SURVEYING AND HARNESSING THE GENETIC, (META)GENOMIC, AND METABOLIC POTENTIAL OF THE DEEP CARBONATED BIOSPHERE

by

Adam Joshua Ehrich Freedman

B.S. Stanford University (2006)
M.S. Stanford University (2007)

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the Degree of

Doctor of Philosophy in Environmental Biology

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Author..... Signature redacted
Department of Civil and Environmental Engineering
May 19, 2016

Certified by..... Signature redacted
Jarelle R. Thompson
Visiting Assistant Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by..... Signature redacted
Heidi Nepf
Donald and Martha Harleman Professor of Civil and Environmental Engineering
Chair, Graduate Program Committee

Surveying and harnessing the genetic, (meta)genomic, and metabolic potential of the deep carbonated biosphere

by

Adam Joshua Ehrich Freedman

Submitted to the Department of Civil and Environmental Engineering
On May 19th, 2016, in partial fulfillment of the requirements for the Degree of
Doctor of Philosophy in Environmental Biology

ABSTRACT

The interaction between microbes and supercritical (sc) carbon dioxide represents an increasingly compelling area of research due to use of $scCO_2$ in geologic carbon sequestration (GCS) and as a sustainable chemical solvent. To investigate the long-term effects of GCS on the *in situ* deep subsurface biosphere, I conducted a taxonomic, geochemical and metagenomic survey of the McElmo Dome $scCO_2$ reservoir, which serves as a natural analog for GCS environments. Through 16S rRNA amplicon and metagenome sequencing, I identified *Sulfurospirillum*, *Rhizobium*, *Desulfovibrio* and members of the Clostridiales family associated with reservoir fluids. Annotations of complete genomes extracted from metagenomes predict diverse mechanisms for growth and nutrient cycling in deep subsurface $scCO_2$ microbial ecosystems at McElmo Dome.

Supercritical CO_2 is frequently used as a solvent for compound extraction and *in vitro* biocatalysis. However, due to its lethal effects, $scCO_2$ has previously been considered inaccessible for *in vivo* microbial bioproduct stripping. Utilizing a bioprospecting approach, I isolated strain *Bacillus megaterium* SR7 through enrichment culture and serial passaging of McElmo Dome $scCO_2$ reservoir fluids. I then initiated process improvements including media and culturing optimization under 1 atm CO_2 that increased SR7 growth frequency under $scCO_2$. After developing a genetic system enabling inducible heterologous enzyme expression, $scCO_2$ incubations of SR7 transformed with a two-gene isobutanol biosynthesis pathway generated up to 93.5 mg/l isobutanol. 5.2% of the total isobutanol was directly extracted by the $scCO_2$ headspace. This finding demonstrates for the first time the feasibility of active bioproduct synthesis and extraction in a single $scCO_2$-exposed bioreactor.

Thesis Supervisor: Janelle R. Thompson
Visiting Assistant Professor of Civil and Environmental Engineering

ACKNOWLEDGEMENTS

I am gratefully indebted to all of the people who have provided me personal, professional and technical support over the course of my Ph.D. dissertation research. This work would not have been possible without the tremendous network of friends, family and colleagues that I am so lucky to have depended upon over the last six years.

First, I would like to thank my advisor, Prof. Janelle R. Thompson, for her steady guidance and invaluable mentorship. I am indebted to Janelle for accepting and welcoming me into her research group and showing patience as I evolved from geochemist into microbiologist during the first few years of my Ph.D. Janelle provided me with the freedom and support to define the nature of my research projects, while ensuring that my work remained focused on generating impactful results based on rigorous discussion and review. One of the most important lessons that I learned from Janelle is that one cannot be sentimental about data; careful consideration of results must dictate the path forward rather than the amount of time invested. I appreciate Janelle's high expectations and the level of detail with which she approaches research, writing, and editing. I am a significantly more qualified and skilled scientist directly on account of the lessons I have learned from Janelle and I cannot thank her enough.

I would also like to thank my committee members, Professors Ed Delong, Martin Polz, and Kristala Prather for all of their valuable insights and advice over the past several years. Committee meetings always felt like a crucial course correction to ensure that I remained on an effective path for achieving my goals in the most scientifically sound manner as possible. The feedback from all committee members, as well as students and post-docs each in their respective labs contributed significantly to this work. I am especially thankful to Prof. Prather for her willingness to take me on as part-time member of the Prather Lab. It did not take very long for me to feel at home at Prather Lab meetings and amongst its members. I feel very lucky that Kris enabled me to learn so much about the areas of genetics and biotechnology that I find so compelling. I am extremely appreciative for her unwavering enthusiasm and support over the past few years, despite the fact that she is a Dallas Cowboys fan.

Without funding, this thesis research would not have been possible. I am thankful to the MIT Energy Initiative (MITEI), which funded the first year of my dissertation research through a BP-MIT Fellowship. I would like to thank Lawrence Berkeley National Laboratory (LBNL) for providing me with a

Molecular Foundry User Grant to conduct research in 2011, and for the advising and feedback by Caroline and Jonathan Ajo-Franklin during my stay at Berkeley. Next, I must thank the National Institutes of Health (NIH) for supporting three years of my research through the NIH Biotechnology Training Program Graduate Fellowship. In addition to funding, BTP exposed me to the tremendous research being done by MIT graduate students across a range of departments. BTP also enabled me to do an internship at Joule Unlimited, where my work with Jess Leber and Brian Green helped me improve significantly as a scientist and biological engineer. Significant funding was also provided by a U.S. Department of Energy (DOE) Office of Biological and Environmental Research Grant, which crucially enabled the final two years of my Ph.D. research. I also relied on funding from the Singapore-MIT Alliance for Research and Technology (SMART) Center for Environmental Sensing and Modeling (CENSAM) for metagenome bioinformatics analysis. Lastly, the Ralph M. Parsons Lab in the MIT Department of Civil and Environmental Engineering has provided innumerable resources throughout my Ph.D., including the MIT Ippen Fund, which funded travel for presentation of this work at several conferences, including American Geophysical Union (AGU) and American Society for Microbiology (ASM).

I am also greatly indebted to KinderMorgan for allowing me to sample well fluids at their McElmo Dome facility in Cortez, Colorado. In particular, I must thank Rick Gersch and Coy Bryant for helping with all of the planning and logistics associated with my visit.

The members of Thompson Lab over the past six years have been wonderful mentors, friends and colleagues to me. I learned a range of foundational protocols and how to ask insightful questions from post-docs Hector Hernandez and Samodha Fernando. Kyle Peet was with me nearly every step of the way and someone on whom I depended for knowledge and advice on a daily basis. He is a wonderful scientist and a great friend, and I cannot thank him enough for traveling the path with me. Two additional members of Thompson Lab who contributed directly to my dissertation research are Kevin Penn and BoonFei Tan, who helped with genomic sequencing and metagenomic analysis, respectively. Other members of the lab who I must thank for their friendship and input are Jia Yi Har, Jia Wang, Jean Pierre Nshimyimana, Tim Helbig, Hanny Rivera, Ju Young Lim, Luciane Chimetto, Tzipora Wagner, and Carolina Bastidas. A frequent visitor to Thompson Lab, and eventual collaborator, Prof. Mike Timko, deserves many thanks for his research vision and chemistry expertise.

The Prather Lab has served a second research home to me. I must thank Jason Boock for his wealth of knowledge and constant source of thoughtful collegiality and friendship. His work contributed significantly to strain and genetics development – and without him would not have been possible. Other members of the Prather Lab merit thanks for their advice, support and upbeat demeanor, including Irene Brockman, Aditya Kunjapur, Michael Hicks, Chris Reisch, Lisa Anderson, Apoorv Gupta, Sue Zanne Tan, Yekaterina Tarasova, Lisa Guay, Stephanie Doong, and Gwen Wilcox.

The Parsons community is unique in having provided such a warm and helpful work environment over the past six years. In particular, members of the Delong, Polz, Alm, Chisholm, Kroll and Gschwend labs have been invaluable sources of knowledge and provided access to research instruments and facilities crucial to my research. A special thank you to the Chisholm Lab for allowing me such extensive use of their microscope, microplate reader, and qPCR thermocycler. I also want to thank Prof. Ben Kocar and his lab members for their patience and hospitality over the past year in continuing to provide me with a laboratory home in Parsons. Sheila Frankel, Jim Long, Darlene Strother, Vicky Murphy, Kris Kipp, Kiley Clapper and all of the members of the custodial staff have kept Parsons and CEE running smoothly and were always there with a smile and a happy conversation, the importance of which cannot be overstated. Most significantly, I would like to thank the other students and post-docs of the Parsons community for their sustained friendship and willingness to take sanity-restoring coffee breaks with me, especially Kaighin, Dave, Kelsey, Diana, Mason, Cherry, Natasha, Kyle D., Anthony, Hamed, and Kelly. Also, big thanks to the Microbiology Program Class of 2010 that adopted me as one of their own.

It goes without saying that I relied significantly upon my friends and family over the past six years in order to persevere. In particular, Jacobs Ruben, Alan Lee, Scott Bürger, Brian Goodman, Graham Brant-Zawadzki, Alex Broad, Ian Fein, and Colin McNair – thank you for all of the happy times together in Cambridge and beyond. My family – my parents, brothers, cousins, aunts, uncles, grandparents – you have been the greatest support network I could have imagined. Thank you for believing in me and keeping me level. And lastly, Adi, my Sweet Dee – you are my rock. I could not have done this without you. I'm sorry I got home late from lab...every single time. Thank you for your unwavering support, inspiration and love.

TABLE OF CONTENTS

1

Review of supercritical carbon dioxide microbiology in natural and engineered systems

1.1 Introduction

The interaction between the microbial biosphere and supercritical (sc) carbon dioxide (above its critical point: T_c = 31.0°C, P_c = 72.8 atm; Figure 1) represents an increasingly compelling area of research due to the unique biochemistry, cellular lethality, and industrial utility afforded by the $scCO_2$ phase. Enabling properties associated with $scCO_2$ are derived from its hybrid gas and liquid type behavior, including high diffusivity and solubilizing capacity, respectively. Due to its non-polar chemistry, $scCO_2$ is capable of serving as a solvent for hydrophobic compounds including both liquids and gases, which demonstrate complete miscibility (Matsuda *et al.*, 2005). As a result, $scCO_2$ has

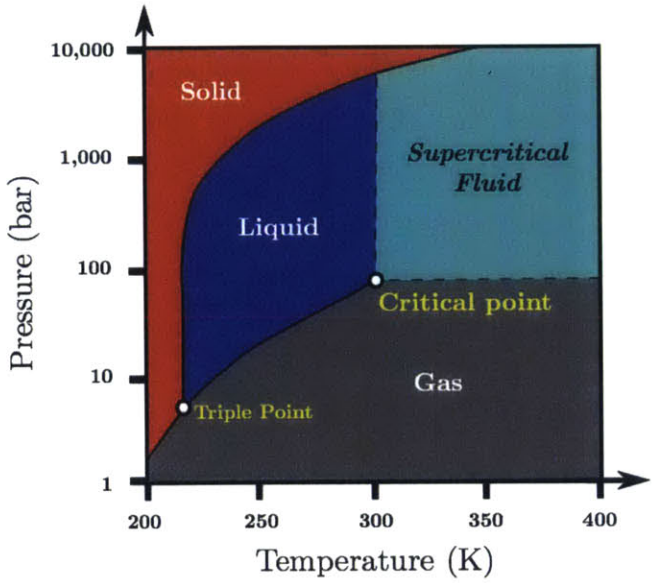


Figure 1. P-T phase diagram of CO_2. Above the T_c and P_c is the supercritical fluid phase.

been utilized for many different industrial processes, including compound extraction (e.g. caffeine), catalyzed organic synthesis, and chromatography (Desimone *et al.*, 2003; Kiran *et al.*, 2000; Beckman, 2004; Leitner, 2002). $ScCO_2$'s unique properties are also tunable, as they may be optimized for each specific use case by manipulating pressure, temperature and co-solvent conditions, which elevate its utility relative to many conventional organic solvents. $ScCO_2$ is also an increasingly attractive relative to typical solvents due to its being environmentally benign, non-toxic, broadly available, inexpensive, and non-flammable due to its fully oxidized state.

Proper facility siting adjacent to point source emitters like power stations would potentially enable flue gas CO_2 to be captured, purified, and utilized for productive processes, which in turn could limit the release of the greenhouse gas to the atmosphere and relieve global stresses on the environment and climate. The study of $scCO_2$ behavior in the deep subsurface has become increasingly crucial, as deep subsurface $scCO_2$ injection for geologic carbon sequestration (GCS) has been proposed as one of the foundational methods for reducing anthropogenic greenhouse gas releases to the atmosphere. The extent to which biological processes play a role in the fate, transport and long-term storage of CO_2 injected into geological formations remains an open question. Therefore, in an attempt to further understand the manner in which microbial community content and structure responds to $scCO_2$ exposure in the deep subsurface, this thesis investigated the microbial biosphere diversity in a natural $scCO_2$ reservoir that serves as an analog for the long-term effects of geologic carbon sequestration. Furthermore, this thesis researched the role that subsurface-derived $scCO_2$-resistant bacteria may play in harnessing the unique properties of sustainable solvent $scCO_2$ for microbial-catalyzed bioproduct generation and extraction through culturing optimization, genetic system development, and heterologous pathway engineering.

1.2 Microbial response to $scCO_2$ and high pCO_2 exposure

Supercritical CO_2 is generally regarded as a sterilizing agent of vegetative cells and a high-level disinfectant of most bacterial endospores (White *et al.*,

2006; Ortuño *et al.*, 2012, Mitchell *et al.*, 2008, Zhang *et al.*, 2006). When $scCO_2$ is introduced to a system, significant pH decreases (i.e. to pH ~3 in unbuffered systems, pH ~5-6 in buffered systems) occur on a timescale of several days (Kharaka *et al.*, 2006), a lethal scenario for most microbes. However, the population numbers of certain species that can tolerate acidic conditions may in fact rise, as previously inaccessible nutrients are released upon carbonic acid-driven mineral dissolution (Kharaka *et al.*, 2006). Furthermore, CO_2 itself may be used as a mineral oxidant or metabolic substrate, potentially enabling methanogens, sulfate-reducing bacteria (SRBs) and other CO_2-fixing autotrophs to thrive (Morozova *et al.*, 2010).

In addition to acidifying effects, $scCO_2$ in direct contact with microbes may introduce a range of potentially toxic stresses. Due to its predominantly non-polar solvent chemistry, $scCO_2$ penetrates bacterial cell walls and membranes, extracting fatty acids, lipids, and other intracellular materials that preferentially partition into the $scCO_2$ from the cytosol (Ulmer *et al.*, 2002). Inside the cell, $scCO_2$ may decrease intracellular pH, disable enzymes, disrupt protein synthesis, and cause cellular desiccation, ultimately resulting in cell death (Spilimbergo and Bertucco, 2003; Kirk, 2011; Zhang *et al.*, 2006).

The interaction of $scCO_2$ and microbial cells has been studied extensively within the context of sterilization for the food and drug industries. Though most microbial species are rapidly inactivated in the presence of $scCO_2$, several microbes have demonstrated the ability to limit the rate and extent of lethality upon exposure (Mitchell *et al.*, 2008; Oulé *et al.*, 2010). The rigidity of gram-positive cell walls afforded by dense layers of peptidoglycan (comprising up to 90% of the thickness) confers enhanced tolerance to exposure by reducing the rate of $scCO_2$ penetration into the cell (Oulé *et al.*, 2010).

In addition to inherent physiological traits like cell wall composition, microbes are thought to employ three major adaptive mechanisms in the presence of $scCO_2$ to maintain viability: 1) the dense matrix of extrapolymeric substances (EPS) composed of carboxylic acids, polysaccharides, amino acids, and other components that are commonly found in biofilms is thought to limit $scCO_2$ cellular envelope penetration through chemical interaction with CO_2 (Mitchell *et al.*, 2008; Braissant *et al.*, 2003); 2) modifications of microbial membrane structure (e.g. branching and chain length, fatty acid saturation) enables a cell to

calibrate its membrane fluidity and permeability in response to solvent, environmental and nutrient conditions (Spilimbergo *et al.*, 2009; Isenschmid *et al.*, 1995; Mitchell *et al.*, 2008; Spilimbergo and Bertucco, 2003; Klein *et al.*, 1999; Mangelsdorf *et al.*, 2009; Kieft *et al.*, 1994; Mukhopadhyay *et al.*, 2006); and 3) expression of alternative transcription factors triggers the general stress response, acid stress response, and sporulation cascade, each of which induces physiological adaptations to offset $scCO_2$-related stresses (Ogasawara *et al.*, 2012; Liao *et al.*, 2011; Martin-Galiano *et al.*, 2001; Richard and Foster, 2004; Foster, 1999; Gaidenko and Price, 1998).

Studies on rapid microbial responses to $scCO_2$ exposure have characterized cells during $scCO_2$ sterilization. While cells ultimately are killed during these experiments, this work implicates membrane adjustments as a short-term acclimation response mechanism to stresses associated with $scCO_2$ (Ogasawara *et al.*, 2012). Furthermore, when *E. coli* cells were exposed to $scCO_2$ sterilization, of the 15 known proteins that demonstrated significantly elevated expression, those with predicted functions for regulation of cell membrane composition and global stress regulation proteins were both included (Liao *et al.*, 2011). Several physiological changes have been documented in response to pH decreases, a stress associated with $scCO_2$: microbes upregulate proton and solute pumps (Martin-Galiano *et al.*, 2001) and pathways that generate and import pH buffering compounds (Richard and Foster, 2004; Foster 1999), and express alternative transcription factors that trigger stress responses and the sporulation cascade (Gaidenko and Price, 1998). Lastly, when $scCO_2$ is removed from the cell, microbes are able to synthesize proteins that repair damage, enabling microbes to grow again (Oulé *et al.*, 2006; Oulé *et al.*, 2010). Therefore, some $scCO_2$ damage is reversible, and the stress response to exposure appears to depend temporally on the direction of $scCO_2$ flux.

