redox cycling. For example, wastewater digester sludge and wetland sediments are both environments where phosphite and DPOM have been independently identified, jointly providing a signal for environmental P redox activity (Table 3.6). The prevalence of *Desulfomonilia_A* DPOM in MWTFs is a pronounced observation because it is observed across continents[117], implying a role for P redox activity within the human waste management stream. Modern waste management practices have been criticized for unsustainably flushing limited P resources into the ocean[27], and this challenge could be accentuated by highly soluble reduced P compounds such as phosphite (1000x more soluble than phosphate)[149].

Meanwhile, wetlands cover about 7% of the Earth's surface, and their microbial activity plays a critical role in nutrient cycling[150]. For example, rice paddies are a form of artificial and temporal wetland with high methanogenic activity and significant implications in global climate change[151]. They also serve as a prime example of active P redox cycling, as they have been repeatedly observed to generate the reduced-P compound phosphine[152,153]. Phosphine can indirectly enhance the greenhouse effect by increasing the residence time of greenhouse gases, thus potentially amplifying the environmental impacts of wetland and paddy biogeochemical activity[51]. While phosphine is not the substrate for DPO, phosphine formation and degradation are likely to include the intermediates of phosphite and hypophosphite, which could support DPO activity[33]. While biogeochemical data suggest P redox is an active feature of wastewaters and wetlands, much work remains to better quantify that role. Meanwhile, novel sites of DPOM activity are likely to be found in environments with detectable levels of phosphite where DPOM have not yet been identified (e.g., rivers, lakes, geothermal water) (Table 3.6). Because the detection of phosphine frequently indicates the presence of other reduced-P compounds (e.g., phosphite), any reduced-P compound might serve as an informative signal for the potential presence of DPOM.

Subsurface groundwater is the only environment to date in which DPOM have been detected but reduced P has not, and this environment may exemplify an example of cryptic redox cycling of P (Table 3.6). As is descriptive of reduced P compounds, the substrates of cryptic cycles remain undetectably low or static due to microbial turnover and frequently represent geochemical intermediates with poorly understood redox chemistry[154]. The subsurface was highlighted by Figueroa et al. to be a likely location for phosphite neogenesis and a promising ecosystem for DPOM[29], and Ruttenberg et al. have emphasized the subsurface as an important black box for the transport of P to marine and oceanic systems[28]. DPOM in the subsurface may consequently be critical in subsurface P cycling while also performing primary production in this aphotic and energy-limited environment[29].

Figures

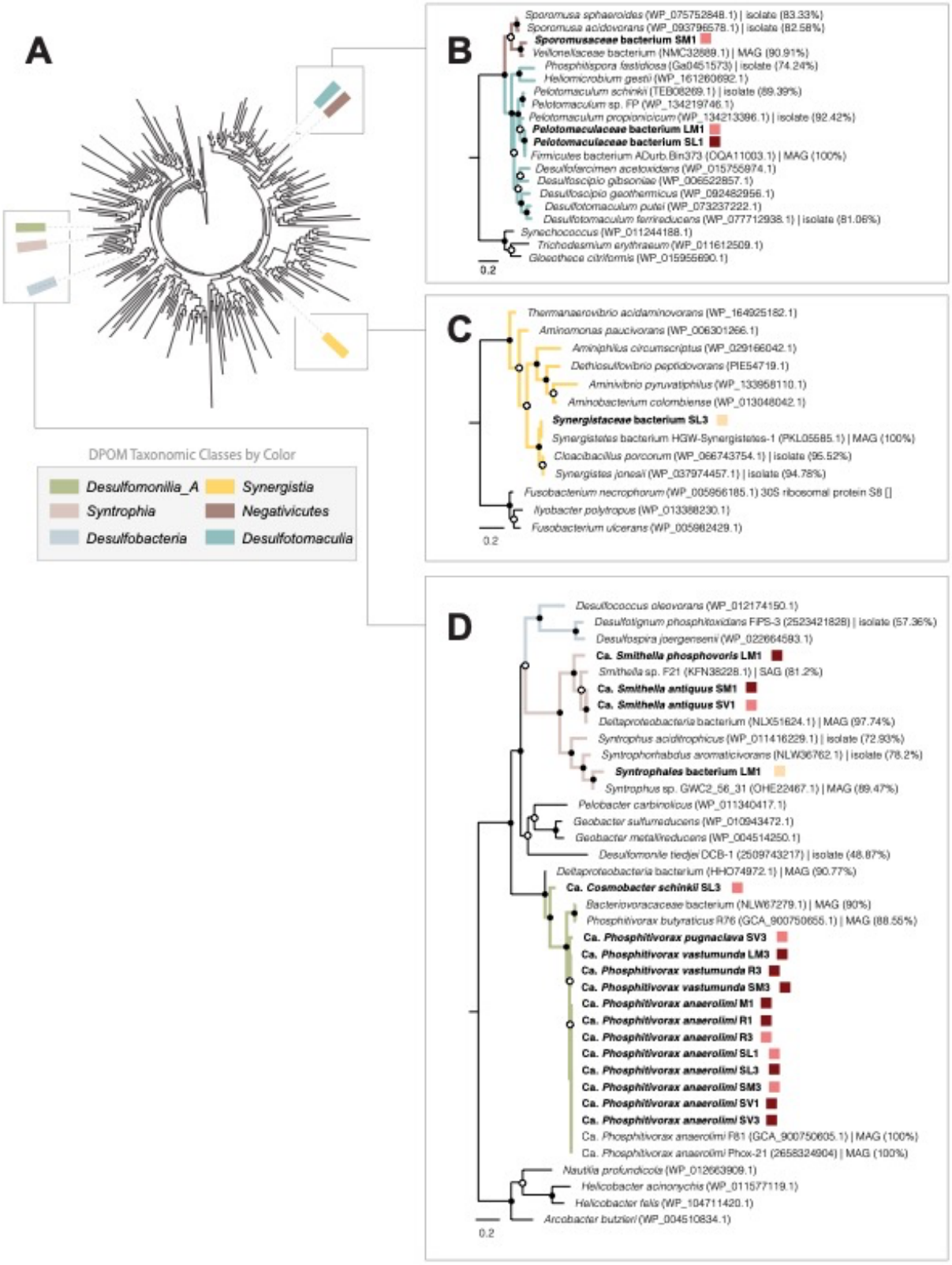

A
B
Sporomusa sphaeroides (WP_075752848.1) | isolate (83.33%)
Sporomusa acidovorans (WP_093796578.1) | isolate (82.58%)
Sporomusaceae bacterium SM1
Veillonellaceae bacterium (NMC32889.1) | MAG (90.91%)
Phosphitispora fastidiosa (Ga0451573) | isolate (74.24%)
Heliomicrobium gestii (WP_161260692.1)
Pelotomaculum schinkii (TEB08269.1) | isolate (89.39%)
Pelotomaculum sp. FP (WP_134219746.1)
Pelotomaculum propionicicum (WP_134213396.1) | isolate (92.42%)
Pelotomaculaceae bacterium LM1
Pelotomaculaceae bacterium SL1
Firmicutes bacterium ADurb.Bin373 (OQA11003.1) | MAG (100%)
Desulfofarcimen acetoxidans (WP_015755874.1)
Desulfoscipio gibsoniae (WP_006522857.1)
Desulfoscipio geothermicus (WP_092482956.1)
Desulfotomaculum putei (WP_073237222.1)
Desulfotomaculum ferrireducens (WP_077712938.1) | isolate (81.06%)
Synechococcus (WP_011244188.1)
Trichodesmium erythraeum (WP_011612509.1)
Gloeothece citriformis (WP_015955690.1)
0.2
C
Thermanaerovibrio acidaminovorans (WP_164925182.1)
Aminomonas paucivorans (WP_006301266.1)
Aminiphilus circumscriptus (WP_029166042.1)
Dethiosulfovibrio peptidovorans (PIE54719.1)
Aminivibrio pyruvatiphilus (WP_133958110.1)
Aminobacterium colombiense (WP_013048042.1)
Synergistaceae bacterium SL3
Synergistetes bacterium HGW-Synergistetes-1 (PKL05585.1) | MAG (100%)
Cloacibacillus porcorum (WP_066743754.1) | isolate (95.52%)
Synergistes jonesii (WP_037974457.1) | isolate (94.78%)
Fusobacterium necrophorum (WP_005956185.1) 30S ribosomal protein S8 []
Ilyobacter polytropus (WP_013388230.1)
Fusobacterium ulcerans (WP_005982429.1)
0.2
D
Desulfococcus oleovorans (WP_012174150.1)
Desulfotignum phosphitoxidans FIPS-3 (2523421828) | isolate (57.36%)
Desulfospira joergensenii (WP_022664583.1)
Ca. Smithella phosphovoris LM1
Smithella sp. F21 (KFN38228.1) | SAG (81.2%)
Ca. Smithella antiguus SM1
Ca. Smithella antiguus SV1
Deltaproteobacteria bacterium (NLX51624.1) | MAG (97.74%)
Syntrophus acidotrophicus (WP_011416229.1) | isolate (72.93%)
Syntrophorhabdus aromaticivorans (NLW36762.1) | isolate (78.2%)
Syntrophales bacterium LM1
Syntrophus sp. GWC2_56_31 (OHE22467.1) | MAG (89.47%)
Pelobacter carbinolicus (WP_011340417.1)
Geobacter sulfurreducens (WP_010943472.1)
Geobacter metallireducens (WP_004514250.1)
Desulfomonile tiedjei DCB-1 (2509743217) | isolate (48.87%)
Deltaproteobacteria bacterium (HHO74972.1) | MAG (90.77%)
Ca. Cosmobacter schinkii SL3
Bacteriovoracaceae bacterium (NLW67279.1) | MAG (90%)
Phosphitivorax butyraticus R76 (GCA_900750655.1) | MAG (88.55%)
Ca. Phosphitivorax pugnaclava SV3
Ca. Phosphitivorax vastumunda LM3
Ca. Phosphitivorax vastumunda R3
Ca. Phosphitivorax vastumunda SM3
Ca. Phosphitivorax anaerolimi M1
Ca. Phosphitivorax anaerolimi R1
Ca. Phosphitivorax anaerolimi R3
Ca. Phosphitivorax anaerolimi SL1
Ca. Phosphitivorax anaerolimi SL3
Ca. Phosphitivorax anaerolimi SM3
Ca. Phosphitivorax anaerolimi SV1
Ca. Phosphitivorax anaerolimi SV3
Ca. Phosphitivorax anaerolimi F61 (GCA_900750605.1) | MAG (100%)
Ca. Phosphitivorax anaerolimi Phox-21 (2658324904) | MAG (100%)
Nautilia profundicola (WP_012663909.1)
Helicobacter acinonychis (WP_011577119.1)
Helicobacter felis (WP_104711420.1)
Arcobacter butzleri (WP_004510834.1)
0.2
DPOM Taxonomic Classes by Color
Desulfomonilia_A
Synergistia
Syntrophia
Negativicutes
Desulfobacteria
Desulfotomaculia

Figure 3.1: Phylogenetic trees of DPO MAGs. (A) A phylogenetic tree of bacterial genomes from the GTDB was visualized with AnnoTree[155]. Nodes of the tree represent class-level taxonomy, and those nodes with DPO organisms are highlighted according to the key. (B-D) Phylogenetic trees of the rpS8 marker gene showing the relationship of DPO MAGs to their closest relatives. Panels depict DPO MAGs belonging to the same phyla: (B) *Firmicutes*; (C) *Synergistota*; (D) *Desulfobacterota*. The DPO MAGs from this study are bolded. Colored squares represent their dominance rank from Fig. 2.3. Each close relative is annotated with its species name, accession number and genome-source type (isolate vs. MAG), as well as its percent identity to the most closely related DPO MAG from this study. Clades are colored and labeled by taxonomic class. Internal nodes with bootstrap support of >90% are indicated by closed circles and those with support of >70% by open circles. Scale bars: 0.2 change per amino acid residue.

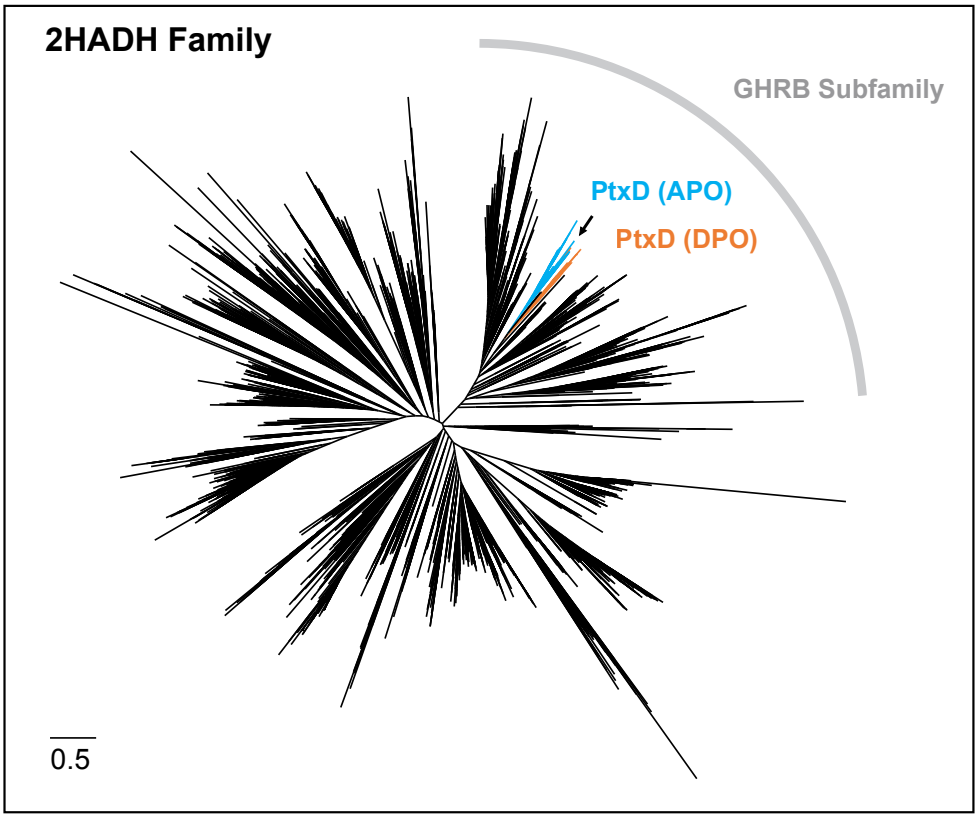


Figure 3.2: Phylogenetic tree of 2HADH family of proteins. The PtxD from IMG/M metagenomes and DPO MAGs were aligned with proteins from the 2-hydroxyacid dehydrogenase (2HADH) family (Pfam PF00389, set representative proteomes to 15%). Protein subfamilies were assigned based on Matelska et al.[140]. An arrow indicates the location of PtxD proteins that are associated with DPO PtdC but clade with assimilatory phosphite oxidation PtxD (APO). Scale bar: 0.5 change per amino acid residue. GHRB: glyoxylate/hydroxypyruvate reductase B protein family.

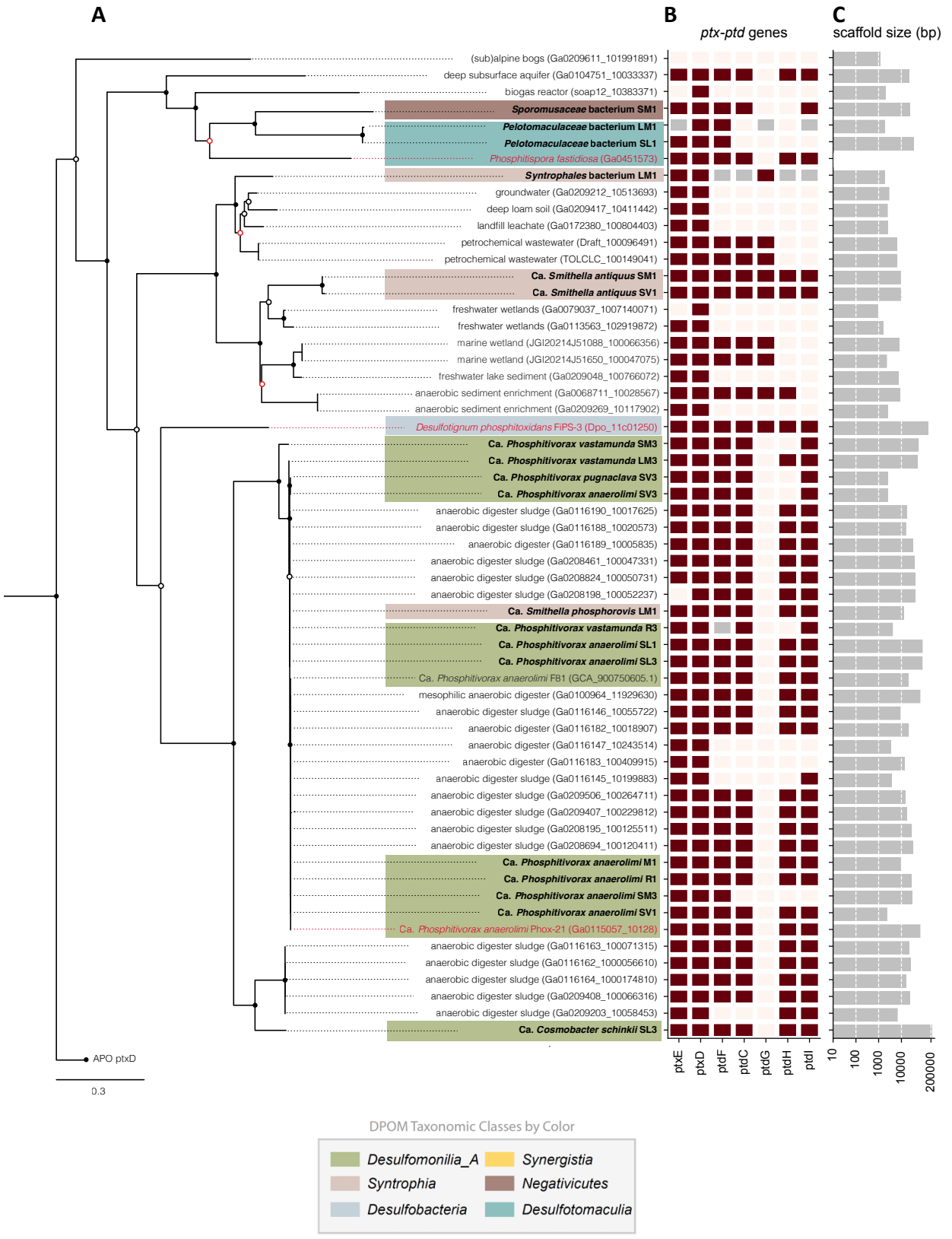

A
B
ptx-ptd genes
C
scaffold size (bp)
(sub)alpine bogs (Ga0209611_101991891)
deep subsurface aquifer (Ga0104751_10033337)
biogas reactor (soap12_10383371)
Sporomusaceae bacterium SM1
Pelotomaculaceae bacterium LM1
Pelotomaculaceae bacterium SL1
Phosphitispora fastidiosa (Ga0451573)
Syntrophales bacterium LM1
groundwater (Ga0209212_10513693)
deep loam soil (Ga0209417_10411442)
landfill leachate (Ga0172380_100804403)
petrochemical wastewater (Draft_100096491)
petrochemical wastewater (TOLCLC_100149041)
Ca. Smithella antiquus SM1
Ca. Smithella antiquus SV1
freshwater wetlands (Ga0079037_1007140071)
freshwater wetlands (Ga0113563_102919872)
marine wetland (JGI20214J51088_100066356)
marine wetland (JGI20214J51650_100047075)
freshwater lake sediment (Ga0209048_100766072)
anaerobic sediment enrichment (Ga0068711_10028567)
anaerobic sediment enrichment (Ga0209269_10117902)
Desulfotignum phosphitoxidans FiPS-3 (Dpo_11c01250)
Ca. Phosphitivorax vastamunda SM3
Ca. Phosphitivorax vastamunda LM3
Ca. Phosphitivorax pugnaclava SV3
Ca. Phosphitivorax anaerolimi SV3
anaerobic digester sludge (Ga0116190_10017625)
anaerobic digester sludge (Ga0116188_10020573)
anaerobic digester (Ga0116189_10005835)
anaerobic digester sludge (Ga0208461_100047331)
anaerobic digester sludge (Ga0208824_100050731)
anaerobic digester sludge (Ga0208198_100052237)
Ca. Smithella phosphorovis LM1
Ca. Phosphitivorax vastamunda R3
Ca. Phosphitivorax anaerolimi SL1
Ca. Phosphitivorax anaerolimi SL3
Ca. Phosphitivorax anaerolimi F81 (GCA_900750605.1)
mesophilic anaerobic digester (Ga0100964_11929630)
anaerobic digester sludge (Ga0116146_10055722)
anaerobic digester (Ga0116182_10018907)
anaerobic digester (Ga0116147_10243514)
anaerobic digester (Ga0116183_100409915)
anaerobic digester sludge (Ga0116145_10199883)
anaerobic digester sludge (Ga0209506_100264711)
anaerobic digester sludge (Ga0209407_100229812)
anaerobic digester sludge (Ga0208195_100125511)
anaerobic digester sludge (Ga0208694_100120411)
Ca. Phosphitivorax anaerolimi M1
Ca. Phosphitivorax anaerolimi R1
Ca. Phosphitivorax anaerolimi SM3
Ca. Phosphitivorax anaerolimi SV1
Ca. Phosphitivorax anaerolimi Phox-21 (Ga0115057_10128)
anaerobic digester sludge (Ga0116163_100071315)
anaerobic digester sludge (Ga0116162_1000056610)
anaerobic digester sludge (Ga0116164_1000174810)
anaerobic digester sludge (Ga0209408_100066316)
anaerobic digester sludge (Ga0209203_10058453)
Ca. Cosmobacter schinkii SL3
APO ptxD
0.3
ptxE
ptxD
ptdF
ptdC
ptdG
ptdH
ptdI
10
100
1000
10000
200000
DPOM Taxonomic Classes by Color
Desulfomonilia_A
Synergistia
Syntrophia
Negativicutes
Desulfobacteria
Desulfotomaculia

Figure 3.3: Phylogenetic tree of the phosphite dehydrogenase PtxD. (A) Refined tree of all PtxD within the DPO-PtxD clade. PtxD from the IMG/M are in light black font and labeled with their source environment and scaffold ID. PtxD from the enriched DPO MAGs are bolded and labeled with their bacterial host name. PtxD that belong to a binned organism are highlighted based on their taxonomic class. Published organisms with validated DPO activity are in red font. Only genes adhering to the IMG/M data usage policy are shown. Internal nodes with bootstrap support of >90% are indicated by closed circles; those with support of >70% by open circles; those with a support of less than 70% by open red circles. Scale bar: 0.2 change per amino acid residue. (B) The presence (maroon) or absence (light pink) of *ptx-ptd* genes in each genome was determined using custom pHMMs. Genes that were absent from a DPO MAG but present in the assembly are in grey, where phylogeny, tanglegrams, and synteny were collectively used to predict the most likely host. (C) Horizontal grey bars display the size (bp) of the contig on which each PtxD was found and are in logarithmic scale to visualize the full range of contig lengths. The minimum length for all seven *ptx-ptd* genes to be present is 7137 bp, based on FiPS-3 sequences. bp, base pair.

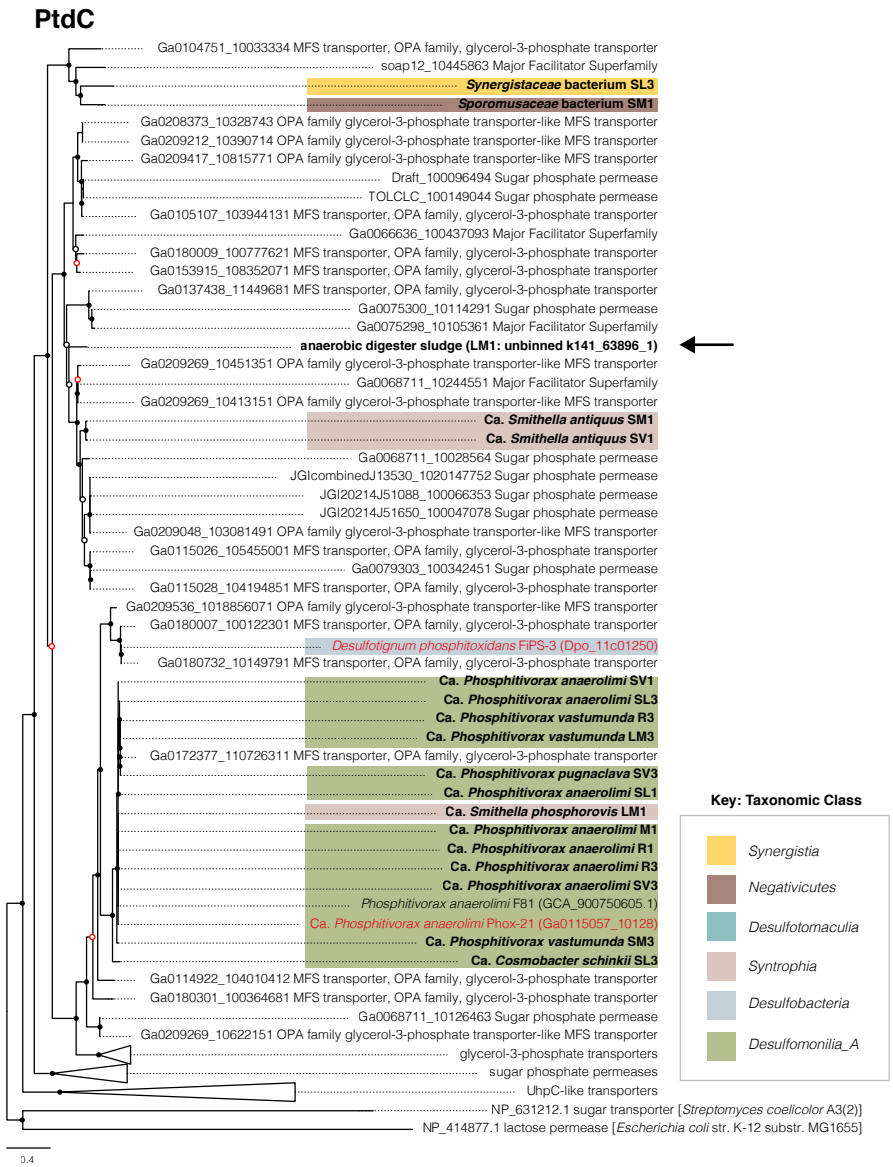


Figure 3.4: **Phylogeny of PtdC in relation to host taxonomy**. The PtdC from FiPS-3 was BLAST searched in the IMG/M database, and protein sequences were aligned with genes from DPO MAGs using MUSCLE[121]. A maximum likelihood tree was constructed with 1000 bootstrap resamplings using FastTree[122]. Genes from the IMG/M are in light black font; genes from the enriched DPO MAGs are bolded. Genes that belong to binned organisms are highlighted based on their taxonomic class. Arrows indicate genes from my enrichments that were identified in an assembly, but not binned in a DPO MAG. Published organisms with validated DPO activity are in red font. *Phosphitispora fastidiosa* was identified during the writing of this dissertation and was not included here[75]. Only genes adhering to the IMG/M data usage policy are shown. Internal nodes with bootstrap support of >70% are indicated by closed black circles, those with support of >50% by open black circles, and those with support of <50% by open red circles. Scale bars: 0.4 change per amino acid residue.

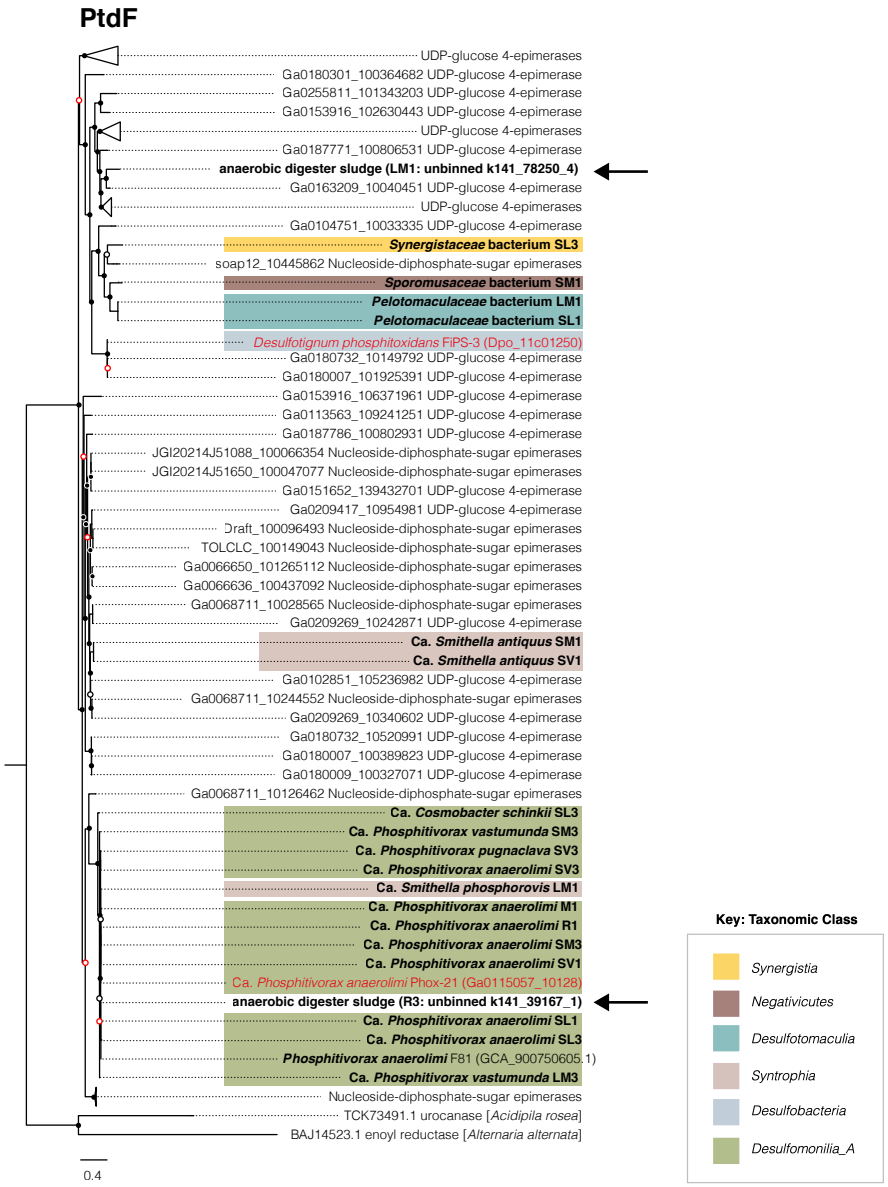


Figure 3.5: **Phylogeny of PtdF in relation to gost taxonomy**. The PtdF from FiPS-3 was BLAST searched in the IMG/M database, and protein sequences were aligned with genes from DPO MAGs using MUSCLE[121]. A maximum likelihood tree was constructed with 1000 bootstrap resamplings using FastTree[122]. Genes from the IMG/M are in light black font; genes from the enriched DPO MAGs are bolded. Genes that belong to binned organisms are highlighted based on their taxonomic class. Arrows indicate genes from my enrichments that were identified in an assembly, but not binned in a DPO MAG. Published organisms with validated DPO activity are in red font. *Phosphitispora fastidiosa* was identified during the writing of this dissertation and was not included here[75]. Only genes adhering to the IMG/M data usage policy are shown. Internal nodes with bootstrap support of >70% are indicated by closed black circles, those with support of >50% by open black circles, and those with support of <50% by open red circles. Scale bars: 0.4 change per amino acid residue.

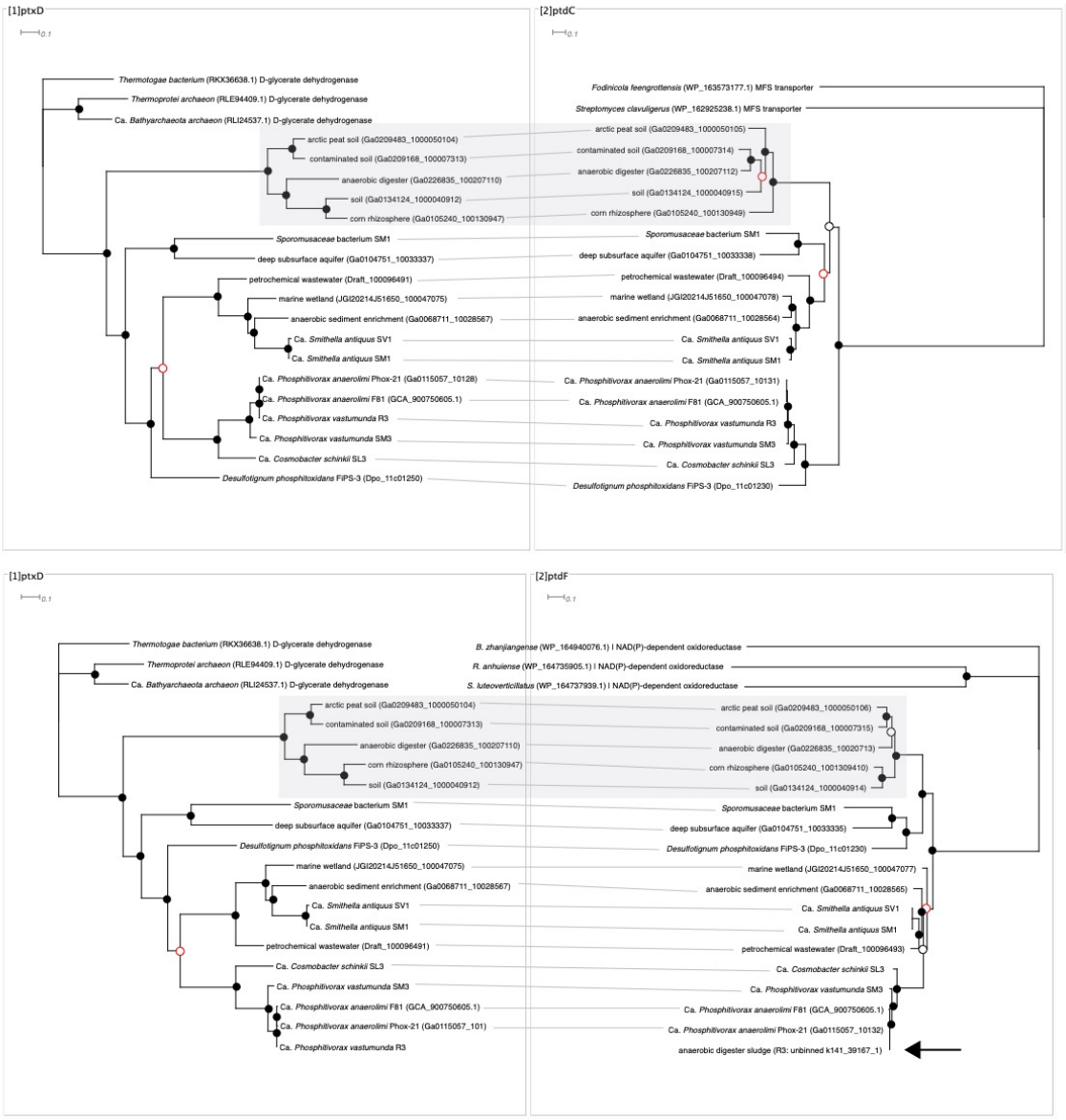


Figure 3.6: Tanglegrams of PtxD, PtdC and PtdF. Tanglegrams were created to compare the evolutionary history of PtxD to that of PtdC and PtdF. Each gene was BLAST searched in the IMG/M database, and protein sequences were aligned with genes from DPO MAGs using MUSCLE[121]. Maximum likelihood trees were constructed with 1000 bootstrap resamplings using FastTree[122]. Only genes adhering to the IMG/M data usage policy are shown. Trees were simplified by de-replication, and genes occurring on the same contig are connected by a line. Only those genes whose tanglegram partner was present are included. *Phosphitispora fastidiosa* was identified during the writing of this dissertation and was not included here[75]. Divergent ptx-ptd genes without confirmed DPO activity are highlighted in grey. Internal nodes with bootstrap support of >70% are indicated by closed black circles, those with support of >50% by open black circles, and those with support of <50% by open red circles. Scale bars: 0.1 change per amino acid residue.

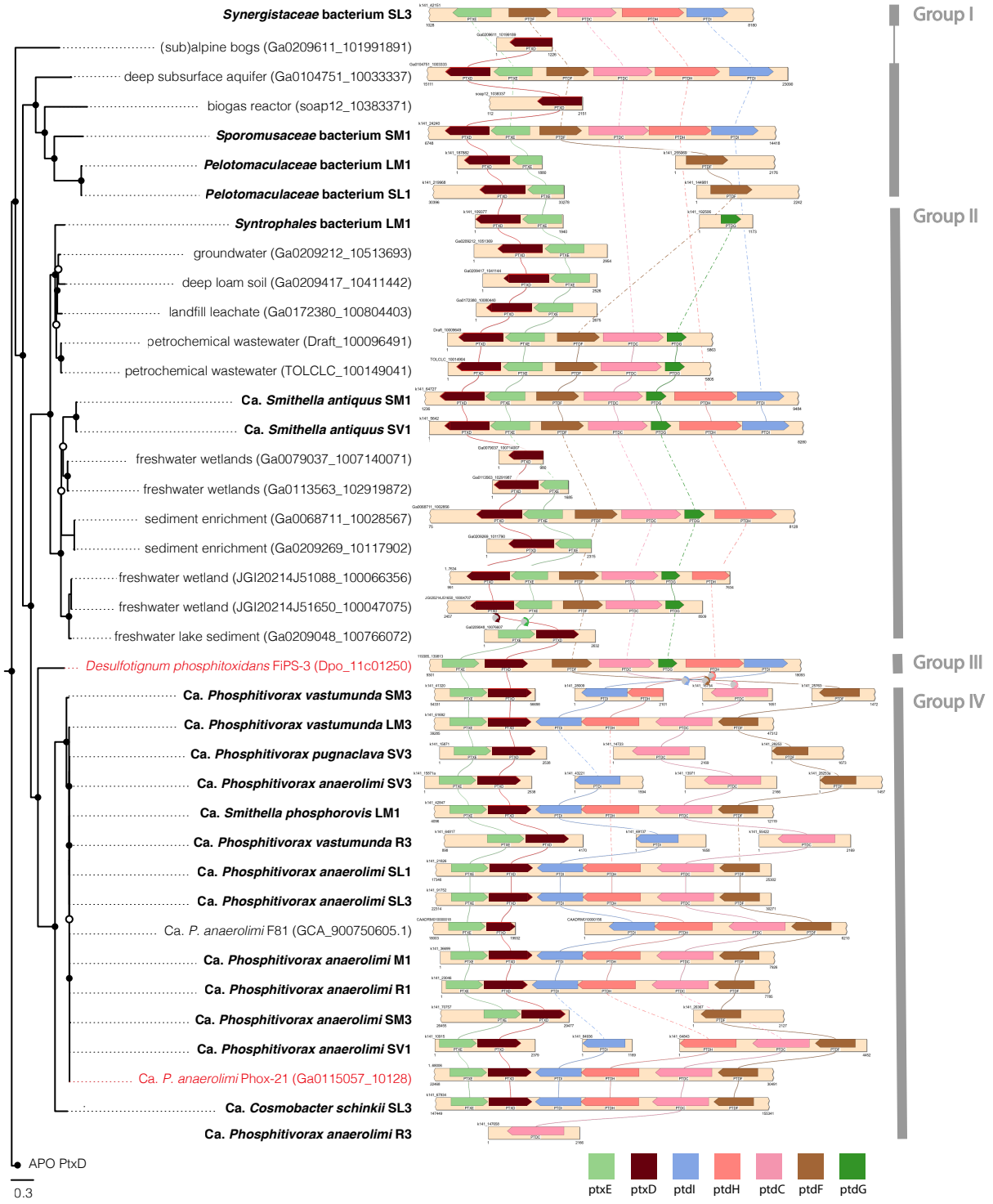


Figure 3.7: **Synteny of *ptx-ptd* gene cluster.** The *ptx-ptd* genes from DPO MAGs and contigs from the IMG/M were identified using custom built HMMs. Synteny was visualized with SimpleSynteny, a platform which displays genes (colored arrows) on their respective contigs

(beige rectangles) and connects homologous genes by a line[128]. Gene inversions are marked by circular arrows. Contigs were aligned with the PtxD phylogenetic tree from Figure 3.3. *Phosphitispora fastidiosa* was identified during the writing of this dissertation and was not included here[75]. Two DPO contigs without the PtxD are visualized outside the bounds of the PtxD tree and in accordance with their host-taxonomy and synteny patterns. Groups with identical synteny are labelled as Group I-IV by gray bars on the right.