Recent work by Peet *et al.*, (*in review*) investigated the extent to which $scCO_2$-resistant *Bacillus* strains alter their protein expression and cell wall and membrane compositions in response to culturing under headspaces of 1 and 100 atm of CO_2 and N_2. Results showed that lipid chain lengths increased while fatty acid branching decreased in cultures incubated under $scCO_2$. Proteomic signatures of $scCO_2$ exposure were less distinct, with similar profiles exhibited under low and high pressures of CO_2 and N_2. The high expression of S-layer

proteins in one of the assayed strains (*Bacillus subterraneus* MITOT1) may suggest this aspect of membrane structure may enable the strain to withstand $scCO_2$ exposure during germination and growth. As upregulation of the glycine cleavage system under CO_2 conditions has previously been associated with acid stress responses, this metabolic function is expected to be activated in carbonic acid acidified media. Overall Peet *et al.* (*in review*) showed that cell membrane modifications, amino acid metabolism, and stress response metabolism may be implicated in persistence and growth under $scCO_2$ headspace.

1.3 GEOBIOLOGY OF $scCO_2$

Despite the sterilizing effect of $scCO_2$, it is hypothesized that a subset of microbes will be able to survive and grow in the presence of $scCO_2$ by employing a range of adaptive behaviors that are native to certain taxa or have evolved over geologic timescales. The constant seeding of microbes at the contacts between saline aquifer fluids and high pCO_2 fluid has likely provided a consistent introduction of genetic variation over millions of years, increasing the likelihood of $scCO_2$-tolerant phenotypes. While high pCO_2 conditions and $scCO_2$ exposure have lethal effects on some populations, it also appears to afford exploitable niche conditions for others, as the following studies demonstrate. Oppermann *et al.* (2010) showed that long-term exposure to a virtually anaerobic CO_2 vent (93-96% CO_2, 0.1-1.0% O_2) in near-surface soils resulted in substantially lower microbial population sizes relative to nearby reference soils under ambient conditions. However, several species with low or undetectable population sizes at the ambient condition site were present in significantly higher numbers at the CO_2 vent, including methanogens, *Geobacteraceae*, and sulfate-reducing bacteria (Oppermann *et al.*, 2010). It therefore appears that high pCO_2 concentrations select for communities comprised of microbes demonstrating anaerobic and acidophilic physiologies.

An additional study monitored the microbial community composition in a deep subsurface sandstone formation before, during and after CO_2 injection (Morozova *et al.*, 2011). While the formation pressure was slightly below the supercritical point for CO_2, the results may still inform hypotheses with regard to

expected community composition in natural subsurface $scCO_2$ reservoirs. Morozova *et al.* (2011) observed that injection of CO_2 decreased fluid pH from 7.5 to 5.5 and caused a three-order of magnitude reduction in bacterial density. An initial shift in community composition from chemoorganotrophic populations (i.e. fermentative halophiles and sulfate-reducing bacteria) to chemolithotrophic populations (i.e. methanogenic archaea) was also observed. After five months, the sulfate-reducing bacteria population rebounded and once again dominated the local community, revealing the potential for microbes able to survive initial $scCO_2$ introduction to demonstrate temporal adaptive abilities.

At high enough concentrations, the oxidizing and acidifying influence of CO_2 on the thermodynamic conditions of a system may make new metabolic niches available. Models suggest that when $scCO_2$ causes the pH of saline aquifers to decrease, the amount of energy available for Fe(III)-reduction increases while energy available for sulfate-reduction and methanogenesis remains largely unchanged (Kirk, 2011). Additional thermodynamic analyses suggest that some sulfur oxidation reactions coupled to CO_2 reduction may be energetically beneficial enough to be exploited *in situ* (West *et al.*, 2011). Further, CO_2-consuming hydrogenotrophic reactions will be more energetically favorable as a metabolic strategy than CO_2-generating acetotrophic reactions in $scCO_2$ reservoirs, increasing in energy potential with increasing CO_2 content (Kirk, 2011).

1.4 Geologic sequestration of $scCO_2$

Atmospheric CO_2 concentrations have increased from a pre-industrial level of 280 ppm to 379 ppm in 2005 (IPCC, 2007), with emissions projected to increase by 25 to 90% between 2000 and 2030 (SRES, 2000). While numerous gaseous species contribute to the cumulative global warming trend associated with the greenhouse effect, the CO_2 contribution is most significant, accounting for 77% of total anthropogenic GHG emissions in 2004 (IPCC, 2007). Several key mitigation strategies must be considered and safely enacted to stabilize the growing concentration of atmospheric CO_2. Carbon capture and storage (GCS) has been identified as one of the most important methods available for

significantly limiting point source emissions to the atmosphere (Orr, 2009).

The GCS process entails capturing CO_2 from the flue gases of power plants or other emitters, compressing the CO_2 into a supercritical fluid, and injecting it below a low-permeability sealing layer or "caprock" into a permeable formation in the deep subsurface for long-term storage. Three onshore geological environments are considered suitable for industrial-scale CO_2 storage: oil and gas reservoirs, unminable coal seams, and saline aquifers. While oil and gas reservoirs have demonstrated effective long-term sealing of their native deposits (Orr, 2009), saline aquifers are more likely injection targets, as they are more ubiquitously distributed and afford the highest projected storage capacity domestically and globally (IPCC, 2007; Orr, 2009).

Natural subsurface accumulations of CO_2 are excellent analogs for studying the long-term effects, implications and benefits of GCS. Massive CO_2 deposits, the largest of which contain the amount of CO_2 emitted at a 1000 MW coal-fired plant over 20 years, have been discovered, characterized and industrially produced around the world (Figure 2; Baines and Worden, 2004; Stevens et al., 2001). Subsurface accumulations of CO_2 in the United States, Hungary, and Turkey (Table 1) have been industrially produced for a range of

Figure 2. Global distribution of high (>20%) CO_2-content basins and major developed natural CO_2 production fields (Adapted from: Baines and Worden, 2004; Stevens *et al.*, 2001; NETL, 2010). No data available for Canada and Russia.

Table 1. Major developed natural CO_2 production fields (Adapted from Stevens et al., 2001)

Field	Location	Operator	Original CO_2 in place		1998 CO_2 production		Reservoir Lithology	Depth (m)
			10^6 tons	Tcf	10^6 t/y	MMcfd		
McElmo Dome	CO	KinderMorgan	1,600	30	15.9	820	Carbonate	2,300
Bravo Dome	NM	Occidental	1,600	30	7.2	375	Sandstone	700
Sheep Mountain	CO	Arco	780	15	2.9	150	Sandstone	1,500
Dodan	Turkey	Turkish Pet.	27	0.5	1.2	60	Carbonate	1,500
NE Jackson Dome Fields	MS	Denbury	320	6	0.4	20	Sandstone	4,700
St. Johns	AZ, NM	Ridgeway	830	16	0	0	Sandstone	500
Units: **Tcf**: trillion ft^3; **t/y**: tons/year; **MMcfd**: million ft3/day								

applications, including carbonated bottling, horticulture, and chemical manufacturing. By far the most prevalent industrial use for $scCO_2$ is enhanced oil recovery, with an estimated 40 megatons of naturally sourced CO_2 re-injected annually (primarily in Texas and New Mexico) by utilizing over 2,000 km of pipelines specifically dedicated to CO_2 transport and operations (Stevens *et al.*, 2001).

A variety of natural geologic processes have caused CO_2 to accumulate in the subsurface over the course of thousands to millions of years. Carbon isotope signatures ($^{13}C/^{12}C$) have revealed that generally, there are four sources of CO_2 emplacement: 1) high temperature igneous processes at plate rifting and collision zones induce thermal metamorphism, causing carbonate-bearing rocks to devolatilize CO_2 to the fluid phase (Yardley, 1989); 2) atmospheric CO_2 dissolved in meteoric water descends into subsurface aquifers, which equilibrates with surrounding rock, exsolving from the water, forming mineralized springs and geysers (Shipton *et al.*, 2004); 3) sedimentary rocks that contain organic detritus decompose upon burial by a combination of bacterial fermentation, oxidation and reduction, releasing CO_2 as a byproduct (Irwin *et al.*, 1977) coupled with thermal degradation of kerogen to CO_2 with petroleum as a potential intermediate (Ehrenberg and Jacobson, 2001); and 4) carbonate and aluminosilicate minerals contained in the same rock will undergo diagenesis upon heating, reacting with each other to generate CO_2 and sheet silicates (Smith and Ehrenberg 1989). Ultimately, massive CO_2 deposits are often derived from multiple sources (Baines and Worden, 2004), though the most volumetrically abundant CO_2 accumulations, such as McElmo Dome and Bravo Dome fields, CO, appear to be associated with deep volcanic, geothermal, or kerogen-based sources (Wycherley

et al., 1997).

The temperature and pressure regime 800 to 1000 m below surface causes CO_2 to be in a supercritical phase, such that CO_2 has the space-filling character of a gas, but the solvent properties and density of a liquid (White *et al.*, 2006). Supercritical phase character enables efficient subsurface storage due to reduced volume relative to gas, and susceptibility to trapping mechanisms that will limit migration or leakage (Orr, 2009). CO_2 will be subject to four dominant trapping mechanisms upon GCS injection or natural geological emplacement: **1)** structural trapping of buoyant $scCO_2$ by the overlaying caprock, **2)** dissolution trapping as CO_2 dissolves in the formation fluid, increasing its density, causing it to sink, **3)** residual trapping, wherein CO_2 is held in pore spaces by capillary action (Szulczewski *et al.*, 2012), and **4)** mineral trapping, as dissolved CO_2 precipitates out of solution as carbonate minerals (Gilfillan *et al.*, 2009; Haszeldine *et al.*, 2004).

CO_2 that accumulates in geologic formations by natural or industrial processes increases fluid acidity by forming carbonic acid upon dissolution (Baines and Worden, 2004), ultimately reaching equilibrium between the following three reactions:

$$CO_{2(g)} + H_2O \rightarrow H_2CO_{3(aq)} \quad (1)$$

$$H_2CO_{3(aq)} \rightarrow HCO_3^-{}_{(aq)} + H^+{}_{(aq)} \quad (2)$$

$$HCO_3^-{}_{(aq)} \rightarrow CO_3^{2-} + H^+{}_{(aq)} \quad (3)$$

The protons released into solution drive geochemical reactions, including mineral dissolution and precipitation, depending on the thermodynamics and equilibrium state of the local system. For example, carbonate minerals may dissolve by the reaction:

$$M^{(II)}CO_{3(s)} + 2H^+{}_{(aq)} \rightarrow M^{2+}{}_{(aq)} + HCO_3^- \quad (4)$$

where 'M' is a divalent cation (Ca, Mg, Fe), or precipitate out of solution by the reaction:

$$CO_{2(aq)} + 2H_2O + M^{2+} \rightarrow M^{(II)}CO_{3(s)} + 2H^+{}_{(aq)} \quad (5)$$

depending on the CO_2 content and divalent cation activity in solution (Baines

and Worden, 2004).

In addition to abiotic reactions, microbial organic acid production is known to catalyze mineral weathering rates by up to two orders of magnitude relative to abiotic controls (Barker *et al.*, 1998; Ferris *et al.*, 1996; Mitchell *et al.*, 2009), while cell surfaces are known to enhance rates of mineral precipitation by serving as mineral nucleation sites (Aloisi *et al.*, 2006). Therefore, characterizing the diversity and function of microbes that are native to formations with high concentrations of $scCO_2$ would dramatically improve our ability to more thoroughly model $scCO_2$ physicochemical behavior, and the evolution of caprock and injection zone integrity post-injection.

As a geoengineering strategy intended to mitigate the migration of injected $scCO_2$ during GCS, members of the natural *in situ* $scCO_2$ microbial biosphere or laboratory-developed synthetic diversity may be injected in the deep subsurface to induce carbonate mineral precipitation (i.e. serving as mineral nucleation sites; Anbu *et al.* 2016) or to generate biofilm and cellular surface EPS substances that may reduce permeability in the injection zone by clogging pores between mineral grains (Mitchell *et al.*, 2008). In Ca^{2+}-rich, neutral to alkaline fluids, most bacterial species are able to facilitate carbonate mineral precipitation (Zamarreno *et al.*, 2009). Specifically, the interaction between positively charged Ca^{2+} ions and negatively charged bacterial cell walls enables bacteria and minerals to aggregate, together serving as mineral nucleation sites (Zamarreno *et al.*, 2009). EPS content, including peptide and nucleic acid matrices, have been shown to serve as mineral nucleation sites as well (Geesey and Jang, 1990).

Due to the natural interactions and geoengineering implications of the deep biogeosphere, this thesis's characterization of microbial population diversity and function in a natural GCS analog system will provide crucial insights into the types of microbes that CO_2 injection may select for and the potential biological effects of the resulting community on the fate, transport, and long-term storage of $scCO_2$ in a GCS context. Moreover, since microbial activity has been responsible for pipeline and structural corrosion in subsurface wells and tunnels (West *et al.*, 2011), the results of thorough community and activity characterizations will inform future GCS siting and infrastructure materials decisions. Though the scope of Chapter 2 in this thesis is limited to a taxonomic, metagenomic and geochemical survey of a carbonate formation community, the

results will provide a reference by which to compare communities isolated from other geologic contexts, including saline aquifer $scCO_2$ reservoirs in sandstone formations (e.g. Bravo Dome). This comparison would clarify whether only certain taxa are equipped to survive in CO_2 accumulation systems or whether geochemical context is a significant selective driver as well.

1.5 Bioprospecting the deep subsurface for $scCO_2$-resistant strains

Successful culturing of isolates from deep subsurface $scCO_2$ reservoir fluids further enables laboratory studies on the extent to which microbial activity may limit $scCO_2$ migration and leakage in GCS environments. However, in a broader context, $scCO_2$-tolerant microbes hold significant potential for additional bioengineering opportunities not only in the GCS space, but also where $scCO_2$ is used for bioindustrial purposes, broadly expanding the scope of impact associated with members of the deep carbonated biosphere. Specifically, biocompatibility with $scCO_2$ would enable novel bioproduction capacity on account of the unique solvent and sterilization properties of $scCO_2$. In a recent study comprising early work associated with this thesis (Peet *et al.*, 2015), we reported the isolation of bacteria from three sites targeted for geologic carbon dioxide capture and sequestration (GCS) that demonstrated active growth in biphasic bioreactors loaded with aqueous media and pressurized with supercritical carbon dioxide. Enrichment cultures seeded with fluids and rock cores from subsurface formations subjected to serial passaging under a $scCO_2$ headspace resulted in the isolation of six strains of spore-forming facultative anaerobes: *Bacillus cereus*, *Bacillus subterraneus*, *Bacillus amyloliquefaciens*, *Bacillus safensis*, and *Bacillus megaterium*. When inoculated as prepared endospores in high-pressure reactors, isolates as well as several *Bacillus* type strains demonstrated growth under $scCO_2$. Therefore, it appears that the capacity for germination may be facilitated by physiological traits associated either with the *Bacillus* genus or the unique protective character of spore coats, which may limit $scCO_2$ membrane penetration. These results generated evidence in support of the notion that microbes may persist and grow at the interface between $scCO_2$ and an aqueous

phase, either in natural or engineered systems.

The demonstration of culturable diversity under $scCO_2$ represents a significant step forward towards the goal of developing a bioproduction platform strain that is able to generate bioproducts through active growth and metabolism under $scCO_2$. Reported growth of several *Bacillus* isolates in Peet *et al.* (2015) was often stochastic and low frequency, behavior that precluded further biotechnological development of the investigated isolates. Therefore, it appears that in order to have a viable bioproduction strain it is crucial to improve $scCO_2$ growth frequencies by modifying incubation conditions and/or attempting to isolate additional strains from the environment with better evolved capacity for growth under $scCO_2$.

1.6 Microbial biotechnology harnessing supercritical CO_2

Many enzymes remain functional in the $scCO_2$ phase, and are even capable of demonstrating unique characteristics that are otherwise not possible in aqueous solutions. In close proximity to the CO_2 critical point, minor conditional changes in pressure or temperature may further modify enzymatic solubility and partition coefficients, in addition to solvent phase conductivity, dielectric constant, and dipole moment (Budisa *et al.*, 2014). Therefore, $scCO_2$ has been extensively explored as solvent for both *in vitro* and *in vivo* biocatalysis reactions that are difficult or expensive in aqueous phase reactors with a specific focus on semi-hydrophobic compounds due to $scCO_2$'s non-polar chemistry. Research and development has led to the design of unique biocatalyzed substrate transformations using a variety of multiphase reactor schemes (Knez *et al.*, 2005; Laudani *et al.*, 2007; Matsuda *et al.*, 2000; Matsuda *et al.*, 2005; Salgin *et al.*, 2007). Industrially relevant biochemical transformations that have previously been demonstrated in $scCO_2$ largely include *in vitro* reactions using purified enzymes, include amidation, esterification (Nakamura *et al.*, 1986; Marty *et al.*, 1992), acetylation, transglycosylation and reduction (Wimmer and Zarevúcka (2010). Of particular interest are biocatalyzed reactions with a single "handedness" or chirality, which in $scCO_2$ generate enantiopure chemical compounds that are otherwise difficult to synthesize (Matsuda *et al.*, 2000; Matsuda *et al.*, 2008;

Matsuda *et al.*, 2004; Salgin *et al.*, 2007). As a result, while CO_2 is typically considered a waste product with a dangerous global impact potential, it also represents a useful, abundant resource that may be employed as a solvent and/or substrate with a broad range of biotechnological applications.

CO_2-fixing carboxylation reactions are of particular interest with regard to the future of sustainable microbial-facilitated bioproduct generation. A broad array of *in vitro* studies support the notion that autotrophic or mixotrophic growth may in the future enable direct fixation of the $scCO_2$ solvent, reducing the need for conventional carbohydrate-based feedstocks. Previous demonstrations of biocatalyzed CO_2-fixation reactions utilizing $scCO_2$ as a substrate include the synthesis of urethane (Yoshida *et al.*, 2000), dimethyl carbonate (Ballivet-Tkatchenko *et al.*, 2006), styrene carbonate (Kawanami and Ikushima, 2000), and methyl acetate (Sowden *et al.*, 1999). A variety of additional *in vitro* CO_2-fixation reactions hold potential for use under supercritical conditions, including catalysis by ribulose-1,5-diphosphate carboxylase (Hartman and Harpel, 1994), isocitrate dehydrogenase (Sugimura *et al.*, 1989), malate dehydrogenase (Sugimura *et al.*, 1990), and the fixation of CO_2 on pyrrole by purified decarboxylases from *Bacillus megaterium* (Wieser *et al.*, 1998; Yoshida *et al.*, 2000; Wieser *et al.*, 2001).

The ability to conduct stereo-specific, carboxylation, and other industrially relevant chemical reactions by accessing $scCO_2$ as the solvent is limited by the availability of purified enzymes (which are progressively degraded by $scCO_2$ exposure) and the need to replenish small molecule reductants (i.e. NAD(P)H) to regenerate the active state of the enzymatic catalyst. In a first demonstration of how to potentially overcome these challenges, Matsuda *et al.* (2000, 2001) reported biocatalyzed reduction of ketones by immobilized *Geotrichim candidum* cells and carboxylation of pyrrole using living *Bacillus megaterium* cells. In both cases, the ability of the cellular envelope to protect biocatalysts and co-factors enabled reactions to proceed without the amendment of purified enzymes. The *B. megaterium* carboxylation report (Matsuda *et al.*, 2000) claimed a doubling of cell concentrations during the course of the incubation, which in tandem with negative results from growth-free controls appeared to confirm the necessity for live cells in catalyzing the $scCO_2$-reducting reaction (Matsuda *et al.*, 2000; Matsuda *et al.*, 2001). While these studies demonstrated the proof of concept

whereby $scCO_2$ exposed whole cells may help facilitate biochemical transformations, cultures demonstrating robust growth under $scCO_2$ would represent a significant improvement in the overall bioplatform development process by continually regenerating relevant enzymes and co-factors in a protected environment. Therefore, the demonstration of enzymatic biocatalysis and active growth under $scCO_2$ reported in both Peet *et al.* (2015) and this thesis indicate the establishment of a novel method for microbial bioproduction that merits further investigation and development.

1.7 Development of advanced biofuel production for *in situ* extraction

A significant factor limiting the capacity of environmental microbes to be metabolically modified for industrial applications is the challenge of genetic intractability. In order to introduce foreign DNA into environmental isolates as a means for generating non-native bioproducts or to investigate gene function by knockouts, a genetic system must be developed. Previous work establishing genetic systems has enabled successful investigation and exploitation of unique biochemical capacities associated with bioprospected strains, including adaptations to extreme deep-sea environments in *Pseudoalteromonas* sp. SM9913 (Yu *et al.*, 2014), bioremediation of organic and metal contaminants by *Geobacter sulfurreducens* (Coppi *et al.*, 2001), and mineral electron donor oxidation by *Thiobacillus denitrificans* (Letain *et al.*, 2007). Bioprospecting of marine microbial natural products over the last few decades has enabled the discovery of vast pharmaceutical and agricultural agents through strain isolation, compound screening, and metagenomics (Xiong *et al.*, 2013). One of the most well known successful bioprospecting outcomes was the identification and purification of DNA polymerase (Chien *et al.*, 1976) from a strain of *Thermus aquaticus* isolated from Yellowstone hot springs (Brock and Freeze, 1969), which enabled the industrial and commercial development of thermophilic *Taq* polymerase for use in polymerase chain reaction (PCR). Analogously, the isolation and genetic system development in environmentally isolated strains demonstrating $scCO_2$-resistance holds a similar potential for enabling new approaches to biotechnological challenges.

Given the available genetic tools and established methods for strain modification, considerable progress has been made toward developing microbes for the production of liquid biofuels (Connor and Liao, 2009). Biofuels are compelling with regard to $scCO_2$ harvesting systems because the moderately hydrophobic chemistry of hydrocarbons like butanol readily causes compound partitioning from the aqueous phase into $scCO_2$ (i.e. octanol-water partition coefficient, K_{ow} >4); Timko *et al.*, 2004). Such partitioning could be harnessed for *in situ* product extraction and recovery during biosynthesis of long-chain/branched alcohols from aqueous environments. Advanced biofuel production is further motivated by performance improvements over ethanol, which to this point has been the common gasoline additive (US ethanol production reached 6.5 billion gallons in 2007 (Dinneen, 2008). However, since ethanol has several practical drawbacks, including low energy density (~70% of gasoline), high hygroscopicity (ability to hold water), and high corrosiveness relative to longer chain hydrocarbons (Connor and Liao, 2009), advanced biofuels (C3-C5 alcohols) that display higher energy density and diminished hygroscopic character are better suited for integration with current infrastructure (Nigam and Singh, 2011). Beyond fuel applications, higher chain alcohols may be biocatalytically dehydrated to alkenes, at which point they may be processed to sustainably generate commodity chemicals including paints, surface coatings, solvents, plastics, and resins (Connor and Liao, 2009).

1.8 Challenges and improvements in biofuel extraction

Process limitations associated with end product sensitivities in bioproduction hosts emphasizes the importance of implementing efficient end-product extraction methods in order to strip bioproducts before they reach toxic concentrations. Many species, including *Bacillus subitilis*, suffer from short-to-medium chain alcohol toxicity (Nielsen *et al.*, 2009; Liu and Qureshi, 2009; Ezeji *et al.*, 2010). While end product tolerance has been improved by acclimatization (Kataoka *et al.*, 2011) and media modifications (Lam *et al.*, 2014), the ability to strip bioproduced compounds *in situ* without inhibiting microbial growth represents a promising approach to relieving toxic effects while facilitating

continued production. Further work using synthetic biology, metabolic engineering, and directed evolution will help guide researchers toward more successful and efficient biochemical tolerance.

Despite broad process improvements, advanced biofuels remain difficult to purify and dehydrate by distillation. For example, the 1-butanol–water system has an azeotrope at 93°C containing 42.4 wt% water (Laitinen and Kaunisto, 1999), preventing separation by conventional distillation processing. Unlike ethanol, an advanced biofuel like butanol has a low vapor pressure and high boiling point (118°C), which poses further challenges for *in situ* distillation due to high energy requirements associated with inducing volatility and downstream dehydration processing to reach fuel-grade specifications (Vane, 2008). As a result, studies have shown that product recovery by distillation is the most energy-intensive step in the microbial production of biobutanol (Ezeji *et al.*, 2004, 2007). Furthermore, elevated temperatures that would be required for distillation of higher chain alcohols concomitant with active fermentation would result in microbially lethal conditions, leading to undesired bioreactor cessation, sterilization and reinoculation.

Due to these myriad challenges, alternative separation technologies such as gas stripping, pervaporation or $scCO_2$ partitioning represent process improvements that can be used in concert with actively fermenting cultures to increase reactor productivity by relieving end product toxicity. However, of these non-lethal extraction methods, only $scCO_2$ provides a pathway for generating purified, dehydrated biofuel products by simple depressurization due to the low solubility of water in $scCO_2$ (Sabirzyanov *et al.*, 2002) and high partitioning coefficient of hydrophobic biofuels in the $scCO_2$ phase (Timko *et al.*, 2004). Unlike $scCO_2$ extraction, gas stripping and pervaporation still require energy intensive product dehydration steps, diminishing the sustainable nature of these methods.

Abiotic experimental studies testing the extraction of 1-butanol from aqueous solutions with $scCO_2$ solvent demonstrated that up to 99.7% of the initial amount of 1-butanol was removed from the feed stream (Laitinen and Kaunisto, 1999). $ScCO_2$ thus represents a technically feasible solvent for continuous advanced biofuel extraction from growing cultures in the event that strains capable of withstanding $scCO_2$ exposure are isolated or otherwise

acclimated to the stressful culturing conditions. Moreover, a bioreactor inoculated with a $scCO_2$-resistant strain should thus not be subject to typical bioreactor contamination challenges as $scCO_2$ exposure represents a lethal condition for the vast majority of microbial species. As described in this thesis, the recovery of an environmental strain with the capacity for growth and heterologous pathway expression in pure culture under $scCO_2$ with direct *in situ* product recovery demonstrates that several of the significant practical barriers to implementing the overall proposed microbial bioproduction system have been solved.

1.9 Research questions

Despite the physiological and metabolic challenges associated with microbial growth while exposed to supercritical carbon dioxide, our previous demonstration of several isolate growth under a $scCO_2$ headspace (Peet *et al.*, 2015) prompted several additional questions about both natural and engineered $scCO_2$ systems. In this thesis, I sought to survey the taxonomic and genomic diversity associated with a natural subsurface $scCO_2$ reservoir and attempted to harness the biotechnological potential of culturable diversity isolated from the reservoir in order to access the sustainable solvent $scCO_2$ as a means for improving existing industrial methods for biocatalyzed product generation and purified extraction. The questions that I address in this thesis are:

Chapter 2

Questions: What is the microbial taxonomic diversity associated with subsurface accumulation of $scCO_2$? Does community genomic content indicate the metabolic potential for nutrient cycling within the local ecosystem?

I hypothesized that low, but non-zero, ecosystem diversity would include taxa from other deep subsurface environments and that taxonomic diversity and abundance would vary as a function of local geochemical nutrient availability and CO_2/H_2O ratios due to stresses associated with $scCO_2$ solvent effects. While subsurface environments have been associated with a wide range of OTU diversity, e.g. n = 1 in deep gold mine fluids (Chivian *et al.*, 2008), n = 445 in

mesothermal petroleum reservoirs (Pham *et al.*, 2009), and n >14,000 in the subseafloor sediments (Biddle *et al.*, 2008), $scCO_2$ reservoirs are a highly selective environment that will likely relegate *in situ* species diversity numbers to the lower end of previously characterized subsurface communities based on several studies investigating the response of the microbial biosphere to injected $scCO_2$ at geologic carbon sequestration pilot sites (Morozova *et al.*, 2011; Mu *et al.*, 2014). The detection of a metabolically active microbial biosphere in deep sea systems exposed to liquid and, in some cases, supercritical phase CO_2 (Yanagawa *et al.*, 2012) indicates that terrestrial analog systems should, in theory, also be able to support microbial life. Previous studies examining the persistence of life under temperatures (Takai *et al.*, 2008), pressures (Boonyaratanakornkit *et al.*, 2007) and acidities (Johnson, 1998) more extreme than conditions associated with $scCO_2$ saline reservoirs indicate that active metabolic outgrowth is fundamentally possible. Lastly, geochemical modeling suggests that the exposure of $scCO_2$ to certain petrological and mineralogical conditions should increase the likelihood that microbial growth is supported *in situ* by making certain redox reactions more thermodynamically favorable (Onstott, 2005; Kirk, 2011).

While genomic content is expected to generally indicate the capacity for anaerobic growth by chemolithoheterotrophy based on previous studies (Mu *et al.*, 2014; Emerson *et al.*, 2015), genes specifically associated with autotrophy or alternative CO_2-fixation reactions are expected due to the ubiquitous presence of substrate CO_2. To test these hypotheses, I collected deep subsurface fluids from the McElmo Dome CO_2 field in southwestern Colorado, which were then prepared for 16S rRNA and metagenomic surveys. Fluids were also analyzed for geochemical content by inductively coupled plasma (ICP) methods. Taxonomic and functional annotations by RDP/Silva and RAST/IMG, respectively, enabled a reconstruction of *in situ* microbial-mediated nutrient cycling, which was further used in conjunction with geochemical measurements to propose signatures of active metabolic growth and inter-species syntrophic relationships. Overall, the previous detection of sustained microbial activity in high pCO_2 environments in conjunction with diversity observed at McElmo Dome have significant implications for understanding the manner in which microbes interact physically, geochemically, and biochemically with their host mineralogy and aqueous and supercritical fluids upon long-term $scCO_2$ emplacement.

CHAPTER 3

Questions: Can microbes isolated from a natural subsurface $scCO_2$ system be maintained in lab culture under $scCO_2$ conditions? Can culturing be optimized to enable natural product generation under a $scCO_2$ solvent headspace?

I hypothesized that by selecting a range of enrichment passaging media it would be possible to cultivate $scCO_2$-adapted microbes under proposed microbial bioreactor system conditions. I further expected that genomic sequencing of isolated strains would reveal which central carbon metabolic pathways may be exploited for improved growth outcomes, including in the development of defined minimal media. Detected genomic signatures were also hypothesized to reveal aspects of isolate strains with biotechnological implications, including promoter systems, plasmid content, secretion systems, and carbon storage (e.g. PHA/PHB metabolism).

To test these hypotheses, I used fluids collected from McElmo Dome wells to inoculate enrichment cultures as a means for isolating strains by serial passaging under $scCO_2$. We previously demonstrated the ability to isolate and culture bacteria from subsurface environments in Peet *et al.* (2015), which thus served as a successful precedent for the overall bioprospecting approach. The isolation of six microbial strains from McElmo Dome fluid-inoculated cultures enabled me to proceed with biotechnological development of the single isolate that demonstrated the most robust and consistent growth under $scCO_2$. As a result, I sequenced and annotated the genome of isolate *Bacillus megaterium* SR7, which despite similarities to previously sequenced *B. megaterium* genomes, contained approximately 10% unique gene content. Functionally annotated genes revealed the presence of a well-established xylose-inducible promoter system, canonical central metabolic pathways (i.e. glycolysis, TCA Cycle), and the capacity for fermentative growth.

CHAPTER 4

Questions: Can a genetic system be developed for an isolate strain to enable heterologous protein expression and biofuel production under $scCO_2$ conditions? Can generated biofuels be recovered directly from the $scCO_2$ phase upon partitioning from the aqueous media?

Previous genetic system development across broad taxonomic classes of environmental bacteria supported the hypothesis that it would be possible to introduce heterologous DNA to strains isolated from the McElmo Dome deep subsurface environment. I hypothesized that upon successful development of a genetic system for *B. megaterium* SR7, it would be possible to demonstrate inducible protein expression of a single gene as well as increasingly complex multi-gene pathways under $scCO_2$. Heterologous LacZ expression was chosen for initial proof of concept. Isobutanol was chosen as a two-gene pathway product due to its ability to readily partition into $scCO_2$ and because its enzymatic production required the introduction of only two genes (isoketovalerate decarboxylase (*kiv*D) and alcohol dehydrogenase (*adh*)) into SR7 when feeding valine biosynthesis pathway intermediate α-isoketovalerate. Since valine is an essential amino acid, I anticipated that by upregulating flux through this pathway, we would generate observable concentrations of desired metabolites. I hypothesized that this genetic modification would be successful based upon the previous successful demonstration of heterologous isobutanol pathway expression in the closely related species *Bacillus subtilis* (Choi *et al.*, 2014), as well as multi-enzyme pathways in alternative strains of *Bacillus megaterium* (e.g. cobalamin in Biedendieck *et al.*, 2010). By assaying several homologs of the alcohol dehydrogenase gene to identify the best performing variant, I anticipated that the optimized isobutanol pathway in SR7 should generate enough product to be extracted directly from the $scCO_2$ phase, proving out the overall concept of single reactor growth, production and extraction system without fermentative disruption. Despite low $scCO_2$-extraction efficiencies associated with my pressurization system (based on abiotic controls), isobutanol produced by the genetically engineered *Bacillus megaterium* SR7 environmental isolate and

extracted directly into $scCO_2$ confirmed my overarching hypothesis that a microbial bioproduction system under $scCO_2$ was feasible.

Overall, my results indicate that the McElmo Dome natural $scCO_2$ reservoir harbors a "deep carbonated biosphere" adapted for growth and persistence despite extreme stresses associated with $scCO_2$ exposure. Bioprospecting this unique environment through enrichment passaging demonstrated the capacity to exploit natural evolutionary adaptations for biotechnological industrial applications. Laboratory process development established that culturing improvements are required to optimize wild type strains for product generation via genetic engineering of metabolic pathways. The detection of taxonomic diversity in the deep subsurface that was not cultured in the laboratory indicates that additional utility may be derived from further attempts at environmental strain isolation. Our results also suggest the existence of a microbial ecosystem associated with the natural McElmo Dome $scCO_2$ reservoir indicates that potential impacts of the deep biosphere on CO_2 fate and transport should be taken into consideration as part of GCS planning and modeling. The successful development of a genetic system in *B. megaterium* SR7, albeit at lower efficiencies than previously developed *B. megaterium* strains, in tandem with the identification of an effective promoter system ultimately enabled inducible heterologous protein expression under $scCO_2$ headspace conditions. Confirmation of this hypothesis was evaluated by scoring for xylose-induced exogenous LacZ activity and isobutanol production in transformed strains of SR7 that were cultured under $scCO_2$ and demonstrated robust growth.

The significance of demonstrable microbial growth and engineered biocatalysis under $scCO_2$ lies in its potential for "greening" biochemical production by streamlining energy costs and bioreactor maintenance. This thesis therefore represents an interdisciplinary research effort that synthesizes taxonomic and genomic data with improvements in bioprocess engineering, microbial physiology, and genetic engineering to overcome long-standing questions about microbial $scCO_2$ resilience and challenges associated with advanced biofuel production.

2

LIFE IN THE DEEP CARBONATED BIOSPHERE: MICROBIAL 16S rRNA GENE AMPLICONS AND GENOMES FROM THE MCELMO DOME SUPERCRITICAL CO_2 RESERVOIR INDICATE MICROBIAL POTENTIAL FOR CARBON AND BIOGEOCHEMICAL CYCLING

Adam J.E. Freedman, BoonFei Tan, and Janelle R. Thompson
(Adapted from a manuscript in preparation for journal submission)

ABSTRACT

Microorganisms catalyze carbon cycling and biogeochemical reactions in the deep subsurface and thus may be expected to influence the fate of injected supercritical (sc) CO_2 following geological carbon dioxide sequestration (GCS). We hypothesize that natural subsurface accumulations of $scCO_2$ that may serve as analogs for the long-term fate of injected $scCO_2$ harbor a "deep carbonated biosphere" adapted to high *in situ* pCO_2 conditions. To examine microbial community structure and biogeochemical potential within a natural $scCO_2$ reservoir we sampled subsurface fluids from CO_2-water separators at McElmo Dome, Colorado for analysis of 16S rRNA gene amplicon diversity and metagenome content. Amplicon and metagenome sequences were dominated by seven bacterial groups, including *Sulfurospirillum*, *Rhizobium*, *Desulfovibrio* and members of the Clostridiales family. Complete genomes from high abundance taxa were extracted from metagenomes using homology and compositional approaches. Annotated genome content revealed diverse mechanisms for growth and nutrient cycling including pathways for CO_2 and N_2 fixation, anaerobic respiration, sulfur oxidation, fermentation and potential for metabolic syntrophy. Differences in biogeochemical potential between two well communities were consistent with differences in fluid chemical profiles, suggesting a potential link between microbial activity and geochemistry. The existence of a microbial ecosystem associated with the McElmo Dome $scCO_2$ reservoir indicates that potential impacts of the deep biosphere on CO_2 fate and transport should be taken into consideration as a component of GCS planning and modeling.