Table 3.1: GTDB classification of DPOM genomes. As part of a multi-method taxonomic novelty and naming schema, each DPO MAG was subjected to taxonomic classification via the GTDB-Tk, rpS8, and 16S rRNA gene[118,120,124]. This table shows the taxonomic classification assigned to each DPO MAG using the GTDB-Tk. Each row represents a DPO MAG. The second column lists the full phylogeny assigned by the GTDB-Tk, where taxonomic ranks in red do not yet have an isolated representative. The last column identifies that taxonomic rank at which no cultured representatives have been identified in the GTDB database.

Full ID	GTDB-tK classification	GTDB novelty rank
LM1-Bin_11_1	d__Bacteria;p__Desulfobacterota;c__Syntrophia;o__Syntrophales;f__UBA5619;g__UBA5619;s__	species
LM1-Bin_12_1	d__Bacteria;p__Firmicutes_B;c__Desulfotomaculia;o__Desulfotomaculales;f__Pelotomaculaceae;g__UBA4789;s__	species
LM1-Bin_42_D	d__Bacteria;p__Desulfobacterota;c__Syntrophia;o__Syntrophales;f__Smithellaceae;g__UBA8904;s__	species
LM3-Bin_1_merge	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__	species
M1-Bin_6_1	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__UBA1062 sp001896555	strain
R1-Bin_1	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__UBA1062 sp001896555	strain
R3-Bin_35_1	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__UBA1062 sp001896555	strain
R3-Bin_35_2	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__	species
SL1-Bin_31_2	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__UBA1062 sp001896555	strain
SL1-Bin_34_2	d__Bacteria;p__Firmicutes_B;c__Desulfotomaculia;o__Desulfotomaculales;f__Pelotomaculaceae;g__UBA4789;s__	species
SL3-Bin_22_1	d__Bacteria;p__Synergistota;c__Synergistia;o__Synergistales;f__Synergistaceae;g__Syner-03;s__Syner-03 sp002839185	strain
SL3-Bin_22_2	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__;s__	genus
SL3-Bin_41_1	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__UBA1062 sp001896555	strain
SM1-Bin_11_1	d__Bacteria;p__Firmicutes_C;c__Negativicutes;o__UBA1444;f__UBA1444;g__UBA1444;s__	species
SM1-Bin_3_1	d__Bacteria;p__Desulfobacterota;c__Syntrophia;o__Syntrophales;f__Smithellaceae;g__UBA8904;s__UBA8904 sp002304995	strain
SM3-Bin_1	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__UBA1062 sp001896555	strain
SM3-Bin_3_merge	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__	species
SV1-Bin_1_1	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__UBA1062 sp001896555	strain
SV1-Bin_3_4	d__Bacteria;p__Desulfobacterota;c__Syntrophia;o__Syntrophales;f__Smithellaceae;g__UBA8904;s__UBA8904 sp002304995	strain
SV3-Bin_1_1	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__	species
SV3-Bin_3_1	d__Bacteria;p__Desulfobacterota;c__Desulfomonilia_A;o__UBA1062;f__UBA1062;g__UBA1062;s__UBA1062 sp001896555	strain

Table 3.2: rps8 classification of DPOM genomes. As part of a multi-method taxonomic novelty and naming schema, each DPO MAG was subjected to taxonomic classification via the GTDB-Tk, rpS8, and 16S rRNA gene[118,120,124]. This table shows the taxonomic classification assigned to each DPO MAG using the rpS8 marker gene. Each row represents a DPO MAG. The second column lists the closest relative as determined by BLAST searching the NCBI database with the rpS8 marker gene. The fourth column indicates the percent sequence identity between the DPO MAG rpS8 gene and its closest relative in the NCBI. The last column identifies that taxonomic rank at which no representatives have been identified in the NCBI database, based on the standard threshold for rpS8 (novel species < 98.3%).

Full ID	rpS8 closest relative	rps8 closest relative % id	rpS8 novelty (S < 98.3%)
LM1-Bin_11_1	OHE22467.1	89.47	species +
LM1-Bin_12_1	OQA11003.1	100	strain
LM1-Bin_42_D	KFN38228.1	81.2	species +
LM3-Bin_1_merge	Ca. Phosphitivorax anaerolimi Phox-21	97.69	species +
M1-Bin_6_1	Ca. Phosphitivorax anaerolimi Phox-21	100	strain
R1-Bin_1	Ca. Phosphitivorax anaerolimi Phox-21	100	strain
R3-Bin_35_1	Ca. Phosphitivorax anaerolimi Phox-21	100	strain
R3-Bin_35_2	Ca. Phosphitivorax anaerolimi Phox-21	97.69	species +
SL1-Bin_31_2	Ca. Phosphitivorax anaerolimi Phox-21	100	strain
SL1-Bin_34_2	OQA11003.1	100	strain
SL3-Bin_22_1	PKL05585.1	100	strain
SL3-Bin_22_2	HHO74972.1	90.77	species +
SL3-Bin_41_1	Ca. Phosphitivorax anaerolimi Phox-21	100	strain
SM1-Bin_11_1	NMC32889.1	90.91	species +
SM1-Bin_3_1	NLX51624.1	97.74	species +
SM3-Bin_1	Ca. Phosphitivorax anaerolimi Phox-21	100	strain
SM3-Bin_3_merge	Ca. Phosphitivorax anaerolimi Phox-21	97.69	species +
SV1-Bin_1_1	Ca. Phosphitivorax anaerolimi Phox-21	100	strain
SV1-Bin_3_4	NLX51624.1	97.74	species +
SV3-Bin_1_1	Ca. Phosphitivorax anaerolimi Phox-21	95.38	species +
SV3-Bin_3_1	Ca. Phosphitivorax anaerolimi Phox-21	100	strain

Table 3.3: 16S rRNA gene classification of DPOM genomes. As part of a multi-method taxonomic novelty and naming schema, each DPO MAG was subjected to taxonomic classification via the GTDB-Tk, rpS8, and 16S rRNA gene [118,120,124]. This table shows the taxonomic classification assigned to each DPO MAG using the reconstructed 16S rRNA marker gene. (A) Each row represents a DPO MAG. The 16S rRNA gene fragments from each DPO MAG were reconstructed using EMIRGE, and statistics for the sequence reconstruction are listed for each DPO MAG. (B) Each row represents a DPO MAG. The second column lists the full phylogeny assigned by the SILVA taxonomic database. The third column indicates the percent sequence identity between the reconstructed 16S rRNA gene from each DPO MAG and its closest relative in SILVA. The fourth column identifies the culture status of the closest relative. The last column represents the taxonomic rank at which no representatives have been identified in the SILVA database, based on the standard threshold for 16S rRNA gene (16S, novel species < 96.7%, novel genus < 94%).

A

Full ID	16S Bin Fragment Length	16S EMIRGE reconstructed length	% identity to Fragment
LM1-Bin_11_1	#N/A	#N/A	#N/A
LM1-Bin_12_1	#N/A	#N/A	#N/A
LM1-Bin_42_D	1190	1527	100
LM3-Bin_1_merge	1440	1526	99.93
M1-Bin_6_1	1457	1526	99.86
R1-Bin_1	1432	1526	99.79
R3-Bin_35_1	435	1526	100
R3-Bin_35_2	438	1526	97.47
SL1-Bin_31_2	1431	1533	99.09
SL1-Bin_34_2	#N/A	#N/A	#N/A
SL3-Bin_22_1	114	1440	99.07
SL3-Bin_22_2	231	1542	100
SL3-Bin_41_1	1447	1526	99.93
SM1-Bin_11_1	65	1511	100
SM1-Bin_3_1	1257	1530	99.76
SM3-Bin_1	382	1526	97.64
SM3-Bin_3_merge	1207	1526	98.76
SV1-Bin_1_1	1435	1526	99.93
SV1-Bin_3_4	1464	1530	98.64
SV3-Bin_1_1	372	1526	97.31
SV3-Bin_3_1	371	1526	100

B

Full ID	16S EMIRGE SILVA Taxonomy	16S SILVA %id	status	16S novelty (S < 96.7%, G < 94%))
LM1-Bin_11_1	#N/A	#N/A	#N/A	#N/A
LM1-Bin_12_1	#N/A	#N/A	#N/A	#N/A
LM1-Bin_42_D	Bacteria;Desulfobacterota;Syntrophia;Syntrophales;Smithellaceae;Smithella;	97.6463	uncultured	strain
LM3-Bin_1_merge	Bacteria;SAR324 clade(Marine group B);	90.4545	uncultured	genus
M1-Bin_6_1	Bacteria;SAR324 clade(Marine group B);	90.5844	uncultured	genus
R1-Bin_1	Bacteria;SAR324 clade(Marine group B);	90.6493	uncultured	genus
R3-Bin_35_1	Bacteria;SAR324 clade(Marine group B);	90.5783	uncultured	genus
R3-Bin_35_2	Bacteria;SAR324 clade(Marine group B);	90.5783	uncultured	genus
SL1-Bin_31_2	Bacteria;SAR324 clade(Marine group B);	90.0585	uncultured	genus
SL1-Bin_34_2	#N/A	#N/A	#N/A	#N/A
SL3-Bin_22_1	Bacteria;Synergistota;Synergistia;Synergistales;Synergistaceae;uncultured;	97.9476	uncultured	strain
SL3-Bin_22_2	Bacteria;SAR324 clade(Marine group B);	91.4174	uncultured	genus
SL3-Bin_41_1	Bacteria;SAR324 clade(Marine group B);	90.8382	uncultured	genus
SM1-Bin_11_1	Bacteria;Firmicutes;Negativicutes;Veillonellales-Selenomonadales;Sporomusaceae;uncultured;	96.36	uncultured	species
SM1-Bin_3_1	Bacteria;Desulfobacterota;Syntrophia;Syntrophales;Smithellaceae;Smithella;	98.8186	uncultured	strain
SM3-Bin_1	Bacteria;SAR324 clade(Marine group B);	90.78	uncultured	genus
SM3-Bin_3_merge	Bacteria;SAR324 clade(Marine group B);	90.78	uncultured	genus
SV1-Bin_1_1	Bacteria;SAR324 clade(Marine group B);	90.65	uncultured	genus
SV1-Bin_3_4	Bacteria;Desulfobacterota;Syntrophia;Syntrophales;Smithellaceae;Smithella;	98.3322	uncultured	strain
SV3-Bin_1_1	Bacteria;SAR324 clade(Marine group B);	90.7792	uncultured	genus
SV3-Bin_3_1	Bacteria;SAR324 clade(Marine group B);	90.7792	uncultured	genus

Table 3.4: Taxonomic naming and etymology. (A) By comparing data between Table 3.1, Table 3.2, and Table 3.3, the lowest rank for which each DPO MAG was novel between all three marker gene databases (GTDB-Tk, rpS8, and 16S rRNA gene) is listed in the 'Lowest Novelty Rank' column. The proposed name is based on the level of novelty across all three databases. (B) Etymology for each DPO MAG were based on all three classification systems.

A

Full ID	Lowest Novelty Rank	Rank of closest cultured representative	Established Genus Name	Established species Name	Proposed Name
LM1-Bin_11_1	species	order	UBA5619	NA	*Syntrophales bacterium* LM1
LM1-Bin_12_1	strain	family	UBA4789	NA	*Pelotomaculaceae bacterium* LM1
LM1-Bin_42_D	strain	family	Smithella	NA	*Ca. Smithella phosphorovis* LM1
LM3-Bin_1_merge	species	class	Phosphitivorax	NA	*Ca. Phosphitivorax vastumunda* LM3
M1-Bin_6_1	strain	class	Phosphitivorax	anaerolimi	*Ca. Phosphitivorax anaerolimi* M1
R1-Bin_1	strain	class	Phosphitivorax	anaerolimi	*Ca. Phosphitivorax anaerolimi* R1
R3-Bin_35_1	strain	class	Phosphitivorax	anaerolimi	*Ca. Phosphitivorax anaerolimi* R3
R3-Bin_35_2	species	class	Phosphitivorax	NA	*Ca. Phosphitivorax vastumunda* R3
SL1-Bin_31_2	strain	class	Phosphitivorax	anaerolimi	*Ca. Phosphitivorax anaerolimi* SL1
SL1-Bin_34_2	strain	family	UBA4789	NA	*Pelotomaculaceae bacterium* SL1
SL3-Bin_22_1	strain	family	Syner-03	Syner-03 sp002839185	*Synergistaceae bacterium* SL3
SL3-Bin_22_2	genus	class	NA	NA	*Ca. Cosmobacter schinkii SL3*
SL3-Bin_41_1	strain	class	Phosphitivorax	anaerolimi	*Ca. Phosphitivorax anaerolimi* SL3
SM1-Bin_11_1	species	class	UBA1444	NA	*Sporomusaceae bacterium* SM1
SM1-Bin_3_1	strain	family	Smithella	UBA8904 sp002304995	*Ca. Smithella antiquus* SM1
SM3-Bin_1	strain	class	Phosphitivorax	anaerolimi	*Ca. Phosphitivorax anaerolimi* SM3
SM3-Bin_3_merge	species	class	Phosphitivorax	NA	*Ca. Phosphitivorax vastumunda* SM3
SV1-Bin_1_1	strain	class	Phosphitivorax	anaerolimi	*Ca. Phosphitivorax anaerolimi* SV1
SV1-Bin_3_4	strain	family	Smithella	NA	*Ca. Smithella antiquus* SV1
SV3-Bin_1_1	species	class	Phosphitivorax	NA	*Ca. Phosphitivorax pugnaclava* SV3
SV3-Bin_3_1	strain	class	Phosphitivorax	anaerolimi	*Ca. Phosphitivorax anaerolimi* SV3

B

Full ID	Proposed Name	Genus Etymology	Species Etymology
LM1-Bin_11_1	*Syntrophales bacterium* LM1	NA	NA
LM1-Bin_12_1	*Pelotomaculaceae bacterium* LM1	NA	NA
LM1-Bin_42_D	*Ca. Smithella phosphorovis* LM1	established	*phosphorovis* ("the energy of phosphorus"); *phosphoro* (L. n.), phosphorus; *vis* (L. n), energy.
LM3-Bin_1_merge	*Ca. Phosphitivorax vastumunda* LM3	established	*vastamunda* ("from wastewater"); *vastum* (L. n.), waste; *unda* (L. n.), wave.
M1-Bin_6_1	*Ca. Phosphitivorax anaerolimi* M1	established	established
R1-Bin_1	*Ca. Phosphitivorax anaerolimi* R1	established	established
R3-Bin_35_1	*Ca. Phosphitivorax anaerolimi* R3	established	established
R3-Bin_35_2	*Ca. Phosphitivorax vastumunda* R3	established	*vastamunda* ("from wastewater"); *vastum* (L. n.), waste; *unda* (L. n.), wave.
SL1-Bin_31_2	*Ca. Phosphitivorax anaerolimi* SL1	established	established
SL1-Bin_34_2	*Pelotomaculaceae bacterium* SL1	NA	NA
SL3-Bin_22_1	*Synergistaceae bacterium* SL3	NA	NA
SL3-Bin_22_2	*Ca. Cosmobacter schinkii SL3*	*Cosmobacte*	*schinkii* (In recognition of Bernhard Schink, for the discovery and isolation of the first DPOM)
SL3-Bin_41_1	*Ca. Phosphitivorax anaerolimi* SL3	established	established
SM1-Bin_11_1	*Sporomusaceae bacterium* SM1	NA	NA
SM1-Bin_3_1	*Ca. Smithella antiquus* SM1	established	*antiquus* ("ancient"); *antiquus* (L), archaic/ancient.
SM3-Bin_1	*Ca. Phosphitivorax anaerolimi* SM3	established	established
SM3-Bin_3_merge	*Ca. Phosphitivorax vastumunda* SM3	established	*vastamunda* ("from wastewater"); *vastum* (L. n.), waste; *unda* (L. n.), wave.
SV1-Bin_1_1	*Ca. Phosphitivorax anaerolimi* SV1	established	established
SV1-Bin_3_4	*Ca. Smithella antiquus* SV1	established	*antiquus* ("ancient"); *antiquus* (L), archaic/ancient.
SV3-Bin_1_1	*Ca. Phosphitivorax pugnaclava* SV3	established	*pugnaclava* ("fight club"); *pugna* (L. n.), fight; *clava* (L. n.), club.
SV3-Bin_3_1	*Ca. Phosphitivorax anaerolimi* SV3	established	established

Table 3.5: Summary of IMG/M 'Genome Sets'. Metagenomes from the IMG/M were saved as genome sets to the IMG/M workspace on August 1, 2018. These metagenomes were sorted into nine different Ecosystem Categories on the IMG/M website. The largest metagenomes were selected from each category so that 90% of the genes for each Ecosystem Category were represented. The statistics for each Ecosystem Category are listed.

Assembly Type	Ecosystem	Ecosystem Category	Total metagenomes (#)	Selected metagenomes (#)	Selected metagenomes (%)	Total genes (#)	Selected genes (#)	Selected genes (%)
Combined	All	All	106	106	100.00%	668676799	668676799	100.00%
Individual	Engineered	All	2138	500	23.39%	501771201	449972347	89.68%
Individual	Environmental	Terrestrial	2677	1000	37.36%	3846539256	3561842957	92.60%
Individual	Environmental	Aquatic	8601	2500	29.07%	4609659111	4155281078	90.14%
Individual	Environmental	Air	114	114	100.00%	3129805	3129805	100.00%
Individual	Host-associated	Human	2489	1500	60.27%	226907004	215312625	94.89%
Individual	Host-associated	Mammal	478	478	100.00%	215255780	215255780	100.00%
Individual	Host-associated	Plant	656	500	76.22%	1495093209	1492802071	99.85%
Individual	Host-associated	Other	629	500	79.49%	561788364	555868534	98.95%
TOTAL	**NA**	**NA**	**17888**	**7198**	**40.24%**	**12128820529**	**11318141996**	**93.32%**

Table 3.6: Literature review of the distribution of DPOM and reduced phosphorus species. For each environment listed in the first column, references for the detection of reduced phosphorus species or DPOM are listed numerically in the subsequent columns. The black box highlights those environments in which reduced phosphorus has been identified, but DPOM have not.

ENVIRONMENT	DPOM	PHOSPHITE	HYPOPHOSPHITE	PHOSPHINE	PHOSPHONATES
DEEP SUBSURFACE AQUIFER/GROUNDWATER	117	n.a. [a]	n.a.	n.a. [a]	n.a.
ENGINEERED SYSTEMS	18,117	156	n.a.	51	157
BIOLOGICAL WASTEWATER DIGESTOR SLUDGE	117	158	n.a.	51,159	160
INDUSTRIAL WASTEWATER (PETROCHEMICAL/LANDFILL)	117	47	47	51,161	44
FRESHWATER SEDIMENTS/WETLANDS/MARSH	117	35	35	51,152,159,162	163
MARINE WETLANDS/SALTMARSH/SEDIMENT	Fig. 3.3	n.a.	n.a.	51,54,162	164
TERRESTRIAL SOILS	117 & Fig. 3.3	37,49	37,49	161,165	166
EXTRATERRESTRIAL	n.a.	167	167	51,167,168	169
GEOTHERMAL/HYDROTHERMAL WATER	n.a.	42	n.a. [b]	n.a.	n.a.
LAKE WATER	n.a.	170	n.a.	51	171
RIVER WATER	n.a.	35	35	51	160
SEA WATER	n.a.	n.a.	n.a.	162	172–174
ANIMAL GUT/ EXCREMENT	n.a.	114	114	51	n.a.

[a] Reduced phosphorus species have been detected from the groundwater of river basins, but not in deep subsurface systems[35]. Bains et al. also propose a model for phosphine generation based on groundwater phosphite[175].

[b] Although hypophosphite was not measured in geothermal or hydrothermal systems, Pech et al. hypothesize that it is likely to be present.

Chapter 4

Metagenomic and Physiological Characterization of DPOM Metabolism

Parts of this chapter have been drawn from the following publication:

Ewens, S. D., Gomberg, A. F. S., Barnum, T. P., Borton, M. A., Carlson, H. K., Wrighton, K. C., Coates, J. D. (2021). The diversity and evolution of microbial dissimilatory phosphite oxidation. PNAS.

Abstract

Three DPOMs (FiPS-3, Phox-21, and *Phosphitispora fastidiosa*) have been physiologically characterized independent of this dissertation. Each of these three representatives is a chemolithoautotroph that uses phosphite as a primary (FiPS-3) or exclusive (Phox-21 and *P. fastidiosa*) electron donor and reduces CO_2 as the preferred (FiPS-3) or exclusive (Phox-21 and *P. fastidiosa*) electron acceptor. Phosphite oxidation is attributed to the *ptx-ptd* gene cluster in all representatives, but the biochemical mechanism for energy conservation remains unclear. The capacity for CO_2 reduction is also a conserved trait between these characterized representatives, however, the mechanisms differ between species. While FiPS-3 uses the Wood-Ljungdahl pathway to generate acetate, CO_2 reduction in Phox-21 is attributed to the reductive glycine pathway with an unknown dissimilatory reduction product, and in *P. fastidiosa* the mechanism of CO_2 reduction has not yet been confirmed. In this chapter, the diverse DPO MAGs that were initially presented in **Chapter 2** are metabolically profiled to characterize the metabolic potential of DPOM more broadly, with particular attention being paid to energy metabolism. Despite their environmental and taxonomic diversity, metagenomic analyses suggest that the typical DPOM is a chemolithoautotroph that occupies low-oxygen environments and specializes in phosphite oxidation coupled to CO_2 reduction. The absence of canonical CO_2-fixation pathways in most DPO MAGs further suggests that DPO is frequently coupled to unique modes of autotrophy. Physiological studies of the HEPO culture confirm the necessity of CO_2 for DPO activity in *Phosphitivorax* species, and analytical chemistry techniques are used to identify an organic compound that is likely the product of dissimilatory CO_2 reduction by *Phosphitivorax*.

Introduction

Phosphite (HPO_3^{2-}) is the electron donor and energy source for dissimilatory phosphite oxidizing microorganisms (DPOM)[18]. Phosphite is a reduced phosphorus (P) compound whose half-cell redox potential ($E^{o\prime}$) is −650 mV for the HPO_4^{2-}/HPO_3^{2-} couple[33], making phosphite the most energetically favorable chemolithotrophic electron donor known. In addition to the DPOM identified in this dissertation, DPO metabolism has been physiologically characterized in three bacterial species: *Desulfotignum phosphitoxidans* FiPS-3[18], *Phosphitivorax anaerolimi* Phox-21[70], and *Phosphitispora fastidiosa*[75]. As is reviewed in **Chapter 3**, DPO activity in all DPOM is attributed to the *ptx-ptd* cluster, which encodes up to seven proteins (PtxDE-PtdCFGHI) but frequently includes only six, as the *ptdG* is typically not present (and presumably not necessary) in the Group I and Group IV syntenic variations of the *ptx-ptd* gene cluster (Fig. 3.7). The PtxD and PtdC were found to be necessary and sufficient to confer phosphite-assimilation activity when heterologously expressed in the closest relative of FiPS-3, *Desulfotignum balticum* SaxT[69]. However, SaxT could not use phosphite as an energy source, and other members of the *ptx-ptd* cluster are likely necessary for energy conservation. As has been summarized previously: *ptxE* is predicted to encode a LysR family transcriptional regulator; *ptxD* is predicted to encode a phosphite dehydrogenase that couples phosphite oxidation to NAD^+ reduction; *ptdC* is predicted to encode a phosphite:phosphate antiporter; *ptdF* is predicted to encode a UDP-glucose 4-epimerase; *ptdG* is predicted to encode a UspA-family nucleotide-binding universal stress protein; *ptdH* is predicted to encode a B12-dependent radical S-adenosylmethionine (SAM) family protein; and *ptdI* is a hypothetical protein with no homologous proteins to support functional predictions[29,69]. While the mechanism of DPO energy conservation has not yet been biochemically determined, these functional predictions have been coupled to genomic analyses and physiological data to model energy conservation in FiPS-3 and Phox-21[29,67].

Desulfotignum phosphitoxidans FiPS-3 is a gram-negative bacterium that, in addition to phosphite, can use hydrogen, formate, and a variety of sugars, carboxylic acids, and amino acids as electron donors[18]. Phosphite oxidation via the *ptx-ptd* gene products is coupled to nitrate reduction to ammonia ($\Delta G^{o\prime}$ = -195 kJ mol^{-1} HPO_3^{2-}), sulfate reduction to sulfide ($\Delta G^{o\prime}$ = -83 kJ mol^{-1} HPO_3^{2-}) and carbon dioxide reduction to acetate ($\Delta G^{o\prime}$ = -69 kJ mol^{-1} HPO_3^{2-}) via the Wood-Ljungdahl pathway (WLP)[18,67,70,75]. For each of these electron acceptors, one phosphite donates two electrons according to:

Equation 1: $HPO_3^{2-} + H_2O \rightarrow HPO_4^{2-} + 2e^- + 2H^+$ ($E^{o\prime}$ = -650 mV)[64]

Genomic analyses coupled to heterologous expression studies suggest that phosphite is transported into the cell via the PtdC, which couples phosphite import to phosphate export in an ATP-independent process[29,67]. This mode of phosphite antiport is unique to DPOM and likely important for energy conservation, since ATP does not need to be expended[29]. The PtxD is then predicted to couple phosphite oxidation to NAD^+ reduction via a mechanism similar to the homologous phosphite dehydrogenase from the APO microorganism, *Pseudomonas stutzeri*[29,64]. As such, each phosphite oxidized donates its two electrons to NAD^+ according to:

Equation 2: $NAD^+ + H^+ + 2e^- \rightarrow NADH$ ($E^{o\prime}$ = -320 mV)[77]

FiPS-3 has higher growth yields (12.1 g/mol) when grown with phosphite ($E^{o'}$ = -650 mV for HPO_4^{2-}/HPO_3^{2-}) as compared to growth (4.85 g/mol) with formate ($E^{o'}$ = -430 mV for $CO_2/COOH^-$), demonstrating that FiPS-3 can harness the free energy associated with phosphite oxidation[18]. Schink et al. consequently predicted that, in addition to NAD^+ reduction, FiPS-3 generates one ATP for every phosphite via substrate-level phosphorylation according to:

Equation 3: $HPO_3^{2-} + 2H^+ + NAD^+ + ADP + P_i \rightarrow HPO_4^{2-} + NADH + ATP$

($\Delta G^{o'}$ = -155 kJ mol^{-1} HPO_3^{2-})

The mechanism by which substrate-level phosphorylation takes place is unknown, but likely depends on the remaining PtdFGHI gene products of the *ptx-ptd* gene cluster[29].

Phosphitivorax anaerolimi Phox-21was identified from an active DPO enrichment, where genome-resolved metagenomics identified it as the only organism harboring the *ptx-ptd* genes[70]. As was proposed for FiPS-3, Phox-21 is predicted to gain one NADH and one ATP for every phosphite oxidized. Like FiPS-3, Phox-21 has an ATPase and an Rnf complex that are likely used for energy conversions between ATP, NADH, and ion gradients. Unlike FiPS-3, no electron donors other than phosphite were identified, and CO_2 was the only electron acceptor used[70]. Furthermore, instead of using the WLP for CO_2 fixation, Phox-21 is predicted to use the reductive glycine pathway based on its physiological requirement for CO_2 and based on the absence of all six of the known carbon-fixation pathways in its genome. This posits Phox-21 to be the first environmental microbe proposed to fix CO_2 via the reductive glycine pathway[70].

Phosphitispora fastidiosa is the most recent DPOM to be isolated[75]. Like Phox-21, it appears to have a limited energy metabolism in which only phosphite and CO_2 can be used as the electron donor and electron acceptor, respectively[75]. Phosphite was shown to be oxidized to stoichiometric amounts of phosphate, and CO_2 was shown to be fixed to biomass and small amounts of acetate[75]. Further analyses will be required to identify the metabolic pathways responsible for phosphite oxidation and CO_2 reduction.

The WLP, which is used by FiPS-3, fixes CO_2 to acetyl-CoA, which is then converted to pyruvate for biomass accumulation or to acetate for energy conservation[67]. When used for biomass accumulation, less than 1 ATP is required as an energy investment, making the WLP the most energy-efficient carbon fixation pathway known[10]. It is also the only carbon fixation pathway known that can be used for energy conservation. The WLP is composed of two branches: the methyl branch and the carbonyl branch. The methyl branch reduces CO_2 to a methyl group that is attached to a methyl carrier at the expense of 1 ATP[176]. In acetogenic bacteria, tetrahydrofolic acid (THF) serves as the methyl carrier, whereas methanogens utilize cofactors such as methanopterin and methanofuran[10]. In the carbonyl branch, CO_2 is reduced to CO and then combined with the methyl group from methyl-THF to generate acetyl-CoA. The carbonyl branch is dependent on the action of CO dehydrogenase/acetyl-CoA synthase (CODH/ACS), which is frequently used as a marker gene for the WLP[176]. The CODH/ACS complex is one of the most oxygen sensitive enzymes known, which results in organisms utilizing the WLP typically occupying anoxic environments[176]. In accordance with this, FiPS-3 was originally characterized as a strict anaerobe[18]. However, subsequent research has shown that FiPS-3 can tolerate some oxygen and might even be microaerophilic[177].

The reductive glycine pathway is likely used by Phox-21 for the accumulation of biomass via CO_2 fixation[70]. This pathway was originally proposed as a synthetic means for autotrophy based on computational analyses and systematic metabolic design[178]. Heterologous expression of the necessary genes subsequently proved the reductive glycine pathway to be functional, first by allowing *Escherichia coli* hosts to partially satisfy their carbon requirements via CO_2 fixation[72], and in a later study by entirely replacing the autotrophic Calvin cycle of *Cupriavidus necator* with the reductive glycine pathway[179]. Genomic analyses of Phox-21 provided the first evidence that the reductive glycine pathway may occur naturally in environmental systems[70], and the pathway was later biochemically confirmed to support carbon fixation in *Desulfovibrio desulfuricans,* designating the reductive glycine pathway as the seventh known carbon fixation pathway in nature[74]. The reductive glycine pathway has been shown to operate as an electron sink and as a means for autotrophic growth, but its role in energy conservation has not yet been proven. Enzymatically, it is not clear how the reductive glycine pathway could conserve energy, as substrate reduction is not coupled to ion transport or the production of reducing equivalents[74]. However, the reductive glycine pathway is the only plausible carbon reduction pathway in Phox-21, whose growth was supported by CO_2 as the only exogenous electron acceptor[70]. It is therefore tempting to speculate that the reductive glycine pathway, or a variation thereof, supports energy conservation via CO_2 reduction in Phox-21. Notably, the reductive glycine pathway is similar to the WLP, sharing all the genes for the methyl branch to generate methyl-THF; but instead of reacting the THF-bound methyl group with CO to generate acetyl-CoA, the reductive glycine pathway feeds methylene-THF (the metabolic precursor of methyl-THF) into the glycine cleavage system to generate glycine following the addition of CO_2, ammonia, and NADH[70,74]. Glycine is then converted to serine and deaminated to pyruvate, which can enter central metabolism for biomass accumulation[74]. Like FiPS-3, Phox-21 and *D. desulfuricans* grow anaerobically[70]. However, oxygen-tolerant variants of the reductive glycine pathway are theoretically possible, suggesting that DPOM using the reductive glycine pathway should not be strictly limited to anoxic environments[74].

The previously characterized DPOM (FiPS-3, Phox-21, and *P. fastidiosa*) are each capable of utilizing CO_2 as a carbon substrate and as an electron acceptor[18,70,75]. However, the mode by which CO_2 is reduced differs between each representative: FiPS-3 uses the WLP[67], Phox-21 is proposed to use the reductive glycine pathway[70], and the pathway used by *P. fastidiosa* is yet to be determined[75]. Independent of the mechanism, the enrichments presented in **Chapter 2** recapitulate the dependence of DPOM on CO_2, as DPO activity in my enrichment campaign was never clearly coupled to any electron acceptor other than CO_2 (Fig. 2.2). This supports the hypothesis that DPO is biochemically linked to CO_2 reduction, a hypothesis that will require extensive biochemical analyses of diverse DPOM to identify an enzymatic linkage, if one exists. In this chapter, I present metabolic analyses of diverse DPOM genomes, with an emphasis on energy conservation metabolism. As part of this work, I present physiological data that demonstrates the production of an unknown organic compound that correlates with DPO activity in the HEPO culture. This unknown compound is hypothesized to be a reduced end-product of CO_2 reduction and could be the first evidence of energy conservation via CO_2 reduction other than the WLP.

Materials and Methods

Metabolic Characterization of DPOM Genomes

The DPO MAGs that had been identified in the enrichments from **Chapter 2** were annotated with DRAM[180], a genome annotation tool that provides metabolic profiles for each input genome. For contigs of interest, these annotations were compared to Prokka v1.14.6 annotations[106]. More detailed DRAM analyses are provided in Shaffer & Borton *et al.*[180]. The raw annotations containing an inventory of all database annotations for every gene from each input genome are reported in Ewens et al.[117]. From the raw annotations, DRAM then summarizes key metabolisms across the genomes, which are also reported in Ewens et al.[117]. All code for running DRAM is available on GitHub: **https://github.com/shafferm/DRAM**.

HEPO Cultivation and Physiology Studies

Physiological studies were performed with the HEPO culture, the cultivation of which is described in **Chapter 2**. Since its original cultivation, a stock culture of HEPO has been continuously maintained via batch-transfers and re-feeding with 10 mM phosphite. Physiological studies were conducted in anaerobic glass tubes (Bellco) with 10% inoculum of the HEPO stock culture in basal bicarbonate media. Anoxic medium was prepared by boiling under N_2/CO_2 (80:20, v/v) to remove dissolved O_2, and dispensed under N_2/CO_2 (80:20, v/v) into anaerobic pressure tubes or serum bottles. These were capped with thick butyl rubber stoppers and sterilized by autoclaving (15 min at 121 ºC). The basal bicarbonate medium was composed of (per 1 L of deionized water): 5 g $NaHCO_3$, 12 g HEPES buffer, 1 g NH_4Cl, 0.5 g KCl, 1.5 g $MgCl_2$, 0.15 g $CaCl_2$ ($2H_2O$), 0.5 g L-cysteine HCl and 10 mL each of vitamins and trace minerals[94]. I will emphasize that CO_2 serves as the only exogenous electron acceptor in the basal bicarbonate medium, but that L-cysteine-HCl is used as the reductant, and any interpretations of carbon metabolism must take into consideration the presence of significant quantities (3 mM) of cysteine. When HEPO was grown without CO_2, bicarbonate ($NaHCO_3$) was excluded from the basal medium, which was degassed under N_2 only. Standard incubations were at 37 ºC. Phosphite (Na_2HPO_3), nitrite ($NaNO_2$), nitrate ($NaNO_3$), sulfite (Na_2SO_3), molybdate (Na_2MoO_4), yeast extract (Sigma-Aldrich), and casamino acids (Sigma-Aldrich) were each added from sterile, anaerobic stocks to the basal medium as needed. Heat killed controls were autoclaved at 121 ºC for 1 h. Samples for ion determination were filtered and stored at 4 ºC prior to ion chromatography (IC) using the method described previously[70]. Cell growth was measured as optical density at 600nm (OD_{600}) using a Genesys™ 20 Visible Spectrophotometer (Thermo Scientific).

Characterization of HEPO Carbon Metabolism

To identify the product of dissimilatory carbon reduction, the metabolism and growth of the HEPO culture was monitored temporally. Basal bicarbonate media with 10 mM phosphite was inoculated in triplicate with 10% inoculum of live HEPO culture or HEPO culture that had been filter sterilized through a 0.22 μm sterile nylon filter (Simsii). Cell growth was measured via optical density as described above. Phosphite consumption and phosphate accumulation were monitored via IC as described above. Carbon products were monitored via high performance liquid chromatography (HPLC), gas chromatography mass spectrometry (GC-MS), and nuclear magnetic resonance spectroscopy (NMR) as described below.

Samples for HPLC analysis were acquired by extracting 1 mL of growth culture using a sterile anaerobic syringe and passing the sample through a 0.22 μm nylon filter (Simsii). Samples were stored at 4 ºC until further processing. Filtered samples were diluted into deionized water in glass GC vials (Agilent). Stock solutions of the following standards were prepared prior to filtration and dilution into GC vials: crotonic acid ($C_4H_6O_2$), butyric acid ($C_4H_8O_2$), calcium lactate ($CaC_6H_{10}O_6$), propionic acid ($C_3H_6O_2$), isovaleric acid ($C_5H_{10}O_2$), 2-methylbutyric acid ($C_5H_{10}O_2$), isobutyric acid ($C_4H_8O_2$), and valeric acid ($C_5H_{10}O_2$). The organic acid analysis standard (BioRad) was rehydrated with 1 mL of deionized water and filtered into a GC vial. The BioRad standard included: sodium oxalate ($Na_2C_2O_4$), sodium citrate ($Na_3C_6H_5O_7$), sodium malate ($Na_2C_4H_4O_5$), sodium succinate ($Na_2C_4H_4O_4$), sodium formate ($NaCHO_2$), and sodium acetate ($NaC_2H_3O_2$). Samples and standards were analyzed with a Shimadzu LC-20AD (flow rate of 0.6 mL/min) equipped with a Shimadzu SIL-10AC HT autosampler (30 μL injection volume), DGU-20A3 degasser, CTO-20A column oven (operated at 30 °C), and an Aminex HPX-87H hydrogen-form column (300 x 7.8 mm) (Bio-Rad, United States), with 0.005 N sulfuric acid (H_2SO_4) as the eluent. Peaks were detected with a Shimadzu SPD-M20A diode array detector with values reported for detection at 210 nm. Peak fractions were manually collected from the waste line, and rerunning collected fractions validated acquisition of the appropriate peak.

Samples for GC-MS analysis were acquired by extracting 4 mL of growth culture using a sterile anaerobic syringe and passing the sample through a 0.22 μm nylon filter (Simsii). Samples were stored in glass blood vials (Becton Dickinson) at 4 ºC until further processing. Organics were extracted from the filtered samples as follows: 1.5 g of anhydrous sodium sulfate (Na_2SO_4) was added to the filtered sample and the pH was adjusted to <2 using 25 μL of 18.3 M sulfuric acid (H_2SO_4). HPLC-grade methyl tert-butyl ether (MTBE) (Fisher Scientific) was spiked with 50 μM of an acetic acid ($C_2H_4O_2$) internal standard, and 1mL of the MTBE-acetic acid mixture was added to the sample. The organic layer was transferred with a glass pipette to silanized GC vials (Agilent) containing ~0.2 g of baked anhydrous sodium sulfate (Na_2SO_4). The dehydrated samples were subsequently transferred to glass GC vials and stored at -20 ºC until submission to the UC Berkeley QB3/Chemistry Mass Spectrometry facility, where samples were analyzed by an AutoSpec Premier mass spectrometer (Waters, Milford, MA), equipped with an Agilent 7890A gas chromatograph (Agilent Technology, Santa Clara, CA), CTC Combi PAL autosampler (CTC Analytics, Switzerland), an electron impact ion source, and Masslynx software. A DB-WAX column (30 m x 0.25 mm x 0.25 μm) was used with helium carrier gas and an inlet temperature of 250 ºC. Samples were split injected (1:2) into the mass spectrometer with a mass detection range of 35 – 400 m/z. Compounds were identified using a National Institute of Standards and Technology (NIST) library and validated using injected standards. The GC-MS peak for 2-methylpentanoic acid was predicted via NIST library but was never confirmed via standards.

Samples for NMR analysis were prepared by mixing 450 μL of filtered growth culture or collected HPLC fraction with 50 μL of deuterium oxide (Sigma-Aldrich) in NMR tubes (Norell). Prepared samples were submitted to the UC Berkeley QB3/NMR facility for analysis on the 900 MHz NMR spectrometer equipped with a CP TCI (^{1}H, ^{13}C,^{15}N,^{2}H) cryoprobe.

Results

The genomes of FiPS-3 and Phox-21 have been used to predict the mechanism for DPO energy conservation[67,70]. In these models, the Ptx-Ptd protein cluster is hypothesized to couple phosphite oxidation to the generation of NADH and ATP through substrate level phosphorylation (Equation 3). Alternative reducing equivalents are likely produced via a Na^+ motive force, ferredoxin, and an electron confurcation mechanism. Carbon is fixed by FiPS-3 via the WLP[67]. In contrast, Phox-21 is predicted to use the reductive glycine pathway for carbon fixation to biomass[70]. The genomes of the enriched DPO MAGs from **Chapter 2** were metabolically profiled and compared to the genomes of FiPS-3 and Phox-21 to generate an updated model for energy conservation in *Desulfomonilia_A* DPOM (Fig. 4.1). The details of the metabolic analyses that contributed to this updated model are described in the following sections, and physiological studies using the HEPO culture are used to supplement genome-derived hypotheses.

Energy Conservation

Like Phox-21, all DPO MAGs were missing a canonical membrane-bound electron transport chain (ETC), as complexes II-IV were either absent or incomplete (Fig. 4.2). *Sporomusaceae* SM1 of the *Negativicutes* class had a complete NADH-quinone oxidoreductase (complex I), including the N, Q, and P-modules for NADH dehydrogenase activity, quinone reduction, and proton translocation, respectively. However, all other DPO MAGs only contained N-module subunits (Fig. 4.2)[117]. The N-module houses the FMN and FeS clusters for electron transport, as well as the NADH binding site. It also chimerically associates with other protein complexes, such as formate dehydrogenases, catalyzing reversible NADH-dependent formate production[181,182]. Poehlein *et al.* suggested that the FiPS-3 N-module may directly transfer electrons from NADH to ferredoxin[67]. However, direct NADH-dependent ferredoxin reduction is thermodynamically unfavorable[183]. Furthermore, the N-module of DPO MAGs is found in various genomic contexts, making it unclear whether the commonality is uniquely associated with DPO activity or with alternative cellular functions.

In Phox-21, ferredoxin reduction by NADH is attributed to a sodium translocating ferrodoxin:NADH oxidoreductase (Rnf) driven by a Na^+ motive force[70]. Consistent with Phox-21, an Rnf complex was present in the *Synergistia* and nearly all *Desulfomonilia_A* DPO MAGs (Fig. 4.2). In contrast, the Rnf was absent from the *Negativicutes, Desulfotomaculia,* and *Syntrophia* DPO MAGs, suggesting that it is dispensable or replaceable for DPO activity. The ion motive force for Rnf activity in Phox-21 is likely provided by a cation-translocating F-type ATPase at the expense of ATP (Fig. 4.1). The F-type ATPase was present in every DPO MAG, except one (*Synergistaceae* SL3) which had the V-type (Fig. 4.2). While two genomes (*Syntrophales* LM1 and *Pelotomaculaceae* LM1) were missing several ATPase subunits, these genomes were only 61% and 69% complete (Table 2.3)[105]. Given the universal absence of an ETC in DPO MAGs, the ATPases are likely involved in ATP hydrolysis with the concomitant generation of a cation motive force.

CO_2 as an Electron Acceptor

No DPO MAGs harbored functional pathways for methanogenesis or common respiratory pathways (oxygen, nitrate, or sulfate), which is consistent with Phox-21 and consistent with the

universal absence of ETC complexes (Figs. 4.2 & 4.3). CO_2 was the only exogenous electron acceptor available to DPOM in sequenced cultures. To test the necessity of CO_2 for DPO activity, the HEPO enrichment, which contains DPOMs of the *Phosphitivorax* genus, was grown with and without CO_2. Consistent with Phox-21[70], CO_2 was necessary to support phosphite oxidation and growth; but unlike Phox-21, CO_2 was also sufficient to support robust growth, as no organic supplement (i.e., rumen fluid) was required to stimulate activity (Fig. 4.4). Despite requiring CO_2 for growth, comparative genomics of DPO MAGs revealed a notable absence of any canonical CO_2-reduction pathways (Fig. 4.2), which is what was observed in Phox-21[70]. While FiPS-3 can use CO_2 as an electron acceptor by reducing it to acetate via the WLP[18], carbon reduction in Phox-21 was attributed to the reductive glycine pathway[70]. This is comprised of the methyl branch of the WLP, combined with the glycine cleavage system, serine hydroxymethyltransferase, and serine deaminase to produce pyruvate as an anabolic intermediate[70,74] (Fig. 4.1). While this demarcates a likely mechanism for carbon assimilation to biomass in Phox-21, the final product of CO_2 reduction for energy conservation remains enigmatic, as the genes for pyruvate conversion to acetate (phosphotransacetylase and acetate kinase) are missing from the genome.

Metabolite Analysis of HEPO Carbon Reduction

The only known CO_2 fixation pathway that is capable of conserving energy from CO_2 reduction is the WLP, which generates acetate in bacteria and generates methane in archaea[184]. Given the absence of a complete WLP in all DPOM but FiPS-3, the characterization of carbon reduction in the HEPO culture could elucidate a novel mode of energy conservation via CO_2 reduction. To identify the reduced end-product of CO_2 reduction in the HEPO culture, anaerobic basal bicarbonate media with 10 mM phosphite was inoculated in triplicate with either live HEPO culture or filter-sterilized HEPO culture. These cultures were temporally monitored across a growth curve for biomass accumulation and metabolic activity (Fig. 4.5). As has been previously shown (Fig. 2.8), growth of the HEPO culture is correlated with DPO activity (Fig. 4.5A & B). HPLC analysis of the HEPO culture further identified an unknown organic compound (Compound X) with a retention time of 17.6 minutes whose accumulation (Fig. 4.5C) correlated with DPO activity ($r = -0.989$) in the HEPO culture (Fig. 4.5D).

To determine the identity of Compound X, samples from the HEPO culture growth curve were subjected to organic extraction and GC-MS analysis. Eight organic acids accumulated in the HEPO culture growth media when compared to filter-sterilized controls (Fig. 4.6). The identity of these compounds was determined using NIST library predictions and confirmed by running pure standards on the GC-MS. However, when standards for these same compounds were run on the HPLC on which Compound X was detected, none of the retention times corresponded with that of Compound X (Table 4.1). This suggests that Compound X was not detected by the GC-MS, which is supported by the accumulation patterns of each compound: exponential accumulation of Compound X, which corresponds with the exponential phase of HEPO culture growth and DPO activity, occurs between day 6 and day 8, and after day 13, concentrations of Compound X begin to decline (Fig. 4.5). In contrast, most compounds that were detected via GC-MS are poorly correlated with DPO activity (Table 4.2), as exponential accumulation instead occurs between day 8 and day 11, and the concentrations of these compounds continue to rise after day 13 (Fig. 4.6). The only compounds to show exponential accumulation in the same time frame as Compound X are acetic acid and butyric acid, whose retention times (14.2 minutes for acetic acid; 20.413 minutes for butyric acid) do not match that of Compound X (17.6 minutes) (Table 4.1). As such,

Compound X was likely not detected in this GC-MS analysis. Based on the correlation analysis, Compound X, butyric acid, and acetic acid are the most likely end-products of carbon reduction by DPO and may be produced directly by the DPOM or by partner organisms that grow in parallel to DPOM. Genome analyses of DPOM suggest the latter, as acetate and butyrate synthesis pathways are incomplete for the DPOM of the HEPO culture (Fig. 4.3). The remaining organic acids that were detected via GC-MS are likely products of the microbial fermentation of Compound X, positing the DPOM of the HEPO culture to be primary producers while simultaneously providing bioavailable phosphate to the surrounding microbial community. The metabolic profile acquired via GC-MS analysis will consequently be useful for future analyses of community dynamics within the HEPO culture.

As discussed, GC-MS identification of Compound X was complicated by the accumulation of alternative organic acids in the culture media. To perform direct identification of Compound X, HPLC fractions of the Compound X peak (17.2 to 18 minutes) were collected. Eight fractions were compiled (3.82 mL total) and subjected to organic acid extraction and GC-MS analysis. When compared to the negative control, no additional peaks were seen in the HPLC fraction (Fig. 4.7), again precluding the identification of Compound X. The inability to detect Compound X via GC-MS may be due to low concentrations, low volatility, or low extraction efficiency. It is unlikely that the concentration of Compound X in the HEPO culture is too low for GC-MS detection, as the GC-MS method detected organic acids that accumulated in the HEPO culture (e.g., acetate) that were otherwise undetectable by the HPLC. Meanwhile, the HPLC still detected Compound X, suggesting Compound X is likely produced at a higher concentration than the other organic acids. However, it is also possible that Compound X has strong UV absorption that would allow for its detection in the HPLC despite low concentrations. Alternatively, Compound X might be a non-volatile compound that does not vaporize when injected into the GC-MS, preventing its detection. This could be resolved by derivatization prior to injection. Finally, it is possible that the organic extraction method that I used failed to extract sufficient quantities of Compound X from the culture matrix, which would require optimization of the extraction method.

A single HPLC fraction was also analyzed via 1D ^{1}H NMR analysis, but no peaks were detected. This is likely due to insufficient concentrations. It was predicted that HPLC samples contained approximately 250 μM of Compound X, as the peak area for Compound X was approximately half the average peak area of the standards, which were each prepared at 500 μM. Each fraction lasts 0.8 minutes at a flow rate of 0.6 mL/min, yielding 0.48 mL per fraction. Given an injection volume of 30 μL, each fraction is therefore predicted to contain ~15 μM of Compound X, which is an order of magnitude lower than the recommended NMR detection threshold of 100 μM. Consequently, future NMR analyses will minimally require a 10-fold concentration of Compound X prior to analysis.

CO_2 Fixation to Biomass and Carbon Metabolism

In addition to serving as the electron acceptor for DPOM, CO_2 is the primary carbon source available for biomass accumulation (Fig. 4.4). While none of the DPO MAGs contained any of the canonical CO_2 fixation pathways (Fig. 4.2), twelve DPO MAGs in the *Desulfomonilia_A, Negativicutes,* and *Syntrophia* classes had all the genes necessary for CO_2-fixation to pyruvate via the reductive glycine pathway (Fig. 4.2). Of the residual nine DPO MAGs whose reductive glycine pathway was incomplete, four were missing homologs of serine deaminases, preventing the final

conversion of serine to pyruvate (Fig. 4.8). The remaining five DPO MAGs (ranging from 61.3 – 98.7% completion) were missing between one and four genes involved in formate and/or glycine transformations, severely impeding the overall pathway (Fig. 4.8). It is possible that homologous enzymes may perform the reactions of missing genes, as might be the case for one genome (*Syntrophales* LM1) which harbored a serine-glyoxylate transaminase as opposed to the standard serine deaminase (Fig. 4.8). Even if not a universal carbon fixation pathway in DPOM, this analysis suggests the reductive glycine pathway might be an important autotrophic mechanism across diverse DPOM taxa. Carbon-tracing studies will be necessary to understand how individual DPOM use CO_2 to simultaneously generate biomass and capture energy from phosphite oxidation.

Heterotrophic carbon metabolism was also limited across DPO MAGs. Five genomes had the complete Embden-Meyerhof pathway (glycolysis), but in most cases fructose bisphosphate aldolase was missing (Fig. 4.2). The pentose phosphate pathway was also incomplete in all genomes, usually lacking the oxidative branch (Fig. 4.2). The widespread presence of alcohol dehydrogenase would support the metabolism of alcohols in most DPO MAGs, but short-chain fatty acid metabolism was consistently limited to D-Lactate, as pyruvate, acetate, butyrate, propionate, and L-lactate pathways were predominantly incomplete (Fig. 4.3). The tricarboxylic acid (TCA) cycle was also incomplete in all genomes except Ca. *Cosmobacter phosphorovis* SL3, where most DPO MAGs lacked the enzymes necessary for succinate conversion (Fig. 4.2). Consistent with this, CO_2 was necessary and sufficient to support DPO activity, as complex carbon supplements could not support DPO activity in the HEPO culture (Fig. 4.9). Furthermore, virtually all DPO MAGs had the genes necessary to degrade polymers (i.e., polyphenolics, starch, chitin), often in high copy number (Fig. 4.3). This observation was unexpected, given the universally limited genomic repertoire for carbon monomer utilization, but highlights the potential importance of community dynamics in DPOM metabolism, which will be further discussed in **Chapter 5**.

Additional Metabolic Traits

FiPS-3 is capable of using nitrate as an electron acceptor by reducing nitrate to ammonia[67]. However, FiPS-3 is likely the only DPOM capable of dissimilatory nitrate reduction, as no other DPOM genome contained a nitrate reductase (Fig. 4.3). Physiological analyses of the DPOM enrichments in **Chapter 2** confirm this, as no DPO enrichment coupled phosphite oxidation to nitrate reduction (Fig. 2.2). It was hypothesized that nitrate may inhibit DPO activity. However, the addition of equivalent concentrations of nitrate (10 mM) to the HEPO culture did not inhibit growth or activity (Fig. 4.10). Alternatively, metabolic byproducts of community nitrate reduction (e.g., nitrite, nitric oxide) may be inhibitory to DPO. Consistent with this, 1 mM of nitrite inhibited growth and DPO activity in the HEPO culture (Fig. 4.10). Twelve genomes from the *Syntrophia, Negativicutes,* and *Desulfomonilia_A* contain a respiratory (cytochrome c) nitrite reductase complex (*nrfAH*), including *Phosphitivorax anaerolimi* SV3 from the HEPO culture (Fig. 4.2). The NrfAH catalyzes the six-electron reduction of nitrite to ammonia[185]. While this could support dissimilatory reduction of nitrite for energy conservation[186], the observed inhibition of DPO activity at concentrations of 1 mM nitrite suggest otherwise. As such, the cytochrome c nitrite reductases likely function in detoxification of nitrite and/or nitric oxide[185], as both compounds are cytotoxic and are frequently present in anoxic microbial communities[187]. Titration inhibition curves of both nitrite and nitric oxide will be informative to test the level of resistance conferred by *nrfAH* in DPOM.

When grown with phosphite as the electron donor, 1.5 – 2.0 mM of sulfite inhibited growth in FiPS-3 and Phox-21[18]. In contrast, the HEPO culture was capable of growth and DPO activity following the addition of 10 mM of sulfite (Fig. 4.10). This is surprising, as sulfite is cytotoxic at concentrations as low as 600 μM unless a detoxification pathway is present[188], and no genes for sulfite transformations were present in the DPO MAGs from the HEPO culture (*P. anaerolimi* SV3 and *P. pugnaclava* SV3) (Fig. 4.2). Meanwhile, phosphite oxidation in FiPS-3 was inhibited by sulfite despite the presence of functional sulfate and sulfite reduction machinery[18]. Schink et al. suggested this might indicate direct inhibition of DPO machinery by sulfite[18], which is supported by *in vitro* analyses of the homologous APO PtxD from *Pseudomonas stutzeri*, which is known to be inhibited by sulfite[64,189]. Given the likely sensitivity of DPO enzymes to sulfite and the lack of sulfite detoxification machinery, it is unlikely that the DPOM from the HEPO culture are singularly resistant to sulfite. Instead, it is possible that another microbial community member is detoxifying sulfite and subsequently allowing the DPOM to grow. This is supported by the observation that a lag phase of ~4 days follows the addition of sulfite, after which the HEPO culture was able to continue DPO-dependent growth (Fig. 4.10). A temporal analysis of sulfite speciation will be required to determine if this is true.

Molybdate has long been considered a specific inhibitor of sulfate reduction, competitively replacing sulfate in the ATP sulfurylase/sulfate adenosyltransferase and causing the futile consumption of ATP[190]. Yet Phox-21 was inhibited by 5 mM of molybdate despite the absence of sulfate reduction machinery[70]. Like Phox-21, 5 mM of molybdate also inhibited growth and DPO activity in the HEPO culture (Fig. 4.10). Recent evidence suggests that molybdate inhibition extends beyond sulfate-reduction machinery to include inhibition of YcaO-like proteins[191]. YcaO-like proteins are reported to facilitate cyclization of molecules via phosphorylation[192], and molybdate-adapted strains of *Desulfovibrio vulgaris* were found to have mutations in the ATP hydrolyzing domain of the YcaO-like protein[191]. As such, Zane et al. predicted that molybdate may competitively bind with YcaO-like proteins to generate unstable adenosine 5'-molybdophosphate, resulting in the futile consumption of ATP that is likewise observed for the ATP sulfurylase/sulfate adenosyltransferase of sulfate reducing microorganisms[191]. DPOM are predicted to conserve energy from phosphite using substrate-level phosphorylation to generate ATP[18,70]. As such, it is possible that molybdate is targeting enzymes from, or associated with, the *ptx-ptd* gene cluster that are responsible for ATP synthesis. However, biochemical analyses will be required to confirm this mechanism of action.

Discussion

In this chapter, cultivation-based investigations coupled to high-resolution metagenomics are used to explore many of the confounding metabolic features that have precluded understanding of DPO. Results from this chapter describe the metabolic characteristics of energy conservation, CO_2 reduction, CO_2 fixation, and inhibitor response across diverse DPOM.

Phosphite oxidation coupled to CO_2 reduction appears to be the primary metabolic niche occupied by diverse DPOM. No DPOM harbored a canonical electron transport chain, supporting the hypothesis that energy from DPO is conserved via substrate level phosphorylation[18]. Furthermore, FiPS-3 appears to be the only DPOM that can use electron acceptors other than CO_2[18]. The unique metabolic flexibility of FiPS-3 may be because FiPS-3 acquired its *ptx-ptd* genes via horizontal gene transfer, which contrasts with the majority of DPOM whose *ptx-ptd* gene

cluster was vertically inherited (see **Chapter 3**). The remaining DPOM genomes encode sparse electron transport machinery and are devoid of the enzymes required to reduce the electron acceptors of common respiratory pathways (i.e., oxygen, nitrate, or sulfate). This is interesting, because DPO coupled to any known inorganic electron acceptor (oxygen, manganese, perchlorate, nitrate, iron, sulfate etc.) is thermodynamically favorable[29,77]. Yet CO_2 appears to be the preferred or exclusive electron acceptor for all characterized DPOM besides FiPS-3[18,70,75,117]. Physiological studies in *Phosphitivorax anaerolimi* Phox-21, *Phosphitispora fastidiosa*, and the HEPO culture further confirm that CO_2 is necessary and sufficient to support DPO-dependent growth[70,75,117]. Yet genome analyses presented in this chapter show that FiPS-3 is again unique, as it is the only DPOM to harbor the complete WLP for CO_2 fixation; all other DPO MAGs are devoid of canonical CO_2 fixation pathways. Like Phox-21, many DPO MAGs harbor the necessary genes for the reductive glycine pathway, which was recently confirmed as the seventh known CO_2 fixation pathway in nature[70,74]. The novel reductive glycine pathway is consequently the likely means for biomass accumulation in many DPOM. However, nine DPOM genomes did not have any complete CO_2 fixation pathways, reductive glycine pathway or otherwise, highlighting the enigmatic means by which DPOM preferentially utilize CO_2 as a carbon source.

To date, microorganisms have only been shown to use the reductive glycine pathway as a means for biomass accumulation[74] or as an electron sink[10]. The WLP is the only CO_2-fixation pathway known to also function in energy conservation, for which CO_2 serves as an electron acceptor that is reduced to acetate or methane[184]. The reductive glycine pathway does not have any known enzymatic means by which energy could be conserved from CO_2 reduction, and it has not yet been shown to serve as an electron accepting pathway[10]. As such, the mode for energy conservation via CO_2 reduction in any DPOM other than FiPS-3 remains unclear. In this chapter, it's suggested that DPOM utilize a variation of the reductive glycine pathway that may serve as a novel means for energy conservation from CO_2 reduction. In their 2012 paper, Bar-Even et al. present an interesting perspective on CO_2 fixation pathways that may justify the utilization of the reductive glycine pathway by DPOM for energy-conserving CO_2 reduction[10]. To paraphrase from Bar-Even et al.: the energy efficiency of CO_2 fixation is identical between all alternative carbon-fixation pathways, as the same number of electrons must always be transferred from the electron donor to CO_2. However, each pathway differs significantly in its ATP requirements, ranging from < 1 ATP/pyruvate for the WLP, to 9 ATP/pyruvate for the 3-Hydroxypropionate-4-hydroxybutyrate cycle[10]. Greater ATP efficiency is often viewed as beneficial, but Bar-Even et al. point out that this also means these organisms are more energetically constrained, as they do not have the extra energetic drive from ATP to keep the reductive reactions irreversibly moving forward[10]. As such, organisms employing energy-efficient carbon fixation pathways will typically occupy energy-poor environments, but will also require lower reduction potentials and higher concentrations of inorganic carbon to thermodynamically drive the reduction of CO_2 forward[10]. The reductive glycine pathway requires just 1 ATP/pyruvate, and Bar Even et al. suggest that in environments with ambient CO_2, reduction potentials well below -300 mV are likely required to push CO_2 reduction forward in microorganisms utilizing the reductive glycine pathway as an electron sink[10]. DPOM utilize phosphite as an electron donor, which has a reduction potential of -650 mV for the HPO_4^{2-}/HPO_3^{2-} couple. The highly negative reduction potential from phosphite could therefore enable DPOM to accumulate significant reducing power in the cell to drive the reductive glycine pathway forward, without sacrificing the limited ATP that is gained from each phosphite oxidized (Equation 3), as would otherwise be required in the more ATP-expensive

carbon fixation pathways. However, it is still unclear if and how DPOM utilize the reductive glycine pathway and to what product CO_2 is being reduced.

To characterize the CO_2 reduction pathway utilized by DPOM in the HEPO culture, carbon metabolism was temporally monitored, and an unknown compound (Compound X) was found to accumulate with DPO activity and growth. The identity of Compound X was not elucidated during the timeline of this dissertation, but several observations may aid its future identification. Compound X eluted from an HPLC with an HPX-87H hydrogen-form column (300 x 7.8 mm) at 17.6 minutes, between the standards of propionate (16.7 minutes) and isobutyrate (18.842 minutes), and it was not detected on the GC-MS following MTBE extraction. By comparing the peak size to standards in the HPLC, 2.5 mM of Compound X likely accumulates in the HEPO culture with 10 mM of phosphite, stoichiometrically suggesting Compound X is a 4 to 5-carbon compound, which is consistent with its HPLC retention time in relationship to standards. The identification of Compound X is greatly important, as this would allow for i) a quantitative description of DPO energetics when coupled to CO_2 reduction, and ii) the commencement of CO_2-labelling studies that can biochemically characterize the CO_2 reduction mechanism in *Phosphitivorax* species. If energy is conserved via CO_2 reduction in the HEPO culture, this would also represent a novel means for energy conservation from CO_2 reduction, as acetogenesis and methanogenesis are currently the only known energy-conserving CO_2 reduction pathways[10]. Independent of its identity, GC-MS analyses of the HEPO culture demonstrate that Compound X is supporting a rich microbial community that generates diverse organic acids (2C to 5C) from its fermentation, and the significance of this role will be explored further in **Chapter 5**

Figures

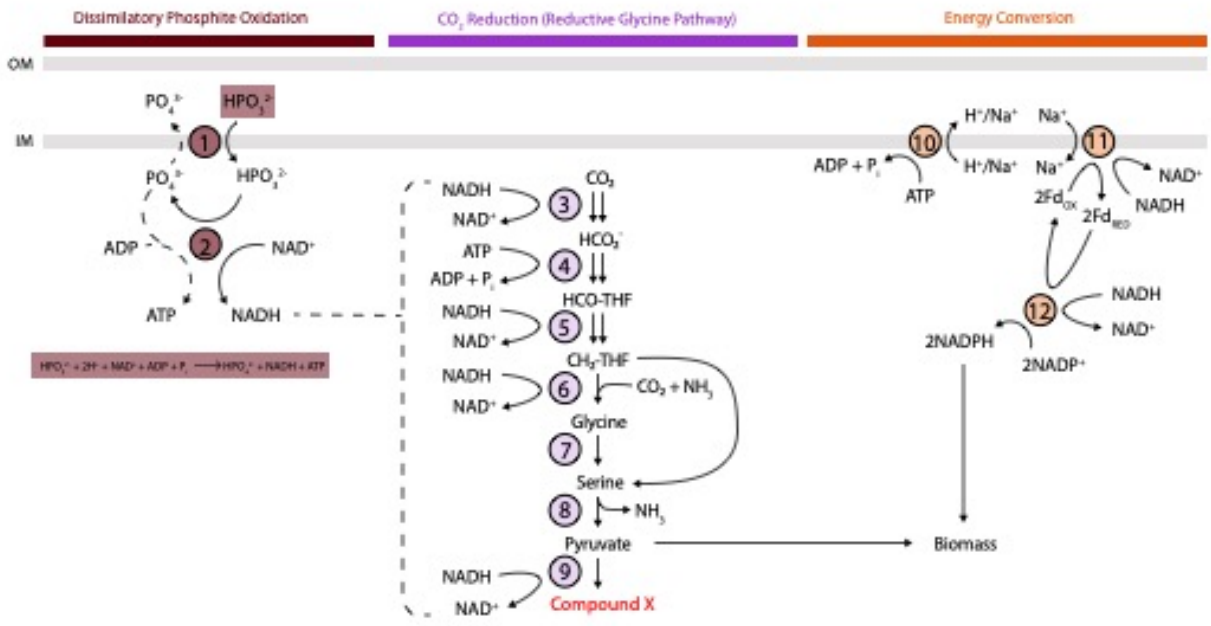


Figure 4.1: Metabolic model of energy conservation by *Desulfomonilia_A* DPOM (adapted from Figueroa, *et al.*)[70]. Dotted lines represent mechanisms that have not been biochemically confirmed. A balanced equation is provided for phosphite oxidation, which is theoretically coupled to CO_2 reduction to 'Compound X'. **Dissimilatory Phosphite Oxidation proteins**: (1) PtdC, phosphite-phosphate antiporter; (2) PtxDE-PtdFHI, putative phosphite dehydrogenase protein complex. **CO_2 Reduction (Reductive Glycine Pathway) proteins**: (3) FdhAB/FdoGHI, formate dehydrogenase; (4) Fhs, formate:THF ligase; (5) FolD, methylene-THF dehydrogenase/methenyl-THF cyclohydrolase; (6) glycine cleavage system (GcvH, lipoyl-carrier protein; GcvPAB, glycine dehydrogenase; GcvT, aminomethyltransferase; Lpd, dihydrolipoyl dehydrogenase); (7) GlyA, serine hydroxymethyltransferase; (8) SdaA/IlvA, serine dehydratase/threonine dehydratase; (9) unknown enzyme for Compound X production. **Energy Conversion proteins:** (10) ATP synthase complex (11) Rnf, sodium-translocating ferredoxin:NAD oxidoreductase complex (12) NfnAB, NAD-dependent ferredoxin:NADP oxidoreductase.

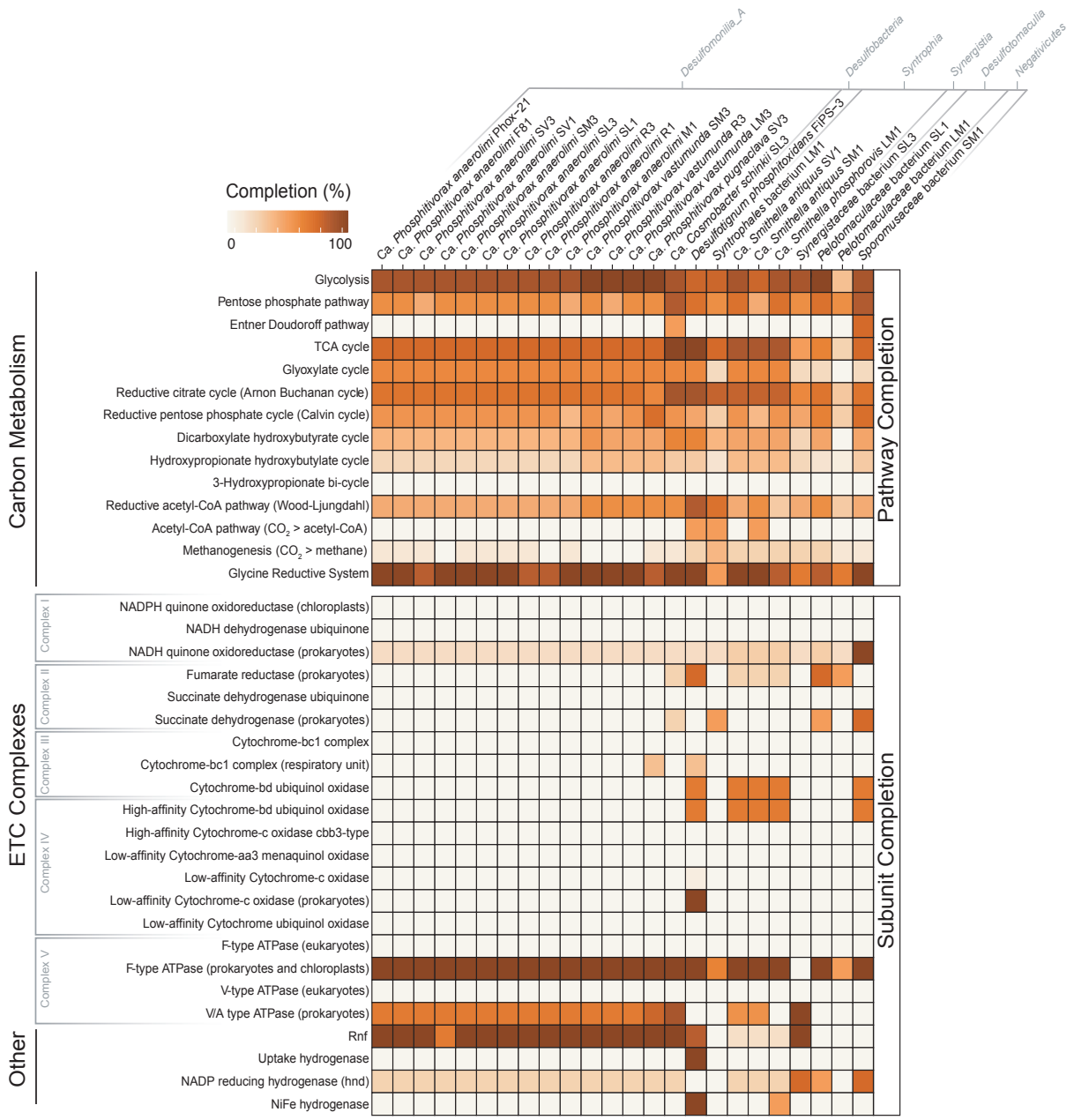


Figure 4.2: Carbon and energy metabolism of DPO MAGs. Each DPO MAG was subjected to metabolic analysis via DRAM[180]. Within this heatmap, each cell represents a metabolic pathway (rows) for each DPO genome (columns). The number of genes for a given pathway is described by percent completion ranging from 0% (white) to 100% (brown). Pathways are organized into modules related to carbon metabolism, electron transport chain (ETC) complexes, and other enzymes referenced in the text. Organisms are annotated with their taxonomic class.

Figure 4.3: Metabolic overview of DPO MAGs. DPO genomes were subjected to metabolic analysis via DRAM, which uses HMM models to identify genes and reconstructs pathways based on KEGG orthology[180,193]. Within this presence-absence matrix, each cell represents a metabolic pathway (rows) for each DPO genome (columns). A metabolic function is considered present

(orange) when all necessary genes for the pathway have been identified within the genome. Incomplete/absent pathways are considered non-functional/absent (light blue). Pathways are organized into modules based on the metabolized substrates. CAZY, carbohydrate active enzymes; SCFA, short chain fatty acids.

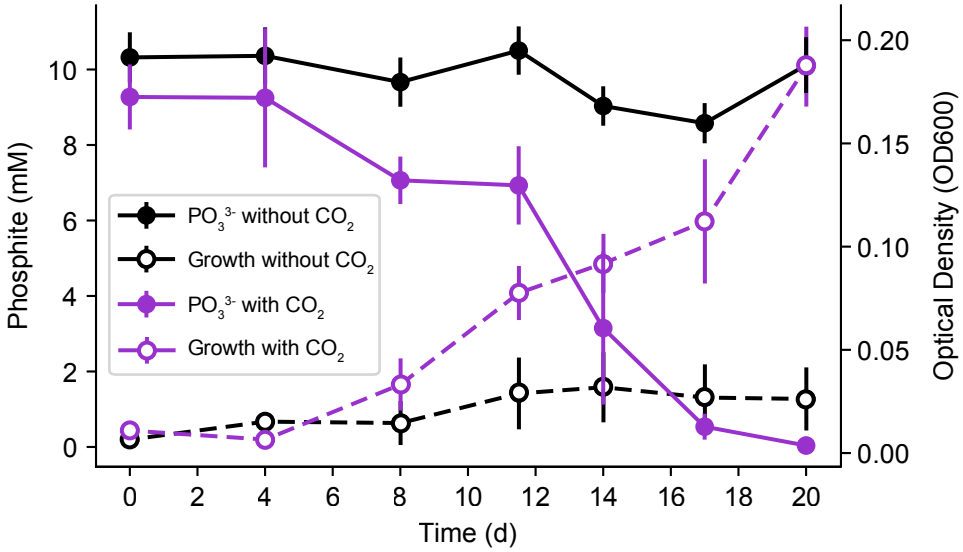


Figure 4.4: CO_2-dependent DPO activity. Growth and phosphite concentrations were temporally monitored in the presence and absence of CO_2 for the HEPO culture. Autoclaved controls showed no activity. Error bars represent standard deviation of triplicate cultures.

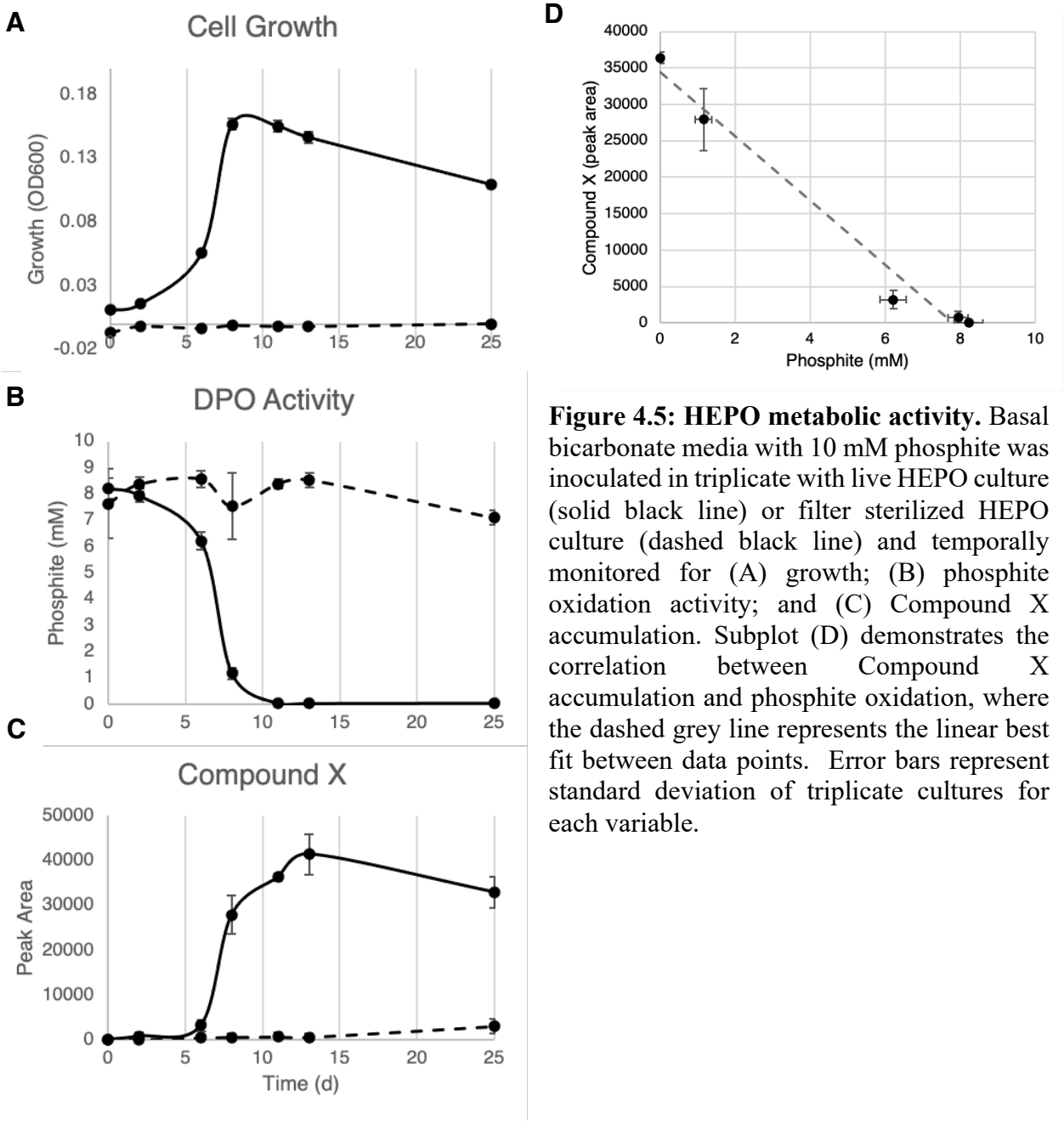


Figure 4.5: HEPO metabolic activity. Basal bicarbonate media with 10 mM phosphite was inoculated in triplicate with live HEPO culture (solid black line) or filter sterilized HEPO culture (dashed black line) and temporally monitored for (A) growth; (B) phosphite oxidation activity; and (C) Compound X accumulation. Subplot (D) demonstrates the correlation between Compound X accumulation and phosphite oxidation, where the dashed grey line represents the linear best fit between data points. Error bars represent standard deviation of triplicate cultures for each variable.