2.1 INTRODUCTION

Natural carbon dioxide reservoirs in the greater Colorado Plateau and Southern Rocky Mountains region serve as models for understanding the long-term fate of CO_2 after subsurface injection for long-term geological carbon sequestration (GCS) (IPCC, 2007; Lal, 2008). While these reservoirs have been geologically and geochemically characterized for commercial CO_2 production (Baines and Worden, 2004; Stevens *et al.*, 2001), the taxonomic and genomic diversity of microbial populations in these systems remain unknown. At reservoir depths (typically >1 km), CO_2 exists in the supercritical (sc) (i.e. ≥31.1°C, ≥72.9 atm) or near critical state. Since $scCO_2$ is regarded as a microbial sterilizing agent (White *et al.*, 2006; Ortuño *et al.*, 2012; Mitchell *et al.*, 2008; Zhang *et al.*, 2006), whether $scCO_2$-bearing reservoirs support microbial life that may catalyze biogeochemical processes remains an open question.

Recent field and laboratory studies indicate that some microbial species may be resilient to stresses associated with $scCO_2$ (Peet *et al.*, 2015; Mu *et al.*, 2014; Mitchell *et al.*, 2008) and near-critical CO_2 (Emerson *et al.*, 2015; de Beer *et al.*, 2013). Due to its predominantly non-polar solvent chemistry, pure $scCO_2$ may penetrate bacterial cell walls and membranes, extracting fatty acids, lipids, and other intracellular materials from the cytosol (Ulmer *et al.*, 2002). High concentrations of dissolved CO_2 may decrease intracellular pH, disable enzymes, disrupt protein synthesis, and cause cellular desiccation, ultimately resulting in cell death (Spilimbergo and Bertucco, 2003; Kirk, 2011; Zhang *et al.*, 2006). Injection of $scCO_2$ during GCS may also indirectly stimulate microbial growth by extracting nutrients from the subsurface organic matrix (Kharaka *et al.*, 2006), releasing redox substrates from dissolved minerals, and potentially serving as a substrate for autotrophic metabolism. Thus, exposure to $scCO_2$ during GCS may represent a major selective agent for microbial diversity (Spilimbergo and Bertucco, 2003; Kirk, 2011; Zhang *et al.*, 2006; Mu *et al.*, 2014).

The 800 km^2 McElmo Dome $scCO_2$ reservoir in southwestern Colorado (Figure S1) is the largest supplier of industrially produced CO_2 in the world (Stevens *et al.*, 2001). Estimates based on stable isotope data suggest CO_2 began to accumulate at McElmo Dome 40 to 72 million years ago (Cappa and Rice, 1995; Gilfillan *et al.*, 2008), a timescale during which assembly and adaptation of

$scCO_2$-tolerant microbial communities may have occurred. CO_2 at McElmo Dome is trapped at depths of 1800 to 2600 m within the 100 m thick dolomite-rich Leadville Formation (Allis *et al.*, 2001; Gilfillan *et al.*, 2009) where the CO_2 exists as a supercritical fluid at an approximate temperature and pressure of 65°C and 135 atm, respectively (Allis *et al.*, 2001). Rich in permeable dolomites ($CaMg(CO_3)_2$), the Leadville Formation has an average porosity of 11%. Above the Leadville Formation is the 400 m thick Paradox Salt Formation, which acts as a low-permeability trapping layer (Stevens *et al.*, 2001). KinderMorgan operates production wells at McElmo Dome where a two-phase gas-brine mixture is produced and separated, after which the CO_2 is further dehydrated and compressed before pipeline delivery, while the brine is re-injected (Stevens *et al.*, 2001). The overall gas content consists of 98.2% CO_2, 1.6% N_2, 0.2% CH_4 (Allis *et al.*, 2001).

We hypothesized that the McElmo Dome formation would harbor a microbial ecosystem adapted to the anoxic, low-nutrient conditions typical of subsurface habitats (Phelps *et al.*, 1994) with adaptations enabling survival and CO_2 utilization under conditions of high dissolved pCO_2 in co-existence with pure-phase $scCO_2$. To examine this, we sampled produced fluids from ten CO_2 production wells and an onsite pond containing water used for well drilling, comparing cell densities with element and nutrient profiles. Taxonomic and genomic diversity were characterized from two wells by analysis of 16S rRNA gene amplicons and microbial genomes extracted from metagenomes. Our results suggest formation fluids harbor a low-density microbial community that includes *Sulfurospirillum*, *Rhizobium*, *Desulfovibrio* and members of the Clostridiales family. Analysis of complete and nearly complete genomes suggests that CO_2 and N_2 fixation, sulfur and arsenic redox processes, and metabolic syntrophy play roles in biomass production, biogeochemistry and carbon cycling in the $scCO_2$ reservoir.

2.2 Methods

Collection of formation water and characterization of biomass and geochemistry

Fluid-gas separators at ten CO_2 production wells from two areas of the McElmo Dome system (Yellow Jacket and Hovenweep fields) were decanted and filled 15 hours prior to sample collection. Approximately 40 liters of fluid was collected in acid-washed carboys from the separators and from a surface pond used as a source of well drilling fluids ("drilling pond"). Using a peristaltic pump on site, samples were pre-filtered by a Nucleopore 10 μm filter, concentrated onto Sterivex 0.22 μm filter units and immediately filled with buffer (40 mM EDTA, 750 mM sucrose, 50 mM Tris-HCl). Filters were shipped on dry ice and stored at -80°C. In addition, 100 ml of fluids were fixed on site in 3.7% formaldehyde for microscopy and 1 liter samples were collected in acid-washed containers for geochemical analyses. Samples were shipped on ice and stored at 4°C. Fluid pH and total salinity were measured on site using pH strips and a handheld refractometer, respectively.

For cell enumeration formaldehyde-preserved samples were treated with Syto 9 stain (Life Technologies), collected on 0.22 μm polycarbonate filters (Nucleopore), washed twice with PBS buffer, and mounted on glass slides for visualization by epifluorescent microscopy (Zeis Axioscope). Cell counts were extrapolated based on sample volume to calculate microbial densities.

Samples were prepared for elemental, sulfur, and chloride analysis by acidification with concentrated nitric acid while raw fluids were submitted for nitrate and ammonium analysis. Soluble element profiles were analyzed using inductively coupled plasma optical emission spectrometry (ICP-OES), while chloride, nitrate and ammonium were analyzed using distinct ICP methods (supplementary methods) on an iCAP 6300 (Thermo Fischer) at the Utah State University Analytical Laboratories (Hill Logan, UT).

DNA extraction

Nucleic acids were extracted from Sterivex filters after removal of DNA protective buffer using two methods: the MoBio UltraClean Soil DNA Isolation

Kit and a hot phenol chloroform method (Crump *et al.*, 2003; supplementary methods). When using the MoBio Kit, half filters and 500 µl of removed DNA-protective buffer were added to bead-beating tubes prior to initial vortexing, but otherwise followed manufacturer's instructions. Measured DNA concentrations were near detection limits (<20 ng/µl) or unable to be accurately measured due to sample discoloration.

Sequencing and analysis of 16S rRNA amplicons

Microbial diversity in well fluids and the drilling pond was characterized by 16S rRNA amplicon clone libraries and next-generation Illumina sequencing. Detailed protocols are provided in supplementary methods. In brief, clone libraries were constructed by ligating amplicons generated using bacterial small subunit 16S rRNA primers SSU_357_F and SSU_1100_R (Table S1) into TOPO TA pCR4 cloning vectors (Invitrogen). Cloned DNA was amplified, purified and submitted for Sanger sequencing at Genewiz (Cambridge, MA). Templates for Illumina sequencing consisted of 1) genomic DNA from wells and the drilling pond, 2) vector pCR4 containing a library of SSU_357_F/SSU_1100_R 16S rRNA amplicons, 3) a no DNA template control, and 4) an archived false positive (AFP) from a discarded PCR run (i.e. a contaminated PCR negative control). The AFP was sequenced to identify potential background laboratory contamination, which has recently been shown to be common and pervasive, especially in low-biomass samples (Salter *et al.*, 2014). A modified version of the Preheim *et al.* (2013) protocol was used to add barcodes and Illumina flow-cell sequencing adaptors, while utilizing a 16S rRNA primer set (PE_357_F/PE_806_R; Table S1) targeting the complete V3-V4 hypervariable regions of the 16S rRNA gene (Mizrahi-Man *et al.*, 2013; Sinclair *et al*, 2015). Multiplexed samples were sequenced as paired end reads (300 bp + 300 bp) at the MIT BioMicro Center via Illumina MiSeq (v.3 chemistry).

Demultiplexed FASTQ files containing 16S rRNA gene sequences prepared by Illumina MiSeq were processed using the UPARSE pipeline (Edgar, 2013) for quality filtering and trimming, merging of paired reads, chimera removal, operational taxonomic units (OTUs) clustering with greater than 97% sequence identity, and phylogenetic annotation (using the SILVA aligner, SINA (Quast et

al., 2013; http://www.arb-silva.de/aligner/) and RDP 16S rRNA database sequences (Wang et al., 2007; http://rdp.cme.msu.edu/). Sample OTUs were screened against OTUs from the AFP to identify and remove sequences likely to be laboratory contaminants. After QC in CLC Genomics Workbench 7, clone library 16S rRNA sequences were chimera-checked and annotated by the same pipeline as Illumina OTUs.

Qiime script *alpha_diversity.py* was used to calculate Chao1 statistics for individual sample libraries. Abundance-weighted, normalized UniFrac (Hamady *et al.*, 2009) distances were subjected to hierarchical clustering and principle coordinate analysis (PCoA) to compare OTU distributions between samples from different sites and prepared by alternative methods. Illumina 16S rRNA OTU sequences present at ≥1% abundance in Wells 3 or 10 were aligned using SINA with default settings were used to construct a bootstrapped (100X) neighbor-joined phylogenetic tree (CLC Genomics Workbench 7), and visualized in FigTree v1.4.2.

Metagenome sequencing and analysis

Genomic DNA samples from Wells 3 and 10 purified using the MoBio Kit (K) and Phenol Chloroform (P) protocols (Table S2) were submitted to the MIT BioMicro Center for metagenome preparation using the Nextera XT DNA Library Preparation Kit (Illumina) followed by sequencing on the Illumina MiSeq v.3 as paired end reads (300 bp + 300 bp). FASTQ files containing metagenome reads from Well 3 (P3, K3) and Well 10 (P10, K10) were demultiplexed using custom scripts. Sequences were trimmed, quality filtered (Quality limit = 0.05; length >50bp), pooled and *de novo* co-assembled into contigs using default kmer size (CLC Genomics Workbench 7). Filtered reads mapped back to assembled contigs (97% similarity over 80% read length) yielded scaffolds that were subsequently binned by analysis of nucleotide frequencies and single copy genes.

Scaffold tetranucleotide frequencies were calculated using a Perl script (Albertsen *et al.*, 2013) prior to principal component analyses (PCA) and plotting in R. Open reading frames (ORFs) were predicted using Metaprodigal, after which single copy genes were identified using a Hidden Markov Model list (Albertsen *et al.*, 2013). All single copy genes were subjected to Blastx searches

against the NCBI NR database, followed by taxon assignment using MEGAN 5.0. Taxon affiliations were overlaid on a PCA-tetranucleotide plot to guide extraction of scaffolds represented by individual microbial genomes ("crude bin"; Figure 5A). To ensure that no scaffolds have been misplaced into unrelated taxonomic bins, scaffolds within each extracted "crude bin" were fragmented *in silico* to 1000 bp, subjected to six-frame sequence translations compared to the NCBI NR database using Diamond (Buchfink *et al.*, 2015) and taxonomically assigned using MEGAN 5.0 (Figure S4). Scaffolds comprised of fragments assigned to unexpected taxa were removed and placed in taxonomic bins based on cluster analyses of the tetranucleotide frequencies and homology. Binning of several strains of the same species (*Sulfurospirillum*) is discussed in supplementary methods (Figure S5). All bins were verified for the presence of potential sequence contaminants by checking for the taxonomic affiliation of single copy genes, 16S and 23S rRNA genes, as well as the taxonomic distribution of ORFs based on best Blastp hits. Furthermore, tetranucleotide frequencies for contigs (*in silico* fragmented to 2000 bp) in each bin were subjected to Emergent Self Organizing Map analyses (default settings except K-Batch training algorithm in 80x120) to visually ascertain contig separation (Figure 5B). Each genomic bin was subsequently submitted to RAST and IMG for automated annotation.

2.3 Results

2.3.1 Fluid geochemistry and suspended cell numbers

Onsite periodic well testing by KinderMorgan (R. Gersch, 2012; personal communication) revealed well temperatures of 59.8-78.1°C, and produced fluid ratios of 2,021-5,418 liters H_2O per liter of liquid CO_2 (based on measured CO_2 gas volumes). ICP elemental analysis of formation fluids (Table 2) indicated total ionic concentrations per well from 0.32 g/l to 16.7 g/l, consistent with salinity measurements taken on site (Table 1). Hierarchical clustering of normalized ICP signal intensities by Spearman rank correlation (Figure S2) revealed two major clusters, where Cluster 1 corresponded to the most dilute samples with lowest H_2O/CO_2 ratios and enrichment in Mg, Fe, Ca, Al, Cr, and Mn. Cluster 2 contains more concentrated samples with enrichment in Na, B, and As. Well

Table 1. Sample well test data and on site fluid measurement summary.

FIELD	WELL	NAME	WELL TEST DATA (KinderMorgan CO_2)				MEASURED ON SITE	
			H_2O (L * 10^3)	CO_2 (L * 10^3)	(H_2O/CO_2) * 10^3	Temp (°C)	pH[b]	Salinity (ppt)
Hovenweep	1	HA-1	12.1	633.6	19.2	62.0	6.0	18
	2	HB-5	16.4	1875.5	8.7	59.8	6.0	15
	3	HC-2	3.5	1036.9	3.4	67.3	5.0	2
	4	HE-1	1.4	675.3	2.0	68.2	5.0	1
	5	HF-3	52.8	1391.6	37.9	74.5	6.0	15
Yellow Jacket	6	YA-3	1.7	487.3	3.4	78.1	6.0	10
	7	YB-4	42.9	791.2	54.2	74.8	6.0	25
	8	YC-4	4.9	1883.1	2.6	66.0	5.5	2
	9	YD-2	19.4	748.7	25.9	74.0	6.0	20
	10[a]	YF-4	12.9	1530.2	8.4	75.3	6.0	10
-	Pond		-	-	-	27.0	5.5	0

[a]Well 10 test results not available. CO_2, H_2O, and temperature data represent average values for three wells in same YF cluster as Well 10

[b]Following 15-hour period of degassing, fluid pH values may be less acidic than in situ due to reduced carbonic acid content

Table 2. ICP-OES analyte summary for major elements in McElmo Dome formation.

FIELD	WELL	PROFILE CLUSTER	Cl	Na	K	S	Mg	Fe	Ca	B	As	Al	Ba	Cr	Mn	Si	Sr	Zn	P	*NO_3^-	*NH_4^+	Σ(Analytes)
			mg/L																			g/L
Hovenweep	1	2	6180	4283	618	279	56.8	1.07	352	24.5	3.93	0.084	0.051	0.029	<	24.8	3.30	<	<	0.13	1.57	11.83
	2	2	6150	3343	498	240	55.6	0.19	182	20.5	3.17	0.003	0.035	0.007	<	21.2	3.04	0.002	<	<	1.25	10.52
	3	1	3630	262	33.8	13.4	6.66	0.48	31.0	1.24	0.01	<	0.005	0.041	0.023	1.39	0.13	0.005	<	<	0.11	3.98
	4	1	131	71.7	19.2	13.5	12.7	3.11	65.3	0.43	0.02	0.015	0.006	0.066	0.062	4.10	0.13	0.037	0.332	<	<	0.32
	5	2	6480	3710	613	312	77.4	0.09	213	22.5	3.12	0.014	0.035	0.004	<	33.1	3.15	<	<	<	1.85	11.47
Yellow Jacket	6	1	4780	1263	187	61.7	189	6.88	626	4.21	0.31	0.137	0.074	0.145	0.110	8.99	8.21	0.113	<	<	1.57	7.14
	7	2	8870	6125	893	522	94.2	0.59	163	22.3	4.83	0.063	0.031	0.011	<	26.5	6.61	<	<	<	3.07	16.73
	8	1	521	334	45.3	46.2	31.8	0.46	134	1.00	0.07	0.059	0.008	0.042	0.018	2.96	0.43	0.039	<	<	0.15	1.12
	9	2	8160	4622	711	330	81.5	0.41	332	26.6	5.56	0.068	0.054	0.015	<	31.1	7.55	0.002	<	<	1.62	14.31
	10	2	5370	2572	415	218	49.3	0.43	347	15.9	2.12	0.041	0.047	0.023	<	21.3	2.68	<	<	0.10	1.12	9.02
Pond		3	57.8	24.6	4.34	14.5	20.3	0.05	30.6	0.11	0.01	0.022	0.062	<	0.030	1.42	0.35	0.003	<	0.40	<	0.15
Detection Limit ('<' below detection)			3.0	0.001	0.005	0.01	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.03	0.10	0.05	-

*Raw fluids submitted for NO_3^- and NH_4^+ ICP-OES analysis; all other elements acidified with concentrated nitric acid to dissolve precipitated particulate prior to ICP-OES analysis

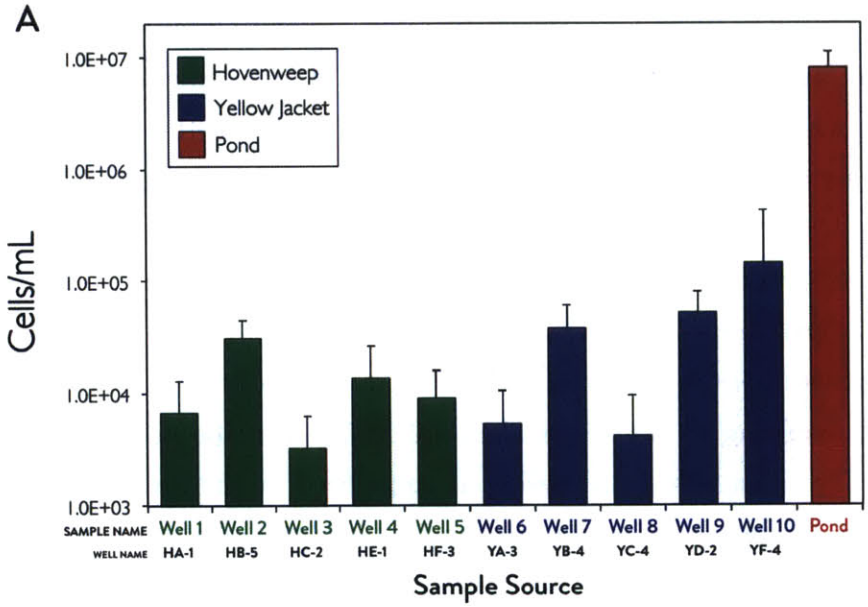


Figure 1. Cell concentrations/mL of ten McElmo Dome wells (five each from the Hovenweep and Yellow Jacket fields, respectively) and nearby drilling pond.

location (i.e. Yellow jacket vs. Hovenweep field, Figure S1) did not emerge as a significant driver of sample clustering (Figure S2). The drilling pond, which was most dilute overall (0.15 g/l total analytes) clustered separately from well samples.

Epifluorescent microscopy of formaldehyde-preserved samples revealed $3.2x10^3$ to $1.4x10^5$ microbial cells/ml of produced fluid and $8.0x10^6$ cells/ml of pond water (Figure 1). No trend in biomass density emerged by t-test with respect to spatial distribution (Hovenweep vs. Yellow Jacket fields; $p = 0.20$) or geochemical clustering (Cluster 1 vs. 2; $p = 0.18$).

2.3.2 Taxonomic diversity in formation fluids and pond water

16S rRNA gene amplicons targeting V3-V5 (clone libraries) and V3-V4 (Illumina) hypervariable regions were generated from genomic DNA prepared from Wells 3 and 10, the drilling pond, and an archived false positive (AFP). No amplification was observed from no-template controls or from genomic DNAs from Wells 1-2 and 4-9 likely due to PCR inhibition and/or insufficient template concentration. V3-V5 clone libraries generated 55 sequences from Wells 3 and 10 and the drilling pond, with an average length of 692 bp (Table S3). Blastn of cloned sequences identified *Sulfurospirillum deleyianum, Sulfurospirillum multivorans, Desulfovibrio marrakechensis, Desulfosporosinus orientis, Oscillibacter valericigenes, Desulfitibacter alkalitolerans, Acetobacterium carbinolicum,* and *Desulfitobacterium metallireducens* as members of the Well 3 assemblage, while two taxa were recovered from Well 10, *S. deleyianum* and *Rhizobium petrolearium* (Table S3). Illumina V3-V4 amplicon sequencing generated a total of 436,318 individual paired-end reads from Well 3 and 10 samples, 96,626 reads from the drilling pond, and 103,714 reads from the AFP, all of which were trimmed to a final length of 385 bp. Following OTU clustering at 97% sequence identity, sequence libraries were decontaminated by comparison to the AFP, resulting in the removal of 5.2-23.5% of reads per library (Table S2, Table S4). The most highly represented sequences in the archived false positive control sample were *Cloacibacterium, Acidovorax, Brevundimonas* and *Halomonas* and thus, likely represent reagent or laboratory contaminants. After decontamination, sequences from the two wells (Well 3 = 130,824; Well 10 = 62,334) formed 290 total OTUs, with Well 3 and Well 10 clustering into 187 and

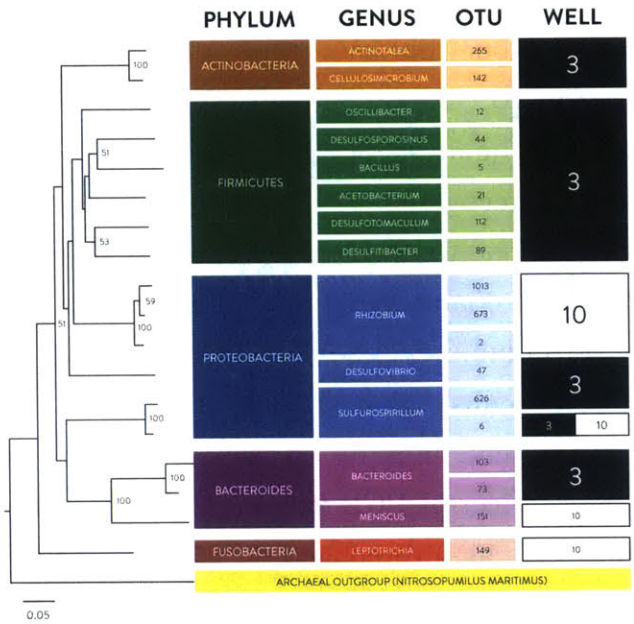


Figure 2. Phylogenetic tree of 16S rRNA Illumina OTUs from McElmo Dome at great than 1% abundance in Wells 3 and/or Well 10 displaying the phylum and genus level RDP/Silva annotations. SINA-aligned sequences were constructed into a neighbor-joined, bootstrapped (100) tree in CLC Genomics Workbench 7, and visualized FigTree. Tree rooted on outgroup Archaeal species *Nitrosopumilus maritmus*.

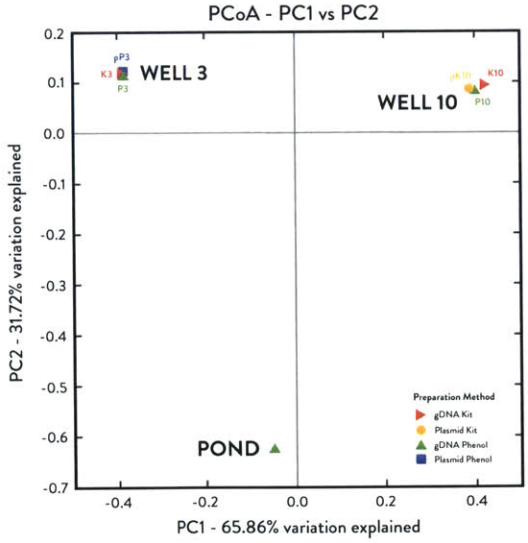


Figure 3. Beta diversity of 16S rRNA gene sequences from McElmo Dome well and drilling pond. The first two axes of the UniFrac Principal Coordinate Analysis (PCoA) explain 97.58% cumulative percent variation. Samples cluster according to origin rather than DNA preparation method.

199 OTUs respectively, and the drilling pond (39,851 reads) clustering into 236 OTUs (Table S2). All clone library genus annotations were represented in Illumina OTUs (Figure 4B; Table S3). PCoA analysis of UniFrac distances (Figure 3) and rarefaction analysis (Figure S3) indicated that diversity in well and pond samples was recovered to near-completion and reflected distinct microbial communities unique to each sample-type (i.e. Well 3, Well 10 or Pond) and independent of DNA preparation method (Figures 3 and 4).

The majority of Illumina OTUs were assigned taxa by SINA/RDP (99.5% to phylum and 68.7% to genus). 11 and 13 Bacterial phyla were recovered from Well 3 and 10, respectively (Figures 2 and 4A). No Archaeal phyla were recovered despite use of universal PCR primers. Most sequences from Well 3 and 10 samples were classified as Proteobacteria (57.7% and 90.1%, respectively), which were dominated by two OTUs corresponding to *Sulfurospirillum* (OTU6 Well 3 31.8%, Well 10 7.3%; OTU626 Well 3 20.8%, Well 10 0.1%), and two OTUs corresponding to *Rhizobium* in Well 10 only (OTU2 65.1%; OTU673 14.1%), and one OTU from *Desulfovibrio* in Well 3 (OTU47 4.1%) (Figures 2 and 5B). Both the *Sulfurospirillum* and *Rhizobium* sequences share the highest nucleotide identity with strains previously detected in oil fields or subsurface environments (Tan and Fogt, 2014; Zhang *et al.*, 2012). Firmicutes sequences recovered in high abundance in Well 3 (39.9%) correspond to several Clostridiales OTUs (*Oscillibacter*, OTU12: 17.0%), *Acetobacterium* (OTU21 10.8%), *Desulfosporosinus* (OTU44 6.0%) and *Desulfitibacter* (OTU89 1.1%)). In contrast, in Well 10 Firmicutes sequences were rare (1.0%) and low levels of the Actinobacteria (4.1%) and Fusobacteria (3.2%) were also recovered. Microbial diversity in the drilling pond was dominated by genera typical of freshwater surface environments (e.g. *Sediminibacterium*, *Limnohabitans*, *Polynucleobacter*, *hgcI_clade*, *Fluviicola*; Figure 4B) and enriched in Bacteroidetes and Actinobacteria. Overall, the pond community was highly distinct (Figures 3 and 4), suggesting drilling fluids are not a major source of microbial diversity in recovered formation fluids.

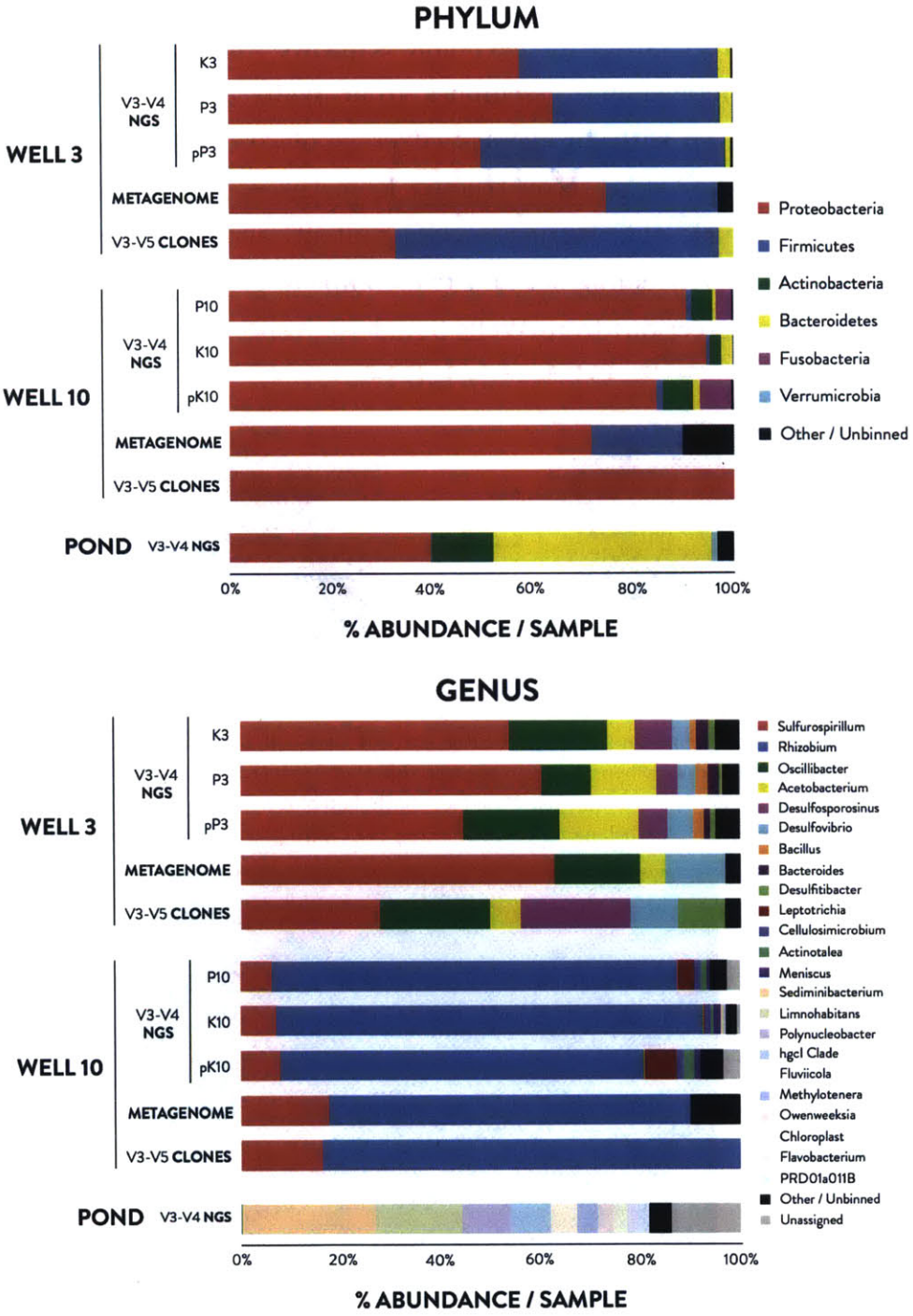


Figure 4. Taxonomic summary of the McElmo Dome microbial community constructed using RDP/Silva-annotated 16S rRNA sequencing of clone (CL) and Illumina (NGS) libraries, and Illumina binned metagenome (MG) read frequencies on the **A)** phylum and **B)** genus levels.

2.3.3 Diversity of genomes recovered from metagenome sequences

A total of 18.7 Gb of metagenomic sequence was generated from McElmo Dome samples. Metagenome assemblies from Wells 3 and 10 had an N50 of 14,942 and 36,674 bp with an average coverage of 20.7X and 35.4X, respectively (Table S2). Genomes were reconstructed from the sequence assembly based on binning by tetranucleotide frequency, GC content, and single-copy gene content (Figures 5A and 5B) yielding six and two complete genomes (>99% complete by single copy gene detection) from Wells 3 and 10, respectively (Table 3).

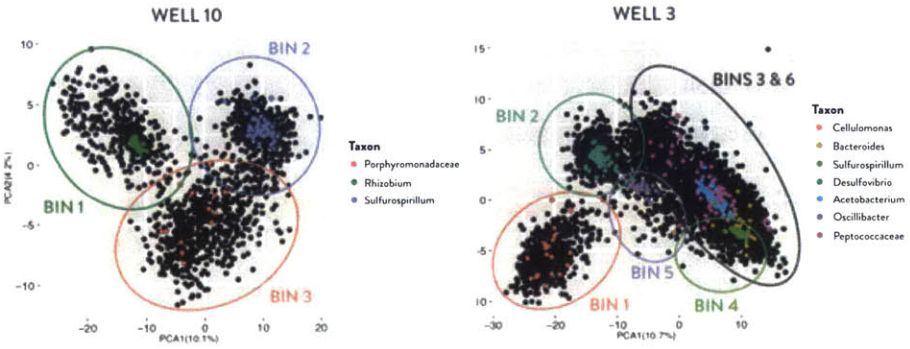


Figure 5A. Principal component analyses of the tetranucleotide frequencies of metagenomes from Wells 3 and 10. Each black dot represents a single contig/scaffold of 1-500 kbp. Contigs containing single copy gene(s) are overlaid with dots where colors represent different bacterial taxa. In circles are contigs (crude bin) extracted for further decontamination based on both homology and compositional-based approaches, after which contigs in crude bins belonging to unexpected taxa were placed into their correct bin.

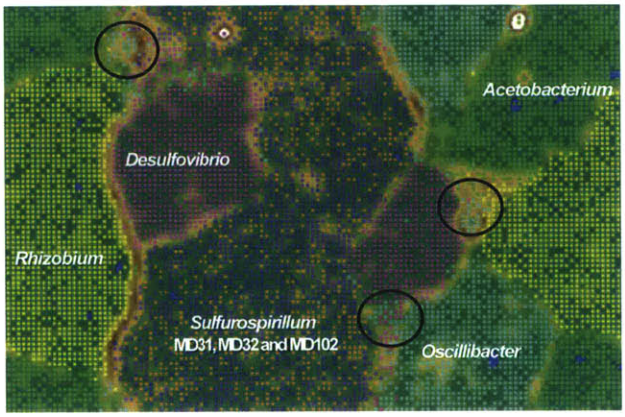


Figure 5B. Emergent Self Organizing Map (ESOM) of the tetranucleotide frequencies of contigs in binned genome. Each dot represents a 2000 bp fragments of scaffold/contig. Dots are colored according to genomic bins presented in Table 3. Region labeled as *Sulfurospirillum* contains two and one strain of *Sulfurospirillum* detected in Metagenome 3 and 10, respectively. In area where dots with different colors appeared to be "mixed" (circles in plot), the entire contig (ORFs) of each fragment was examined for their consistency in sequence homology (Blastp against NCBI NR-database). In all cases, these fragments were part of a long contig (of which other fragments were located in the correct region of the ESOM map). In some cases, fragments that are "mixed" were related to mobile genetic elements, which may have resulted in differences in GC content and therefore tetranucleotide frequencies. The three strains of *Sulfurospirillum* were subsequently separated using homology-based approaches (supplementary methods).

Table 3. Overview of genomes detected in the metagenome.