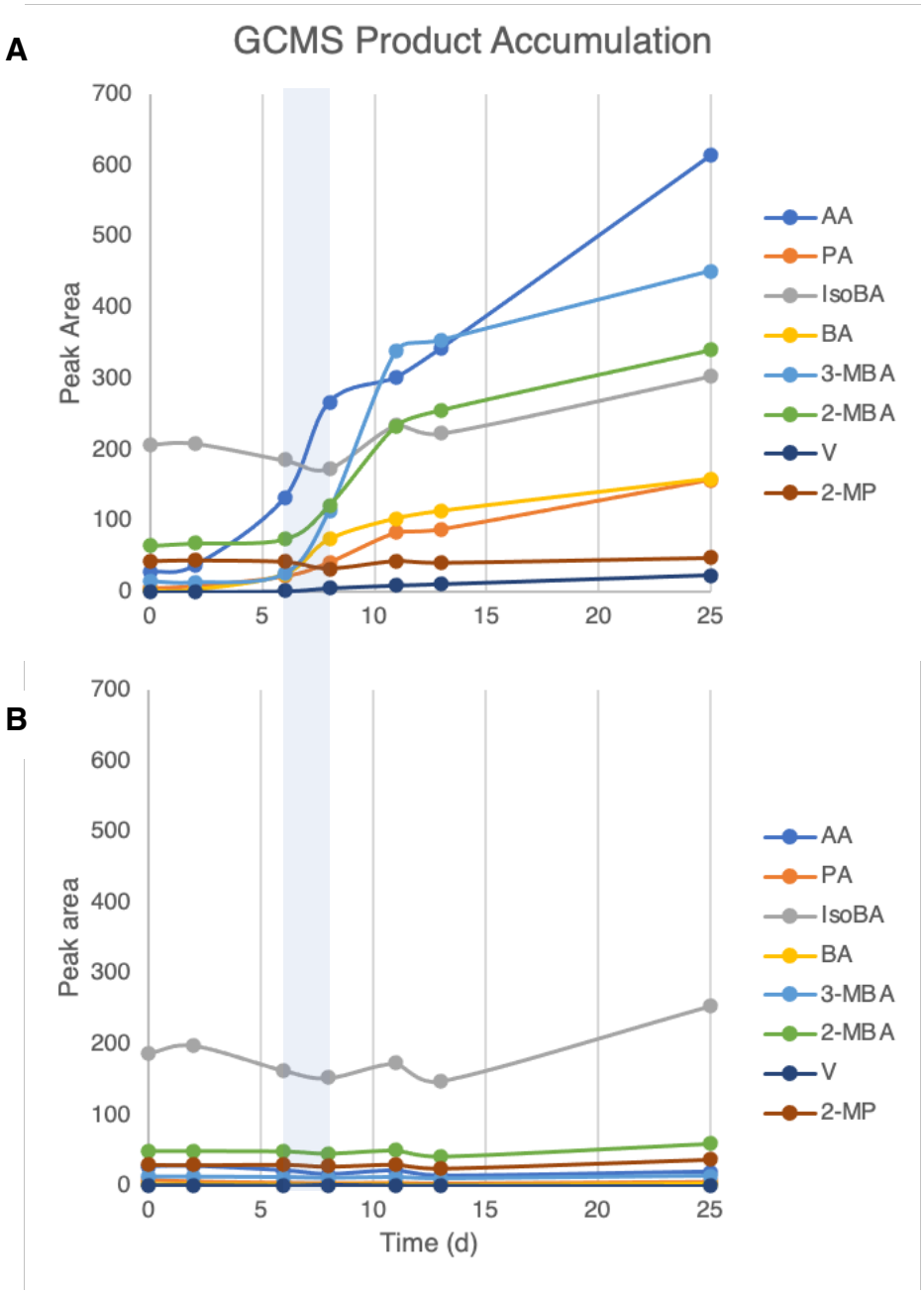


Figure 4.6: GC-MS temporal analysis of HEPO organic metabolites. Basal bicarbonate media with 10 mM phosphite was inoculated in triplicate with (A) live HEPO culture or (B) filter-sterilized HEPO culture, and samples across the growth curve were subjected to organic extraction and GC-MS analysis. The temporal profile of each of the detected organic compounds is shown: AA, acetic acid; PA, propanoic acid; IsoBA, isobutyric acid; BA, butyric acid; 3-MBA, 3-methylbutyric acid; 2-MBA, 2-methylbutyric acid; V, valerate; 2-MP, 2-methylpentanoic acid. The light blue shaded region demarcates the time frame (days 6 to 8) of exponential growth in the HEPO culture. The relatively higher starting concentration of isobutyric acid is due to the presence of a 500μM internal standard.

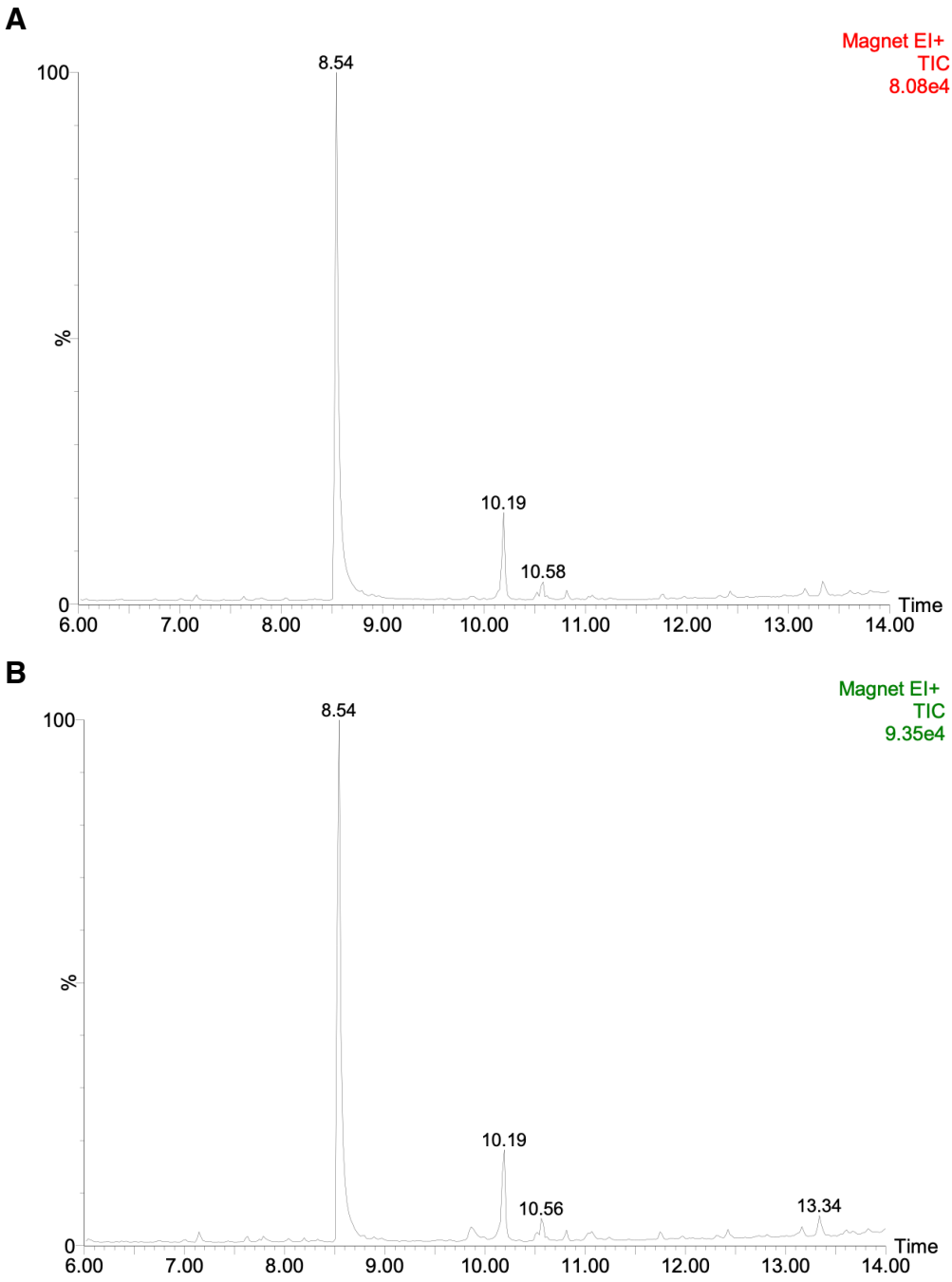


Figure 4.7: Retention times of HPLC fraction in GC. Fractions of Compound X were collected from the HPLC and run on the GC-MS following organic acid extraction. Peaks in the (A) HPLC fraction, were compared to peaks from (B) day 25 of the filter-sterilized control.

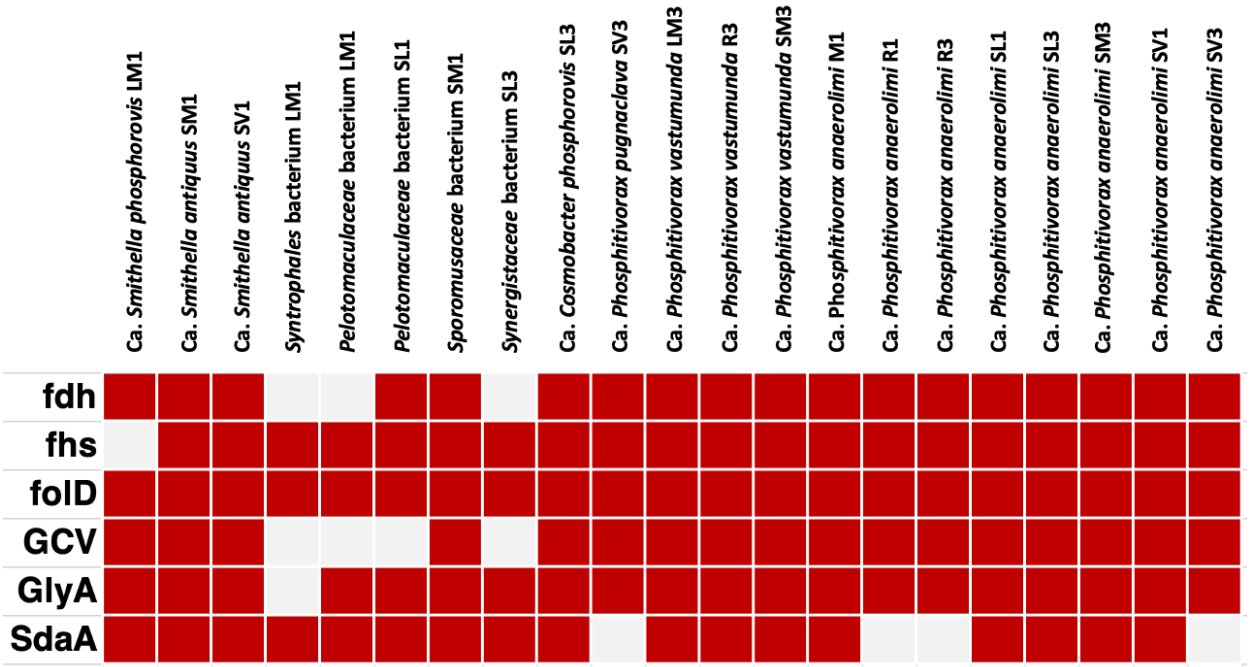


Figure 4.8: Inclusion matrix for the reductive glycine pathway genes. DPO genomes were subjected to metabolic analysis via DRAM, which uses HMM models to identify genes and reconstructs pathways based on KEGG orthology[180]. Within this presence-absence matrix, each cell denotes the presence (red) or absence (grey) of an essential gene (or gene cluster) that is required for the reductive glycine pathway for each DPO genome (columns). The reductive glycine pathway is predicted to be functional if the genes for each row are present. Fdh, formate dehydrogenase; fhs, formate:THF ligase; folD, methyltetrahydrofolate dehydrogenase; GCV, glycine cleavage system; GlyA, serine hydroxymethyl transferase; SdaA, serine deaminase*.

* SdA was also considered present if the following genes were available, as they can catalyze the conversion of serine to pyruvate and ammonia: threonine synthase, D-serine dehydratase, serine ammonia-lyase, and serine-glyoxylate transaminase.

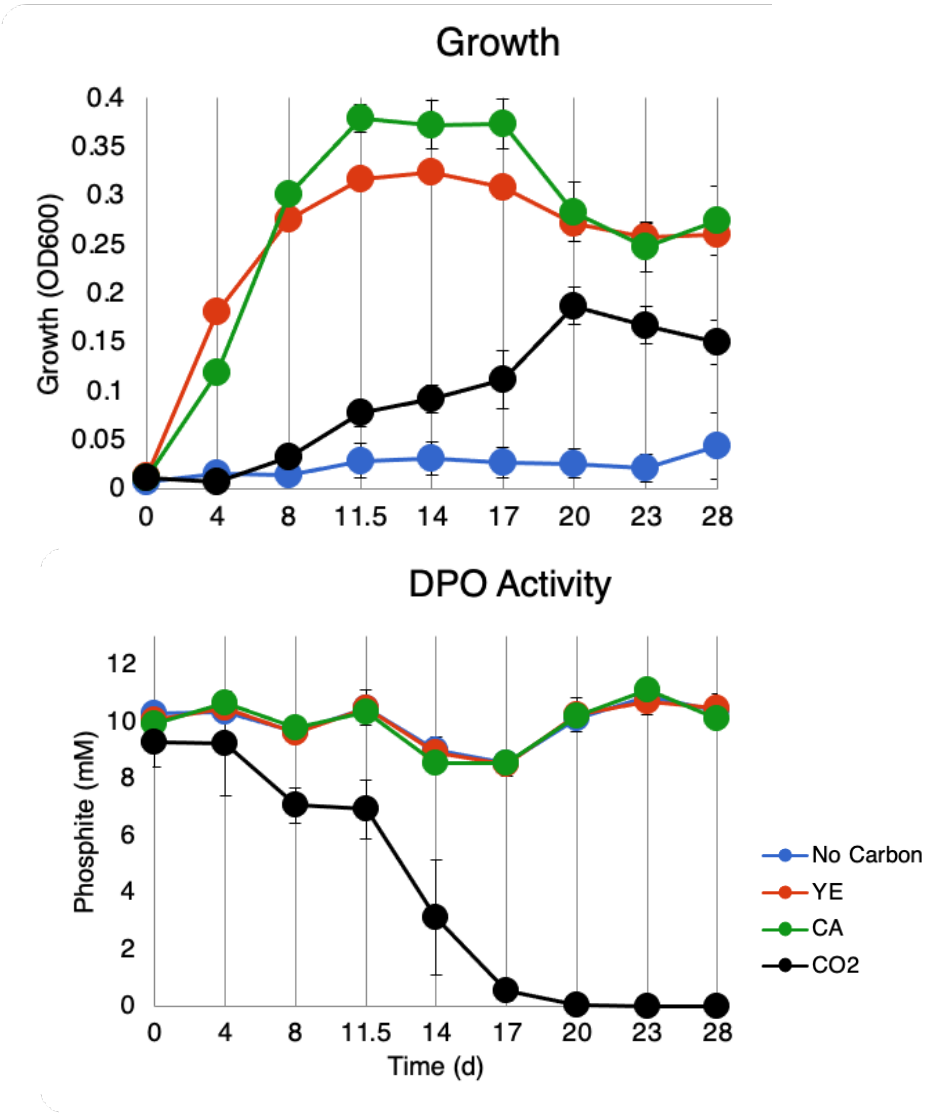


Figure 4.9: Complex carbon utilization in the HEPO culture. Growth and DPO activity were temporally monitored in the HEPO culture following the addition of 10 mM phosphite and alternative carbon sources: No Carbon, CO_2 and bicarbonate were excluded from the media; YE, yeast extract (0.1 g/L); CA, casamino acids (0.1 g/L); CO2, carbon dioxide only. Autoclaved controls showed no activity. Error bars represent standard deviation of triplicate cultures.

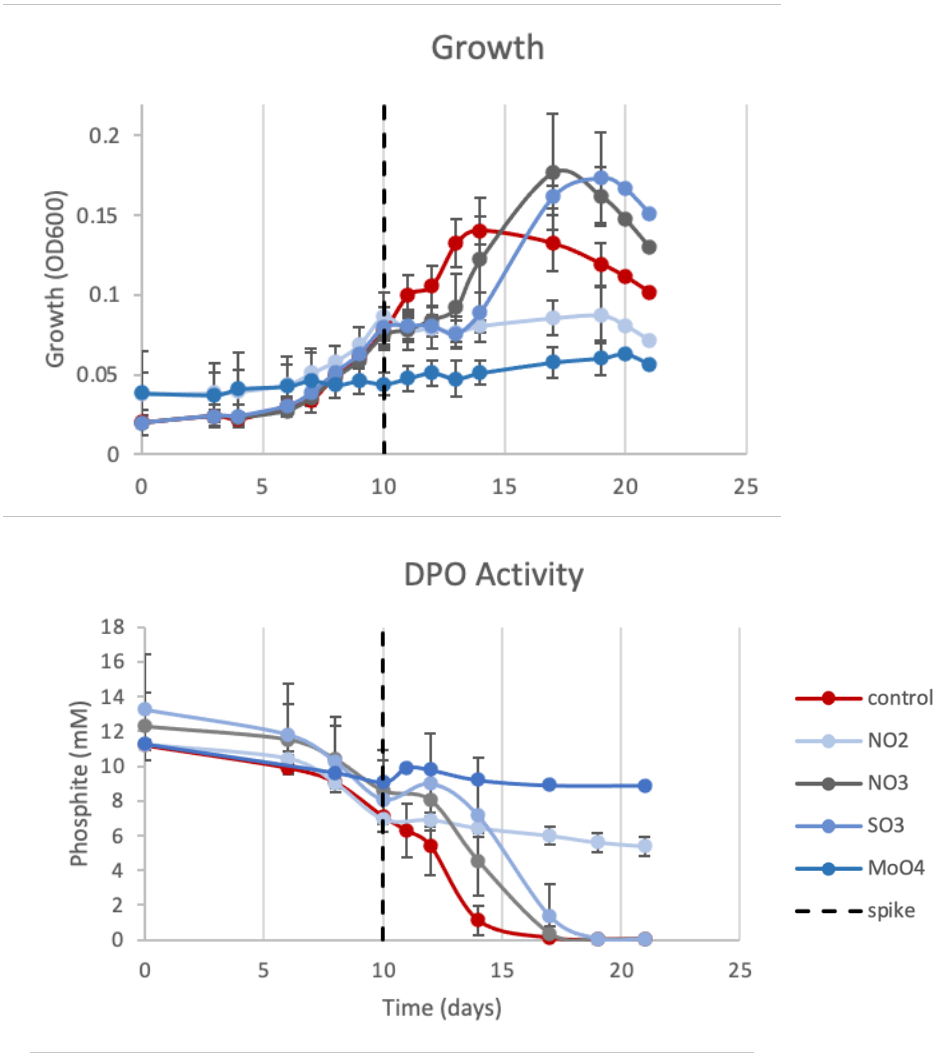


Figure 4.10: Genome-informed inhibition of the HEPO culture. Basal bicarbonate media with 10 mM phosphite was inoculated in triplicate with the HEPO culture and monitored for growth and DPO activity. On day 10 (demarcated by the black dashed line), cultures were spiked with various potential inhibitors of DPO activity: NO2, nitrite (1 mM); NO3, nitrate (10 mM); SO3, sulfite (10 mM). MoO4, molybdate (5 mM) was spiked at the beginning of the growth curve to be consistent with previous studies[70].

Table 4.1: HPLC standard retention times. Standards for HPLC analysis were run with Compound X under identical conditions and their identities were compared with respect to their retention time. Compounds in green were found to accumulate in the HEPO culture using GC-MS analysis.

Sample	Retention Time (min)
oxalate	5.66
citrate	6.885
malate	8.351
succinate	10.706
lactate	11.188
formate	12.532
acetate	**14.209**
propionate	**16.723**
Compound X	17.568
isobutyrate	**18.842**
butyrate	**20.413**
isovalerate (3-MB)	**23.247**
2-methyl butyrate	**23.461**
crotonate	24.4
valerate	**28.152**

Table 4.2: Correlation coefficients of GC-MS detected organic products. Pearson's correlation coefficients (r) were calculated by comparing organic product accumulation to phosphite depletion in the HEPO culture from day 0 to day 11 in the growth curve, as represented by Figure 4.5 and Figure 4.6.

Symbol	Compound	Correlation Coefficient (r)
Compound X	unknown	-0.989
AA	Acetic acid	-0.992
PA	Propionic acid	-0.931
IsoBA	Isobutyric acid	-0.1
BA	Butyric acid	-0.992
3-MBA	3-Methyl-butyric acid	-0.865
2-MBA	2-Methyl-butyric acid	-0.876
V	Valeric acid	-0.966
2-MP	2-methylpentanoic acid	-0.511

Chapter 5

Investigations on the Community Interactions of DPO Enrichments

(unpublished)

Abstract

Metagenomic analyses of environmental enrichments suggest that *Phosphitivorax* species may be the dominant environmental DPOM. Yet efforts to isolate *Phosphitivorax* species have repeatedly been unsuccessful, and *Desulfotignum phosphitoxidans* FiPS-3 and *Phosphitispora fastidiosa* remain the only DPOM isolates acquired to date[18,75]. The HEPO culture, whose enrichment is discussed in **Chapter 2,** is occupied by *Phosphitivorax* DPOM. As is suggested in previous chapters, these *Phosphitivorax* likely sustain the complex microbial community of the HEPO culture by supplying fixed carbon and orthophosphate from CO_2 and phosphite. This chapter will explore the hypothesis that *Phosphitivorax* are also nutritionally dependent on the microbial community that they sustain. Observations from previous studies and from earlier chapters in this dissertation are first reviewed to demonstrate the genomic and physiological qualities that suggest DPO activity by *Phosphitivorax* is dependent on a symbiotic partnership. Thermodynamic analyses supplement this hypothesis by predicting which types of symbioses are likely serving the DPOM. The studies presented in this chapter attempt to i) identify the symbiotic partner of *Phosphitivorax*, and ii) identify the exchanged metabolite that supports DPO activity. Efforts to identify the symbiotic partner of *Phosphitivorax* include 16S rRNA gene sequencing to describe the community structure of the HEPO culture, and inhibition studies to test the resilience of DPO activity in the absence of select community members. Efforts to identify the exchanged metabolite include hypothesis-driven physiology studies, including the characterization of corrinoids in the HEPO culture. In addition to cobalamin (vitamin B12), which is a constituent of the growth media, three corrinoids are identified that are synthesized by the HEPO culture. One of these biosynthesized corrinoids is uncharacterized, representing a potentially novel corrinoid. The implications of these findings are explored in the context of DPO metabolism.

Introduction

Desulfotignum phosphitoxidans FiPS-3 and *Phosphitispora fastidiosa* are currently the only DPOM that have been isolated[18,75]. According to the GTDB, FiPS-3 belongs to the *Desulfobacteria* taxonomic class, and *P. fastidiosa* belongs to the *Desulfotomaculia* taxonomic class. Meanwhile, it was shown in **Chapter 3** that 13 of the 21 DPOM enriched in this study (62%) belong to the *Desulfomonilia_A* taxonomic class, and 21 of the 35 additional PtxD identified in public metagenomes (60%) belong to the *Desulfomonilia_A* associated clade of PtxD. As such, *Desulfomonilia_A* represent the majority of known DPOM and may also be the most environmentally prevalent DPOM. Yet no *Desulfomonilia_A* DPOM have been isolated, as all representatives exist exclusively in DPO enrichments and metagenomic datasets.

Phosphitivorax anaerolimi Phox-21 was the first *Desulfomonilia_A* DPOM to be identified, and it was enriched from anaerobic digester sludge in the San Francisco Bay area[70]. Despite dedicated isolation attempts using the same techniques that were used to isolate FiPS-3 and *P. fastidiosa*, Phox-21 was never isolated from its enrichment community[177]. Metagenomic analyses of the Phox-21 enrichment community identified the presence of 12 additional non-DPOM binned genomes, including the archaea *Methanoculleus* and *Methanococcoides,* and the bacteria *Coprothermobacter, Spirochaeta, Aminobacterium,* two *Thermotogales,* two *Proteiniphilum*, two *Tepidanaerobacter*, and an unclassified bacterium[70]. In a 16S rRNA gene amplicon survey of this community, Phox-21 dominated the microbial population at 39% relative abundance, and a *Tepidanaerobacter* strain represented the next dominant OTU, at 34% relative abundance[70]. Phox-21 was the only organism to house the *ptx-ptd* gene cluster for DPO activity, and it also exhibited the greatest positive fold difference in the presence of phosphite as compared to controls without phosphite. It is unclear if the non-DPO organisms metabolically interact with Phox-21 or if their growth is strictly due to the consumption of substrates in rumen fluid, which was added to the enrichments to stimulate DPO activity[70]. For example, metabolic analyses suggested that the *Tepidanaerobacter, Methanoculleus,* and *Methanococcoides* are all capable of coupling growth to acetate consumption, while the remaining community was likely fermenting carbon substrates from the rumen fluid. To better understand the influence of the surrounding microbial community on DPOM, this chapter will explore the community structure of the HEPO culture, which contains two *Phosphitivorax* DPOM that were enriched in the absence of rumen fluid.

While symbiotic interactions between the *Desulfomonilia_A* DPOM and the surrounding microbial community have not been explicitly demonstrated, mounting evidence suggests that these interactions may be necessary to support DPO activity. In **Chapter 2**, the HEPO culture was hyper-enriched from wastewater sludge, and two species of *Phosphitivorax* DPOM were identified in the primary enrichment metagenome. Like Phox-21, all isolation attempts proved unsuccessful, despite using the same techniques reported for the isolation of *Desulfotignum phosphitoxidans* FiPS-3 and *Phosphitispora fastidiosa*[18,75]. Unlike Phox-21, rumen fluid was not necessary for growth and DPO activity in the HEPO culture, resulting in its subsequent omission from the growth media (Fig. 2.6). This suggests that the remaining community is sustained by the reduced carbon product of DPO-driven CO_2 reduction (Fig. 4.5), as opposed to being sustained by exogenous fermentable carbon substrates. The persistence of a complex community in these highly selective enrichment conditions posits the *Phosphitivorax* DPOM to be primary producers by fixing CO_2 while also supplying bioavailable phosphate from phosphite oxidation. Furthermore, the failure to isolate the *Phosphitivorax* species from the community (Fig. 2.5) suggests that the

Phosphitivorax DPOM may also depend on select community members to support DPO activity. This hypothesis is supported by the observed limited growth and limited DPO activity in the presence of excess phosphite, suggesting another nutrient is limiting in the media (Fig. 2.8). This nutrient could be supplied by another community member, whose growth is limited by DPOM carbon fixation. Phase-contrast microscopy of the HEPO culture also revealed the prevalence of microbial aggregates (Fig. 2.7), which might suggest that close proximity between cells facilitates nutrient exchange to potentially support DPO activity[194]. In **Chapter 3,** phylogenetic analyses revealed that the closest cultured relatives of several DPOM taxa are canonically dependent on syntrophic associations[71]. In fact, *Phosphitivorax anaerolimi* F81 of the *Desulfomonilia_A* clade was identified in an enrichment series designed to cultivate syntrophic bacteria, where the *P. anaerolimi* strains F81, Phox-21, and R76 were each found to contain a complete butyrate beta-oxidation pathway as well as multiple additional features common to syntrophic bacteria[71]. The DRAM analyses in **Chapter 4** corroborate these findings, demonstrating that *Desulfomonilia_A* DPOM have highly simplified genomes with characteristic features of symbionts, including: i) the absence of any canonical electron transport chain; ii) the absence of most genes required for utilizing common electron acceptors (i.e., sulfate, nitrate); iii) the presence, in high copy number, of genes for enzymatic polymer degradation (i.e., chitin, starch), and iv) the presence of the sodium-translocating ferredoxin:NAD^+ oxidoreductase (Rnf) which is central to proton reduction in syntrophic hydrogen generation[134]. Yet, while each of the aforementioned observations might imply a metabolic interaction between DPOM and the surrounding community, none conclusively affirm or deny the existence of such an interaction. As such, the studies discussed in this chapter aim to further explore this possibility.

The HEPO community contains more than two members, inferring the existence of multiple types of interactions between the various members. The classifications of potential interactions include: parasitism, predation, amensalism, competition, commensalism, and mutualism, which have been thoroughly reviewed elsewhere[195]. Studies of the HEPO culture are principally interested in those interactions which benefit DPO activity and growth of the *Phosphitivorax* species. As such, negative or neutral interactions (i.e., parasitism, predation, amensalism, and competition) are of limited interest. While it could be theorized that the *Phosphitivorax* species are the beneficiary of a negative interaction (such as in parasitism or predation), this would be a weak classification, as the metabolic activity of *Phosphitivorax* species inherently supports the remaining microbial community by supplying fixed carbon and phosphate. Likewise, the role of 'primary producer' negates the possibility of *Phosphitivorax* being beneficiaries of a commensalistic relationship, leaving mutualism as the final classification most apt to describe the anticipated relationship between *Phosphitivorax* DPOM and a supportive symbiont in the HEPO community.

Defining a mutualistic relationship between the *Phosphitivorax* species and community members of the HEPO culture will require i) identifying which non-DPOM community member, if any, is required to support DPO activity, and ii) identifying the physiological role played by this community member in supporting DPO activity. Syntrophy is one potential sub-class of mutualism, defined by Schink et al. as:

> *"[A] mutual dependence of partner bacteria with respect to energy limitation … [such] that neither partner can operate without the other and that together they exhibit a metabolic activity that neither one could accomplish on its own."*[136].

A syntrophic relationship supporting DPO would therefore require the energetics of phosphite oxidation coupled to its electron acceptor (likely protons or CO_2) to be thermodynamically unfavorable, thereby requiring a partner organism to utilize DPO end-products as a substrate to thermodynamically drive DPO metabolism forward. However, the phosphate/phosphite redox couple is so energetically favorable ($E^{o'}$ = -650 mV), that DPO coupled to common syntrophic proton-reduction or CO_2-reduction processes should thermodynamically not require a partner organism (Fig. 5.1). Alternatively, DPOM may depend on the community to remove toxic products of DPO metabolism. The immediate product of phosphite oxidation is phosphate, which was shown in **Chapter 2** to be non-toxic to DPO, despite its stoichiometric accumulation (Fig. 2.9). Therefore, phosphate removal is likely not required for functional DPO metabolism. However, it was shown in **Chapter 4** that DPO metabolism is coupled to the accumulation of an unknown organic compound (Compound X). The capacity of Compound X to inhibit DPO metabolism has not yet been studied and will require its purification and identification. Another possibility is that DPOM participate in cross-feeding with the surrounding microbial community. For example, the *Phosphitivorax* DPOM may support the HEPO culture with phosphate and fixed carbon in exchange for essential cofactors and/or amino acids that it is incapable of synthesizing. This chapter uses 16S rRNA gene amplicon analyses, physiological analyses, and analytical chemistry to identify and characterize the nature of a mutualistic relationship that may support DPO activity in the HEPO culture.

Materials and Methods

Cultivation Conditions

Physiological studies were conducted in anaerobic glass tubes (Bellco) with 10% inoculum. Studies of the HEPO culture were inoculated using the HEPO stock culture. The cultivation of HEPO is described in **Chapter 2,** and since its cultivation, a stock of the HEPO culture has been continuously maintained via batch-transfers and re-feeding with 10 mM phosphite. Studies of *Desulfotignum phosphitoxidans* FiPS-3 were inoculated from anaerobic freezer stocks. Anoxic medium was prepared by boiling under N_2/CO_2 (80:20, v/v) to remove dissolved O_2, and dispensed under N_2/CO_2 (80:20, v/v) into anaerobic pressure tubes or serum bottles. These were capped with thick butyl rubber stoppers and sterilized by autoclaving (15 min at 121 ºC). The basal medium was composed of (per 1 L of DI water): 5 g $NaHCO_3$, 12 g HEPES buffer, 1 g NH_4Cl, 0.5 g KCl, 1.5 g $MgCl_2$, 0.15 g $CaCl_2$ ($2H_2O$), 0.5 g L-cysteine HCl and 10 mL each of vitamins and trace minerals[94]. I will emphasize that CO_2 serves as the only exogenous electron acceptor in the basal bicarbonate medium, but that L-cysteine-HCl is used as the reductant, and any interpretations of carbon metabolism must take into consideration the presence of significant quantities (3 mM) of cysteine. HEPO incubations were at 37 ºC. FiPS-3 incubations were at 30 ºC.

Phosphite ($Na_2HPO_3 \bullet 5$ H2O), sulfate (Na_2SO_4), acetate ($NaC_2H_3O_2$), hydrogen (H_2), trimethylamine (TMA; $N(CH_3)_3$), 2-bromoethanesulfonate (BES; $BrCH_2CH_2SO_3Na$), ethyl-2-butynoate (E2B; $C_6H_8O_2$), lumazine ($C_6H_4N_4O_2$), and S-(5'-Adenosyl)-L-methionine (SAM; $C_{15}H_{23}ClN_6O_5S \bullet 2HCl$) (Sigma Aldrich) were each added from sterile, anaerobic stocks to the basal medium as needed. Sterile, anaerobic antibiotic stocks were prepared in the following solvents: ampicillin (100 mg/mL in H_2O), erythromycin (10 mg/mL in EtOH), kanamycin (10 mg/mL in H_2O), chloramphenicol (10 mg/mL in H_2O). Heat killed controls were autoclaved at 121

ºC for 1 h. Samples for ion determination were filtered and stored at 4 ºC prior to ion chromatography (IC) using the method described previously[70]. Cell growth was measured as optical density at 600nm (OD_{600}) using a Genesys^{TM} 20 Visible Spectrophotometer (Thermo Scientific).

Cell filtrates and lysates were prepared from the HEPO stock culture and from a culture of *Methanosarcina acetivorans.* The *M. acetivorans* culture was grown with 50 mM TMA in basal bicarbonate-buffered high salt (HS) liquid medium, as has been described previously[196]. Filtrate was introduced to growth cultures via passing stock cultures through a 0.22 μm sterile nylon filter (Simsii) attached to an anaerobic syringe. Lysate was prepared from 100 mL of HEPO stock culture, or 2 mL of *M. acetivorans* stock culture, which were cell pelleted prior to bead beating at 5.5 m/sec for 30 seconds 5 consecutive times in degassed PowerBiofilm Bead Tubes (Quiagen). Heat pasteurization was performed for 60 seconds at 80 °C.

16S rRNA Gene Sequencing

To monitor the community dynamics of the HEPO culture, DNA was extracted from growth cultures utilizing the DNeasy PowerBiofilm Kit (Qiagen). 5 μL of gDNA was added to a 96-well PCR plate with 1 μM 515F oligo primer. Hot Start Platinum PCR master mix with 806R or 926R indexed oligo primers was then added prior to thermocycling with the standard EMP_PCR method. PCR samples were prepared for sequencing using the ZymoClean PCR clean-up kit and eluted in 40 μL of DNase/RNase free water. Samples were sequenced on an Illumina MiSeq (300 bp paired-end reads) at the UC Berkeley Vincent J. Coates Genomics Sequencing Laboratory.

Corrinoid Analyses

To computationally predict the corrinoid utilization and biosynthesis capacity of DPOM, the DRAM output from the DPO MAGs of active DPO enrichments was analyzed. The preparation of the metagenomes is described in **Chapter 2.** The methods for DRAM profiling and annotation are described in **Chapter 4**. The genes required for corrinoid utilization and biosynthesis were acquired from computational predictions made previously in Shelton et al.[197], and these genes were then manually searched for in the DRAM output for each DPO MAG.

The HEPO culture was cell pelleted to generate 0.3 g of cells, from which corrinoids were extracted as described previously[198,199]. Briefly, harvested cells were resuspended in a mixture of methanol and cyanide for cell lysis and cyano-replacement of the upper corrinoid ligand. Following removal of cellular debris, rotary evaporation was used to remove the solvent. Corrinoids present in the dried sample were resuspended in water, desalted in a C_{18} column, eluted in methanol, dried overnight, and stored in water at -80C until further processing. The corrinoid extracts were then separated via HPLC as described previously[198], from which the UV/vis spectra was utilized to distinguish corrinoid peaks from contaminants in the extract. Corrinoid fractions were submitted to the UC Berkeley QB3/Chemistry Mass Spectrometry facility, where high-resolution mass spectra were acquired using a 1200 series liquid chromatography system (Agilent Technologies, Santa Clara, CA) that was connected in line with an LTQ-Orbitrap-XL mass spectrometer equipped with an electrospray ionization source and operated in the positive ion mode (Thermo Fisher Scientific, Waltham, MA). The instrumentation has been described

previously[200]. Compounds were manually identified based on mass calculations for various protonation states of known corrinoids.

Results

Descriptions of DPO Enrichment Communities

In **Chapter 2**, sludge from MWTFs was used to inoculate 11 selective DPO enrichments, from which 21 DPOM were identified by coupling physiological studies to genome-resolved metagenomics. In previous chapters, the resulting metagenomic data has been primarily analyzed for the identification and characterization of the DPOM from each of these enrichments. However, each enrichment contains a complex microbial community beyond just the DPOM. An adaptation of Figure 2.3 demonstrates that 35 taxonomic classes of bacteria and archaea are found across the 11 active DPO enrichments (Fig. 5.2). Each DPO enrichment community is dominated by the bacterial domain, but also always includes methanogenic archaea. When comparing the community composition across all enrichments at exponential phase (P1), we find that 4 phyla are always present: the bacterial phyla of *Desulfobacterota*, *Bacteroidota,* and *Synergistota*, and the archaeal phyla of *Halobacterota*.

Of the *Desulfobacterota* found across enrichments, 61% are DPOM from the *Desulfomonilia_A* and *Syntrophia* taxonomic classes. The remaining, non-DPOM *Desulfobacterota* are undefined genera of the taxonomic classes *Desulfobacterales, Syntrophia, Syntrophobacteria,* and *Syntrophorhabdia.* As was discussed in **Chapter 3**, the *Syntrophia, Syntrophobacteria,* and *Syntrophorhabdia* are characterized by their ability to perform syntrophic fermentations in the presence of methanogens[135,138,139]. The average relative abundance of non-DPOM *Desulfobacterota* (0.1 ± 1.6%) is two orders of magnitude lower than the average relative abundance of DPOM *Desulfobacterota* (36 ± 28%) across enrichment communities. At such low abundance and without the *ptx-ptd* gene cluster, the non-DPOM *Desulfobacterota* are likely sustained by organic carbon supplied by either rumen fluid or the DPOM, and their growth is likely coupled to the activity of methanogens within the community.

All *Bacteroidota* are of the *Bacteroidales* taxonomic order. The most common taxonomic families to occur are the *Dysgonomonadaceae* (19% of *Bacteroidales* across enrichments) and the undefined family UBA932 (24% of *Bacteroidales* across enrichments). The *Dysgonomonadaceae* family includes the genera *Petrimonas, Proteiniphilum,* and *Fermentimonas*, while the UBA932 family includes the undefined genera DMER64 and Bact-08. Despite the ubiquity of *Bacteroidota* across all enrichments, their relative abundance within their respective communities is low (1.9 ± 2.5%). The *Bacteroidales* are a metabolically heterogenous group of bacteria, and the closest cultured relatives to the genera enriched here are broad-range fermenters whose metabolic activity results in acidogenesis via the production of acetate, propionate, succinate, butyrate, isovaleric acid, hydrogen, and CO_2 from diverse carbon substrates[201]. Though physiologically ambiguous, the metabolic activity of these fermenters likely supports the fatty-acid profile that was observed in the HEPO culture via GC-MS analysis in **Chapter 4**.

All *Synergistota* are of the *Synergistales* taxonomic order. The most common taxonomic family is the *Aminobacteriaceae* (52% of *Synergistota* across enrichments). These *Aminobacteriaceae* are represented by the genus *Aminobacterium* or the undefined genus Syner-01. Like the *Bacteroidota*, the *Synergistota* exhibit low relative abundance (1.3 ± 1.9%) within

their respective communities, despite their ubiquity across enrichments. In the literature, *Aminobacterium* are represented by the cultures of *Aminobacterium colombiense, Aminobacterium mobile,* and *Aminobacterium thunnarium*[202,203]. *Aminobacterium* are nonsaccharolytic and aminolytic, fermenting pyruvate and the amino acids serine, glycine, and threonine in pure culture, with acetate as the primary fermentation product. When in the presence of the hydrogenotrophic methanogen *Methanobacterium formicicum*, additional amino acids are also fermented, including alanine (to acetate), valine (to isobutyrate), leucine (to isovalerate), and isoleucine (to 2-methylbutyrate). Short chain organic acids (i.e., lactate, fumarate, succinate) are not utilized by known *Aminobacterium*[202]. The sustenance of *Aminobacterium* across DPO enrichments suggests that DPO communities generate a basal level of amino acids which can sustain these aminolytic bacteria. However, it is unclear whether basal concentrations of amino acids are generated from metabolic activity or from proteolysis of dead cell matter or from rumen fluid substrates.