Well Sample	N50	Taxonomic affiliation based on 16S rRNA and single copy gene assignment	Binned genome size, Mbp (contig size range, bp) [Reference genome size][c]	Blastn 16S rRNA taxonomic affiliation of against NCBI NR database (full length unless stated)	Blastn comparison of binned 16S rRNA gene to Illumina generated OTU	# Contigs	# cds	Taxonomic distribution of ORF[b]	Genome Completeness[a]	Frequency in Metagenome (%)	Frequency of Corresponding Illumina OTU (%)
10	36,674	*Sulfurospirillum* MD102	2.6 (1,000-24,067) [2.3-3.2]	*Sulfurospirillum deleyianum* DSM 6946 (98%)	OTU 6 (100%)	590	2559	Campylobacterales (92%); Epsilonproteobacteria (95%)	100%	18	7
		Rhizobium MD101	6.8 (1,015-292,738) [4.9-7.5]	*Rhizobium petrolearium* strain SL-1 (100%)	OTU 2 (100%)	164	6755	Rhizobiales (88%); Alphaproteobacteria (93%)	100%	72	65
3	14,942	*Desulfovibrio* MD33	5.0 (1,001-68,886) [3.2-5.25]	*Desulfovibrio marrakechensis* strain EMSSDQ4 (98%)	OTU47 (99.2%)	767	4545	Desulfovibrionales (92%)	100%	12	4
		Acetobacterium MD34	4.0 (1,011-114,001) [4.04 – 4.05]	*Acetobacterium carbinolicum* (100%)	OTU21 (100%)	315	3847	Eubacteriaceae (73%); Clostridiales (86%); Firmicutes (94%)	100%	5	11
		Sulfurospirillum MD31[d]	2.6 N/A [2.3-3.2]	*Sulfurospirillum deleyianum* DSM 6946 (97%); Sulfurospirillum sp. M10 (100%); 523 bp fragment	Binned fragment does not overlap with Illumina OTU; OTU 6[e]	34	2776	Campylobacterales (92%); Epsilonproteobacteria (95%)	100%	36	32
		Sulfurospirillum MD32	3.2 N/A [2.3-3.2]	*Sulfurospirillum* sp. JPD-1 (99%); 676 bp fragment	Binned fragment does not overlap with Illumina OTU; OTU 626[e]	29	3095	Campylobacterales (88%); Epsilonproteobacteria (92%)	100%	27	21
		Oscillibacter MD34	3.2 (1,013-160,132) [3.1-4.4]	*Oscillibacter valericigenes* strain Sjm18-20 (96%)	OTU 12 (100%)	130	3278	Oscillospiraceae (42%); Clostridiales (66%); Firmicutes (88%)	100%	17	17
		Desulfosporosinus	Not binned	*Peptococcaceae* *Desulfosporosinus orientis* (OTU affiliation)	No fragment available; OTU 44[e]	N/A	N/A	*Peptococcaceae*	N/A	N/A	6

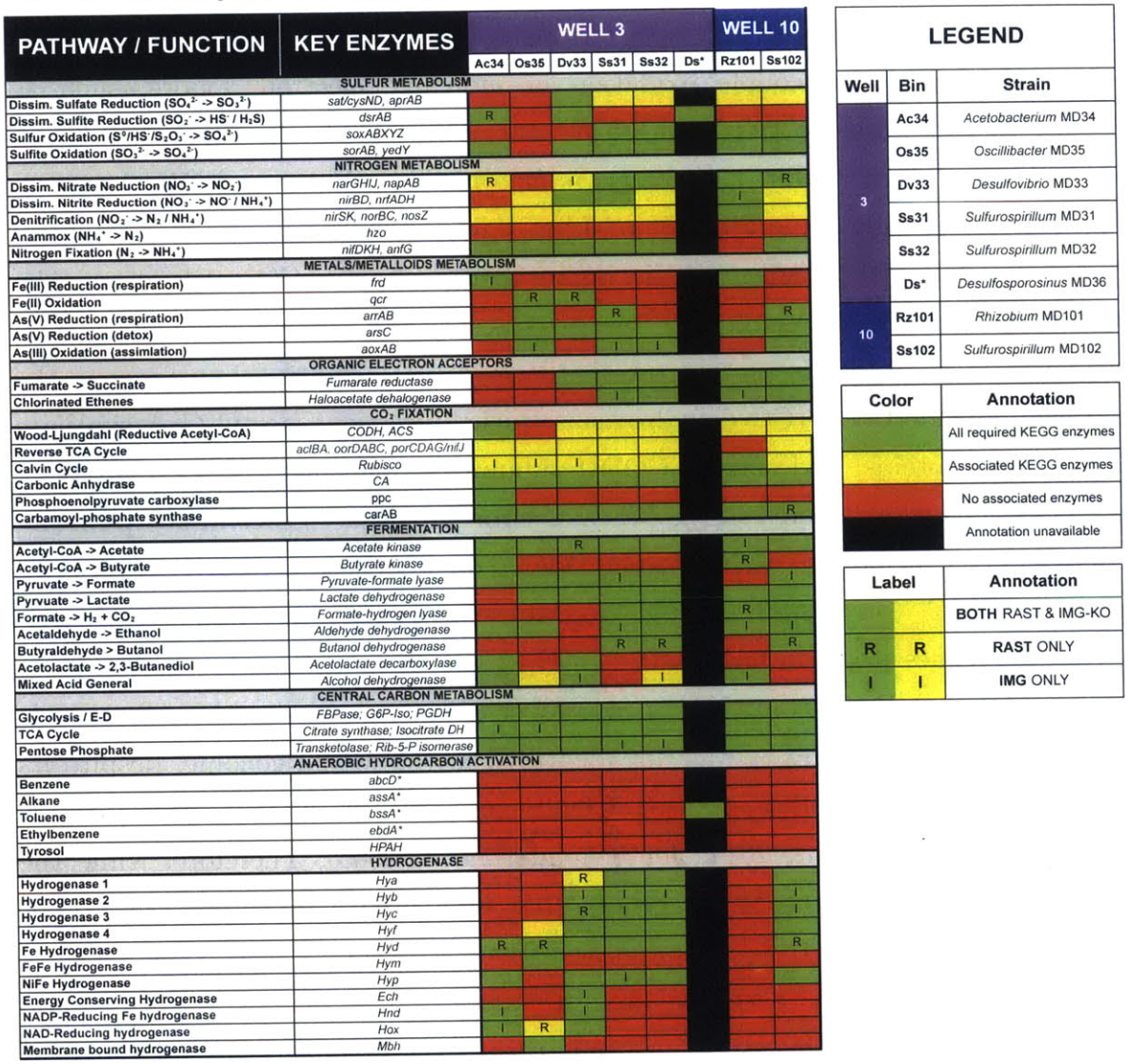
Those genomic bins representing the dominant bacterial groups as a majority of metagenomics reads (97% and 90% for Wells 3 and 10, respectively) were mapped back onto reconstructed genomes. Genomic bins from Well 3 were assigned as *Sulfurospirillum* MD31, *Sulfurospirillum* MD32, *Desulfovibrio* MD33, *Acetobacterium* MD34, *Oscillibacter* MD35, while those from Well 10 were *Rhizobium* MD101 and *Sulfurospirillum* MD102, based on best Blastn hit of the 16S rRNA gene and/or single copy genes. For the remaining ~3% of the metagenomics reads obtained from Well 3, contigs were assigned to *Desulfosporosinus* (>95% complete), *Bacteroides* (30% complete), and *Cellulomonas* (24% complete) based on taxonomically-informative genes. These genomes were unable to be binned due to insufficient contig length (<2kb), had low coverage (<20X; Figure S4) and were clustered with other Firmicutes-affiliated contigs (Figure S4). The detection of the *Desulfosporosinus* population in the metagenome was supported by phylogenetic placement of genes recovered from unbinned sequences (Figure S6, *dsr*AB; Figure S7, *bss*A). For the ~10% of metagenomics reads in Well 10 that were not associated with *Rhizobium* and *Sulfurospirillum*, these were assigned to Porphyromonadaceae (19% complete) and other unknown microbes (<5% complete). Members of the Porphyromonadaceae family have previously been detected in deep subsurface formations (Rastogi et al., 2010). Emergent Self Organizing Map (ESOM) showed clear separation of contigs between complete genomic bins (Figure 5B). The taxonomic affiliation and sequence distribution of binned genomes were consistent with abundance profiles of Illumina 16S rRNA OTUs and cloned 16S rRNA genes (Table 3; Figure 4).

2.3.4 Functional capacity of microbial genomes

Sequences recruited to the three *Sulfurospirillum* binned genomes made up the major portion of the Well 3 metagenome (63%) and was the second most abundant bin of the Well 10 metagenome (18%). *Sulfurospirillum* genomes MD31 and MD102 were most closely related to *S. deleyanium* (98% 16S ID), while the third (MD32) more closely resembled *S. multivorans* (98% 16S ID) (Tables 3 and S3). All *Sulfurospirillum* genomes harbor predicted genes for chemotactic motility (*che* and *fla* genes), nitrogen fixation (*nif*DKH) and access to multiple electron

acceptors for respiration including nitrate (*nap*AB), arsenate (*arr*AB), fumarate (fumarate reductase) and chlorinated ethenes (haloacetate dehalogenase) (Table 4), consistent with activities observed in closely related strains (Goris *et al.*, 2014; John *et al.*, 2009; Magnuson *et al.*, 1998). *Sulfurospirillum* genomes predict organic carbon transporters for lactate, formate, gluconate, peptides, and amino acids, indicating potential uptake for heterotrophic carbon metabolism via

Table 4. Metabolic potential of binned genomes in Wells 3 and 10.



*Desulfosporosinus genes were assigned function based on GC content and sequence coverage using HMM based on best Blastp hit. When a key gene or subunit is not detected in a genomic bin based on RAST annotation, tBlastn and HMM were used to screen the entire metagenome, followed by Blastp searches (for taxonomic affiliation) against the NCBI database to verify the absence of such function in the metagenome.

glycolysis, the TCA cycle, and several fermentation pathways (Tables 4 and S5C). Mixotrophic growth fueled by inorganic electron donors hydrogen and sulfur is documented in cultured strains of *Sulfurospirillum* (Finster *et al.*, 1997) and is indicated in the three McElmo Dome *Sulfurospirillum* genomes (MD31, MD32 and MD102) by annotation of *sox*ABXYZ, *sor*AB and *yed*Y genes for the sulfur oxidation pathway, (which can be coupled to nitrate reduction for anaerobic energy conservation (Eisenmann *et al.*, 1995)), and *hyp*ABCDEF (for NiFe hydrogenase) and *hyf*ABCEF (for Hydrogenase 4) which can catalyze hydrogen oxidation.

The ability to couple oxidation of inorganic electron donors, such as hydrogen and sulfur, to CO_2 fixation for chemolithoautotrophic growth has not been demonstrated within the genus *Sulfurospirillum* (Kelly and Wood, 2006). However a partially annotated Wood-Ljungdahl pathway within the genome of *Sulfurospirillum* MD32 suggests the potential capacity for CO_2 utilization. The Wood-Ljungdahl pathway requires two enzymes, carbon monoxide dehydrogenase (CODH) (AcsA), and an acetyl-CoA synthase (AcsB). Additional enzymes in the pathway include CooC (CO dehydrogenase nickel insertion), acetyl-CoA synthase iron-sulfur proteins, and several CODH subunit components (e.g. CooALHUX). An *acs*B gene was detected in the MD32 genome adjacent to several key pathway components, including carbon monoxide dehydrogenase (CODH) accessory protein, CooC (on RAST contig 10/IMG contig 128), as well as CODH-associated genes *coo*A, *coo*L, *coo*X, *coo*U, *coo*H and *coo*F on a single contig (Rast contig 6/IMG contig 125). However, the gene for carbon monoxide dehydrogenase (*coo*S/*acs*A) was not identified in the MD32 genome, despite its previous detection (and associated CODH activity) in other *Sulfurospirillum* strains (Tan and Foght, 2014; Jensen and Finster, 2005). Therefore, while the *acs*B gene present in the MD32 genome has been used as a marker for Wood-Ljungdahl-mediated CO_2 fixation and acetogenesis (Gagan, 2010), the absence of *acs*A suggests that a partially or alternatively functional Wood-Ljungdahl pathway is present in this strain. All *Sulfurospirillum* genomes were annotated as containing genes for carbonic anhydrase that facilitates conversion of CO_2 to bicarbonate (Smith and Ferry, 2000) and the bicarbonate incorporating enzyme carbamoyl-phosphate synthase for pyrimidine and arginine biosynthesis (Arioli *et al.*, 2009). Taken together the McElmo Dome *Sulfurospirillum* genomes are

predicted to be capable of growth under conditions of low nitrogen, using bicarbonate, potentially CO_2, and fermentative substrates as carbon and/or energy sources, with access to diverse substrates for anaerobic respiration including arsenate and nitrate, both of which are detected in formation fluids.

In addition to *Sulfurospirillum* MD31 and MD32, Metagenome 3 yielded three additional putatively complete genome bins (*Acetobacterium* MD34, *Desulfovibrio* MD33, and *Oscillibacter* MD35). Notably, these genomes are also associated with predicted genes for N_2 fixation (*nif*), chemotactic motility (*che* and *fla*), carbonic anhydrases (CA) for CO_2 conversion to bicarbonate, and bicarbonate-utilizing enzymes such as carbamoyl-phosphate synthase, which allows entry of inorganic carbon into central metabolism (Tables 4 and S5B). In addition, all genomes harbored predicted transporters or pathways for uptake of amino acids and peptides suggesting the ability to recycle organic materials released by cell lysis or exudation (Table S5C).

Chemolithoautotrophy within the Well 3 community is supported by annotation of a complete Wood-Ljungdahl CO_2 fixation pathway in the *Acetobacterium* MD34 genome, together with genes associated with FeFe and NiFe hydrogenases for H_2 oxidation (Tables 4 and S5B) consistent with the growth physiology of close relative, homoacetogen *A. carbinolicum* (100% 16S rRNA) (Eichler and Schink, 1984). In addition, annotation of a phosphoenolpyruvate carboxylase gene, which catalyzes the irreversible addition of bicarbonate to phosphoenolpyruvate and is hypothesized to promote tolerance of high pCO_2 conditions (Santillan *et al.*, 2015, Arioli *et al.*, 2009), indicates a possible entry point for inorganic carbon to central metabolism. *Acetobacterium* MD34 does not appear to be an obligate autotroph as the capacity for heterotrophy is indicated by annotation of transporters and permeases for uptake of lactate, sugar monomers, and ethanol with energy conservation by mixed acid fermentation (Table 4). In addition, annotation of a respiratory sulfite reduction pathway and a non-energy generating As(V) reduction pathway associated with detoxification may represent adaptations to high As and S levels in formation fluids (Table 2).

The genome of *Desulfovibrio* MD33 is most closely related to the sulfate reducing bacteria *D. marrakechensis* strain EMSSDQ4 (98% 16S ID) and contains a complete pathway for respiration by sulfate reduction and nitrite

reduction (*nrf*AH) as well as an incomplete denitrification pathway (*nir*S, *nor*B; Tables 4 and S5A). Heterotrophy in *Desulfovibrio* MD33 is indicated by annotation of transporters for lactate, glycerol, fructose, and other sugar monomers (Table S5C), which can be metabolized via central carbon metabolism (glycolysis, the TCA cycle and several fermentation pathways (Table 4). Annotations further suggest *Desulfovibrio* MD33 may grow mixotrophically by accessing inorganic electron donors Fe(II) (Fe(II)-cytochrome c reductase) and molecular hydrogen (FeFe and NiFe hydrogenases) as previously demonstrated for cultured *Desulfovibrio* strains (Timóteo *et al.*, 2012; Romao *et al.*, 1997; Roseboom *et al.*, 2006). A partially annotated Wood-Ljungdahl pathway (presence of carbon-monoxide dehydrogenase catalytic subunit CooS/AcsA clustered with several accessory proteins including CooF, CooC, and CooA on Rast contigs 75/819 and IMG contig 128)) may indicate additional capacity for C1 metabolism.

Annotations from the Well 3 binned genome of *Oscillibacter* MD35 (closest relative *O. valericigenes* Sjm18-20 with 96% 16S rRNA; Katano *et al.*, 2012) are indicative of a heterotroph able to uptake and metabolize lactate, glycerol, glutamate, and several sugars. Energy conservation may occur by fermentation or by anaerobic respiration by utilizing As(V) as an electron acceptor or through incomplete denitrification (Tables 4 and S5A). Predicted fermentation pathways generate formate, lactate and ethanol end products (Table 4), which may contribute to an anaerobic food web. A predicted Fe(II)-cytochrome c reductase suggests *Oscillibacter* MD35 may also be able to access Fe(II) as an electron donor for mixotrophic growth.

The presence of a sulfate-reducing *Desulfosporosinus* population in Well 3 is supported by both 16S rRNA data and detection of dissimilatory sulfite reductase alpha and beta subunits *dsr*AB that are phylogenetically related to *Desulfosporosinus* spp. (Figure S6). In addition, a single copy of the *bss*A gene encoding the alpha-subunit of benzylsuccinate synthase (BssA) affiliated with *Desulfosporosinus* was also recovered from the metagenome (Figure S7). The BssA enzyme is involved in the anaerobic degradation of monoaromatic hydrocarbons via fumarate addition, and *Desulfosporosinus* has been routinely implicated in subsurface degradation of toluene coupled to sulfate reduction (Lee *et al.*, 2009; Liu *et al.*, 2004). As McElmo Dome overlies minor hydrocarbon

deposits (Rabinowitz and Janowiak, 2005) it is possible that these reduced compounds represent accessible sources of carbon and energy.

The Well 10 metagenome was dominated (72% of sequences) by the *Rhizobium* MD101 genome, which displays 100% 16S rRNA homology to an oil-contaminated soil isolate, *Rhizobium petrolearium* strain SL-1. Annotation of a sulfur oxidation pathway (soxABXYZ) and a complete Calvin Cycle in genome MD101 indicates the potential for autotrophic growth and chemosynthesis in this strain, as has been previously described for *Rhizobium* isolates from calcareous desert soil (El-Tarabily et al, 2006). The RuBisCO gene detected in the *Rhizobium* MD101 phylogenetically clusters with type IC and ID forms (Figure S8), which are typically associated with mixotrophs, including members of the Rhizobiales order (Badger and Bek, 2007). Additional genes associated with the Wood-Ljungdahl pathway (CO dehydrogenase) and carbonic anhydrase may also serve as enzymatic mechanisms for utilization of CO_2 and other C1 compounds including carbon monoxide as an electron donor (*cox*LMS; Cunliffe, 2011). The genome does not appear to represent an obligate autotroph as annotated transporters for formate, glucose, xylose, fructose, lactose, peptides and amino acids (Table S5C) suggest the ability to take up organic compounds from the environment. Energy from heterotrophic or autotrophic metabolism may be conserved by predicted pathways for fermentation or anaerobic respiration, utilizing iron, nitrate (*nap*AB), nitrite (*nir*BD) or fumarate as electron acceptors (Table S5A), while also containing a full denitrification pathway (*nir*K, *nor*BC, *nos*Z; Table 4). As with other McElmo Dome genomes, this genome predicts chemotactic motility (*che* and *fla* genes), however *Rhizobium* MD101 notably lacks genes for nitrogen fixation ubiquitous among other McElmo Dome genomes.