All active DPO enrichments also contain methanogenic archaea from the *Halobacterota* phylum in low relative abundance (2.5 ± 3.8%). The *Halobacterota* present across enrichments are from the taxonomic classes of either *Methanosarcina* (54% of *Halobacterota* across enrichments) or *Methanomicrobia* (46% of *Halobacterota* across enrichments). Of the *Methanosarcina* genera, 20% are *Methanomethylovorans*, 20% are *Methanosarcina mazei,* 27% are *Methanothrix*, and 33% are *Methanothrix_A*. The *Methanomicrobia* are predominantly of the genus *Methanoculleus* (69% of *Methanomicrobia* across enrichments), but also include *Methanospirillum* and the undefined genus SD8. The inoculum for all DPO enrichments was digester sludge from MWTFs, which explains the phylogenetic profile of enriched methanogens because *Methanosarcina, Methanothrix,* and *Methanoculleus* are all abundant members of wastewater digester sludge[204]. However, their sustenance across selective DPO enrichments suggests substrates are consistently available to support their survival. The *Methanosarcina* genera that are present across enrichments are typically heterotrophic, preferentially utilizing methylated substrates such as methanol, methylamines, and acetate[205–207]. Meanwhile, the *Methanoculleus* of *Methanomicrobia* are typically hydrogenotrophic, generating methane from H_2 and CO_2. Some species of *Methanoculleus* may also use formate, 2-propanol, and 2-butanol coupled to CO_2 reduction[208]. The maintenance of both heterotrophic and hydrogenotrophic populations of methanogens suggests that DPO enrichments supply methylated substrates as well as hydrogen. However, it is unclear whether these substrates are directly generated by DPOM, if they are secondary products of fermenters acting on the reduced carbon from DPOM CO_2 fixation, or if they are merely supplied from the addition of rumen fluid.

The microbial phyla conserved across all DPOM enrichments in this study (*Desulfobacterota, Bacteroidota, Synergistota,* and *Halobacterota*) were also conserved in the Phox-21 culture. In the Phox-21 enrichment, the *Desulfobacterota* were represented by the DPOM *Phosphitivorax anaerolimi* Phox-21; the *Bacteroidota* were represented by the genera *Petrimonas* and *Proteiniphilum;* the *Synergistota* were represented by the genera *Aminobacterium* and *Aminomonas*; and the *Halobacterota* archaea were represented by *Methanoculleus* and *Methanococcoides* (also known as *Methanomethylovorans*)[70]. It is tempting to speculate that the conserved community structure across DPO enrichments, including that of the Phox-21 enrichment, is the result of DPOM activity generating selective conditions that support predictable and defined community structures. However, also common between all the aforementioned enrichments is the addition of rumen fluid, which was found to stimulate DPO activity[70]. It remains unclear which substrate in rumen fluid stimulates DPO activity, and the complex organic

composition of rumen fluid could instead be the primary selective pressure that maintains common community structures across DPOM enrichments[109].

Community Composition of the HEPO Culture

The HEPO culture exhibits growth and DPO activity without rumen fluid (see **Chapter 2**), offering a platform to analyze the microbial community of a DPO enrichment without the confounding presence of complex organic matter. The HEPO culture has been maintained and transferred for over three years (inoculated in June 2018) in conditions where phosphite is the only exogenous electron donor and CO_2 is the only exogenous electron acceptor. Carbon dioxide is also likely to be the primary carbon source, as the growth media contains 60 mM bicarbonate under a headspace of 80:20 N_2:CO_2 and is devoid of any complex organics (e.g. rumen fluid). Beyond the trace quantities of vitamins that are included in the media (see Materials and Methods), L-cysteine HCl (3 mM) is the only additional organic compound added, serving as a reducing agent to protect the anaerobic culture from oxygen contamination. Microorganisms can metabolize L-cysteine to hydrogen sulfide, ammonia, and pyruvate via the activity of L-cysteine desulfhydrase[209]. While L-cysteine desulfhydrase is not present in any DPOM genomes, cysteine-S-conjugate β-lyase is capable of L-cysteine desulfhydrase activity[209] and is present in some DPOM, including those of the HEPO culture. These DPOM should be capable of utilizing L-cysteine as a sulfur, nitrogen, and/or carbon source. However, physiological analyses in **Chapter 4** have shown that i) DPO-dependent growth of the HEPO culture is only possible in the presence of CO_2 (Fig. 4.4), suggesting L-cysteine HCl cannot replace CO_2 as an electron acceptor or carbon source, and ii) no growth is observed in the absence of phosphite (Fig. 2.8A), suggesting L-cysteine HCl cannot serve as a sole electron donor to the HEPO culture. These data coupled with additional metagenomic and physiological analyses in **Chapter 4** suggest that the DPOM in the HEPO culture utilize electrons from phosphite to reduce CO_2 to an unknown organic compound (Compound X). Since growth of the HEPO culture is only ever observed with the onset of phosphite oxidation (Fig. 2.8), primary productivity by the DPOM is likely the driving force that sustains the remaining microbial community of the HEPO culture.

Metagenome sequencing of the parental SV3 enrichment (inoculated in June 2018), estimated the community contained 24 members and 2 species of *Phosphitivorax* DPOM. The HEPO culture was subsequently cultivated from the SV3 enrichment in the absence of rumen fluid, and metagenome sequencing in March 2019 predicted the community contained 16 members and 2 species of *Phosphitivorax* DPOM, suggesting the removal of rumen fluid had simplified the community but had maintained the original DPOM members. 16S rRNA gene sequencing of the HEPO microbial community in September 2021, revealed that despite over three years of highly selective growth conditions, 28 microbial OTU's persist at > 0.1% relative abundance, and four of these OTU's are strains of *Phosphitivorax* DPOM (Fig. 5.3). The higher number of species identified in this recent sequence analysis of the HEPO culture is likely due to the fact that 16S rRNA gene amplicon sequencing consistently detects a greater number of taxa than metagenome sequencing of the same sample[210]. The diversity of *Phosphitivorax* 16S rRNA gene sequences also suggests that DPOM strains in the HEPO culture are diversifying to occupy alternative niches within the community[211].

Of the 28 OTU's detected via 16S rRNA gene amplicon sequencing of the HEPO culture, just 10 OTU's comprise 94% of the community. Four of these OTU's are strains of *Phosphitivorax* DPOM that are most closely related to *Phosphitivorax anaerolimi*. The *Phosphitivorax* DPOM

collectively comprise 64 ± 0.6% of the community. The remaining six OTU's (comprising 30% of the community) include: an unclassified bacterium (16 ± 0.9%), *Mesotoga prima* (8 ± 0.2%), *Desulfomicrobium baculatum* (1 ± 0.4%), *Methanoculleus* (3 ± 0.2%), *Porphyromonadaceae* (1 ± 0.5%), and *Aminovibrio pyruvatiphilus* (1 ± 0.1%). The community profile of the HEPO culture conserves the four major phyla seen across DPO enrichments: the *Desulfobacterota* phylum is represented by the *Phosphitivorax* and *Desulfomicrobium baculatum*; the *Bacteroidota* phylum is represented by the *Porphyromonadaceae,* which is most closely related to the *Petrimonas* genera; the *Synergistota* phylum is represented by the *Aminovibrio pyruvatiphilus*; and the *Halobacterota* phylum is represented by *Methanoculleus.* Unlike previous enrichments, the HEPO culture has been maintained in the absence of rumen fluid for over 3 years, suggesting that DPOM activity, rather than exogenous carbon substrates from rumen fluid, selectively maintains this conserved community profile.

In addition to the four phyla that are ubiquitously seen across DPOM enrichments, *Mesotoga prima* from the *Thermotogota* phylum and an unclassified bacterium are present at significant relative abundance in the HEPO culture. At 8 ± 0.2% and 16 ± 0.9% relative abundance respectively, these two OTU's are also the most prevalent non-DPOM (Fig. 5.3). Interestingly, the Phox-21 enrichment community was also characterized by the presence of a *Thermotogota* bacterium and an uncultured bacterium[70]. *Mesotoga prima* from the HEPO culture shared 100% 16 rRNA gene sequence identity with *Thermotogota* from the Phox-21 enrichment. Furthermore, *Thermotogota* were present at low relative abundance (1.2 ± 1.1%) in 9 of the 11 DPO enrichments, and 87% of these *Thermotogota* representatives were of the genus *Mesotoga*. While the physiological role of *Thermotogota* in DPO communities is unclear, their conservation across the majority of DPO enrichments and their relative prevalence in both the HEPO culture and the Phox-21 culture, irrespective of the presence of rumen fluid, suggest favorable selection in the presence of *Phosphitivorax* DPOM. Meanwhile, no 16S rRNA gene sequence in the Phox-21 enrichment shared more than 81% similarity with the 16S rRNA gene sequence of the uncultured bacteria of the HEPO culture. It is therefore unclear if the two uncultured bacteria that are prevalent in the HEPO and Phox-21 enrichments are physiologically selected for. However, future work must prioritize characterizations of these shared community members to define their role and understand their persistence in the presence of *Phosphitivorax* DPOM.

Methanogens in the HEPO Culture

The ubiquity of methanogens across DPO enrichments and within the HEPO culture is of particular interest, as this could indicate a physiological relationship between DPO metabolism and methanogenesis, which is an environmentally and industrially important metabolic process[12]. Though not all-inclusive, three major classes of electron donors are typically utilized for methanogenesis: hydrogen, acetate, and methylated C_1 compounds (e.g., trimethylamine; TMA)[212]. As expected, the maximum cell density of the HEPO culture increased from the control (max OD_{600}: 0.14 ± 0.01) following the addition of methanogenic growth substrates of hydrogen (max OD_{600}: 0.17 ± 0.02), acetate (max OD_{600}: 0.17 ± 0.02), and TMA (max OD_{600}: 0.27 ± 0.02) (Fig. 5.4). Hydrogen and TMA support increased growth yields without altering the rate of phosphite oxidation or DPO-dependent growth (Fig. 5.4). In contrast, the addition of acetate slows DPO activity and growth, suggesting that the accumulation of acetate, which is also endogenously generated during DPOM growth (Fig. 4.6), may inhibit DPO activity. However, while hydrogen, acetate, and TMA can each be utilized by methanogens, augmented growth of the HEPO culture

in the presence of these substrates cannot be definitively attributed to methanogenic growth, as methanogenesis and community composition were not temporally monitored.

To test the relationship between methanogenesis and DPO activity, specific inhibitors of methanogens were introduced to the HEPO culture, including 2-bromoethanesulfonate (BES), ethyl-2-butynoate (E2B), and lumazine (Fig. 5.5).

BES is an analogue of coenzyme M (CoM)[213] and competitively inhibits methanogenesis[214]. BES is widely considered a specific inhibitor of methanogenesis, since methanogens are the principal producers and utilizers of CoM[94]. However, recent work has shown that bacteria also utilize CoM in the synthesis of alkenes[215] and can be susceptible to BES inhibition[216].

E2B is an unsaturated ester that has shown mixed success in suppressing methanogenesis of microbial cultures[217]. E2B is hypothesized to serve as an electron sink, thereby withdrawing electrons from hydrogen and preventing the growth of hydrogenotrophic methanogens[217]. E2B is also directly toxic to methanogens, but different methanogenic species vary in their sensitivity to E2B toxicity[218].

Lumazine is a pterin compound that was found to competitively inhibit methanogenesis[214,219]. Lumazine acts as a partial structural analogue to folates and flavins, including coenzyme F_{420}, which is found in all methanogens and some actinomycetes[220]. While lumazine has been found to specifically inhibit methanogens, it is less effective in mixed communities where it may be consumed as a substrate by other organisms, which would consequently lower the concentrations of lumazine to which methanogens are exposed[218]

While lumazine and E2B addition briefly inhibited growth in the HEPO culture, after two to three days, growth rebounded, and cells reached optical densities equivalent to the control (Fig. 5.5). The relatively minor effects of lumazine and E2B on the growth of HEPO suggest that DPOM are largely unaffected by these inhibitors and would further indicate that DPO activity acts independently of methanogenesis. However, DPO activity and growth were immediately inhibited by the addition of BES (Fig. 5.5). If BES directly inhibits methanogens in the HEPO community, the pronounced inhibition of DPO activity would suggest that DPOM depend on methanogens. However, lumazine and E2B are also methanogenic inhibitors[214,218], so DPO should have been similarly inhibited following their addition. Alternatively, BES may be a direct inhibitor of the DPOM in HEPO, which would suggest that DPOM rely on a CoM analogue for their growth and activity.

To understand the role of the tested inhibitors at the community level, DNA was extracted temporally for 16S rRNA gene sequencing from HEPO cultures that were grown in the presence and absence of BES and lumazine (Fig. 5.6). As was previously observed (Fig. 5.5), BES inhibits growth of the HEPO culture immediately upon addition, while lumazine shows a slight growth defect but is largely tolerated (Fig. 5.6A). Lysate from a *Methanosarcina acetivorans* culture does not rescue BES inhibition (Fig. 5.6A). At day 0, the community contains 12 OTU's with relative abundances >1%, but it is dominated (>86% relative abundance) by just four OTU's: ZOTU1, ZOTU2, ZOTU3, and ZOTU4 (Fig. 5.6 B-E). ZOTU1 is an unclassified bacterium and is the dominant OTU (34 ± 0.2% relative abundance) at day 0. ZOTU2 is classified as *Mesotoga prima* and occupies 18 ± 0.4% of the community at day 0. ZOTU3 and ZOTU4 are DPOM, as their 16S rRNA genes share 99.7% and 98.9% sequence identity, respectively, with the reconstructed 16S rRNA gene of *Phosphitivorax anaerolimi* SV3. At day 0, each *P. anaerolimi* strain is present at

similar relative abundance (ZOTU3, 17 ± 0.4%; ZOTU4, 17 ± 1.1%). ZOTU8 is the only methanogen detected in the community and is classified as *Methanoculleus*.

In the control condition (no inhibitors added), one strain of *P. anaerolimi* (ZOTU4) blooms, dominating the culture at 62 ± 4% relative abundance in early stationary phase and outcompeting the other *P. anaerolimi* strain (ZOTU3) by the same timepoint (0.9 ± 0.3% relative abundance) (Fig. 5.6B, F, & G). In exponential phase, *Desulfomicrobium baculatum* (ZOTU7) and two strains of *Clostridiales* (ZOTU19 and ZOTU26) bloom with the DPOM but diminish again by stationary phase. *Methanoculleus* also blooms with DPOM during exponential phase but maintains its maximum relative abundance (3 ± 0.1%) well into stationary phase (Fig. 5.6B & H). Meanwhile, the unclassified bacterium (ZOTU1) and *Mesotoga prima* (ZOTU2) maintain relatively consistent populations through exponential phase and into stationary phase (Fig. 5.6B).

The temporal community dynamics of the lumazine condition are nearly identical to the control condition. Like the control, the DPOM ZOTU4 dominates the culture by early stationary phase (Fig. 5.6C), though the maximum relative abundance (49 ± 5%) is lower than in the control condition (62 ± 4%). This appears to be due to the growth during stationary phase of a third *Clostridia* strain (ZOTU37). The unclassified bacterium (ZOTU1) and *Mesotoga prima* (ZOTU2) exhibit similar temporal abundance values as the control, while *Methanoculleus* (ZOTU8) appears to be inhibited by the addition of lumazine, as its maximum relative abundance (2 ± 0.2%) is reduced relative to the control (3 ± 0.1%) (Fig. 5.6H). This suggests lumazine is acting as an effective methanogen inhibitor in these culture conditions.

The addition of BES significantly changes the temporal community dynamics of the HEPO culture (Fig. 5.6E). As expected, the addition of BES appears to inhibit *Methanoculleus* (ZOTU8) (Fig. 5.6E & H). Notably, BES addition also causes an immediate and lasting decline in the DPOM population as well (Fig. 5.6E & G). Consistent with growth data (Fig. 5.6A), *M. acetivorans* cell lysate does not rescue the DPOM population and instead appears to support the outgrowth of the unclassified bacterium (ZOTU1), *Mesotoga prima* (ZOTU2), and *Desulfomicrobium baculatum* (ZOTU7) (Fig. 5.6D). The precipitous decline of *P. anaerolimi* (ZOTU4) in response to BES contrasts with *Methanoculleus* (ZOTU8), whose inhibition is delayed for one day after BES addition and results in stasis rather than cell lysis (Fig. 5.6G & H). This suggests that *P. anaerolimi* is not metabolically linked to *Methanoculleus* (ZOTU8) inhibition but is instead directly inhibited by BES. This interpretation is supported by the fact that *P. anaerolimi* (ZOTU4) continues to grow in the presence of lumazine, despite inhibition of *Methanoculleus* (ZOTU8) by lumazine.

The direct inhibition of DPOM growth by BES suggests that DPOM in the HEPO culture utilize CoM as a cofactor in energy metabolism. However, the DPOM genomes do not harbor any CoM synthesis genes (*comABCDE*)[117]. It is possible that the DPOM contain a bacterial CoM synthesis pathway whose sequence homology is too distant to be computationally detected[215]. Alternatively, it is possible that DPOM import CoM that is synthesized by the community methanogens. While biochemical evidence exists to support the existence of CoM transporters in methanogens[221–223], their mechanism, sequence motifs, and prevalence in bacteria have not yet been demonstrated.

Inhibition Physiology of FiPS-3

To test the direct effects of methanogen inhibitors on DPOM, a pure culture of the DPOM *Desulfotignum phosphitoxidans* FiPS-3 was monitored for growth and activity in the presence of lumazine and BES (Fig. 5.7). Neither DPO activity nor growth of FiPS-3 are inhibited by BES or lumazine, regardless of whether FiPS-3 is grown with CO_2 (Fig. 5.7A & C) or sulfate (Fig. 5.7B, D & E) as the electron acceptor. However, as was observed in the HEPO culture, lumazine does cause a slight growth limitation by reducing the maximum cell density to ~80% of the control (Fig. 5.7A & B). Lumazine is a pterin compound that may competitively inhibit the WLP of FiPS-3 and the reductive glycine pathway of *P. anaerolimi*, as both autotrophic pathways utilize tetrahydrofolate (THF), a compound which includes a pterin ring[117,224]. Future studies will be needed to test the concentration range and mechanism of lumazine inhibition in DPOM.

Preceding studies suggest that DPO-dependent growth of *P. anaerolimi* in the HEPO culture is directly inhibited by BES. BES is an analogue of CoM, which functions as a carrier molecule in alkene oxidation and methanogenesis[225]. Given the integral role of CoM in carbon redox chemistry, it is tempting to speculate that *P. anaerolimi* utilizes CoM during carbon reduction. The mechanism for energy conservation via carbon reduction in *P. anaerolimi* has not yet been characterized but is predicted to be a variant of the reductive glycine pathway[117]. Unlike *P. anaerolimi,* FiPS-3 utilizes the WLP for carbon reduction, which does not require CoM and might be the reason that FiPS-3 can evade BES inhibition[67]. Future studies will be needed to elucidate the mechanism and targeted pathway for BES inhibition of DPO-dependent growth in the HEPO culture.

Antibiotic Profile of the HEPO Culture

The HEPO culture was exposed to standard concentrations of ampicillin, erythromycin, kanamycin, and chloramphenicol and was monitored for growth and DPO activity (Fig. 5.8). When compared to the control, kanamycin (25 μg/mL) initially delayed growth, but the HEPO culture adapted and was able to oxidize all the phosphite to reach control cell densities (Fig. 5.8), suggesting that DPOM harbor mechanisms for kanamycin resistance[226]. Ampicillin (100 μg/mL) limits growth of the HEPO culture, but it does not alter DPO activity. This may suggest that non-DPOM organisms in the HEPO culture show greater susceptibility to ampicillin than the DPOM. Growth and DPO-activity respond negatively to increasing concentrations of ampicillin, and concentrations > 300 μg/mL appear to be necessary for full inhibition (Fig. 5.9). Both erythromycin (50 μg/mL) and chloramphenicol (50 μg/mL) fully inhibit DPO-dependent growth of the HEPO culture (Fig. 5.8). Erythromycin stocks are solubilized in ethanol, and controls spiked with an equivalent volume (50 μL) of ethanol show enhanced growth and DPO activity, demonstrating that i) erythromycin inhibition is due to the antibiotic and not the solvent, and ii) ethanol is utilized as a substrate by the HEPO culture. Erythromycin and chloramphenicol are both protein synthesis inhibitors that bind the 50S ribosomal subunit, suggesting that DPOM in the HEPO culture are sensitive to this mechanism of inhibition.

Stimulation of the HEPO Culture by Spent Media and Cell Lysates

DPOM in the HEPO culture utilize phosphite oxidation to reduce CO_2, thereby supporting a complex microbial community with the resulting fixed carbon and soluble phosphate. Evidence

suggests that members of the non-DPOM community also support DPOM activity and growth, likely by supplying essential nutrients and/or cofactors. It was therefore hypothesized that spent media from retired HEPO cultures might stimulate DPO activity of an active HEPO culture. To test this, the HEPO culture was grown in the presence of filtered supernatant (filtrate) and lysed cells (lysate) from an expended HEPO culture (Fig. 5.10). When compared to controls, filtrate was inhibitory, as it reduced cell yields and slowed the rate of DPO activity. This suggests that the HEPO community generates a product that inhibits DPOM growth and activity. Fermentation products, such as the fatty acids shown to accumulate in **Chapter 4,** might elicit this response by creating acidic or thermodynamically unfavorable conditions for growth. Acetate is one such product that was shown to inhibit the HEPO culture upon addition (Fig. 5.4). In contrast, lysate was stimulatory. While lysate did not increase cell yields, it reduced the lag phase of DPO-dependent growth by three to four days. Lysate that had been heat pasteurized was also stimulatory, but less-so than untreated lysate (Fig. 5.10). This suggests that cells of the HEPO community synthesize a heat-sensitive substrate that stimulates DPOM.

To test the stimulatory effects of methanogens on DPO activity and growth, the HEPO culture was grown in the presence of filtered supernatant (filtrate) and lysed cells (lysate) from a *Methanosarcina acetivorans* culture (Fig. 5.11). Unlike the HEPO filtrate, *M. acetivorans* filtrate was stimulatory to the HEPO culture, reducing the lag phase by 4 days and increasing cell yields (0.17 ± 0.01 ΔOD_{600} units) when compared to the control (0.12 ± 0.01 ΔOD_{600} units) (Fig. 5.11A). Phosphite oxidation also began two days earlier, with a slightly increased rate (1.8 ± 0.66 mM HPO_3^{2-}/day) when compared to the control (1.0 ± 0.67 mM HPO_3^{2-}/day) (Fig. 5.11B). Whole cells and cell lysates also improved growth yields and phosphite oxidation rates, but heat-treated cell lysates were less effective, as had been observed for HEPO cell lysates (Fig. 5.11). This suggests that *M. acetivorans* cultures contain a heat sensitive nutrient that is stimulatory to the DPOM in the HEPO culture.

Corrinoid Profile of DPOM and the HEPO Culture

In methanogens, the coenzyme M-methyltransferase mediates a methyl transfer from methyltetrahydromethanopterin to CoM via a corrinoid-dependent ping-pong mechanism[227]. Genome analyses suggest that DPOM utilize THF (the bacterial equivalent of archaeal tetrahydromethanopterin) for methylene transfer in the reductive glycine pathway[117]. Mild inhibition by the pterin, lumazine, supports this genomic hypothesis. Physiological analyses further show that DPOM are directly inhibited by BES, suggesting they also utilize CoM in DPO-dependent carbon reduction (Fig. 5.6). Given these observed relationships between DPOM and the homologous cofactors used by methanogens in coenzyme M-methyltransferase, corrinoids were next investigated for their role in the HEPO culture.

The DPO MAGs generated in **Chapter 2** were computationally analyzed for their predicted capacity to utilize (Fig. 5.12) and biosynthesize (Fig. 5.13) corrinoids, based on the genomic requirements described previously[197]. DPOM were universally predicted to utilize corrinoids based on the presence of two corrinoid-dependent genes (Fig. 5.11). All but one DPO MAG contained the corrinoid-dependent methionine synthase (*metH*), which catalyzes the synthesis of methionine from homocysteine via corrinoid-mediated methyl-transfer[228]. No DPO MAGs contained the corrinoid-independent methionine synthase (*metE*), suggesting DPOM are dependent on corrinoid cofactors to synthesize methionine (Fig. 5.12). Additionally, *ptdH* is

considered a core gene of the *ptx-ptd* gene cluster for DPO metabolism and is present in diverse DPOM (Fig. 5.12). PtdH is predicted to be a corrinoid-dependent radical S-adenosyl L-methionine (SAM) family protein[29]. As such, corrinoids are likely to be directly necessary for DPO metabolism. In addition to utilizing a corrinoid, PtdH is predicted to utilize the SAM cofactor. Methionine is a biosynthetic precursor of SAM, and as described above, methionine biosynthesis in DPOM is likely corrinoid-dependent, reinforcing the requirement of corrinoids for DPO metabolism.

Despite the predicted requirement of corrinoids by all DPOM, genome analysis shows that DPOM are largely incapable of endogenously biosynthesizing complete corrinoids (Fig. 5.13). Corrinoid synthesis follows a complex biochemical pathway that has been eloquently described previously[197] and in summary, requires the biosynthesis of a tetrapyrrole precursor which is converted, either aerobically or anaerobically, into a corrin ring. Attached to the corrin ring is the nucleotide loop and the lower ligand, and the lower ligand largely defines the structural variability of corrinoids seen across bacterial taxa[229]. All but three DPOM MAGs contain the complete set of genes required to synthesize the tetrapyrrole precursor, which is a common precursor to the biosynthesis of heme, chlorophyll, and other tetrapyrroles involved in antibiotic resistance[197]. However, the genes for corrin ring biosynthesis are only found in FiPS-3 and *Sporomusaceae* bacterium SM1 (Fig. 5.13), and the genes for nucleotide loop assembly are only found in FiPS-3, *Sporomusaceae* bacterium SM1, and the *Pelotomaculaceae* (Fig. 5.13). As such, only FiPS-3 and *Sporomusaceae* bacterium SM1 are likely capable of endogenous corrinoid biosynthesis.

Since the remaining DPOM appear unlikely to synthesize corrinoids, their genomically-predicted corrinoid requirements are likely satisfied by community nutrient exchange[230]. Eight DPOM genomes contain genes associated with corrinoid remodeling, suggesting they may import corrinoids from the community and structurally alter them for the biochemical needs of the cell[198]. Seven DPOM genomes contain the BtuB, which is associated with corrinoid transport[231]. A *Phosphitivorax* DPOM from the HEPO culture is one of the genomes that harbors BtuB, suggesting corrinoid exchange might benefit DPO activity in the HEPO culture. Lastly, most DPOM have the enzymatic machinery required to import cobalt (Fig. 5.13), which is the catalytic metal ion of corrinoids[198]. However, cobalt can be used for a variety of non-corrinoid, cobalt-dependent enzymes[232] that are also found throughout the DPOM genomes (Fig. 5.13), making this feature minimally predictive of corrinoid utilization.

Genomic analyses predict that all DPOM require corrinoids and that the majority of DPOM import corrinoids from the environment (Fig. 5.12 & 5.13). To begin testing this hypothesis, corrinoids were extracted from a 0.3 g cell pellet of the HEPO culture. Corrinoids in the dried extract were separated and analyzed via HPLC (Fig. 5.14A). Corrinoids have a distinct UV/visible light (UV/vis) spectrum that is characterized by a signal maximum near 360 nm and a shoulder that extends to approximately 550 nm. The exact UV/vis spectrum for a given corrinoid is influenced by the upper ligand, cobalt oxidation state, and lower base coordination[233]. This UV/vis spectral profile was used to distinguish corrinoids (Fig. 5.14C – F) from those compounds that were extracted but unlikely to be a corrinoid (Fig. 5.14G). Based on UV/vis spectral analyses, four corrinoids are present in the cell lysate of the HEPO culture (Fig. 5.14).

The retention time (RT) of extracted corrinoids was compared to chemical standards to predict their identity (Fig. 5.14B). Three of the four corrinoids in the HEPO culture were identified based on their retention time as follows: 8.064 min, 2-methyladenine cobamide ([2-MeAde]Cba); 8.442 min, 5-hydroxybenzimidazole cobamide ([5-OHBza]Cba); 12.1 min, cobalamin (B12) (Fig. 5.14A

& B). Chemical standards were unable to identify the corrinoid peak at 15.65 min (Fig. 5.14A). The HPLC retention pattern of the 15.65 min peak is unique because its signal is composed of a single peak (Fig. 5.14A). This is in contrast to the known corrinoids that elute after cobalamin (>12.1 min), which are typically "base-off" corrinoids that exhibit a split peak due to chiral cyanide binding[234]. As such, the 15.65 min corrinoid peak is likely a "base-on" corrinoid with an unusually high RT (Fig. 5.14A).

To confirm the RT-based assignments, and to assign identity to peak 15.65 min, HPLC fractions of each peak were collected and analyzed via direct infusion electrospray ionization mass spectrometry (MS) (Fig. 5.15). MS analysis confirmed that the peak at 8.064 min is [2-MeAde]Cba (Fig. 5.15D), and the peak at 8.442 min is [5-OHBza]Cba (Fig. 5.15F). The peak at 12.1 min was not analyzed via MS because its RT identification was unambiguous (Fig. 5.15A & B). However, MS analysis was still unable to identify the peak at 15.65 min, as its signal did not match any known corrinoids in the mass spectral databases that were searched (Fig. 5.15E).

Four corrinoids are present in the HEPO culture (Fig. 5.14). B12 is detected because it is included in the vitamin mix of the basal bicarbonate media in which the HEPO culture is grown[94]. The remaining corrinoids ([2-MeAde]Cba, [5-OHBza]Cba, and the unknown corrinoid with RT 15.65 min) are likely to be biosynthesized by the HEPO community. [5-OHBza]Cba is strongly associated with methanogens and might be a product of the *Methanoculleus* in the HEPO culture[229,235]. However, the *Phosphitivorax* DPOM in the HEPO culture are unlikely to synthesize any of these corrinoids, as they do not contain corrinoid biosynthesis genes, and future studies will be necessary to identify which organisms synthesize each of the detected corrinoids. Meanwhile, genomic analyses suggest that the *Phosphitivorax* DPOM in the HEPO culture require corrinoids, but it is currently unclear which, if any, of the four detected corrinoids support DPO activity. Future studies monitoring the fate and association of labelled corrinoid variants will be required to identify their relationship to DPO, with the important prerequisite of characterizing the unknown corrinoid (RT 15.65 min).

Stimulation of the HEPO Culture via the Addition of SAM Cofactor

The aforementioned studies have demonstrated the potential importance of unique cofactors (i.e., CoM, pterin analogues, and corrinoids) for DPO activity and growth. Sequence alignments also predict that the SAM cofactor is utilized by the PtdH protein from the DPO *ptx-ptd* cluster. Radical SAM proteins are responsible for catalyzing diverse biochemical reactions, of which methylation reactions are common[236]. It remains unclear what biochemical function is served by the PtdH of DPOM, but it was hypothesized that the addition of the SAM cofactor might stimulate DPOM growth and activity in the presence of excess phosphite. As was previously observed (Fig. 2.8), phosphite concentrations > 20 mM were growth limiting, indicating that 30 mM phosphite is in excess (Fig. 5.16A & C). The HEPO culture was grown in excess phosphite (30 mM) and spiked with variable concentrations of SAM cofactor to see if SAM could enhance growth and phosphite oxidation (Fig. 5.16B & D). SAM concentrations below 500 μM showed no physiological change when compared to the controls. However, SAM concentrations of 500 μM and 1000 μM showed a stepwise increase in DPO-dependent growth (Fig. 5.16B & D).

To monitor the community composition, DNA was extracted for 16S rRNA gene amplicon sequencing on day 15 and 27 of the growth curves (Fig. 5.16E – H). Differing concentrations of phosphite did not significantly alter the community structure of the HEPO culture (Fig. 5.16E & G). However, the relative abundance of *Phosphitivorax* in 500 μM and 1000 μM SAM on day 27

(70 ± 3%) is increased when compared to the controls on the same day (62 ± 2%) (Fig. 5.16H). This suggests that SAM is stimulating the DPOM, as opposed to supporting growth of fermenters, which would have instead caused a decrease in the relative abundance of the *Phosphitivorax*. Furthermore, SAM minimized *Phosphitivorax* diversity by stimulating the outgrowth of one *Phosphitivorax* strain (Fig. 5.16H). The stepwise, increased growth and DPO activity by a community dominated by *Phosphitivorax* suggests that SAM is an important cofactor for the DPOM of the HEPO culture.

Discussion

Comparisons of DPO enrichments show that DPOM are typically found in complex, methanogenic communities. Four phyla (*Desulfobacterota, Bacteroidota, Synergistota,* and *Halobacterota*) are conserved across all DPOM enrichments. These bacterial phyla are predicted to perform fermentations, consuming complex organics (e.g., carbohydrates and amino acids) to generate organic acids, hydrogen, and CO_2. The archaeal *Halobacterota* are represented by both hydrogenotrophic and methylotrophic methanogens. The methanogens are predicted to convert fermentative products into methane, but future studies are needed to validate and characterize methanogenesis in DPOM enrichments. Due to the universal presence of rumen fluid in the primary enrichments whose communities were analyzed, it was originally unclear if these phyla were carried over from wastewater inoculum, selectively conserved by components of rumen fluid, or if they were selected for by DPOM-generated metabolites. Growth of the HEPO culture is supported by phosphite and CO_2 in basal bicarbonate media that is not amended with any complex organic mixtures (i.e., rumen fluid). Despite the selectivity of the media broth, the HEPO culture has sustained a complex microbial community after over three years of selective enrichment, including all four phyla that were conserved across the primary DPOM enrichments, as well as *Mesotoga prima* and an unclassified bacterium whose relative abundances suggest a significant role. Given the temporal distance from the original wastewater inoculum, and given the absence of rumen fluid, the complex HEPO community is likely supported by the metabolism of *Phosphitivorax* DPOM, who are predicted to serve as primary producers that supply fixed carbon and soluble phosphate (**Chapter 4)**. It has been previously suggested that this role would be critically important in aphotic environments like subsurface aquifers, where geochemically synthesized phosphite could supply the energy and reducing equivalents that support microbial communities via the action of DPOM[29].

Evidence is presented throughout this dissertation and is summarized in this chapter's introduction that suggests DPOM metabolically rely on symbiotic nutrient exchange with non-DPOM community members of the HEPO culture. However, it is not yet clear which community members might support this hypothetical interaction nor what nutrients are being shared. Inhibition studies show that BES, a CoM analogue and a specific inhibitor of methanogens, immediately and completely inhibits DPO activity and growth of the HEPO culture. This initially suggested that DPOM in the HEPO culture have a tight relationship with methanogens in the community, but temporal analysis of 16S rRNA gene amplicons across a growth curve suggest instead that DPOM are directly inhibited by BES. The *Phosphitivorax* DPOM of the HEPO culture are predicted to conserve energy by coupling phosphite oxidation to CO_2 reduction. Direct inhibition by BES should therefore target the metabolic machinery of either phosphite oxidation or carbon reduction. When compared to the *Phosphitivorax* DPOM of the HEPO culture, FiPS-3 shares homologous *ptx-ptd* genes for DPO enzymatic machinery but utilizes the WLP for CO_2 reduction to acetate.

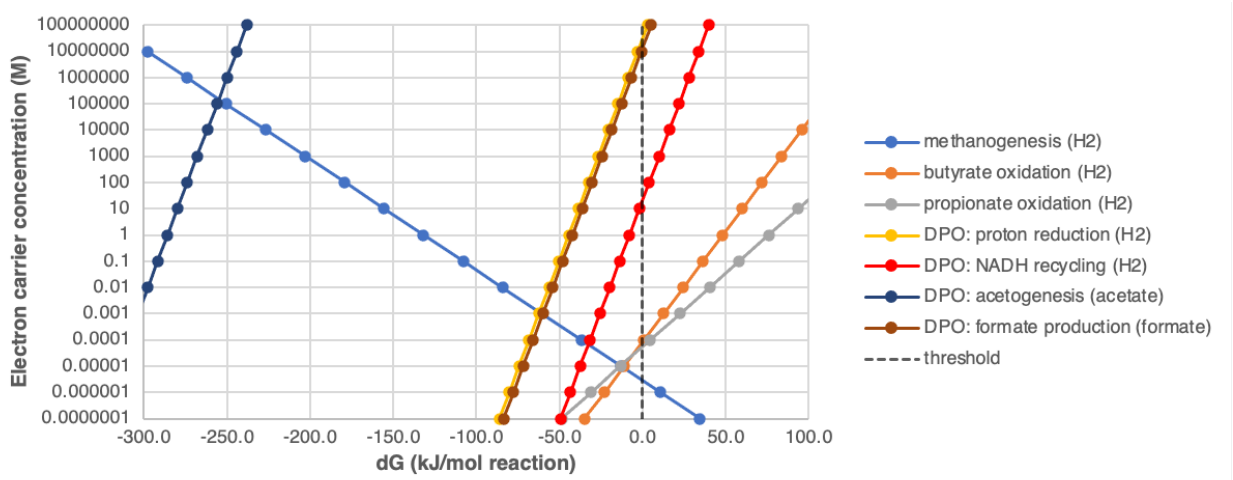
This contrasts with the *Phosphitivorax* DPOM, which do not harbor all the genes for a complete WLP and are instead predicted to reduce CO_2 via an alternative carbon reduction pathway that shares genes with the reductive glycine pathway (**Chapter 4**). We found that FiPS-3 was not inhibited by BES, suggesting that phosphite oxidation machinery is immune to BES inhibition and that BES is more likely to target the carbon reduction pathway of *Phosphitivorax.* Inhibition of the HEPO culture by BES implies that the *Phosphitivorax* DPOM utilize CoM, where CoM would likely serve as either a methyl carrier (as is seen in methanogens[235]) or as a nucleophile and carbon carrier (as seen in bacteria[213]). However, genes for CoM synthesis were not present in the *Phosphitivorax* genomes of the HEPO culture. It is possible that *Phosphitivorax* import CoM that is synthesized by non-DPOM HEPO community members[221,223]. Computationally searching the HEPO metagenome for CoM synthesis genes would identify CoM-producers that may be supporting *Phosphitivorax* metabolism.

The mechanism of BES inhibition in *Phosphitivorax* is unclear. In methanogens, BES targets the methyl-CoM reductase by binding and inactivating the nickel tetrapyrrole F_{430} and competitively excluding CoM[237]. BES inhibition of alkene-oxidizing bacteria competitively excludes CoM by targeting the redox-active cysteine pair in 2-ketopropyl-CoM carboxylase/oxidoreductase[213]. Based on these mechanistic studies, we might then predict that a BES-inhibited enzyme in *Phosphitivorax* includes either a redox-active cysteine pair, or a nickel-bearing tetrapyrrole analogous to F_{430}. It is unlikely that *Phosphitivorax* contain the coenzyme F_{430}, as it has only been detected in methanogens[235], and *Phosphitivorax* genomes do not contain genes (*cfbABCDE*) for its biosynthesis. However, of note is that the Ni(II)/Ni(I) redox couple for the nickel center of F_{430} has a standard reduction potential range of -630 to -650 mV[81], making F_{430} the only electron transfer cofactor whose reduction potential is within such close range of that for phosphite oxidation ($E^{o'}$ = -650 mV)(Table 1.1)[77,78,81,83,84,238,239]. A low-potential cofactor like F_{430} would be essential for DPOM to efficiently capture energy from phosphite oxidation. It is consequently tempting to speculate that *Phosphitivorax* DPOM utilize a cofactor with a nickel center whose reduction potential mirrors that of F_{430}.