2.4 Discussion

The first insights into the taxonomic and genomic content within a natural $scCO_2$ reservoir biosphere have been revealed through 16S rRNA and metagenome sequence analysis. Despite known biases associated with amplification-based 16S rRNA gene surveys, agreement in taxonomic distribution was observed for the most highly represented 16S rRNA gene amplicons and

sequences from amplification-independent metagenomes. Based on 16S rRNA gene and metagenome sequences the microbial assemblage in formation fluids appear to be dominated by two microbial taxa from Well 10 (*Rhizobium petrolearium* MD101 and *Sulfurospirillum deleyianum* MD102) and six taxa from Well 3 *Sulfurospirillum deleyianum* MD31, *Sulfurospirillum multivorans* MD32, *Desulfovibrio marrakechensis* MD33, *Acetobacterium carbinolicum* MD34, *Oscillibacter valericigenes* MD35, and *Desulfosporosinus orientis*. The abundance of these taxa and their previous detection in deep subsurface anoxic environments (Marshall *et al.*, 2013; LaBelle *et al.*, 2014; Itävaara *et al.*, 2011; Suess *et al.*, 2005; Engelhardt *et al.*, 2013) suggest a biological presence in the $scCO_2$-exposed formation fluids.

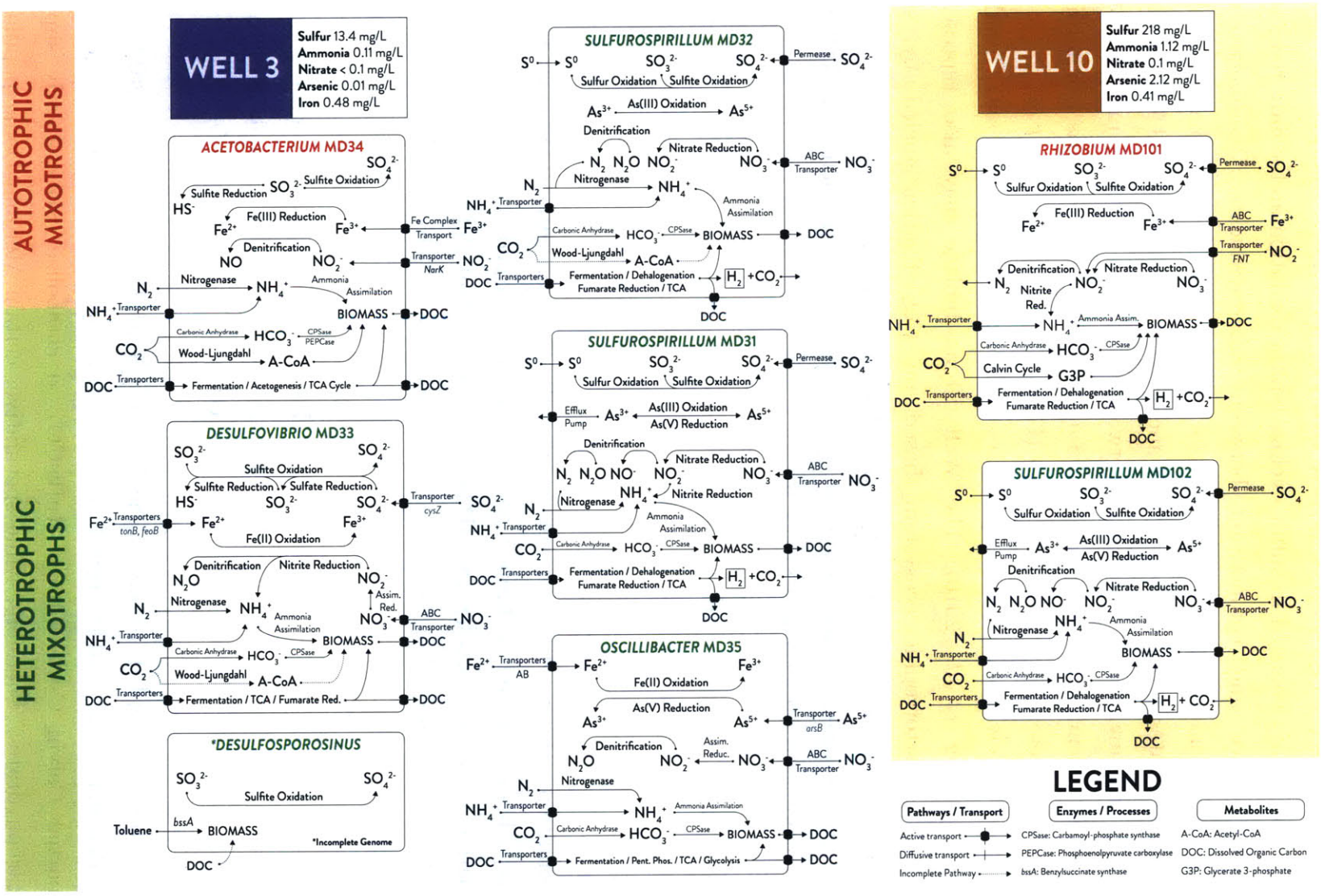
Metabolic annotations of recovered genomes suggest a potential microbial food web based on remineralization of organic carbon or primary production via chemolithoautotrophy where inorganic electron donors may include reduced sulfur (*sox*ABXYZ genes in *Rhizobium* MD101 and *Sulfurospirillum* MD31, MD32 and MD102), hydrogen (FeFe and NiFe hydrogenases in *Desulfovibrio* MD33, NiFe hydrogenase in all *Sulfurospirillum* strains) or iron (Fe(II)-cytochrome c reductase in *Desulfovibrio* MD33 and *Oscillibacter* MD35). Annotation of pathways for anaerobic respiration using diverse electron acceptors (e.g. nitrate, sulfate, As(V)) or fermentation indicates mechanisms for energy conservation. Challenges associated with low dissolved nitrogen appear offset by the near ubiquitous capacity for nitrogen fixation (*nif*DHK in all genomes except *Rhizobium* MD101).

Ecosystem interaction among members of McElmo Dome assemblages is suggested by their metabolic potential and recovery of similar microbial assemblages from multiple locations around the world. The Well 3 assemblage resembles a chemolithoautotrophic ecosystem in anoxic subglacial volcanic-fed lakes in Iceland that is comprised of *Sulfurospirillum, Acetobacterium,* and *Desulfosporosinus* (Gaidos *et al.*, 2008, Marteinsson *et al.*, 2013). A similar assemblage of *Sulfurospirillum, Acetobacterium,* and *Desulfovibrio* was enriched from brewery wastewater in the USA as members of an electrosynthetic microbiome capable of CO_2 fixation and H_2 or acetate production using electrodes as an electron donor (Marshall *et al.*, 2013; LaBelle *et al.*, 2014). In both cases, *Acetobacterium* is proposed to play a major role in fixing CO_2 into biological

chemicals such as formate or acetate that can be accessed by other members of the community engaged in sulfur cycling (sulfur oxidation and sulfate reduction). Co-occurrence and competition between sulfate-reducing *Desulfovibrio* and nitrate-reducing *Sulfurospirillum* spp. for organic electron donors has been shown to reduce rates of sulfide production and associated souring in oil fields (Hubert and Voordouw, 2007). Similarly, the two taxa that make up the majority of the Well 10 assemblage, *Rhizobium* and *Sulfurospirillum* are noted as ubiquitous in petroleum reservoir formation waters (Gao *et al.*, 2016; Zhang *et al.*, 2012) and are also widely associated with high arsenic environments (Lloyd and Oremland, 2006; Chang *et al.*, 2010). The Well 10 assemblage may also represent a chemolithoautotrophic ecosystem based on CO_2 fixation fueled by sulfur oxidation by *Rhizobium* MD101 and carbon remineralization by *Sulfurospirillum* MD102. A schematic overview of ecosystem metabolism predicted by Well 3 and 10 genomes is presented in Figure 6.

We hypothesized that autotrophic bacteria would be present in McElmo Dome and able to fix the abundantly available CO_2 for carbon cycling (98.2% total gas content; Allis *et al.*, 2001). Models suggest CO_2-consuming metabolic reactions (e.g. acetogenesis) would occur with increased rate and favorability under conditions associated with supercritical CO_2 relative to atmospheric conditions (Kirk, 2011; Onstott, 2005; West *et al.*, 2011). Support for this hypothesis is found in binned genomes, which revealed complete CO_2 fixation pathways in *Acetobacterium* MD34 (Wood-Ljungdahl Pathway) and *Rhizobium* MD101 (Calvin Cycle). Several partially annotated CO_2-fixation Wood-Ljungdahl pathways (Tables 4 and S5) including annotation of CO dehydrogenase (*acs*A) in *Desulfovibrio* MD33 and acetyl-CoA synthase (*acs*B) in *Sulfurospirillum* MD32 (in conjunction with contig-clustered CODH accessory proteins in both genomes) may reflect additional CO_2 fixation capacity with sequencing/annotation gaps, or may represent distinct metabolic functions. For example *Dehalococcoides mccartyi* leveraged an incomplete Wood-Ljungdahl pathway for one-carbon metabolism coupled to amino acid biosynthesis (Zhuang *et al.*, 2014) while Nakayama *et al.* (2014) argued that endosymbiont *Epithemia turgida* utilizes a partial Calvin Cycle (absent RuBisCO) to catabolize extracellularly supplied carbohydrates. Therefore, while speculative, certain McElmo Dome populations may have repurposed autotrophy-affiliated enzymes for alternative metabolic

Figure 6. Model of potential metabolic interactions and dependencies among populations in the deep biosphere Well 3 and Well 10 scCO$_2$ systems.



activities. Notably, metagenome sequences from the Crystal Geyser high pCO_2 system also contained signatures of CO_2-fixation via the Calvin Cycle (Emerson *et al.*, 2015), suggesting a similar microbial capability to participate in carbon cycling within high pCO_2 systems.

Sulfur is dissolved in fluids from both Well 3 (13.4 mg/l) and Well 10 (218 mg/l) although the oxidation state *in situ* is not known. Genomes from *Sulfurospirillum* MD31, MD32, MD102, and *Rhizobium* MD101 carry the full suite of *sox* genes for sulfur or sulfide oxidation while sulfite has the potential to be oxidized via sulfite oxidase/hydrogenase (Table 4) by all taxa except *Oscillibacter* MD35. Because Well 3 populations (represented by *Desulfovibrio* MD33, *Acetobacterium* MD34, and *Desulfosporosinus*) may also access oxidized sulfur compounds as electron acceptors for respiration through dissimilatory sulfate and sulfite reduction, the order of magnitude lower sulfur concentration in Well 3 relative to Well 10 may be due to the biogenic conversion of oxidized species to sulfide and subsequent loss of H_2S during fluid degassing or abiotic reaction and precipitation. We therefore hypothesize that sulfur-bearing redox substrates provide energy to fuel the persistence of the carbonated biosphere, and may also strongly impact local geochemical conditions.

The nitrogen cycle is accessible to microbial communities at McElmo Dome as dissolved N_2 gas (1.6% of total gas content) and ionic species, including nitrate (Well 3: b.d.; Well 10: 0.10 mg/l) and ammonium (Well 3: 0.11 mg/l; Well 10: 1.12 mg/l) (Table 2). The genomic capacity for N_2 fixation is nearly ubiquitous, as the full suite of nitrogenase genes (*nif*DHK) are present in all binned genomes except *Rhizobium* MD101, increasing feasibility for diverse growth under low N_2 conditions. In Well 10, where nitrate and ammonia are ~10X more concentrated than Well 3, sequences from a non-nitrogen-fixer (*Rhizobium* MD101) predict nitrite and ammonia transporters and a full pathway for nitrate/nitrite reduction and denitrification, suggesting that the bioavailability of dissolved nitrogen species reduces the necessity for N_2 fixation and promotes the ability to use nitrate as a terminal electron acceptor. In Well 3, where nitrate is below detection, the genes for dissimilatory nitrate reduction are found only in *Sulfurospirillum* MD31 and MD32. However, the detection of nitrite transporters and incomplete denitrification pathways (Table 4) suggest the potential for intermediate nitrogenous redox species to accumulate *in situ* as

syntrophic electron acceptors (Carlson and Ingraham, 1983) similar to activities observed in geothermal springs (Dodsworth *et al.*, 2011) and coastal sediments (Fernandes *et al.*, 2010).

Many energy-generating processes may be coupled to redox reactions involving iron Fe(II/III) or arsenic As(III/V). Thermodynamic models of saline aquifers suggest that $scCO_2$-induced acidity increases available energy for Fe(III)-reduction (Kirk, 2011). Fe(III) reduction is potentially exploited in Well 10 (0.41 mg/l) by *Rhizobium* MD101, which contains ferric-chelate reductase and Fe(III) transporters. Arsenic present in Well 10 (2.12 mg/l) may enable dissimilatory arsenate reduction by *Sulfurospirillum* MD102, a trait commonly observed in the *Sulfurospirillum* genus (Stoltz et al., 1999). Though the lower concentration of arsenic in Well 3 (0.01 mg/l) may limit its involvement in redox coupling, genes for dissimilatory arsenate reductase (*arr*AB) detected in *Oscillibacter* MD35 and *Sulfurospirillum* MD31 raise the possibility that arsenic reactions causing a change in redox state may reduce its solubility and thus dissolved concentration. Arsenate reductase detoxification genes (*ars*C) and efflux pumps (*asr*AB, ACR3) detected in all binned genomes suggests the element's presence in formation fluids requires the capacity for redox processing and export.

In the absence of terminal electron acceptors or redox couples, the utilization of organic carbon from fermentation end products (e.g. butanol, succinate, lactate, etc.) in anoxic systems supports the growth of secondary fermenters and acetogens (Table 4), which can lead to the production of electron carriers typical of anaerobic aqueous environments: acetate and formate (Hattori *et al.*, 2001; De Bok *et al.*, 2004; Kaden *et al.*, 2002). These simple carbon compounds may in turn be metabolized to produce H_2, which can serve as an essential reductant for CO_2 fixers (Morris *et al.*, 2013). Annotation of acetate kinase (catalyzing acetate formation from acetyl-CoA) in all genomes and pyruvate-formate lyase (catalyzing formate production from pyruvate) in all Well 3 genomes suggest that both acetate and formate may be generated by central carbon metabolism. The presence of the TCA cycle in several taxa (Table 4) indicates the capacity to consume acetate (as acetyl-CoA), while the presence of formate-hydrogen lyase genes in *Rhizobium* MD10 and all *Sulfurospirillum* genomes suggests that formate may be oxidized to H_2 and CO_2. FeFe hydrogenase in *Desulfovibrio* MD33 may also aid in H_2 production. Since

bioavailable H_2 can serve as an electron donor for numerous respiratory and CO_2 fixation pathways including denitrification, sulfate reduction, and the Wood-Ljungdahl Pathway, its rapid utilization may explain the very low *in situ* H_2 partial pressures observed at McElmo Dome (Morris *et al.*, 2013; Nedwell and Banat, 1981).

Low levels of accumulated hydrocarbons (Rabinowitz and Janowiak, 2005) may also provide reduced carbon substrates to microbial communities. *Desulfosporosinus* harbors genes for anaerobic activation of monoaromatic hydrocarbons (benzylsuccinate synthase; Figure S7), making them metabolically available as electron donors. Organic carbon transporters annotated across all binned genomes indicate the capacity for uptake of substrates for heterotrophic growth including sugars, amino acids, organic acids, and alcohols in McElmo Dome well fluids.

Despite the lethal effects of $scCO_2$ exposure and temperatures >65°C, exploitable niches for microbial growth may have emerged at McElmo Dome, although rates of cell division and nutrient cycling are expected to be diminished relative to surface environments (Lovley and Chapelle, 1995). The capacity for persistence and growth at McElmo Dome may partially be due to the pH buffering capacity of dissolved carbonate from the dolomitic Leadville Formation. In fact, the two wells that yielded amplifiable DNA (Wells 3 and 10) have among the highest reported CO_2 to water ratios at McElmo Dome (Table 1), suggesting that high CO_2 content does not preclude recovery of microbial biomass. Previous studies (Oppermann *et al.*, 2010; Morozova *et al.*, 2011) reveal the resilience of sulfate-reducing bacteria to high pCO_2 exposure and additional studies (Mu *et al.*, 2014; Emerson *et al.*, 2015) report *Proteobacteria* and *Firmicutes* are the dominant observed phyla under high pCO_2 conditions following GCS and at the CO_2-venting Crystal Geyser formation, respectively, consistent with high-level taxonomic identification at McElmo Dome. Metagenomes from Crystal Geyser fluids revealed genes for anaerobic respiration, nitrogen fixation, CO_2 fixation, and fermentation (Emerson *et al.*, 2015), which together with our results from a $scCO_2$ dominated system, suggest that these metabolic strategies support survival of a biosphere under high pCO_2 conditions.

2.5 Conclusions

Genome content of the bacterial community in McElmo Dome formation fluids suggest that a microbial ecosystem fueled by inorganic electron donors and nutrients may catalyze carbon cycling through CO_2 fixation and remineralization of organic matter. Therefore, future studies attempting to model the behavior of $scCO_2$ injected for GCS should take into account the potential geochemical and physical effects of the deep biosphere. Though the scope of this study was initially limited to a dolomitic carbonate formation, the results have provided a reference by which to compare communities isolated from other geologic contexts in future work, including $scCO_2$ and near-critical CO_2 reservoirs in unbuffered sandstone formations (e.g. Bravo Dome, CO) and from surface venting of near critical CO_2 reservoirs (e.g. Crystal Geyser). These comparisons will help clarify the extent to which geochemical context is a selective driver for diversity or whether certain taxa are specifically adapted to survive in high pCO_2 conditions.