Four corrinoids were identified in the HEPO culture: [2-MeAde]Cba, [5-OHBza]Cba, B12, and an unknown corrinoid. While B12 was likely residual from the growth media, the remaining three corrinoids must be synthesized by the HEPO community. Genome analyses revealed that, of the known DPOM, only FiPS-3 and *Sporomusaceae* bacterium SM1 are likely capable of synthesizing corrinoids. All other DPOM, including the *Phosphitivorax* of the HEPO culture, have incomplete corrinoid synthesis pathways. This contrasts with the finding that all DPOM likely require corrinoids based on the presence of *metH* and *ptdH.* It comes to no surprise, then, that FiPS-3 is one of just two DPOM that have been isolated. It will be interesting to see if the second isolate, *Phosphitispora fastidiosa*, is also capable of independently biosynthesizing corrinoids, as this might suggest a prerequisite for DPOM isolation. For the majority of DPOM, corrinoids are likely acquired from the environment, as is supported by the presence of the corrinoid transporter *btuB* in several DPOM. Corrinoids are a frequently exchanged cofactor in microbial communities[240], and in the HEPO culture, metagenomic analyses will be predictive of which community members are synthesizing the detected corrinoids. Tracking labelled variants of [2-MeAde]Cba, [5-OHBza]Cba, B12, and the unknown corrinoid will also be crucial in identifying which, if any, of these corrinoids are utilized by the *Phosphitivorax* DPOM, with the prerequisite identification of the unknown corrinoid.

The PtdH from the *ptx-ptd* gene cluster for DPO metabolism is predicted to be a corrinoid-dependent radical SAM family protein. As such, it is likely that one of the corrinoids detected in the HEPO culture is directly required for PtdH functionality. The PtdH-associated corrinoid could also be crucial for energy conservation from phosphite oxidation, as the Co(II)/Co(I) couple of corrinoid cobalt centers has a standard reduction potential below -500 mV (Table 1.1)[82,241]. The PtdH is also predicted to contain the SAM cofactor[29]. This homology-based prediction is supported by the fact that DPO-dependent growth of the HEPO culture is stimulated by the addition of SAM. S-adenosyl L-methionine synthetase catalyzes the synthesis of SAM from ATP and methionine, and it is present in all high-quality DPOM genomes. Genome analysis further reveals that all DPOM utilize the corrinoid-dependent methionine synthase (*metH*), suggesting corrinoids are also required for the synthesis of the SAM cofactor of PtdH. Methionine synthase methylates homocysteine with a methyl donor to generate methionine. Diverse methyl donors can be utilized in this reaction[242], where 5-methyl-THF is most likely used by DPOM given their universal employment of the methyl branch of the WLP, which necessitates and provides the THF cofactor[117]. It has been repeatedly observed that phosphite oxidation is preferentially coupled to CO_2 reduction and is hosted exclusively by autotrophic organisms[18,117]. This inspires the hypothesis that DPO metabolism is metabolically linked to CO_2 reduction. Analyses of the potential cofactors in DPOM corroborate this point, as CoM, corrinoids, and SAM are all associated with methylation reactions. However, protein assays and ^{13}C labelling studies coupled to genome metabolic mapping will all be necessary to elucidate the mechanistic link between phosphite oxidation and carbon reduction.

Figures



METABOLISM	CHEMICAL EQUATION	ΔG (KJ/MOL REACTION)	ΔG (KJ/MOL ELECTRON CARRIER)	Q EQUATION
METHANOGENESIS	$4H_2 + CO_2 \rightarrow CH_4 + 2H_2O$	-131.7	-32.925	1/([H2]^4)
BUTYRATE	$butyrate^- + 2H_2O \rightarrow 2\ acetate^- + H^+ + 2H_2$	48.3	24.15	[H2]^2
PROPIONATE	$propionate^- + 2H_2O \rightarrow acetate^- + CO2 + 3H_2$	76	25.33333333	[H2]^3
DPO: PROTON REDUCTION	$HPO_3^{2-} + H_2O \rightarrow HPO_4^{2-} + H_2$	-44.383	-44.383	[H2]
DPO: NADH RECYCLING	$HPO_3^{2-} + H_2O + ADP + Pi \rightarrow HPO_4^{2-} + H_2 + ATP$	-7.7	-7.7	[H2]
DPO: ACETOGENESIS	$4HPO_3^{2-} + 2CO_2 + 2H_2O \rightarrow 4HPO_4^{2-} + H^+ + C_2H_3O_2^-$	-285.595	-285.595	[acetate]
DPO: FORMATE PRODUCTION	$HPO_3^{2-} + CO_2 + H_2O \rightarrow HPO_4^{2-} + CH_2O_2$	-42.067	-42.067	[formate]

Figure 5.1: Gibb's Free Energy per mole of electron carrier for hypothetical syntrophic metabolisms of DPO. The free energy of reaction for hypothetical syntrophic metabolisms is plotted (top) for a range of electron carrier concentrations using $\Delta G = (\Delta G/mol\ carrier) + RTlnQ$, where ΔG represents the free energy (x-axis) available at differing concentrations of electron carriers (y-axis). The table (bottom) shows the balanced chemical equation for each plotted metabolism, with the hypothesized syntrophic electron carrier in red. R, gas constant, 8.314 J mol^{-1} K^{-1}. T, temperature, 310 K. Q, reaction quotient, calculated for each concentration as demonstrated in final column of the table. The vertical dashed line in the plot demarcates ΔG = 0. Methanogenesis, butyrate oxidation, and propionate oxidation are calculated and plotted for reference. The remaining reactions (DPO: xxx) represent hypothetical variations of DPO, in which the electron acceptor is either protons (to generate exchanged hydrogen) or CO_2 (to generate exchanged acetate or formate).

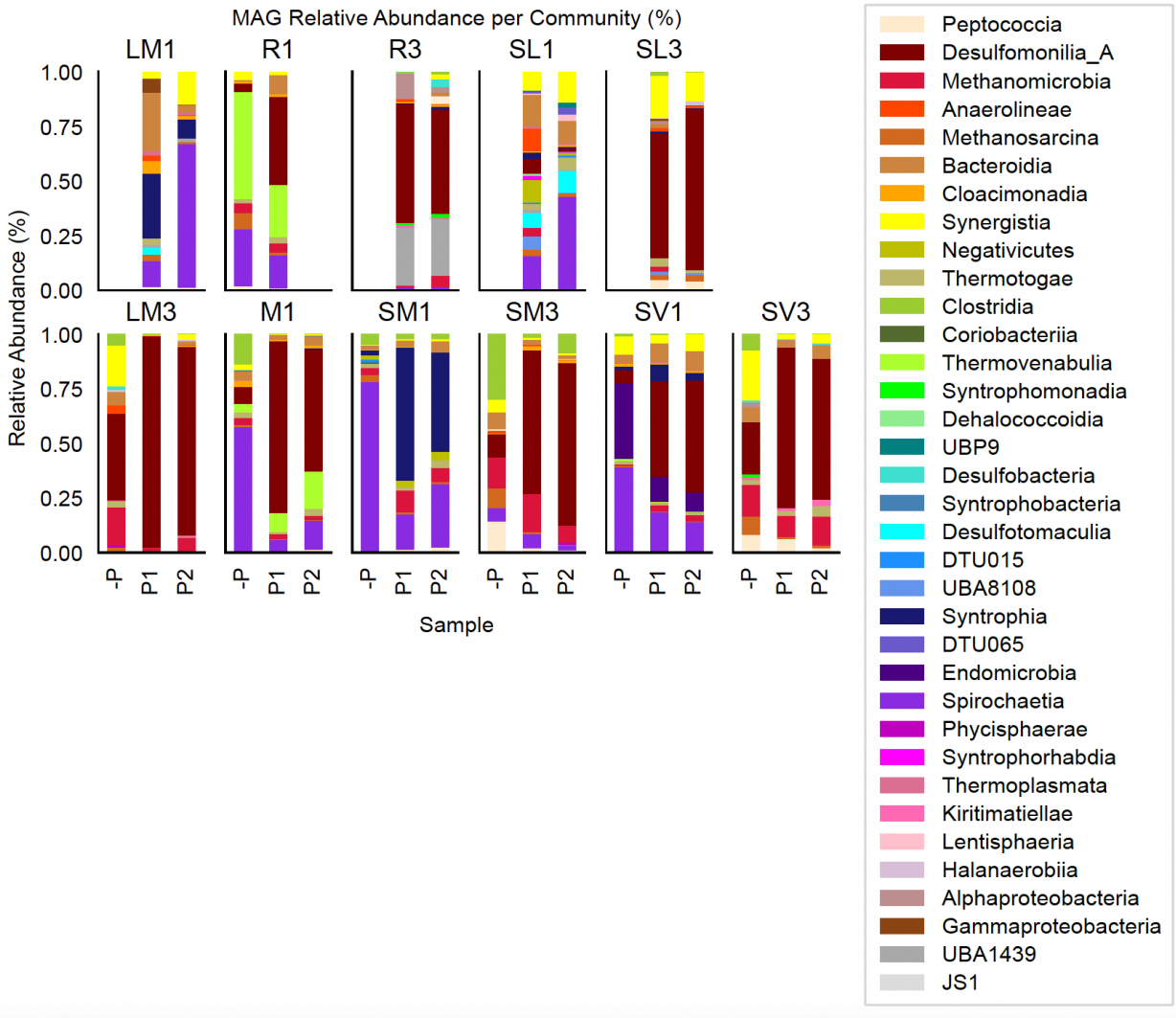


Figure 5.2: Relative abundance of DPO MAGs. Adaptation of Figure 2.3, representing the relative abundance of all MAGs across time. Each subplot represents one community, while each stacked bar represents the community composition of one sample. Colors indicate the taxonomic class of each community member, as specified in the legend on the right. Relative abundance was calculated by dividing the mean coverage of a single MAG by the sum of mean coverages for all MAGs in the respective sample.

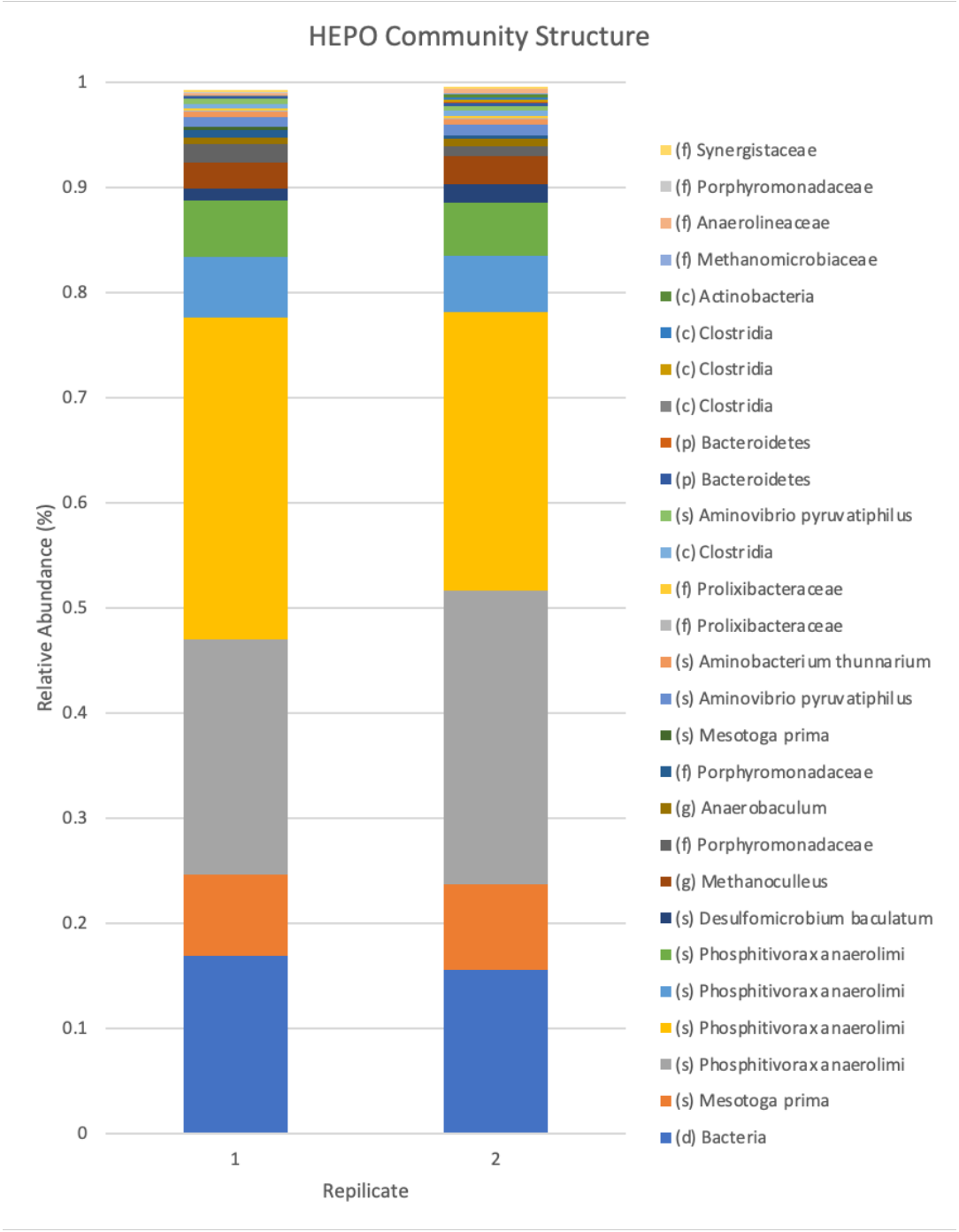


Figure 5.3: Community composition of the HEPO culture. Taxonomic composition of duplicate samples of the HEPO culture, which were sequenced on Sept. 15th, 2021. The HEPO culture has been maintained and transferred in basal bicarbonate buffered medium with 10 mM phosphite since June 26th, 2018. Each OTU is labeled according to the lowest taxonomic rank assigned with greater than 90% sequence identity: d, domain; p, phylum; c, class; f, family; g, genus; s, species.

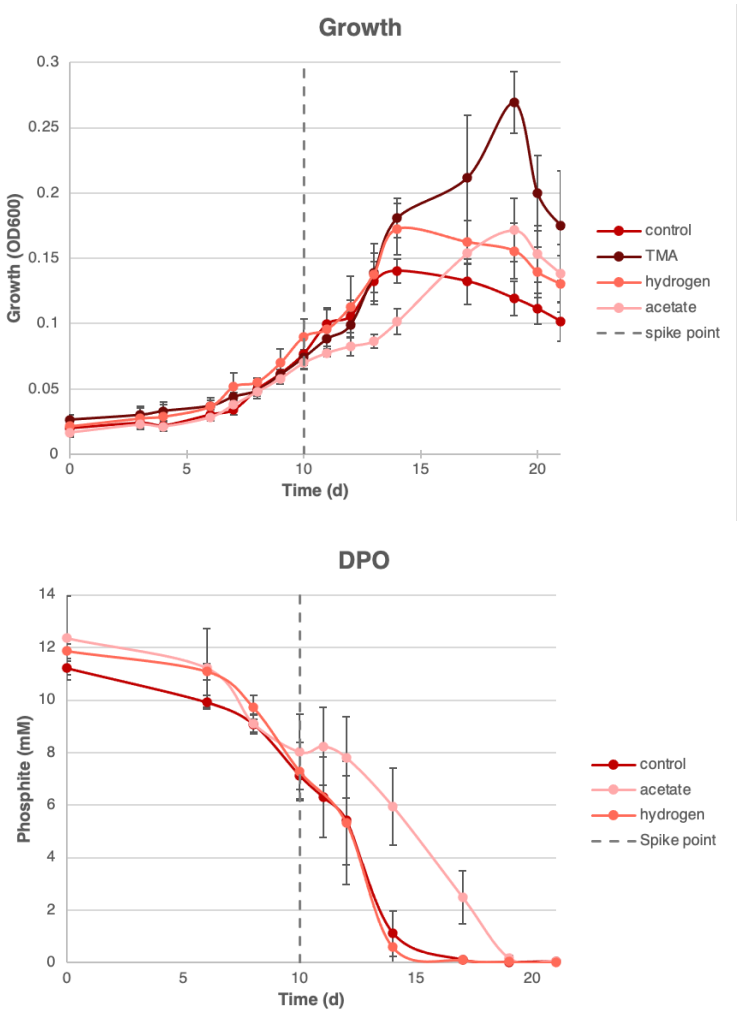


Figure 5.4: Response of the HEPO culture to methanogen substrates. Basal bicarbonate media with 10 mM phosphite was inoculated in triplicate with the HEPO culture and monitored for growth (top) and DPO activity (bottom). On day 10 (demarcated by the grey dashed line), cultures were spiked with methanogenic growth substrates: TMA, trimethylamine (25 mM), acetate (10 mM), and hydrogen (5%). DPO activity was not monitored in the TMA-spiked samples for technical reasons.

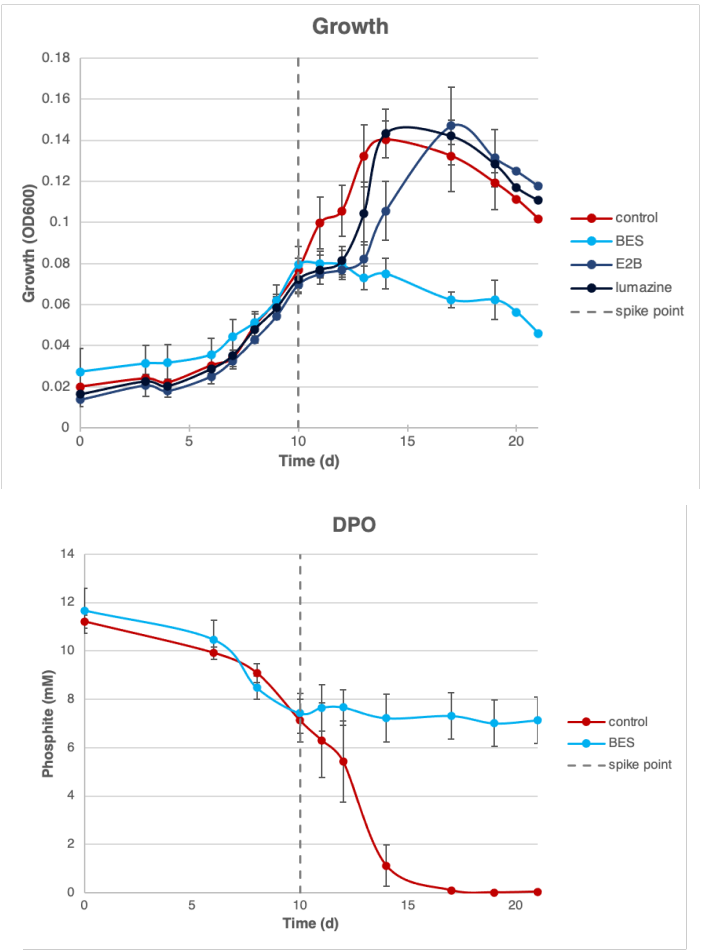


Figure 5.5: Response of the HEPO culture to methanogen inhibitors. Basal bicarbonate media with 10 mM phosphite was inoculated in triplicate with the HEPO culture and monitored for growth (top) and DPO activity (bottom). On day 10 (demarcated by the grey dashed line), cultures were spiked with methanogenic inhibitors: BES, 2-bromoethanesulfonate (10 mM), E2B, ethyl-2-butynoate (10 mM), and lumazine (0.6 mM). DPO activity was not monitored in the E2B and lumazine-spiked samples for technical reasons.

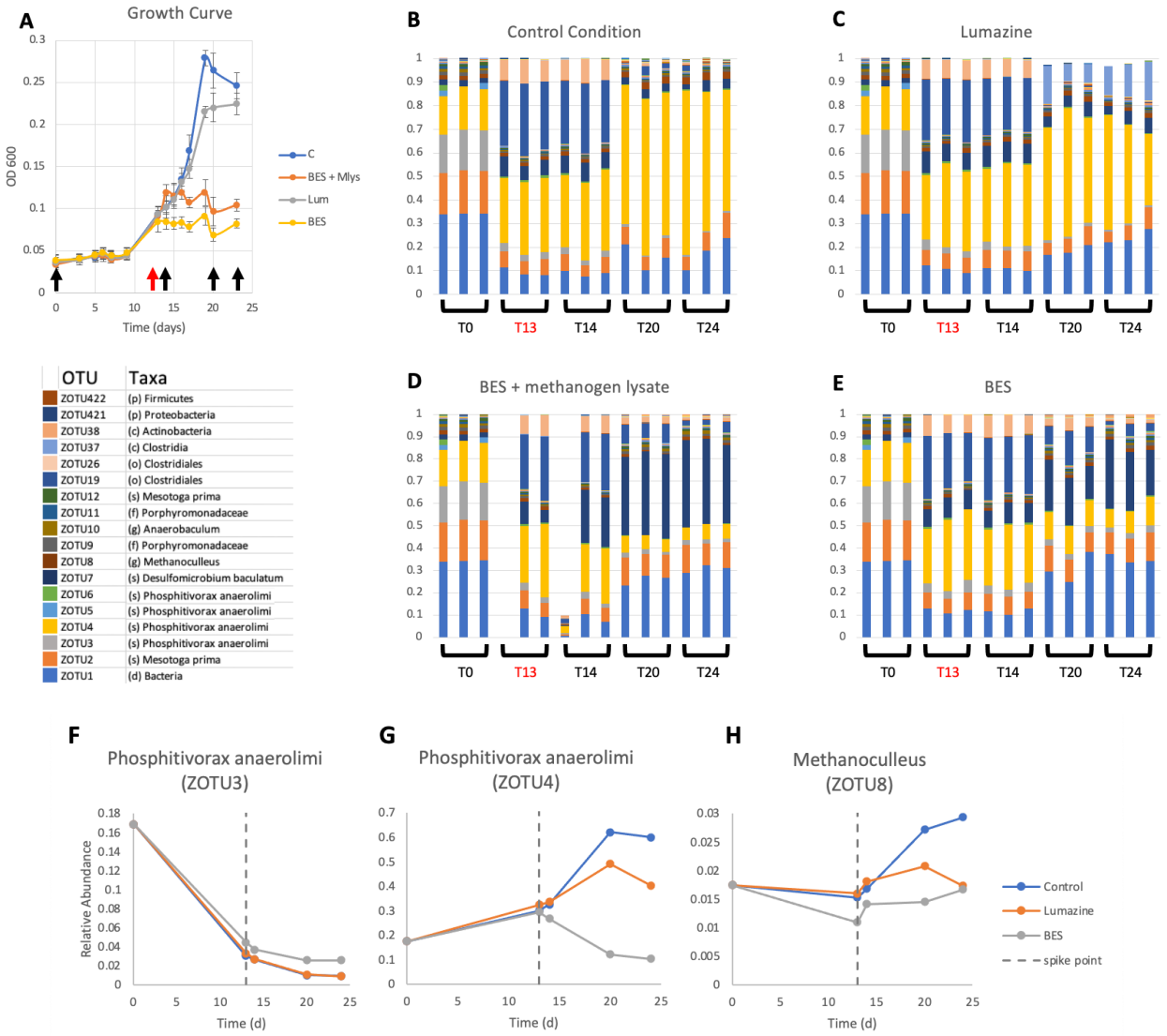


Figure 5.6: Temporal community analysis of the HEPO culture. Basal bicarbonate media with 10 mM phosphite was inoculated in triplicate with the HEPO culture and monitored for (A) growth and (B-E) community composition. On day 13, cultures were spiked with methanogenic inhibitors and supplements: BES, 2-bromoethanesulfonate (10 mM), Lum, lumazine (0.6 mM), Mlys, *M. acetivorans* lysate. The taxonomic composition of triplicate samples across the growth curve are shown for (B) the control condition, (C) cultures exposed to lumazine, (D) cultures exposed to BES and methanogen lysate, and (E) cultures exposed to just BES. The key in the middle-left indicates the colors assigned to each OTU that represents >1% relative abundance in any given community. Each OTU is labeled according to the lowest taxonomic rank assigned with greater than 90% sequence identity: d, domain; p, phylum; c, class; o, order; f, family; g, genus; s, species. (F-H) Subplots demonstrating the average relative abundance over time for (F) *Phosphitivorax anaerolimi* (ZOTU3), (G) *Phosphitivorax anaerolimi* (ZOTU4), and (H) *Methanoculleus* (ZOTU8) in the control condition (blue), lumazine condition (orange), and BES condition (grey).

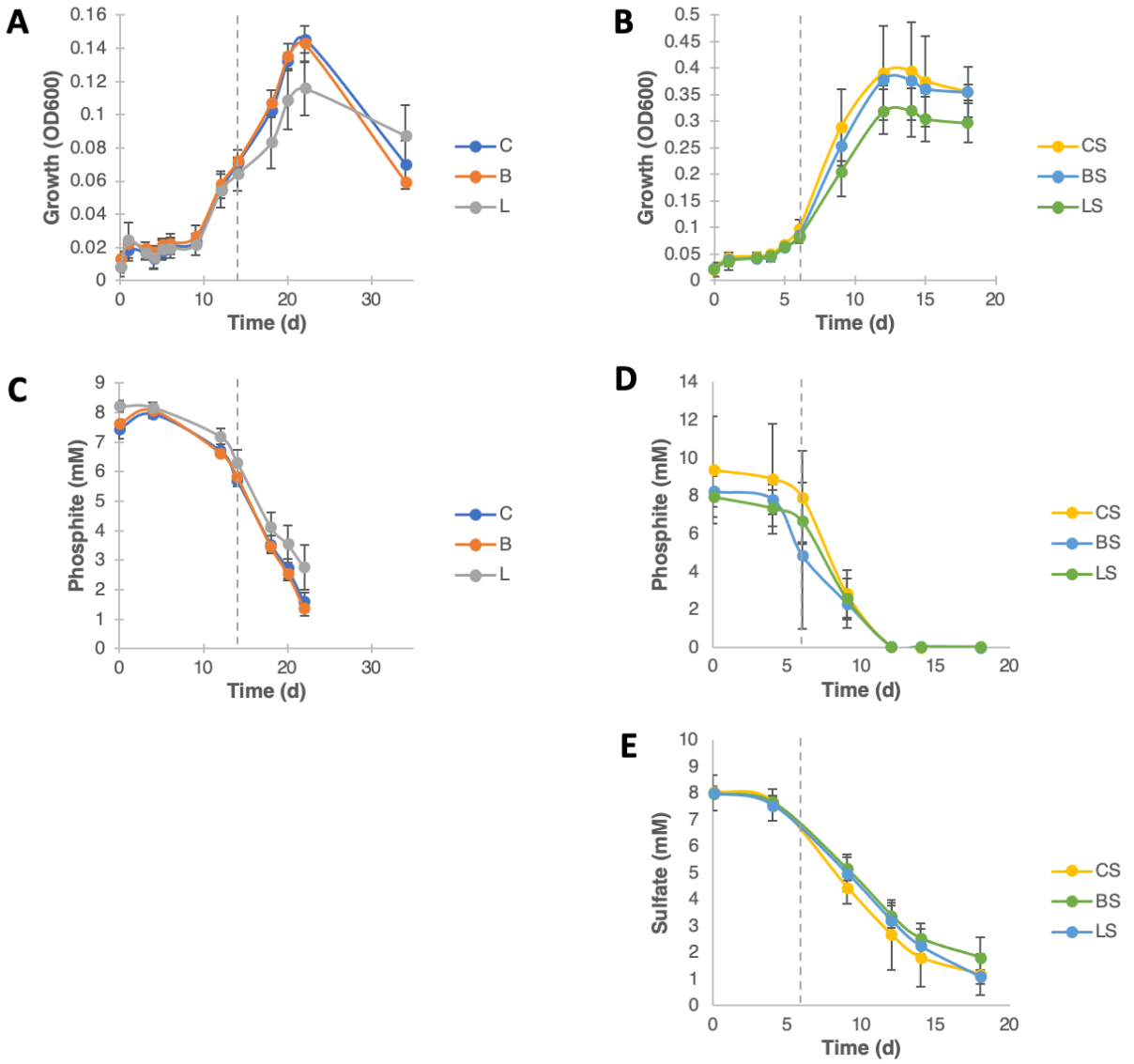


Figure 5.7: FiPS-3 response to methanogen inhibitors. *Desulfotignum phosphitoxidans* FiPS-3 was inoculated into triplicate tubes of basal bicarbonate media with either: (A & C) 10 mM phosphite only, or (B, D & E) 10 mM phosphite and 10 mM sulfate. Once cultures began to exhibit growth, cultures were spiked (grey dashed line) with either 10 mM BES (legend: B or BS) or 0.6 mM lumazine (legend: L or LS). Control cultures (legend: C or CS) were not spiked with inhibitor. Cultures were monitored for (A & B) growth; (C & D) phosphite oxidation; and E) sulfate reduction.

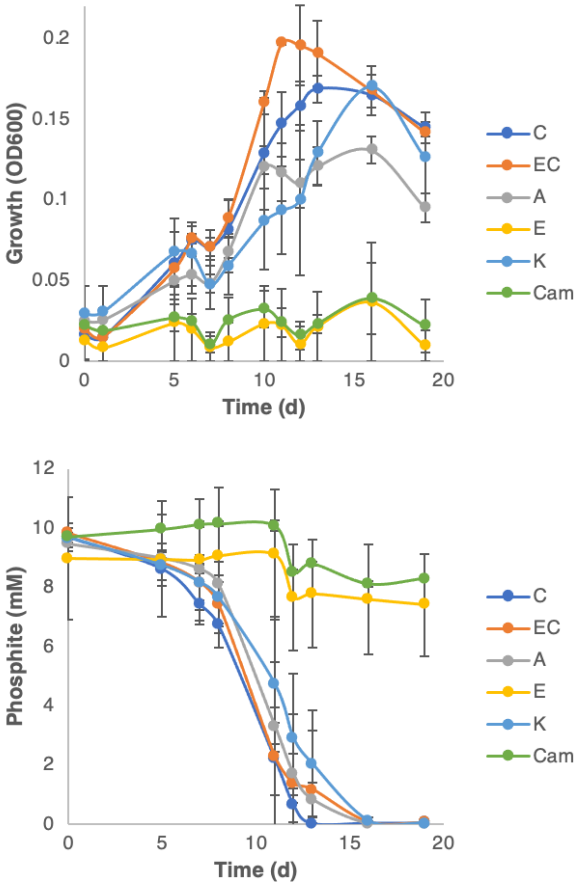


Figure 5.8: Antibiotic profile of the HEPO culture. Basal bicarbonate media with 10 mM phosphite was inoculated in triplicate with the HEPO culture and monitored for growth (top) and DPO activity (bottom). Each condition represents the presence of an alternative antibiotic: C, control (no antibiotic); EC, ethanol control (0.05 mL of 190 proof ethanol); A, ampicillin (100 μg/mL); E, erythromycin (50 μg/mL); K, kanamycin (25 μg/mL); Cam, chloramphenicol (50 μg/mL).

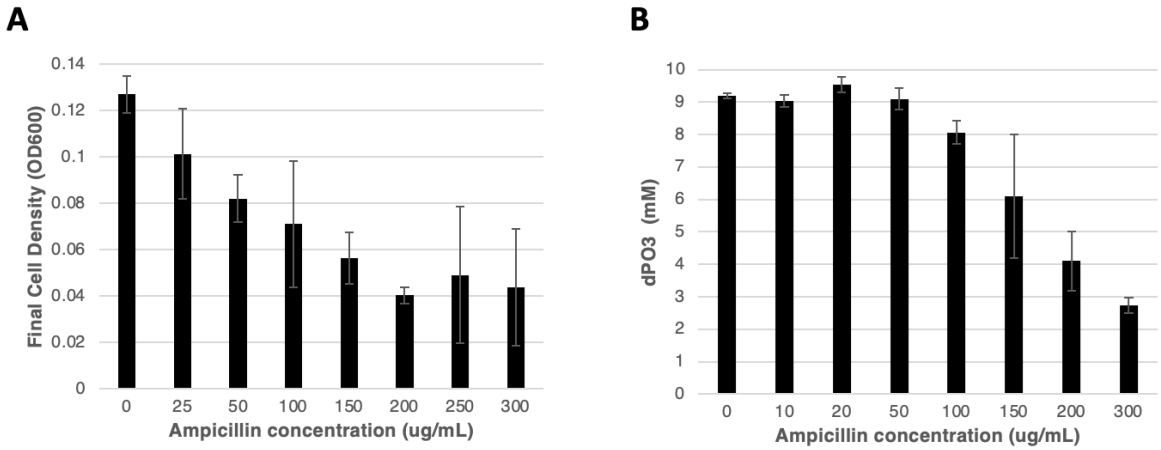


Figure 5.9: Ampicillin inhibition curve of the HEPO culture. The HEPO culture was monitored for (A) net growth and (B) net phosphite consumption following inoculation into triplicate tubes of basal bicarbonate media with 10 mM phosphite and varying concentrations of ampicillin.

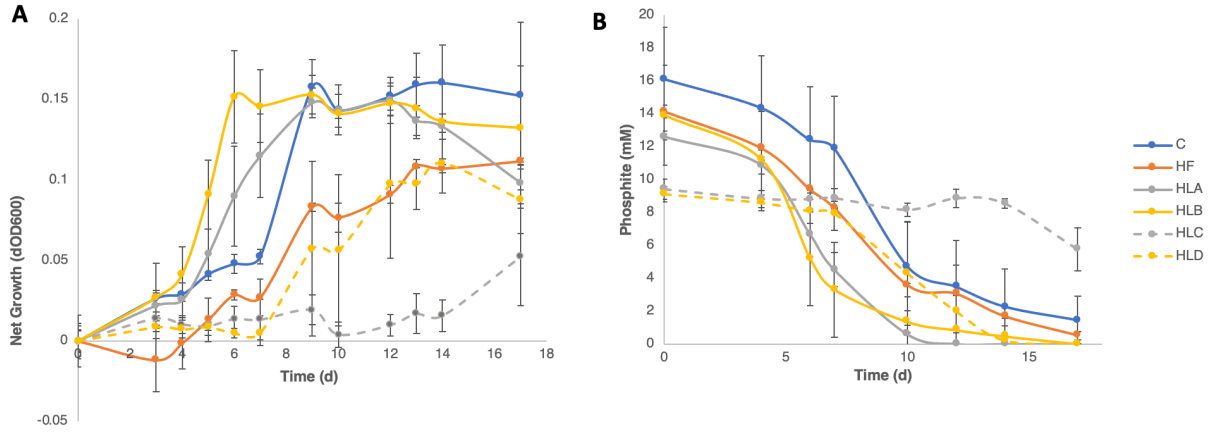


Figure 5.10: HEPO culture response to HEPO filtrates and lysates. The HEPO culture was monitored for (A) net growth and (B) phosphite consumption following inoculation into triplicate tubes of basal bicarbonate media with 10 mM phosphite and alternative supplements: C, control, no additions; HF, HEPO culture filtrate (0.5 mL); HLA, HEPO cell lysate (0.1 mL, heat pasteurized); HLB, HEPO cell lysate (0.1 mL, no heat treatment). The following conditions were not inoculated with the HEPO culture, as sterile controls for the cell lysates: HLC, HEPO cell lysate (0.1 mL, heat pasteurized); HLD, HEPO cell lysate (0.1 mL, no heat treatment).

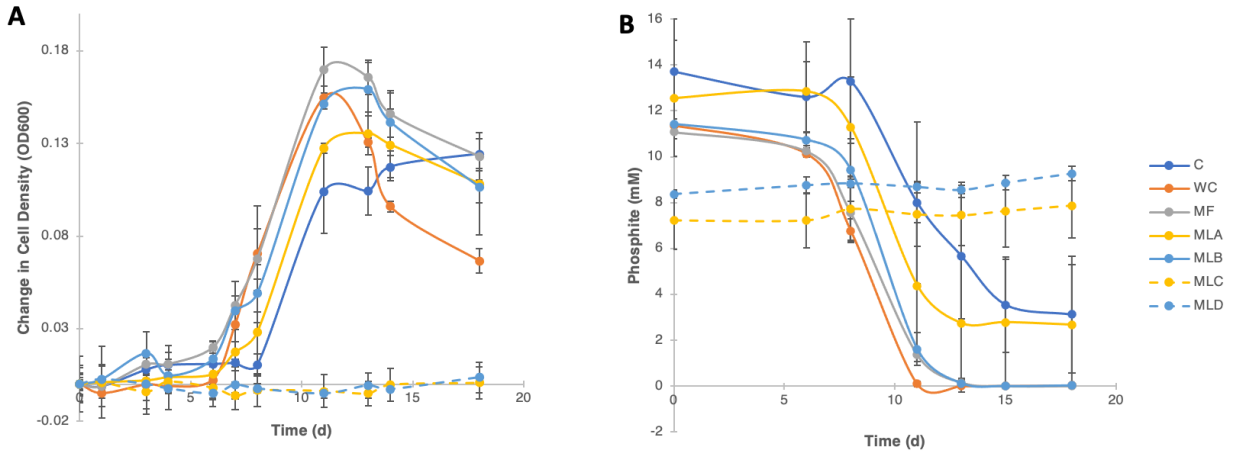


Figure 5.11: HEPO culture response to *M. acetivorans* filtrates and lysates. The HEPO culture was monitored for (A) net growth and (B) phosphite consumption following inoculation into triplicate tubes of basal bicarbonate media with 10 mM phosphite and alternative supplements: C, control, no additions; WC, whole cells (0.5 mL); MF, *M. acetivorans* culture filtrate (0.5 mL); MLA, *M. acetivorans* cell lysate (0.1 mL, heat pasteurized); MLB, *M. acetivorans* cell lysate (0.1 mL, no heat treatment). The following conditions were not inoculated with the HEPO culture, as controls for the cell lysates: MLC, *M. acetivorans* cell lysate (0.1 mL, heat pasteurized); MLD, *M. acetivorans* cell lysate (0.1 mL, no heat treatment).

Figure 5.12: Inclusion matrix for corrinoid-dependent genes. DPOM genomes were subjected to metabolic analysis via DRAM, which uses HMM models to identify genes and reconstructs pathways based on KEGG orthology[180]. Within this presence-absence matrix, each cell denotes the presence (red) or absence (grey) of genes predicted to utilize a corrinoid cofactor based on the computational predictions presented in Shelton et al.[197]

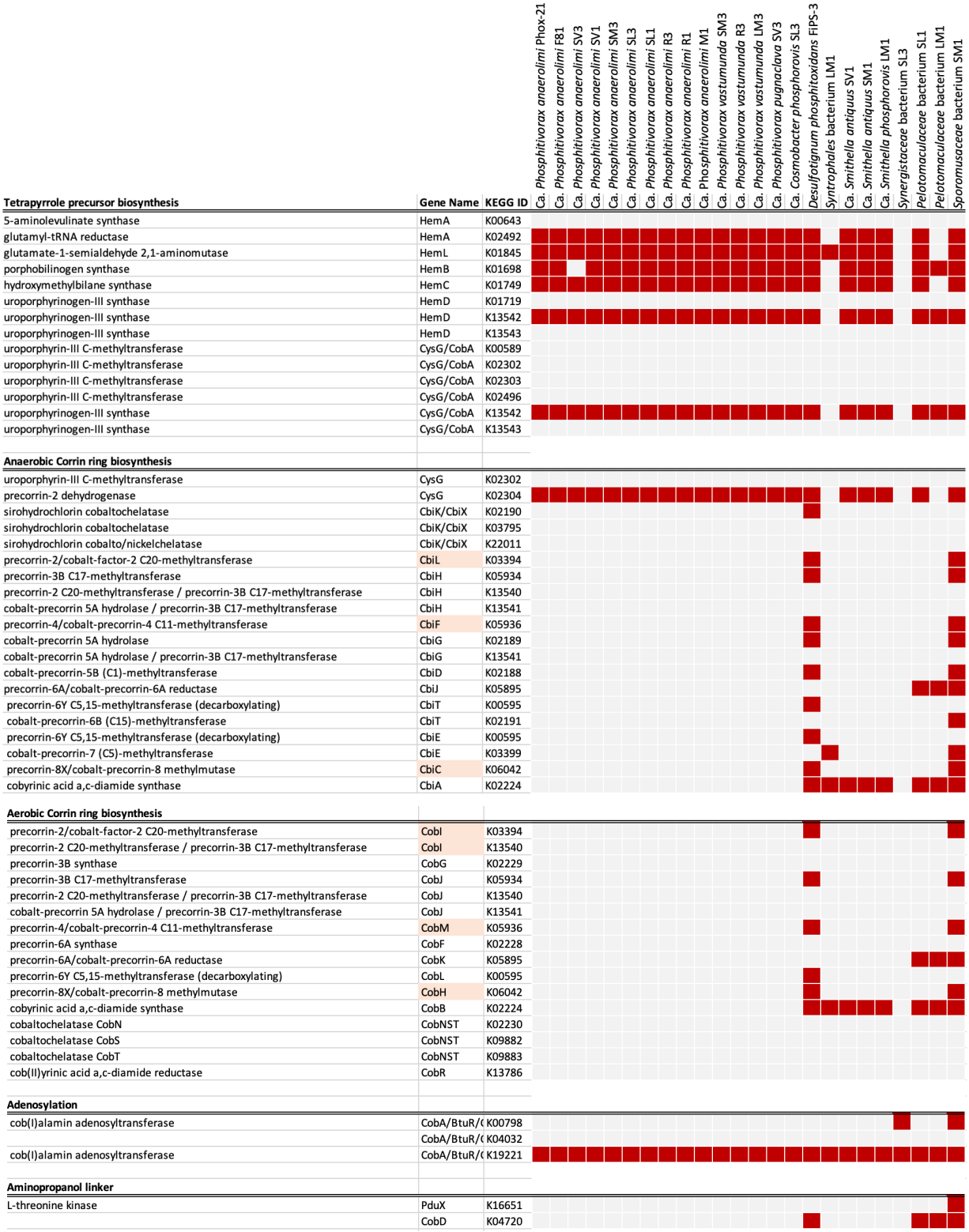


Figure 5.13: *continued on next page.*

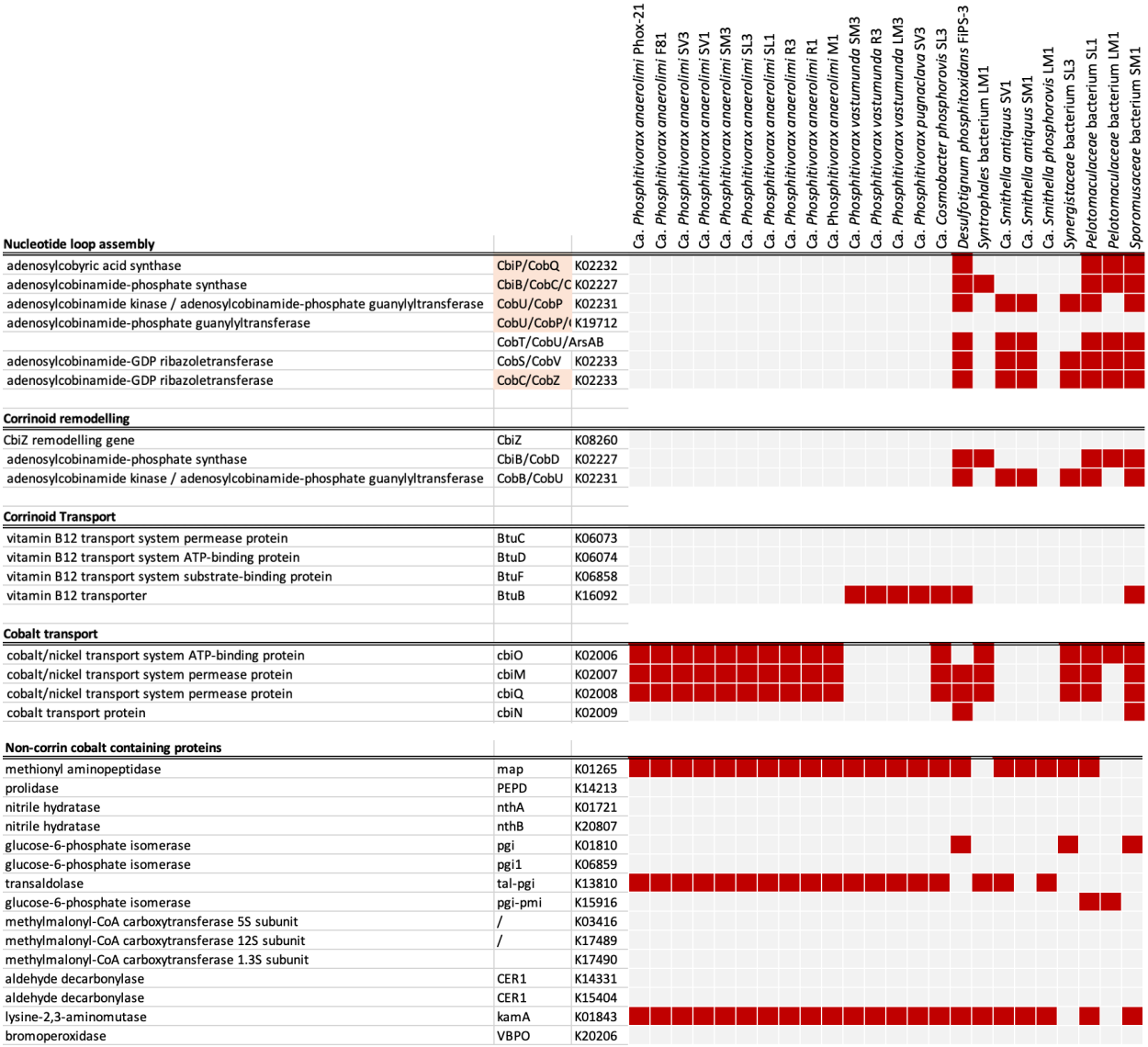


Figure 5.13: Inclusion matrix for corrinoid-biosynthesis genes. DPO genomes were subjected to metabolic analysis via DRAM, which uses HMM models to identify genes and reconstructs pathways based on KEGG orthology[180]. Within this presence-absence matrix, each cell denotes the presence (red) or absence (grey) of genes required to synthesize, modify, or acquire components of a complete corrinoid, based on the computational predictions presented in Shelton et al.[197]. Gene names that are highlighted in orange represent the core genes identified by Shelton et al. to be present in all cobamide-producing microorganisms.

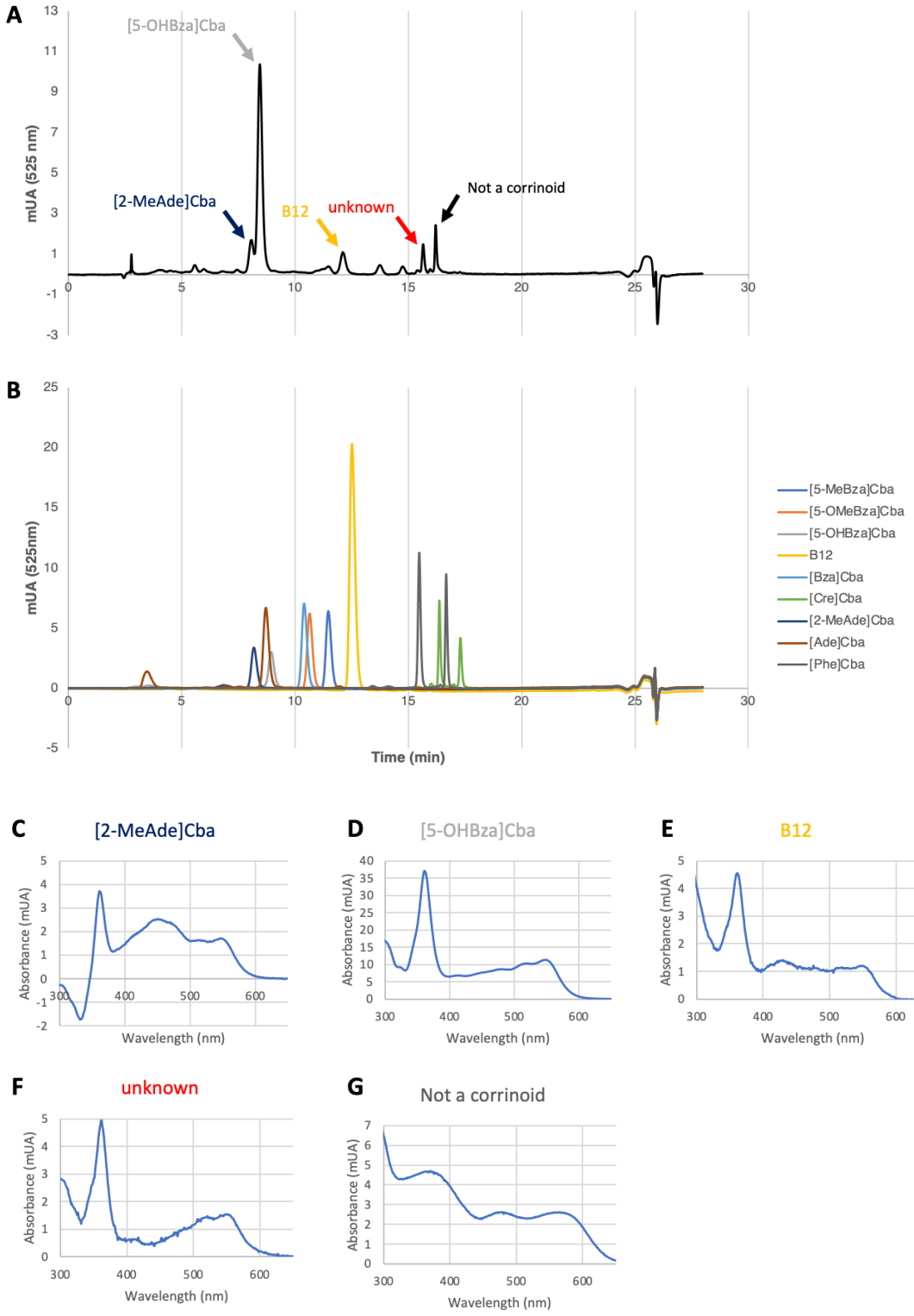


Figure 5.14: Corrinoid profile of the HEPO culture. Corrinoids were extracted from a 0.3 g cell pellet of the HEPO culture. (A) The corrinoid extract from the HEPO culture was fractioned on an HPLC. Peaks were identified (arrows) based on their (C – G) UV/vis spectra, and (B) the retention time of corrinoid standards. (C-F) The UV/vis spectra of each HPLC peak that showed the

signature absorption pattern of a corrinoid. (G) The UV/vis spectra of a peak that did not show the signature absorption pattern of a corrinoid.

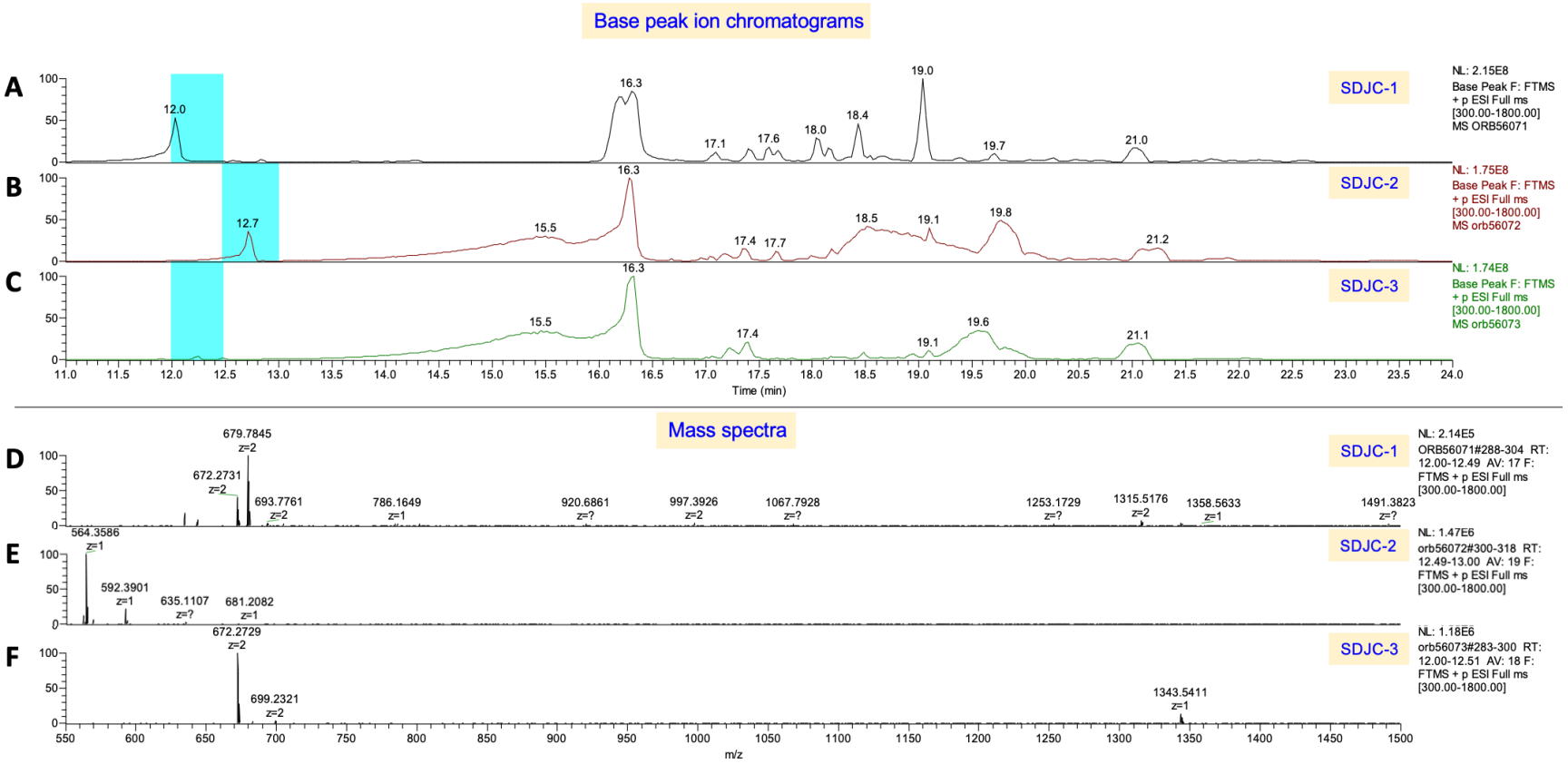

Base peak ion chromatograms
A
B
C
SDJC-1
SDJC-2
SDJC-3
NL: 2.15E8
Base Peak F: FTMS
+ p ESI Full ms
[300.00-1800.00]
MS ORB56071
NL: 1.75E8
Base Peak F: FTMS
+ p ESI Full ms
[300.00-1800.00]
MS orb56072
NL: 1.74E8
Base Peak F: FTMS
+ p ESI Full ms
[300.00-1800.00]
MS orb56073
Time (min)
Mass spectra
D
E
F
NL: 2.14E5
ORB56071#288-304 RT:
12.00-12.49 AV: 17 F:
FTMS + p ESI Full ms
[300.00-1800.00]
NL: 1.47E6
orb56072#300-318 RT:
12.49-13.00 AV: 19 F:
FTMS + p ESI Full ms
[300.00-1800.00]
NL: 1.18E6
orb56073#283-300 RT:
12.00-12.51 AV: 18 F:
FTMS + p ESI Full ms
[300.00-1800.00]
m/z

Figure 5.15: Mass spectra of corrinoid fractions. Three corrinoid fractions (SDJC-1, SDJC-2, and SDJC-3) were submitted for LC-MS analysis. (A-C) Each sample was fractioned on an LC, and time-intervals (highlighted in blue) were injected to the mass-spectrometer. (D-F) Mass spectra of the interval (highlighted in blue) for each sample. Sample names correspond to the following HPLC peak-retention times and predicted corrinoid: SDJC-1, 8.064 min, [2-MeAde]Cba; SDJC-2, 15.65 min, unknown corrinoid; SDJC-3, 8.442 min, [5-OHBza]Cba.

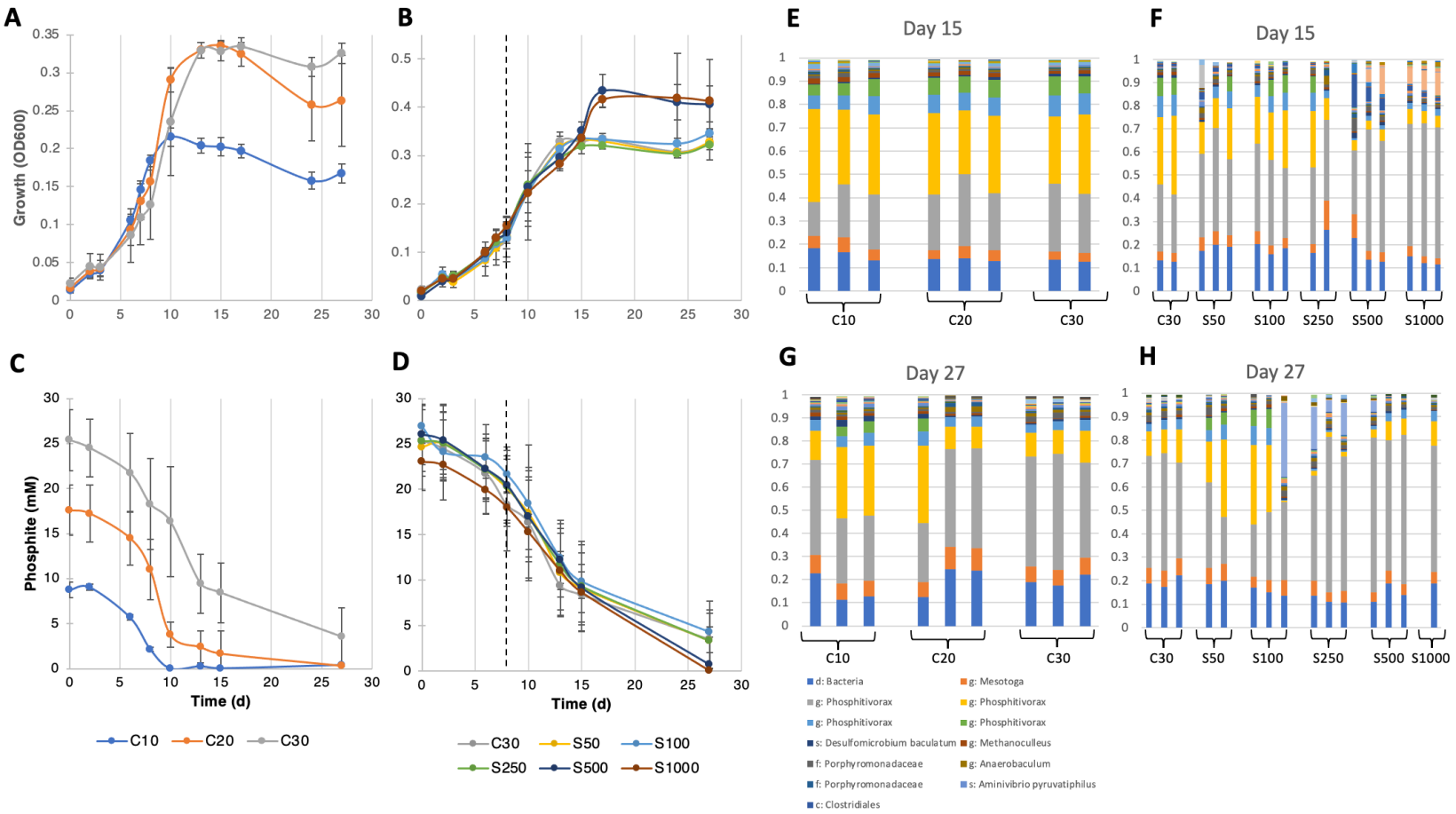

A
Growth (OD600)
B
C
Phosphite (mM)
Time (d)
C10
C20
C30
D
Time (d)
C30
S50
S100
S250
S500
S1000
E
Day 15
C10
C20
C30
F
Day 15
C30
S50
S100
S250
S500
S1000
G
Day 27
C10
C20
C30
H
Day 27
C30
S50
S100
S250
S500
S1000
d: Bacteria
g: Mesotoga
g: Phosphitivorax
g: Phosphitivorax
g: Phosphitivorax
g: Phosphitivorax
s: Desulfomicrobium baculatum
g: Methanoculleus
f: Porphyromonadaceae
g: Anaerobaculum
f: Porphyromonadaceae
s: Aminivibrio pyruvatiphilus
c: Clostridiales

Figure 5.16: SAM-stimulation of the HEPO culture. The HEPO culture was monitored for (A & B) net growth, and (C & D) phosphite consumption following inoculation into triplicate tubes of basal bicarbonate media with (A & C) variable concentrations of phosphite, and (B & D) variable concentrations of SAM. The tested phosphite concentrations were: 10 mM (C10), 20 mM (C20), and 30 mM (C30). The tested SAM concentrations were 50 μM (S50), 100 μM (S100), 250 μM (S250), 500 μM (S500), and 1000 μM (S1000). SAM was spiked into cultures on day 8 (black dashed line). (E-H) DNA samples for each condition were collected on days 15 and 27 and submitted for 16S rRNA gene amplicon sequencing. The key below plot G indicates the colors associated for each OTU that represents > 1% relative abundance in any given community. Each OTU is labeled according to the lowest taxonomic rank assigned with greater than 90% sequence identity: d, domain; p, phylum; c, class; o, order; f, family; g, genus; s, species.

Chapter 6

Conclusions and Future Directions

Parts of this chapter have been drawn from the following publication:

Ewens, S. D., Gomberg, A. F. S., Barnum, T. P., Borton, M. A., Carlson, H. K., Wrighton, K. C., Coates, J. D. (2021). The diversity and evolution of microbial dissimilatory phosphite oxidation. PNAS.

Despite diverse environmental signals that suggest a contemporary P redox cycle (Table 3.6), most biogeochemical models of global P ignore redox transformations[28]. Likewise, the term "phosphorus" in biology is typically synonymous with "orthophosphate", despite sufficient evidence that demonstrates microbial redox cycling of P[29,30]. Dissimilatory phosphite oxidation (DPO) is the only known microbial metabolism that utilizes reduced inorganic P as an electron donor. For DPO microbes (DPOM), phosphite serves as the energy source and represents the most energetically favorable chemotrophic electron donor known, with a half-cell redox potential ($E^{o\prime}$) of −650 mV for the HPO_4^{2-}/HPO_3^{2-} couple (Table 1.1). Yet despite these remarkable energy yields, DPO has been considered an esoteric metabolism unique to few microbial representatives. Since the discovery of the first DPOM in 2000, only two taxonomically disparate species (*Desulfotignum phosphitoxidans* FiPS-3 and *Phosphitivorax anaerolimi* Phox-21) had been identified prior to the writing of this dissertation. The third DPOM (*Phosphitispora fastidiosa*) was isolated during the writing of this dissertation. As such, a false notion of rarity has limited our understanding of DPO prevalence and DPOM diversity. This dissertation aims to challenge that notion and presents cultivation-based investigations, coupled to high-resolution metagenomics, which collectively describe the prevalence, diversity, evolution, physiology, and community dynamics of DPOM.

Prevalence, Diversity, and Evolution of DPOM

The work presented in this dissertation expanded the known diversity of cultured DPOM ten-fold (from 2 to 23 genomes), with an additional DPOM (*P. fastidiosa*) identified during the writing of this dissertation. The DPOM from this study spanned six classes of three bacterial phyla (*Desulfobacterota, Firmicutes* and *Synergistota*). Phylogenetic and syntenic analyses of the *ptx-ptd* gene cluster reveal that DPO metabolism is inherited vertically, and that DPO genes coevolved as a metabolic unit specialized for DPO metabolism. Evolutionary mapping further suggests that DPOM evolved approximately ~3.2 Gya. This prediction is just ~0.8 Gy after the evolution of methanogenesis and consistent with geological predictions of phosphite prevalence on early Earth. Most ancient phosphite reserves are likely to have been oxidized to phosphate following the oxygenation of Earth's atmosphere (~2.5 Gya), yet pockets of phosphite and phosphite oxidizing metabolisms still occur in diverse environments on extant Earth[30,36,43,148]. Environmental metadata from global metagenomes suggest DPOM occupy multiple anoxic environments that represent relics of ancient Earth (i.e., oil reservoirs, deep subsurface aquifers) and serve as potential examples of contemporary phosphite accumulation (i.e., wastewater sludge, freshwater wetlands). Several environments evidently continue to support phosphorus redox cycling.

Future studies should work to validate the metagenomic predictions of DPOM diversity that are presented in this dissertation and begin mapping the influence of DPOM on the environments that they inhabit. All but one DPOM have been derived from wastewater digester sludge; FiPS-3, which was isolated from Venetian canal sediment, represents the only exception. A large-scale sampling campaign could probe diverse environments for DPOM via enrichments and metagenomic surveys using methods from **Chapter 2** of this dissertation as a technical guide. Metadata from the IMG/M computational survey in **Chapter 3** could be used to select those environments (i.e., wetlands, subsurface aquifers, river sediments) that are most likely to contain DPOM (Table 3.6), and the pHMMs presented herein will be a useful tool for identifying DPOM from metagenomic data[117]. DPOM identified from diverse environments could be taxonomically classified, thereby confirming, or expanding the diversity established in this dissertation. Such a sampling campaign could incorporate geochemical analyses to survey the environmental conditions that support DPO activity and determine if specific environmental variables select for distinct DPOM taxa. In situ phosphite measurements could establish the relationship between phosphite distribution and DPO, a relationship that has not yet been established. Extensions of these studies could measure phosphorus distribution and flux, offering a description of those environments that are supporting active P redox cycling. Finally, measurements of environmental P redox dynamics will necessitate explanatory mechanisms for the neogenesis of reduced P species, which continues to be an enigma despite concerted efforts to understand the sources of reduced P.

DPOM Physiology

This dissertation presents the metabolic profile for all known DPOM genomes (except for *P. fastidiosa*). These analyses confirmed prior hypotheses that DPO activity in all DPOM is attributed to the *ptx-ptd* gene cluster, which encodes up to seven proteins (PtxDE-PtdCFGHI). The only gene that appears to be dispensable across DPOM is the *ptdG,* which is predicted to be a transcriptional regulator and is only present in FiPS-3 and the *Syntrophia* class of DPOM. Notably, phosphite oxidation coupled to CO_2 reduction appears to be the primary metabolic niche occupied by DPOM. Although DPO coupled to any known inorganic electron acceptor (oxygen, manganese, perchlorate, nitrate, iron, sulfate etc.) is thermodynamically favorable, DPOM genomes encode sparse electron transport machinery and are largely devoid of the enzymes required to reduce these compounds. These observations support the hypothesis that DPOM conserve energy from phosphite oxidation via substrate level phosphorylation. Furthermore, CO_2 was the only exogenous electron acceptor provided to sequenced enrichments, and physiological experiments demonstrate a CO_2-dependency for DPO-dependent growth. Yet most DPOM also lack canonical carbon reduction or fixation pathways. The reductive glycine pathway was present in many DPOM and may support CO_2 fixation, but the method by which CO_2 is fixed by those DPOM without the reductive glycine pathway is unknown. The cultivation of a highly enriched phosphite oxidizing (HEPO) culture containing *Phosphitivorax* DPOM allowed for DPO-dependent growth in the absence of complex organic supplements, leading to the identification of an organic compound (Compound X) that accumulates in the media temporally with phosphite oxidation. While the identity of Compound X remains unknown, I predict that DPOM couple phosphite oxidation to CO_2 reduction to generate Compound X, which is then metabolized by the community to produce a complex mixture of fermentation products. Preliminary analyses have shown that Compound X is not acetate nor methane, which are the only known products of CO_2 reduction for energy

conservation (as opposed to CO_2 fixation for biomass accumulation). If Compound X is the end-product of dissimilatory CO_2 reduction in *Phosphitivorax* DPOM, it will represent a unique end-product for the reduction of CO_2 as an electron acceptor, indicating a novel pathway for energy conservation via CO_2 reduction. By coupling DPO to primary production via an uncharacterized CO_2 reduction pathway, *Phosphitivorax* DPOM likely play a unique ecological role in any environment they inhabit.

A plethora of work remains to be done to characterize the physiology of DPO metabolism and energy conservation in DPOM. Despite clear evidence that energy from DPO supports the growth of DPOM, the mechanism by which the *ptx-ptd* enzymatic machinery conserves energy from phosphite remains unknown. It is currently hypothesized that phosphite enters DPOM cells via the PtdC phosphite:phosphate antiporter and is subsequently oxidized to phosphate by PtxD, which concomitantly reduces NAD^+ to NADH. The role of the remaining *ptdFGHI* gene products remains unclear but is hypothesized to play an essential role in energy conservation via substrate level phosphorylation. Studies elucidating the biochemical mechanism for phosphite transport, phosphite oxidation, electron transport, and energy conservation by the *ptx-ptd* protein machinery are of high-priority. Such studies might include heterologous expression of the complete *ptx-ptd* gene cluster, or the development of a genetic system allowing for *in situ* gene manipulation. *In vitro* protein assays, crystallization studies, and metabolite analyses will be important supplements to map electron transport. Furthermore, genomic analyses and physiological observations suggest that DPOM preferentially utilize CO_2 as an electron acceptor, but a biochemical link between phosphite oxidation and CO_2 reduction has not been confirmed. FiPS-3 would serve as an excellent model system to test the essentiality of CO_2 reduction machinery, since it is the only cultured DPOM capable of coupling DPO to alternative electron acceptors (i.e., sulfate, nitrate). Finally, exhilarating follow-up studies in *Phosphitivorax* enrichments will be required to confirm the predicted utilization of the novel reductive glycine pathway for autotrophic carbon fixation, and to identify the potentially novel carbon reduction pathway that is utilized for energy conservation. Such studies must first elucidate the identity of Compound X and confirm if and how it is generated from CO_2 using ^{13}C labelling studies. The isolation of *Phosphitivorax,* or further simplification of its host community, would greatly streamline these analyses, though pulse-chase experiments could aid community-based carbon tracking.

Community Dynamics of the HEPO culture

Community analyses suggest that DPOM preferentially occupy methanogenic communities containing four distinct microbial phyla: *Desulfobacterota, Bacteroidota, Synergistota,* and *Halobacterota*. The HEPO community structure reflects this trend, demonstrating that this community structure is maintained even in the absence of complex organics (i.e., rumen fluid). This posits DPOM, such as the *Phosphitivorax* of the HEPO culture, to be primary producers that can support a complex microbial community via fixed carbon and soluble phosphate. Meanwhile, the highly specialized metabolic repertoire of DPOM is analogous to that of syntrophs, corroborating the observation that DPOM frequently belong to known syntrophic taxa[137]. Thermodynamically, phosphite is too energetically favorable an electron donor to require a syntrophic partner, but a symbiotic nutritional dependency could explain their resistance to isolation, especially in the case of *Phosphitivorax* DPOM. In **Chapter 5** of this dissertation, the HEPO community was studied with the aim of i) identifying symbiotic partners of *Phosphitivorax* DPOM in the HEPO culture, and ii) identifying nutrients that *Phosphitivorax* DPOM acquire from

the HEPO community. Three years of selective culturing have enriched *Phosphitivorax* species to >60% relative abundance in the HEPO culture. Yet despite these long-term selective conditions, two non-DPOM still represent over 20% of the HEPO community. These two organisms are *Mesotoga prima* and an unclassified bacterium, and their high relative abundance suggests they may be metabolically important in supporting DPO activity. Inhibition studies also demonstrate that DPO activity is inhibited by BES, a CoM-analogue that is broadly recognized as a specific inhibitor of methanogens. While this evidence might suggest that methanogens are crucial to DPO activity, the community dynamics in the presence of BES instead suggest that DPOM are directly inhibited by BES, implying a potential role for CoM in DPO-dependent growth. Corrinoids were also investigated as a potential exchange intermediate, and extracts from the HEPO culture contain three corrinoids in addition to the B12 supplied in the growth media. These three corrinoids are likely synthesized by the HEPO community, and the identity of one corrinoid remains unknown, possibly representing a novel cofactor. While it has not yet been physiologically shown that corrinoids are required for DPO-activity, genomic evidence suggests that DPOM require corrinoids. Specifically, homology predictions suggest that PtdH (a product of the *ptx-ptd* gene cluster) is corrinoid-dependent, and that corrinoids are likely required for the synthesis of SAM, another cofactor that stimulates DPO-dependent growth of the HEPO culture and is also predicted to be required for PtdH functionality.

Describing the relationship between DPOM and their host communities will be essential to understanding the role of DPOM in the environment. Future studies must confirm the central role of DPOM as primary producers by tracking $^{13}CO_2$ and measuring its fixation to an unknown product and the subsequent catabolism of that product by the microbial community. Bioaugmentation studies may also be informative by adding DPOM to non-DPO communities and monitoring how DPOM change the trajectory of the community structure and influence the metabolite profile. Of particular interest is the relationship between DPO-activity and methanogenesis. The inverse of these relationships is also of interest: studies characterizing how non-DPOM microbes support DPOM growth and activity. Deconstruction of the HEPO community would help identify which organisms are necessary for the growth of *Phosphitivorax* DPOM, and DRAM metabolic analyses of the non-DPOM community would be informative in predicting the metabolic potential of these community members. Metabolic profiling would be particularly informative in describing the role of *Mesotoga prima* and the unclassified bacteria of the HEPO culture, whose relative abundances suggest a significant role. This dissertation also presents CoM, SAM, and corrinoids as potential metabolites that are generated by the community and used by *Phosphitivorax* DPOM. Mechanistic studies will be necessary to confirm and characterize the role of these cofactors in DPO metabolism as well as their microbial source. Purification and identification of the unknown corrinoid is of high priority, as this may represent a novel microbial cofactor. Additionally, studies using labelled corrinoids will be necessary to monitor their fate and identify which, if any, of the corrinoids are utilized by DPOM.

References

1. Ehrlich, H. L., Newman, D. K. & Kappler, A. *Ehrlich's Geomicrobiology*. (2016).
2. Winogradsky, S. Ueber Schwefelbakterien. *Bot. Zeitung* 489–610 (1887).
3. Winogradsky, S. Ueber Eisenbakterien. *Bot. Zeitung* **4**, 261–270 (1888).
4. Winogradsky, S. Recherches sur les organismes de la nitrification. *Ann. Inst. Pasteur (Paris).* **4**–**5**, 213–231 (1890).
5. Dworkin, M. Sergei Winogradsky: a founder of modern microbiology and the first microbial ecologist. *FEMS Microbiol. Rev.* **36**, 364–379 (2012).
6. Lovley, D. R. Dissimilatory Metal Reduction. *Annu. Rev. Microbiol.* **47**, 263–290 (1993).
7. Liermann, L. J., Hausrath, E. M., Anbar, A. D. & Brantley, S. L. Assimilatory and dissimilatory processes of microorganisms affecting metals in the environment. *J. Anal. At. Spectrom.* **22**, 867–877 (2007).
8. Hayes, J. M. & Waldbauer, J. R. The carbon cycle and associated redox processes through time. *Philos. Trans. R. Soc. B Biol. Sci.* **361**, 931–950 (2006).
9. Houghton, R. A. *The Contemporary Carbon Cycle*. *Treatise on Geochemistry: Second Edition* vol. 10 (Elsevier Ltd., 2013).
10. Bar-Even, A., Noor, E. & Milo, R. A survey of carbon fixation pathways through a quantitative lens. *J. Exp. Bot.* **63**, 2325–2342 (2012).
11. Myhre, G. *et al.* Anthropogenic and natural radiative forcing. *Clim. Chang. 2013 Phys. Sci. Basis Work. Gr. I Contrib. to Fifth Assess. Rep. Intergov. Panel Clim. Chang.* 659–740 (2013) doi:10.1017/CBO9781107415324.018.
12. Getabalew, M., Alemneh, T. & Bzuneh, E. Review on Methanogenesis and its Role. *World J. Agric. Soil Sci.* **6**, 1–7 (2020).
13. Stein, L. Y. & Klotz, M. G. The nitrogen cycle. *Curr. Biol.* **26**, R94–R98 (2016).
14. Klotz, M. G., Bryant, D. A. & Hanson, T. E. The microbial sulfur cycle. *Front. Microbiol.* **2**, 3–4 (2011).
15. Jørgensen, B. B., Findlay, A. J. & Pellerin, A. The biogeochemical sulfur cycle of marine sediments. *Front. Microbiol.* **10**, 1–27 (2019).
16. Wang, R. *et al.* Impacts of human activities on the composition and abundance of sulfate-reducing and sulfur-oxidizing microorganisms in polluted river sediments. *Front. Microbiol.* **10**, 1–13 (2019).
17. Yamazaki, C. *et al.* A novel dimethylsulfoxide reductase family of molybdenum enzyme, Idr, is involved in iodate respiration by Pseudomonas sp. SCT. *Environ. Microbiol.* **22**, 2196–2212 (2020).
18. Schink, B., Thiemann, V., Laue, H. & Friedrich, M. W. Desulfotignum phosphitoxidans sp . nov ., a new marine sulfate reducer that oxidizes phosphite to phosphate. *Arch. Microbiol.* **177**, 381–391 (2002).

19. Stolz, J. F., Basu, P. & Oremland, R. S. Microbial transformation of elements: the case of arsenic and selenium. *Int. Microbiol.* **5**, 201–207 (2002).

20. Lloyd, J. R. Microbial reduction of metals and radionuclides. *FEMS Microbiol. Rev.* **27**, 411–425 (2003).

21. Wall, J. D. & Krumholz, L. R. Uranium reduction. *Annu. Rev. Microbiol.* **60**, 149–166 (2006).

22. Ashley, K., Cordell, D. & Mavinic, D. A brief history of phosphorus: From the philosopher's stone to nutrient recovery and reuse. *Chemosphere* **84**, 737–746 (2011).

23. Gusmão, R., Sofer, Z. & Pumera, M. Black Phosphorus Rediscovered: From Bulk Material to Monolayers. *Angew. Chemie - Int. Ed.* **56**, 8052–8072 (2017).

24. Reyhani, R. The Legality of the use of White Phosphorus by the United States Military during the 2004 Fallujah Assaults. *J. Law Soc. Chang.* **10**, (2007).

25. Moore, C. M. *et al.* Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* **6**, 701–710 (2013).

26. Vitousek, P. M., Porder, S., Houlton, B. Z. & Chadwick, O. A. Terrestrial phosphorus limitation: Mechanisms, implications, and nitrogen-phosphorus interactions. *Ecol. Appl.* **20**, 5–15 (2010).

27. Jupp, A. R., Beijer, S., Narain, G. C., Schipper, W. & Slootweg, J. C. Phosphorus recovery and recycling-closing the loop. *Chem. Soc. Rev.* **50**, 87–101 (2021).

28. Ruttenberg, K. C. Phosphorus Cycle. in *Encyclopedia of Ocean Sciences* 447–460 (Elsevier, 2019). doi:10.1016/B978-0-12-409548-9.10807-3.

29. Figueroa, I. A. & Coates, J. D. Microbial Phosphite Oxidation and Its Potential Role in the Global Phosphorus and Carbon Cycles. *Adv. Appl. Microbiol.* **98**, 93–117 (2017).

30. White, A. K. & Metcalf, W. W. Microbial Metabolism of Reduced Phosphorus Compounds. *Annu. Rev. Microbiol.* **61**, 379–400 (2007).

31. Rhodes, C. J. *Peak phosphorus - Peak food? The need to close the phosphorus cycle*. *Science Progress* vol. 96 (2013).

32. Pasek, M. A role for phosphorus redox in emerging and modern biochemistry. *Curr. Opin. Chem. Biol.* **49**, 53–58 (2019).

33. Roels, J. & Verstraete, W. Biological formation of volatile phosphorus compounds. *Bioresour. Technol.* **79**, 243–250 (2001).

34. Pasek, M. A. Rethinking early Earth phosphorus geochemistry. **105**, 853–858 (2008).

35. Pasek, M. A., Sampson, J. M. & Atlas, Z. Redox chemistry in the phosphorus biogeochemical cycle. *Proc. Natl. Acad. Sci.* **111**, 15468–15473 (2014).

36. Pasek, M. & Block, K. Lightning-induced reduction of phosphorus oxidation state. *Nat. Geosci.* **2**, 553–556 (2009).

37. Pasek, M. A., Harnmeijer, J. P., Buick, R., Gull, M. & Atlas, Z. Evidence for reactive reduced phosphorus species in the early Archean ocean. *Proc. Natl. Acad. Sci.* **110**, 10089–10094 (2013).

38. Gulick, A. Phosphorus as a factor in the origin of life. *Am. Sci.* **43**, 479–489 (1955).

39. Schopf, J. W., Kitajima, K., Spicuzza, M. J., Kudryavtsev, A. B. & Valley, J. W. SIMS analyses of the oldest known assemblage of microfossils document their taxon-correlated carbon isotope compositions. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 53–58 (2018).

40. Pasek, M. A., Dworkin, J. P. & Lauretta, D. S. A radical pathway for organic phosphorylation during schreibersite corrosion with implications for the origin of life. *Geochim. Cosmochim. Acta* **71**, 1721–1736 (2007).

41. Han, C. *et al.* Phosphite in sedimentary interstitial water of Lake Taihu, a large eutrophic shallow lake in China. *Environ. Sci. Technol.* **47**, 5679–5685 (2013).

42. Pech, H. *et al.* Detection of geothermal phosphite using high-performance liquid chromatography. *Environ. Sci. Technol.* **43**, 7671–7675 (2009).

43. Yu, X., Geng, J., Ren, H., Chao, H. & Qiu, H. Determination of phosphite in a full-scale municipal wastewater treatment plant. *Environ. Sci. Process. Impacts* **17**, 441–447 (2015).

44. Rott, E. *et al.* Removal of phosphonates from synthetic and industrial wastewater with reusable magnetic adsorbent particles. *Water Res.* **145**, 608–617 (2018).

45. Nagaosa, Y. & Aoyama, E. Catalytic oxidation of phosphite and hypophosphite to phosphate on Pd/activated carbon powder. *Carbon N. Y.* **39**, 2077–2088 (2001).

46. Gao, X., Locke, S., Zhang, J., Joshi, J. & Wang-Pruski, G. Metabolomics Profile of Potato Tubers after Phosphite Treatment. *Am. J. Plant Sci.* **09**, 845–864 (2018).

47. Morton, S. C., Glindemann, D., Wang, X., Niu, X. & Edwards, M. Analysis of reduced phosphorus in samples of environmental interest. *Environ. Sci. Technol.* **39**, 4369–4376 (2005).

48. Britvin, S. N., Murashko, M. N., Vapnik, Y., Polekhovsky, Y. S. & Krivovichev, S. V. Earth's phosphides in Levant and insights into the source of Archean prebiotic phosphorus. *Sci. Rep.* **5**, 10–14 (2015).

49. Herschy, B. *et al.* Archaean phosphorus liberation induced by iron redox geochemistry. *Nat. Commun.* 1–7 (2018) doi:10.1038/s41467-018-03835-3.

50. Zhu, R. *et al.* Tropospheric phosphine and its sources in coastal Antarctica. *Environ. Sci. Technol.* **40**, 7656–7661 (2006).

51. Fu, W. & Zhang, X. Global phosphorus dynamics in terms of phosphine. *Clim. Atmos. Sci.* **3**, 62–65 (2020).

52. Bains, W., Petkowski, J. J., Sousa-Silva, C. & Seager, S. New environmental model for thermodynamic ecology of biological phosphine production. *Sci. Total Environ.* **658**, 521–536 (2019).

53. Ohtake, H. *et al.* Bacterial phosphonate degradation, phosphite oxidation and polyphosphate accumulation. *Resour. Conserv. Recycl.* **18**, 125–134 (1996).

54. Devai, I. & Delaune, R. D. Evidence for phosphine production and emission from Louisiana and Florida marsh soils. *Org. Geochem.* **23**, 277–279 (1995).

55. Cordell, D. & White, S. Peak Phosphorus: Clarifying the Key Issues of a Vigorous Debate about Long-Term Phosphorus Security. *Sustainability* **3**, 2027–2049 (2011).

56. Alori, E. T., Glick, B. R. & Babalola, O. O. Microbial phosphorus solubilization and its potential for use in sustainable agriculture. *Front. Microbiol.* **8**, 1–8 (2017).

57. Repeta, D. J. *et al.* Marine methane paradox explained by bacterial degradation of dissolved organic matter. *Nat. Geosci.* **9**, 884–887 (2016).

58. Sosa, O. A. Phosphorus redox reactions as pinch hitters in microbial metabolism. *Proc. Natl. Acad. Sci.* 201719600 (2017) doi:10.1073/pnas.1719600115.

59. Clark, L. L., Ingall, E. D. & Benner, R. Marine organic phosphorus cycling: novel insights from nuclear magnetic resonance. *Am. J. Sci.* **29999**, 724–737 (1999).

60. Kolowith, L. C., Ingall, E. D. & Benner, R. Composition and cycling of marine organic phosphorus. *Limnol. Oceanogr.* **46**, 309–320 (2001).

61. Karl, D. M. Microbially Mediated Transformations of Phosphorus in the Sea: New Views of an Old Cycle. *Ann. Rev. Mar. Sci.* **6**, 279–337 (2014).

62. Yang, K. & Metcalf, W. W. A new activity for an old enzyme: Escherichia coli bacterial alkaline phosphatase is a phosphite-dependent hydrogenase. *Proc. Natl. Acad. Sci.* **101**, 7919–7924 (2004).

63. White, A. K. & Metcalf, W. W. Two C-P lyase operons in Pseudomonas stutzeri and their roles in the oxidation of phosphonates, phosphite, and hypophosphite. *J. Bacteriol.* **186**, 4730–4739 (2004).

64. Costas, A. M. G., White, A. K. & Metcalf, W. W. Purification and Characterization of a Novel Phosphorus-oxidizing Enzyme from *Pseudomonas stutzeri* WM88. *J. Biol. Chem.* **276**, 17429–17436 (2001).

65. Stone, B. L. & White, A. K. Most probable number quantification of hypophosphite and phosphite oxidizing bacteria in natural aquatic and terrestrial environments. *Arch. Microbiol.* **194**, 223–228 (2012).

66. Schink, B. & Friedrich, M. Bacterial metabolism: Phosphite oxidation by sulphate reduction. *Nature* **406**, 37 (2000).

67. Poehlein, A., Daniel, R., Schink, B. & Simeonova, D. D. Life based on phosphite: a genome-guided analysis of Desulfotignum phosphitoxidans. *BMC Genomics* **14**, 753 (2013).

68. Simeonova, D. D., Susnea, I., Moise, A., Schink, B. & Przybylski, M. "Unknown Genome" Proteomics: A new NAD(P)-dependent epimerase/dehydratase revealed by N-terminal sequencing, inverted PCR, and high resolution mass spectrometry. *Mol. Cell. Proteomics* 122–131 (2009) doi:10.1074/mcp.M800242-MCP200.

69. Simeonova, D. D., Wilson, M. M., Metcalf, W. W. & Schink, B. Identification and heterologous expression of genes involved in anaerobic dissimilatory phosphite oxidation by Desulfotignum phosphitoxidans. *J. Bacteriol.* **192**, 5237–5244 (2010).

70. Figueroa, I. A. *et al.* Metagenomics-guided analysis of microbial chemolithoautotrophic phosphite oxidation yields evidence of a seventh natural CO2 fixation pathway. *Proc. Natl. Acad. Sci.* 201715549 (2017) doi:10.1073/pnas.1715549114.

71. Hao, L. *et al.* Novel syntrophic bacteria in full-scale anaerobic digesters revealed by genome-centric metatranscriptomics. *ISME J.* (2020) doi:10.1038/s41396-019-0571-0.

72. Yishai, O., Bouzon, M., Döring, V. & Bar-Even, A. In vivo assimilation of one-carbon via a synthetic reductive glycine pathway in Escherichia coli. *ACS Synth. Biol.* (2018) doi:10.1021/acssynbio.8b00131.

73. Bar-Even, A. Formate Assimilation: The Metabolic Architecture of Natural and Synthetic Pathways. *Biochemistry* (2016) doi:10.1021/acs.biochem.6b00495.

74. Sánchez-Andrea, I. *et al.* The reductive glycine pathway allows autotrophic growth of Desulfovibrio desulfuricans. *Nat. Commun.* 1–12 (2020) doi:10.1038/s41467-020-18906-7.

75. Mao, Z. *et al.* Phosphitispora fastidiosa gen. nov. sp. nov., a new dissimilatory phosphite-oxidizing anaerobic bacterium isolated from anaerobic sewage sludge. (2021) doi:10.1099/ijsem.0.005142.

76. Bard, A. J., Parsons, R. & Jordon, J. *Standard Potentials in Aqueous Solution*. (1985).

77. Thauer, R. K., Jungermann, K. & Decker, K. Energy Conservation in Chemotrophic Anaerobic Bacteria. *Bacteriol. Rev.* **41**, 100–180 (1977).

78. Walker, L. M., Li, B., Niks, D., Hille, R. & Elliott, S. J. Deconvolution of reduction potentials of formate dehydrogenase from Cupriavidus necator. *J. Biol. Inorg. Chem.* **24**, 889–898 (2019).

79. Reyes-Umana, V., Henning, Z., Lee, K., Barnum, T. P. & Coates, J. D. Genetic and phylogenetic analysis of dissimilatory iodate-reducing bacteria identifies potential niches across the world's oceans. *ISME J.* **16**, 38–49 (2022).

80. Barnum, T. P. & Coates, J. D. Chlorine redox chemistry is not rare in biology. *bioRxiv* 2021.12.08.471835 (2021).

81. Holliger, C., Pierik, A. J., Hagen, W. R. & Reijerse, E. J. A Spectroelectrochemical Study of Factor F430 Nickel(II/I) from Methanogenic Bacteria in Aqueous Solution. *J. Am. Chem. Soc.* **115**, 5651–5656 (1993).

82. Banerjee, R. V., Matthews, R. G., Harder, S. R. & Ragsdale, S. W. Mechanism of Reductive Activation of Cobalamin-Dependent Methionine Synthase: An Electron Paramagnetic Resonance Spectroelectrochemical Study. *Biochemistry* **29**, 1129–1135 (1990).

83. Ney, B. *et al.* The methanogenic redox cofactor F420 is widely synthesized by aerobic soil bacteria. *ISME J.* **11**, 125–137 (2017).

84. Garner, D. K. *et al.* Reduction potential tuning of the blue copper center in Pseudomonas aeruginosa azurin by the axial methionine as probed by unnatural amino acids. *J. Am. Chem. Soc.* **128**, 15608–15617 (2006).

85. Peng, L., Dai, H., Wu, Y., Peng, Y. & Lu, X. A comprehensive review of phosphorus recovery from wastewater by crystallization processes. *Chemosphere* **197**, 768–781 (2018).

86. Liang, S. *et al.* One-step treatment of phosphite-laden wastewater: A single electrochemical reactor integrating superoxide radical-induced oxidation and electrocoagulation. *Environ. Sci. Technol.* acs.est.9b00841 (2019) doi:10.1021/acs.est.9b00841.

87. Lloyd, K. G., Steen, A. D., Ladau, J., Yin, J. & Crosby, L. Phylogenetically Novel Uncultured Microbial Cells Dominate Earth Microbiomes. *mSystems* **3**, (2018).

88. Rosenberg, E. *et al. Principles of Enrichment, Isolation, Cultivation, and Preservation of Prokaryotes*. *The Prokaryotes: Prokaryotic Biology and Symbiotic Associations* (2013). doi:10.1007/978-3-642-30194-0.

89. Lewis, W. H., Tahon, G., Geesink, P., Sousa, D. Z. & Ettema, T. J. G. Innovations to culturing the uncultured microbial majority. *Nat. Rev. Microbiol.* **19**, 225–240 (2021).

90. Wu, L. *et al.* Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nat. Microbiol.* **4**, 1183–1195 (2019).

91. Taylor, S. C. *et al.* The Ultimate qPCR Experiment: Producing Publication Quality, Reproducible Data the First Time. *Trends Biotechnol.* **37**, 761–774 (2019).

92. Sherwood, A. R., Paasch, B. C., Worby, C. A. & Gentry, M. S. A malachite green-based assay to assess glucan phosphatase activity. *Anal. Biochem.* **435**, 54–56 (2013).

93. Overmann, O., Abt, B. & Sikorski, J. Present and Future of Culturing Bacteria. *Ann. Rev. Mar. Sci.* **71**, 711–730 (2017).

94. Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R. & Wolfe, R. S. Methanogens : Reevaluation of a Unique Biological Group. **43**, 260–296 (1979).

95. Larsson, A. AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **30**, 3276–3278 (2014).

96. Joshi, N. & JN, F. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). *[Software]* (2011).

97. Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).

98. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

99. Hyatt, D. *et al.* Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, (2010).

100. Wu, D., Jospin, G. & Eisen, J. A. Systematic Identification of Gene Families for Use as 'Markers' for Phylogenetic and Phylogeny-Driven Ecological Studies of Bacteria and Archaea and Their Major Subgroups. *PLoS One* **8**, (2013).

101. Kim, D., Song, L., Breitwieser, F. P. & Salzberg, S. L. Centrifuge: rapid and accurate classification of metagenomic sequences. *bioRxiv* **26**, 054965 (2016).

102. Eren, A. M. *et al.* Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**, (2015).

103. Alneberg, J. *et al.* Binning metagenomic contigs by coverage and composition. *Nat. Methods* **11**, 1144–1146 (2014).

104. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM : assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).

105. Bowers, R. M. *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **35**, 725–731 (2017).

106. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).

107. Kunhikrishnan, A. *et al. The influence of wastewater irrigation on the transformation and bioavailability of heavy metal(loid)s in soil*. *Advances in Agronomy* vol. 115 (2012).

108. Zhao, Y., Zhuang, X., Ahmad, S., Sung, S. & Ni, S. Q. Biotreatment of high-salinity wastewater: current methods and future directions. *World J. Microbiol. Biotechnol.* **36**, 1–11 (2020).

109. Saleem, F. *et al.* The Bovine Ruminal Fluid Metabolome. *Metabolomics* **9**, 360–378 (2013).

110. Milo, R. & Phillips, R. Cell biology by the numbers. *CRC Press* (2015).

111. Hempfling, W. P. & Mainzer, S. E. Effects of Varying the Carbon Source Limiting Growth on Yield and Maintenance Characteristics of Escherichia coli in Continuous Culture. *J. Bacteriol.* **123**, 1076–1087 (1975).

112. Gubernat, S., Masłoń, A., Czarnota, J. & Koszelnik, P. Reactive materials in the removal of phosphorus compounds from wastewater-A review. *Materials (Basel).* **13**, 1–26 (2020).

113. McDowell, M. M. *et al.* Detection of hypophosphite, phosphite, and orthophosphate in natural geothermal water by ion chromatography. *J. Chromatogr. A* **1039**, 105–111 (2004).

114. Pech, H. *et al.* Elucidating the redox cycle of environmental phosphorus using ion chromatography. *J. Chromatogr. Sci.* **49**, 573–581 (2011).

115. Sadeghi, S., Anderson, T. A. & Jackson, W. A. Determination of phosphite (HPO3−2) by a new IC/MS/MS method using an 18O-labeled HPO3−2 internal standard. *Talanta* **230**, 0–5 (2021).

116. Yoon, B. Hidden Markov Models and their Applications in Biological Sequence Analysis. 402–415 (2009).

117. Ewens, S. D. *et al.* The diversity and evolution of microbial dissimilatory phosphite oxidation. *PNAS* **118**, (2021).

118. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 1–3 (2019) doi:10.1093/bioinformatics/btz848.

119. Parks, D. H. *et al.* GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* **202**, 1–10 (2021).

120. Olm, M. R. *et al.* Consistent Metagenome-Derived Metrics Verify and Delineate Bacterial Species Boundaries. *mSystems* **5**, e00731-19 (2020).

121. Edgar, R. C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

122. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 - Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS One* **5**, (2010).

123. Miller, C. S., Baker, B. J., Thomas, B. C., Singer, S. W. & Banfield, J. F. EMIRGE: Reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol.* **12**, (2011).

124. Pruesse, E. *et al.* SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* **35**, 7188–7196 (2007).

125. Johnson, L. S., Eddy, S. R. & Portugaly, E. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* (2010).

126. Li, W., Fu, L., Niu, B., Wu, S. & Wooley, J. Ultrafast clustering algorithms for metagenomic sequence analysis. *Brief. Bioinform.* **13**, 656–668 (2012).

127. Huson, D. H. & Scornavacca, C. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* **61**, 1061–1067 (2012).

128. Veltri, D., Wight, M. M. & Crouch, J. A. SimpleSynteny: a web-based tool for visualization of microsynteny across multiple species. *Nucleic Acids Res.* **44**, W41–W45 (2016).

129. Galushko, A. & Kuever, J. Desulfomonile. in *Bergey's Manual of Systematics of Archaea and Bacteria* (2019).

130. Kim, A. D., Mandelco, L., Tanner, R. S., Woese, C. R. & Suflita, J. M. Desulfomonile tiedjei gen. nov. and sp. nov., a novel anaerobic, dehalogenating, sulfate-reducing bacterium. *Arch. Microbiol.* **154**, 23–30 (1990).

131. Rainey, F. A. Pelotomaculum. in *Bergey's Manual of Systematics of Archaea and Bacteria* 1–6 (2009).

132. Imachi, H. *et al.* Pelotomaculum thermopropionicum gen . nov., sp . nov., an anaerobic, thermophilic, syntrophic propionate-oxidizing bacterium. *Int. J. Syst. Evol. Microbiol.* **52**, 1729–1735 (2002).

133. Liu, Y., Balkwill, D. L., Henry, C. A., Drake, G. R. & Boone, D. R. Characterization of the anaerobic propionate- degrading syntrophs Smithella propionica gen. nov., sp. nov. and Syntrophobacter wolinii. *Int. J. Syst. Bacteriol.* **49**, 545–556 (1999).

134. McInerney, M. J. *et al.* Physiology, ecology, phylogeny, and genomics of microorganisms capable of syntrophic metabolism. *Ann. N. Y. Acad. Sci.* **1125**, 58–72 (2008).

135. Mouttaki, H., Nanny, M. A. & McInerney, M. J. Cyclohexane carboxylate and benzoate formation from crotonate in Syntrophus aciditrophicus. *Appl. Environ. Microbiol.* **73**, 930–938 (2007).

136. Schink, B. Energetics of Syntrophic Cooperation in Methanogenic Degradation. *Microbiol. Mol. Biol. Rev.* **61**, 262–80 (1997).

137. Mcinerney, M. J., Sieber, J. R. & Gunsalus, R. P. Syntrophy in Anaerobic Global Carbon Cycles. *Curr. Opin. Biotechnol.* **20**, 623–632 (2010).

138. Galushko, A. & Kuever, J. Syntrophus. in *Bergey's Manual of Systematics of Archaea and Bacteria* 1–2 (2009).

139. Galushko, A. & Kuever, J. Smithella. *Bergey's Man. Syst. Archaea Bact.* 1–4 (2019) doi:10.1002/9781118960608.gbm01063.pub2.

140. Matelska, D. *et al.* Classification, substrate specificity and structural features of D-2-hydroxyacid dehydrogenases: 2HADH knowledgebase. *BMC Evol. Biol.* **18**, 1–23 (2018).

141. Gordon, B. R. *et al.* Decoupled genomic elements and the evolution of partner quality in nitrogen-fixing rhizobia. *Ecol. Evol.* **6**, 1317–1327 (2016).

142. Junier, I. & Rivoire, O. Synteny in Bacterial Genomes: Inference, Organization and Evolution. (2013).

143. Sevillya, G. & Snir, S. Synteny footprints provide clearer phylogenetic signal than sequence data for prokaryotic classification. *Mol. Phylogenet. Evol.* **136**, 128–137 (2019).

144. Raymann, K., Brochier-Armanet, C. & Gribaldo, S. The two-domain tree of life is linked to a new root for the Archaea. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6670–6675 (2015).

145. Battistuzzi, F. U., Feijao, A. & Hedges, S. B. A genomic timescale of prokaryote evolution: Insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evol. Biol.* **4**, 1–14 (2004).

146. Lyons, T. W., Reinhard, C. T. & Planavsky, N. J. The rise of oxygen in Earth's early ocean and atmosphere. *Nature* **506**, 307–315 (2014).

147. Wolf, Y. I. & Koonin, E. V. Genome reduction as the dominant mode of evolution. *BioEssays* 829–837 (2013) doi:10.1002/bies.201300037.

148. Pirim, C. *et al.* Investigation of schreibersite and intrinsic oxidation products from Sikhote-Alin, Seymchan, and Odessa meteorites and Fe3P and Fe2NiP synthetic surrogates. *Geochim. Cosmochim. Acta* **140**, 259–274 (2014).

149. Liang, S. *et al.* One-Step Treatment of Phosphite-Laden Wastewater: A Single Electrochemical Reactor Integrating Superoxide Radical-Induced Oxidation and Electrocoagulation. *Environ. Sci. Technol.* **53**, 5328–5336 (2019).

150. Jarvie, H. P. *et al.* Biogeochemical and climate drivers of wetland phosphorus and nitrogen release: Implications for nutrient legacies and eutrophication risk. *J. Environ. Qual.* **49**, 1703–1716 (2020).

151. Alpana, S., Vishwakarma, P., Adhya, T. K., Inubushi, K. & Dubey, S. K. Molecular ecological perspective of methanogenic archaeal community in rice agroecosystem. *Science of the Total Environment* vols 596–597 136–146 (2017).

152. Han, S. H., Zhuang, Y. H., Liu, J. A. & Glindemann, D. Phosphorus cycling through phosphine in paddy fields. *Sci. Total Environ.* **258**, 195–203 (2000).

153. Zhang, J., Geng, J. J., Zhang, R., Ren, H. Q. & Wang, X. R. Matrix-bound phosphine in paddy fields under a simulated increase in global atmospheric CO2. *Environ. Chem.* **7**, 287–291 (2010).

154. Kappler, A. & Bryce, C. Cryptic biogeochemical cycles: unravelling hidden redox reactions. *Environ. Microbiol.* **19**, 842–846 (2017).

155. Mendler, K. *et al.* Annotree: Visualization and exploration of a functionally annotated microbial tree of life. *Nucleic Acids Res.* **47**, 4442–4448 (2019).

156. Johannes, T. W., Woodyer, R. D. & Zhao, H. Efficient Regeneration of NADPH Using an Engineered Phosphite Dehydrogenase. **96**, 18–26 (2007).

157. Tóth, N., Tajti, Á., Ladányi-Pára, K., Bálint, E. & Keglevich, G. Synthesis of phosphonates in a continuous flow manner. *Phosphorus, Sulfur Silicon Relat. Elem.* **194**, 285–286 (2019).

158. Han, C. *et al.* In situ sampling and speciation method for measuring dissolved phosphite at ultratrace concentrations in the natural environment. *Water Res.* **137**, 281–289 (2018).

159. Dévai, I., Felföldy, L., Wittner, I. & Plosz, S. Detection of phosphine: new aspects of the phosphorus cycle in the hydrosphere. *Nature* **336**, 343–345 (1988).

160. Armbruster, D., Rott, E., Minke, R. & Happel, O. Trace-level determination of phosphonates in liquid and solid phase of wastewater and environmental samples by IC-ESI-MS/MS. *Anal. Bioanal. Chem.* **412**, 4807–4825 (2020).

161. Glindemann, D., Edwards, M., Liu, J. A. & Kuschk, P. Phosphine in soils, sludges, biogases and atmospheric implications - A review. *Ecol. Eng.* **24**, 457–463 (2005).

162. Gassmann, G. & van Beusekom, J. E. E. Offshore Atmospheric Phosphine. *Naturwissenschaften* 129–131 (1996).

163. Zhang, R. *et al.* Phosphorus Composition in Sediments from Seven Different Trophic Lakes, China: A Phosphorus-31 NMR Study. *J. Environ. Qual.* **38**, 353–359 (2009).

164. Benitez-Nelson, C. R. *et al.* Phosphonates and particulate organic phosphorus cycling in an anoxic marine basin. *Limnol. Ocean.* **49**, 1593–1604 (2004).

165. Eismann, F., Glindemann, D., Bergmann, A. & Kuschk, P. Soils as Source and Sink of Phosphine. *Chemosphere* **35**, 523–533 (1997).

166. Tate, K. R. & Churchman, G. J. Organo-Mineral Fractions of a Climosequence of Soils in New Zealand Tussock Grasslands. *J. Soil Sci.* **29**, 331–339 (1978).

167. Pasek, M. a & Lauretta, D. S. Aqueous corrosion of phosphide minerals from iron meteorites: a highly reactive source of prebiotic phosphorus on the surface of the early Earth. *Astrobiology* **5**, 515–35 (2005).

168. Greaves, J. S. *et al.* Phosphine gas in the cloud decks of Venus. *Nat. Astron.* (2020) doi:10.1038/s41550-020-1174-4.

169. Pasek, M. & Lauretta, D. Extraterrestrial flux of potentially prebiotic C, N, and P to the early earth. *Orig. Life Evol. Biosph.* **38**, 5–21 (2008).

170. Han, C. *et al.* In situ sampling and speciation method for measuring dissolved phosphite at ultratrace concentrations in the natural environment. *Water Res.* **137**, 281–289 (2018).

171. Yao, M., Henny, C. & Maresca, J. A. Freshwater bacteria release methane as a by-product of phosphorus acquisition. *Appl. Environ. Microbiol.* **82**, 6994–7003 (2016).

172. Kolowith, L. C., Ingall, E. D. & Benner, R. Composition and cycling of marine organic phosphorus. *Limnol. Oceanogr.* **46**, 309–320 (2001).

173. Clark, L. L., Ingall, E. D. & Benner, R. Marine phosphorus is selectively remineralized. *Nature* **393**, 426 (1998).

174. McGrath, J. W., Chin, J. P. & Quinn, J. P. Organophosphonates revealed: New insights into the microbial metabolism of ancient molecules. *Nat. Rev. Microbiol.* **11**, 412–419 (2013).

175. Bains, W., Petkowski, J. J., Sousa-Silva, C. & Seager, S. New environmental model for thermodynamic ecology of biological phosphine production. *Sci. Total Environ.* **658**, 521–536 (2019).

176. Ragsdale, S. W. Enzymology of the Woods-Ljundahl Pathway of Acetogenesis. *Ann. N. Y. Acad. Sci.* **1125**, 129–136 (2008).

177. Figueroa, I. A. A physiological and genomic investigation of dissimilatory phosphite oxidation in Desulfotignum phosphitoxidans strain FiPS-3 and in microbial enrichment cultures from wastewater treatment sludge. *Dissertation* (2016).

178. Bar-Even, A., Noor, E., Flamholz, A. & Milo, R. Design and analysis of metabolic pathways supporting formatotrophic growth for electricity-dependent cultivation of microbes. *Biochim. Biophys. Acta - Bioenerg.* **1827**, 1039–1047 (2013).

179. Claassens, N. J. *et al.* Replacing the Calvin cycle with the reductive glycine pathway in Cupriavidus necator. *Metab. Eng.* **62**, 30–41 (2020).

180. Shaffer, M. *et al.* DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res.* 1–58 (2020) doi:https://doi.org/10.1101/2020.06.29.177501.

181. Friedrich, T., Dekovic, D. K. & Burschel, S. Assembly of the Escherichia coli NADH:ubiquinone oxidoreductase (respiratory complex I). *Biochim. Biophys. Acta - Bioenerg.* **1857**, 214–223 (2016).

182. Young, T. *et al.* Crystallographic and kinetic analyses of the FdsBG subcomplex of the cytosolic formate dehydrogenase FdsABG from Cupriavidus necator. *J. Biol. Chem.* **295**, 6570–6585 (2020).

183. Jouanneau, Y., Jeong, H. S., Hugo, N., Meyer, C. & Willison, J. C. Overexpression in Escherichia coli of the rnf genes from Rhodobacter capsulatus - Characterization of two membrane-bound iron-sulfur proteins. *Eur. J. Biochem.* **251**, 54–64 (1998).

184. Ljungdahl, L. G. & Wood, H. G. Total Synthesis of Acetate From CO2 by Heterotrophic Bacteria. *Annu. Rev. Microbiol.* **23**, 515–538 (1969).

185. Poock, S. R., Leach, E. R., Moir, J. W. B., Cole, J. A. & Richardson, D. J. Respiratory detoxification of nitric oxide by the cytochrome c nitrite reductase of Escherichia coli. *J. Biol. Chem.* **277**, 23664–23669 (2002).

186. Zhu, B., Friedrich, S., Wang, Z., Táncsics, A. & Lueders, T. Availability of Nitrite and Nitrate as Electron Acceptors Modulates Anaerobic Toluene-Degrading Communities in Aquifer Sediments. *Front. Microbiol.* **11**, (2020).

187. Watmough, N. J. *et al.* Nitric Oxide in Bacteria. *Biochim. Biophys. Acta* **1411**, 456–474 (1999).

188. Gould, G. W. & Russell, N. J. *Food Preservatives*. *Elsevier Applied Science* (2003).

189. Relyea, H. A., Vrtis, J. M., Woodyer, R., Rimkus, S. A. & Van Der Donk, W. A. Inhibition and pH dependence of phosphite dehydrogenase. *Biochemistry* **44**, 6640–6649 (2005).

190. Carlson, H. K. *et al.* Monofluorophosphate is a selective inhibitor of respiratory sulfate-reducing microorganisms. *Environ. Sci. Technol.* **49**, 3727–3736 (2015).

191. Zane, G. M., Wall, J. D. & De León, K. B. Novel Mode of Molybdate Inhibition of Desulfovibrio vulgaris Hildenborough. *Front. Microbiol.* **11**, (2020).

192. Burkhart, B. J., Schwalen, C. J., Mann, G., Nalsmith, J. H. & Mitchell, D. A. YcaO-Dependent Posttranslational Amide Activation: Biosynthesis, Structure, and Function. *Physiol. Behav.* **117**, 5389–5456 (2017).

193. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).

194. Weber, S. D., Ludwig, W., Schleifer, K. H. & Fried, J. Microbial composition and structure of aerobic granular sewage biofilms. *Appl. Environ. Microbiol.* **73**, 6233–6240 (2007).

195. Faust, K. & Raes, J. Microbial interactions: From networks to models. *Nat. Rev. Microbiol.* **10**, 538–550 (2012).

196. Gupta, D., Shalvarjian, K. E. & Nayak, D. D. Biogenesis and Function of c-type Cytochromes in the Methanogenic Archaeon, Methanosarcina acetivorans. *bioRxiv* **29**, 197 (2022).

197. Shelton, A. N. *et al.* Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. *ISME J.* **13**, 789–804 (2019).

198. Yi, S. *et al.* Versatility in corrinoid salvaging and remodeling pathways supports corrinoid-dependent metabolism in Dehalococcoides mccartyi. *Appl. Environ. Microbiol.* **78**, 7745–7752 (2012).

199. Allen, R. H. & Stabler, S. P. Identification and quantitation of cobalamin and cobalamin analogues in human feces. *Am. J. Clin. Nutr.* **87**, 1324–1335 (2008).

200. Carlson, H. K. *et al.* Mechanisms of direct inhibition of the respiratory sulfate-reduction pathway by (per)chlorate and nitrate. *ISME J.* **9**, 1295–1305 (2015).

201. Hahnke, S., Langer, T., Koeck, D. E. & Klocke, M. Description of Proteiniphilum saccharofermentans sp. nov., Petrimonas mucosa sp. nov. and Fermentimonas caenicola gen. nov., sp. nov., isolated from mesophilic laboratory-scale biogas reactors, and emended description of the genus Proteiniphilum. *Int. J. Syst. Evol. Microbiol.* **66**, 1466–1475 (2016).

202. Baena, S., Gacia, J.-L., Cayol, J.-L. & Ollivier, B. Aminobacterium. *Bergey's Man. Syst. Bacteriol.* **3**, (1999).

203. Hamdi, O. *et al.* Aminobacterium thunnarium sp. Nov., a mesophilic, amino acid-degrading bacterium isolated from an anaerobic sludge digester, pertaining to the phylum Synergistetes. *Int. J. Syst. Evol. Microbiol.* **65**, 609–614 (2015).

204. Vítězová, M., Kohoutová, A., Vítěz, T., Hanišáková, N. & Kushkevych, I. Methanogenic microorganisms in industrial wastewater anaerobic treatment. *Processes* **8**, 1–27 (2020).

205. Wagner, D. Methanosarcina . *Bergey's Man. Syst. Archaea Bact.* **1987**, 1–23 (2020).

206. Kim, S. & Rhee, S. Methanomethylovorans. *Bergey's Man. Syst. Archaea Bact.* 1–5 (2019) doi:10.1002/9781118960608.gbm01602.

207. Akinyemi, T. S., Shao, N. & Whitman, W. B. Methanothrix. *Bergey's Man. Syst. Archaea Bact.* 1–8 (2020) doi:10.1002/9781118960608.gbm00513.

208. Lai, M. Methanoculleus. *Bergey's Man. Syst. Archaea Bact.* 1–8 (2019) doi:10.1002/9781118960608.gbm00505.pub2.

209. Yokota, A. & Ikeda, M. *On Amino Acid Fermentation*. *Advances in Biochemical Engineering/Biotechnology* vol. 8 (2016).

210. Peterson, D. *et al.* Comparative Analysis of 16S rRNA Gene and Metagenome Sequencing in Pediatric Gut Microbiomes. *Front. Microbiol.* **12**, 1–13 (2021).

211. Goyal, A., Bittleston, L. S., Leventhal, G. E., Lu, L. & Cordero, O. X. Interactions between strains govern the eco-evolutionary dynamics of microbial communities. *Elife* **11**, 1–23 (2022).

212. Ferry, J. G. *Methanogenesis: Ecology, Physiology, Biochemistry, & Genetics*. (1993).

213. Boyd, J. M., Clark, D. D., Kofoed, M. A. & Ensign, S. A. Mechanism of inhibition of aliphatic epoxide carboxylation by the coenzyme M analog 2-bromoethanesulfonate. *J. Biol. Chem.* **285**, 25232–25242 (2010).

214. Liu, H., Wang, J., Wang, A. & Chen, J. Chemical inhibitors of methanogenesis and putative applications. *Appl. Microbiol. Biotechnol.* **89**, 1333–1340 (2011).

215. Partovi, S. E. *et al.* Coenzyme M biosynthesis in bacteria involves phosphate elimination by a functionally distinct member of the aspartase/fumarase superfamily. *J. Biol. Chem.* **293**, 5236–5246 (2018).

216. Boyd, J. M., Ellsworth, A. & Ensign, S. A. Characterization of 2-bromoethanesulfonate as a selective inhibitor of the coenzyme M-dependent pathway and enzymes of bacterial aliphatic epoxide metabolism. *J. Bacteriol.* **188**, 8062–8069 (2006).

217. Zhou, Z., Meng, Q. & Yu, Z. Effects of methanogenic inhibitors on methane production and abundances of methanogens and cellulolytic bacteria in in vitro ruminal cultures. *Appl. Environ. Microbiol.* **77**, 2634–2639 (2011).

218. Patra, A., Park, T., Kim, M. & Yu, Z. Rumen methanogens and mitigation of methane emission by anti-methanogenic compounds and substances. *J. Anim. Sci. Biotechnol.* **8**, 1–18 (2017).

219. Nagar-Anthal, K. R., Worrell, V. E., Teal, R. & Nagle, D. P. The pterin lumazine inhibits growth of methanogens and methane formation. *Arch. Microbiol.* **166**, 136–140 (1996).

220. Allen, K. D. & White, R. H. *Identification of the Radical SAM Enzymes Involved in the Biosynthesis of Methanopterin and Coenzyme F420 in Methanogens*. *Methods in Enzymology* vol. 606 (Elsevier Inc., 2018).

221. Santoro, N. & Konisky, J. Characterization of bromoethanesulfonate-resistant mutants of Methanococcus voltae: Evidence of a coenzyme M transport system. *J. Bacteriol.* **169**, 660–665 (1987).

222. Balch, W. E. & Wolfe, R. S. Transport of coenzyme M (2-mercaptoethanesulfonic acid) in Methanobacterium ruminantium. *J. Bacteriol.* **137**, 264–273 (1979).

223. Dybas, M. & Konisky, J. Transport of coenzyme M (2-mercaptoethanesulfonic acid) and methylcoenzyme M [(2-methylthio)ethanesulfonic acid] in Methanococcus voltae: Identification of specific and general uptake systems. *J. Bacteriol.* **171**, 5866–5871 (1989).

224. Basu, P. & Burgmayer, S. J. N. Pterin chemistry and its relationship to the molybdenum cofactor. *Coord. Chem. Rev.* **255**, 1016–1038 (2011).

225. Krishnakumar, A. M. *et al.* Getting a Handle on the Role of Coenzyme M in Alkene Metabolism. **72**, 445–456 (2008).

226. Suzuki, Y. *et al.* Detection of kanamycin-resistant Mycobacterium tuberculosis by identifying mutations in the 16S rRNA gene. *J. Clin. Microbiol.* **36**, 1220–1225 (1998).

227. Gärtner, P., Weiss, D. S., Harms, U. & Thauer, R. K. N5-Methyltetrahydromethanopterin:Coenzyme M Methyltransferase from Methanobacterium thermoautotrophicum Catalytic Mechanism and Sodium Ion Dependence. *Eur. J. Biochem.* **226**, 465–472 (1994).

228. Gonzalez, J. C., Banerjee, R. V., Huang, S., Sumner, J. S. & Metthews, R. G. Cobalamin-dependent and cobalamin-independent methionine synthases in Escherichia coli: Two solutions to the same chemical problem. *Adv. Exp. Med. Biol.* **338**, 687–692 (1992).

229. Men, Y. *et al.* Identification of specific corrinoids reveals corrinoid modification in dechlorinating microbial communities. *Environ. Microbiol.* **17**, 4873–4884 (2015).

230. Sokolovskaya, O. M., Shelton, A. N. & Taga, M. E. Sharing vitamins: Cobamides unveil microbial interactions. *Science (80-.).* **369**, (2020).

231. Cadieux, N. & Kadner, R. J. Site-directed disulfide bonding reveals an interaction site between energy-coupling protein TonB and BtuB, the outer membrane cobalamin transporter. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 10673–10678 (1999).

232. Sun, X. *et al.* Putative cobalt- and nickel-binding proteins and motifs in Streptococcus pneumoniae. *Metallomics* **5**, 928–935 (2013).

233. Yan, J. *et al.* Purinyl-cobamide is a native prosthetic group of reductive dehalogenases. *Nat. Chem. Biol.* **14**, 8–14 (2018).

234. Matthews, R. G. *Cobalamin- and Corrinoid-Dependent Enzymes*. *Metal-Carbon Bonds in Enzymes and Cofactors* (2009). doi:10.1515/9783110436587-006.

235. Dimarco, A. A., Bobik, T. A. & Wolfe, R. S. Unusual coenzymes of methanogenesis. *Annu. Rev. Biochem.* (1990).

236. Bauerle, M. R., Schwalm, E. L. & Booker, S. J. Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. *Journal of Biological Chemistry* vol. 290 3995–4002 (2015).

237. Goenrich, M. *et al.* Probing the reactivity of Ni in the active site of methyl-coenzyme M reductase with substrate analogues. *J. Biol. Inorg. Chem.* **9**, 691–705 (2004).

238. Zhang, J., Chiu, W., Hydrodynamics, M. & Cryo-microscopy, E. *Encyclopedia of Biophysics*. *Encyclopedia of Biophysics* (2013). doi:10.1007/978-3-642-16712-6.

239. Morita, Y. *et al.* Redox Potentials of Cobalt Corrinoids with Axial Ligands Correlate with Heterolytic Co-C Bond Dissociation Energies. *Inorg. Chem.* **56**, 1950–1955 (2017).

240. Seth, E. C. & Taga, M. E. Nutrient cross-feeding in the microbial world. *Front. Microbiol.* **5**, 1–6 (2014).

241. Ragsdale, S. W. & Pierce, E. Acetogenesis and the Wood-Ljungdahl Pathway of CO2 Fixation. **1784**, 1873–1898 (2009).

242. Price, M. N., Deutschbauer, A. M. & Arkin, A. P. Four families of folate-independent methionine synthases. *PLoS Genet.* **17**, 1–21 (2021).