2.6 Supplementary figures

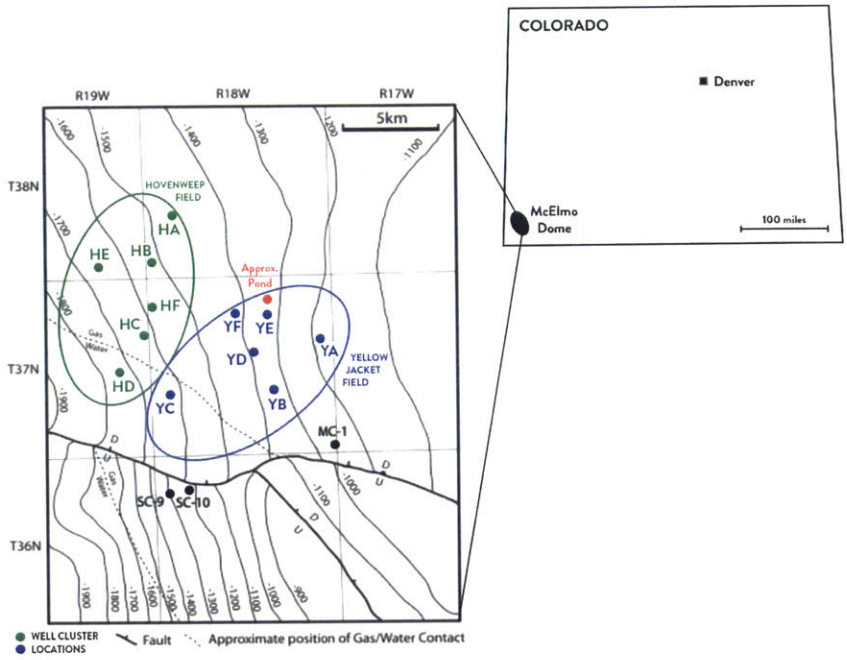


Figure S1. At right, location of McElmo Dome system within the Colorado Plateau in SW Colorado. Inset, approximate well cluster fluid sampling locations in the Hovenweep (green) and Yellow Jacket (blue) fields and drilling pond (red). Adapted from Gilfillan et al., 2008.

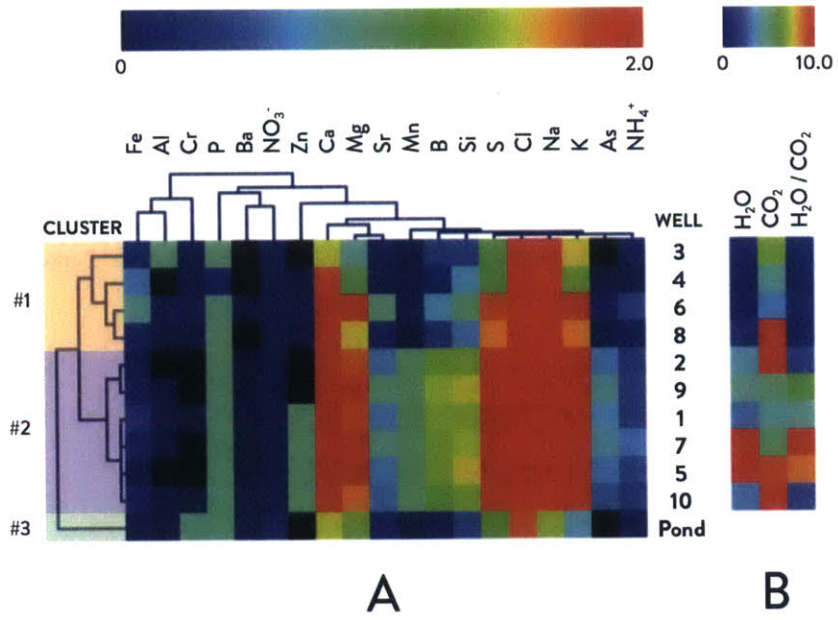


Figure S2. A) Hierarchical clustering by Spearman rank correlation of sample ICP-OES profiles. Heat map displays log transformed (X+1) mg/l concentrations. Clustering reveals three geochemical signature groups. **B)** Base 10 normalized CO_2 and H_2O well test values.

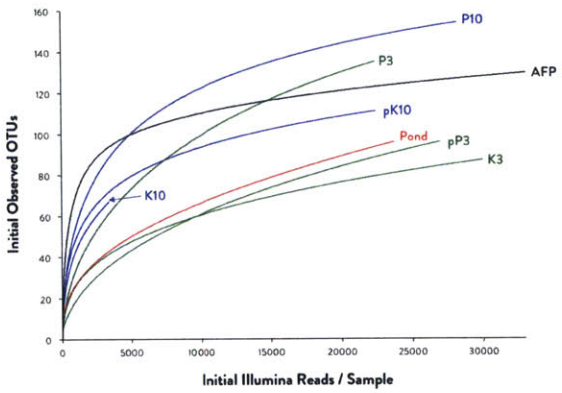


Figure S3. Rarefaction curve generated for initial OTU table based on raw reads demonstrates a sampling of system diversity that nears completion for most wells, the pond and AFP control.

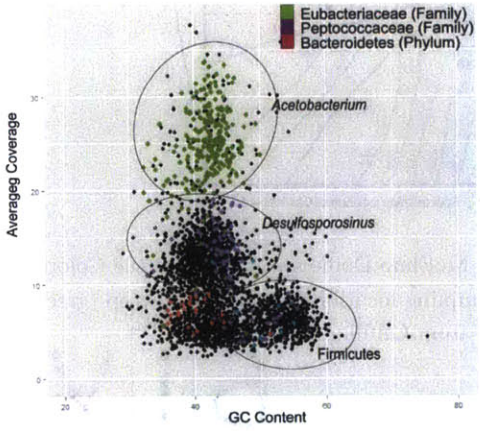


Figure S4. Sequence coverage and GC content of contigs in Well 3 that could not be separated based on tetranucleotide frequencies and sequence homology. Each dot represents a single scaffold/contig with minimum contig length of 1000 bp. All contigs with sequence coverage lower than 40 were fragmented to 500 bp *in silico* followed by Blastx searches against the NCBI NR-database and subsequently assigned a taxon using MEGAN with bitscore of 100. Contigs containing fragments with hits to a single taxon were classified to the same taxonomic group unambiguously (*i.e. Acetobacterium*, *Desulfosporosinus*, Peptococcaceae and Bacteroides) and plotted to guide extraction of genomic bin.

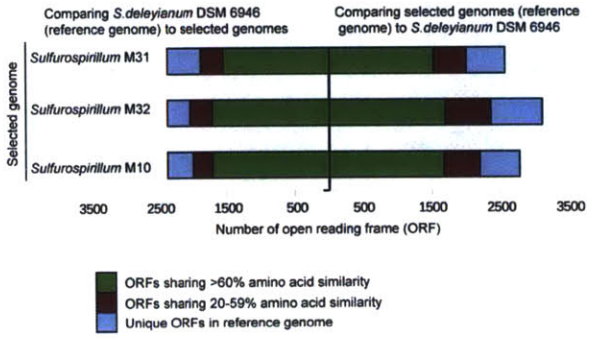


Figure S5. Psi-Blast comparison of the ORFs between binned *Sulfurospirillum* genomes and reference genome *S. deleyanium* DSM 6946 using RAST default settings.

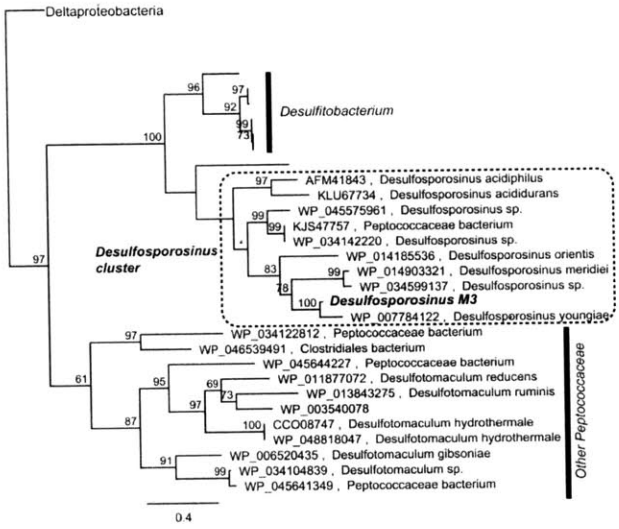


Figure S6. Maximum likelihood tree of reference dsrAB amino acid sequences together with full-length dsrAB recovered from Metagenome 3. A HMM model for dsrAB was used in screening individual genomic bins and complete Metagenomes 3 and 10. No dsrAB was detected in Metagenome 10. The phylogenetic tree was constructed with 1000X bootstrapping and rooted using pyruvate formate lyase gene in Clostridium noyvi (WP_039252367). Bootstrap support values >60 are shown on each branch.

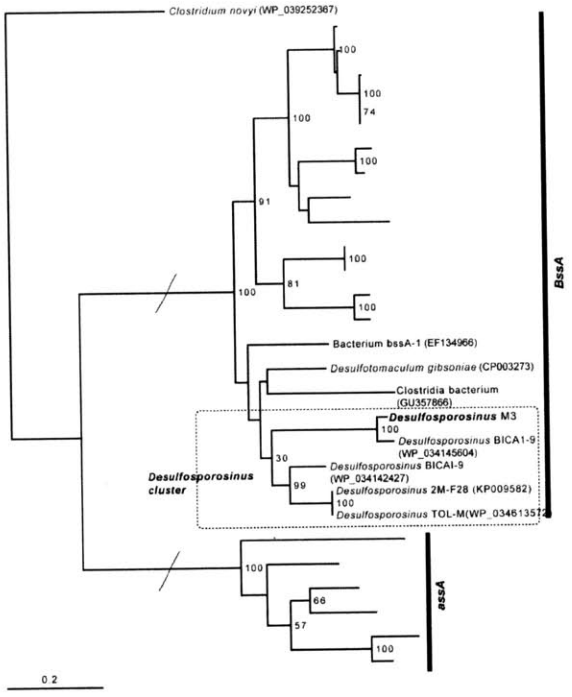
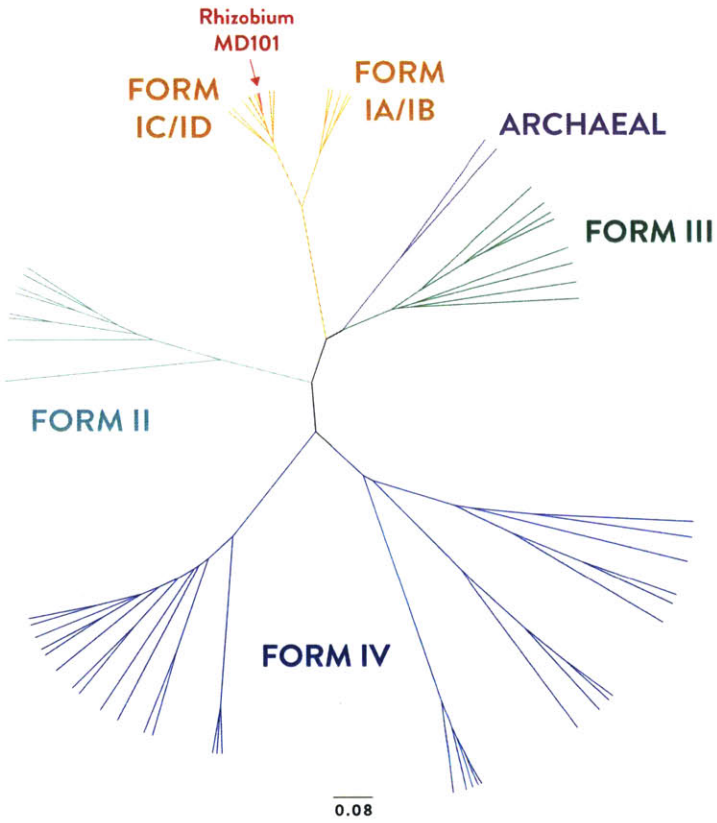


Figure S7. Maximum likelihood tree of reference AssA/BssA amino acid sequences together with full-length bssA (bold) recovered from Metagenome 3. A HMM model for AssA and BssA was used in screening individual genomic bins and complete Metagenomes 3 and 10. No AssA/BssA was detected in Metagenome 10. The phylogenetic tree was constructed with 1000X bootstrapping and rooted using pyruvate formate lyase gene in Clostridium noyvi (WP_039252367). Bootstrap support values >60 are shown on each branch.



Family	Organism	Genbank Accession #
IA	Bradyrhizobium sp. BTAi1	ZP_00862357
	Hydrogenophilus thermoluteolus	Q51856
	Nitrobacter vulgaris	Q59613
IB	Anabaena variabilis ATCC 29413	ABA23512
	Nicotiana tabacum	P00876
IC	Rubrivivax gelatinosus PM1	ZP_00243775
	Xanthobacter autotrophicus Py2	ZP_01200598
ID	Acidiphilium cryptum JF-5	ZP_01146719
	Aurantimonas sp. SI85-9A1	AAB41464
	Burkholderia xenovorans LB400	ZP_00283421
	Nitrosococcus oceani ATCC 19707	ABA56859
	Nitrosospira multiformis ATCC 251966	YP_411385
	Odontella sinensis	NP_043654
	Pleurochrysis carterae	Q08051
II	Hydrogenovibrio marinus	Q59462
	Lingulodinium polyedrum	AAA98748
	Magnetospirillum magnetotacticum AMB-1	YP_422059
	Rhodoferax ferrireducens T118	YP_522655
	Symbiodinium sp.	AAG37859
	Thiomicrospira crunogena XCL-2	ABB41020
	Methanococcoides burtonii DSM 6242	ZP_00563653
Archaea	Methanosaeta thermophila PT	ZP_01153096
	Methanospirillum hungatei JF-1	YP_503739
III	Archaeglobus fulgidus DSM 4304	NP_070466
	Hyperthermus butylicus DSM 5456	YP_001012710
	Methanocaldococcus jannaschii	AAB99239
	Methanosarcina acetivorans C2A	AAM07894
	Natronomonas pharaonis DSM 2160	CAI49476
	Pyrococcus horikoshii OT3	BAA30036
	Thermococcus kodakaraensis KOD1	BAD86479
	Thermofilum pendens Hrk 5	YP_920628

Family	Organism	Genbank Accession #
IV-Non-photo	Acidiphilium cryptum JF-5	ZP_01146529
	Bordetella bronchiseptica RB50	CAE31534
	Burkholderia xenovorans LB400	ZP_00284840
	Chromohalobacter salexigens DSM 3043	ZP_00471249
	Delftia acidovorans SPH1	ZP_01577127
	Fulvimarina pelagi HTCC2506	ZP_01438569
	Jannaschia sp. CCS1	YP_511005
	Mesorhizobium loti	BAB53192
	Polaromonas sp. JS666	ZP_00502320
	Polaromonas sp. JS666	ZP_00502381
	Pseudomonas putida F1	ZP_00900417
	Rhizobium leguminosarum	CAK12105.12
	Roseobacter sp. MED193	ZP_01056409
	Sinorhizobium meliloti 1021	CAC48779
	Xanthobacter autotrophicus Py2	ZP_01199940
IV-DeepYkr	Alkalilimnicola ehrlichei MLHE-1	YP_742007
	Halorhodospira halophila SL1	YP_001002057
	Heliobacillus mobilis	ABH04879
	Ostreococcus tauri	XP_003080289.1
	Ostreococcus tauri	XP_003080830.1
	Rhodopseudomonas palustris BisA53	YP_782588
IV-Photo	Chlorobium chlorochromatii CaD3	ABB28892
	Chlorobium phaeobacteroides DSM 266	ZP_00527577
	Chlorobium tepidum TLS1	AAM72993
	Prosthecochloris aestuarii DSM 271	ZP_00590874
	Rhodopseudomonas palustris BisB18	YP_530146
IV-YkrW	Bacillus cereus E33L	AAU16474
	Bacillus clausii KSM-K16	BAD64310
	Bacillus licheniformis ATCC 14580	AAU23062
	Bacillus subtilis subsp. subtilis str. 168	CAB13232

Figure S8. Maximum likelihood tree (above, left) of reference Ribulose 1,5-bisphosphate (RuBP) carboxylase/oxygenase (RuBisCO) amino acid sequences together with RuBisCO gene (Red) recovered from Well 10 binned genome *Rhizobium* MD101. Phylogenetic tree was constructed with 100X bootstrapping. Clustered sequences listed below, right.

2.7 SUPPLEMENTARY TABLES

Table S1. Summary of PCR primers used in clone library and Illumina sequencing preparation.

Primer Name	Primer Target	Sequncing Method	Sequence (5'-3')	Reference
SSU_357_F	16S rRNA Gene	Clone Library + Sanger	CTCCTACGGGAGGCAGCAG	Turner et al., 1999
SSU_1100_R			AGGGTTGCGCTCGTTG	
M13_F	M13 Vector		GTAAAACGACGGCCAGT	Invitrogen (Carlsbad, CA)
M13_R			AACAGCTATGACCATG	
PE_357_F	16S rRNA Gene	Illumina MiSeq	ACACGACGCTCTTCCGATCTYRYRCTCCTACGGGAGGCAGCAG	Dong et al., 2012 Preheim et al., 2013
PE_806_R			CGGCATTCCTGCTGAACCGCTCTTCCGATCTGGACTACHVGGGTWTCTAAT	
PE_SEQ_F	16S Amplicon		AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT	Illumina (San Diego, CA)
PE_SEQ_R			CAAGCAGAAGACGGCATACGAGAT-BBBBBB-CGGTCTCGGCATTCCTGCTGAACCGCTCT	
Notes: PCR target sequences underlined; six nucleotide multiplexing barcodes denoted in PE_SEQ_R by "BBBBBB"				

Table S2. Sample preparation methods for Illumina sequencing.

Well	Sample	DNA Extraction	PCR Template	Illumina 16S rRNA Statistics									Illumina Nextera Metagenome Statistics						
				Initial Reads	% Merged	% Mapped	% AFP Discard	Final Reads	Reads/OTU	# OTUs	Total OTUs	Chao1	Post-QC Reads	De Novo Assembled	Max Contig	N50	# Contigs >1 kb	Mean Coverage	Median Coverge
3	K3	Kit (K)	gDNA	77961	84.8	72.4	5.2	53487	713.2	75	187	79.1	4,074,221	8,391,774	512.5 kbp	14,942	5,321	20.7	8.9
	P3	Phenol (P)	gDNA	46688	84.5	75.1	7.4	32407	234.8	138		181.8	5,994,660						
	pP3	Phenol (P)	Plasmid (p)	75366	84.0	61.0	2.1	44930	468.0	96		114.5							
10	K10	Kit (K)	gDNA	11503	75.6	44.5	18.7	4133	86.1	48	199	63.0	5,963,526	6,072,097	292.7 kbp	36,674	1,674	35.4	21.0
	P10	Phenol (P)	gDNA	57012	83.0	71.6	23.5	31183	194.9	160		170.3	2,633,670						
	pK10	Kit (K)	Plasmid (p)	92722	74.2	37.5	21.9	27018	355.5	76		81.6							
Pond		Phenol (P)	gDNA	77735	87.3	62.2	17.4	39851	168.9	236		247.8							

Table S3. Clone library taxonomic annotation and abundance summary.

RDP/Silva Phylum	RDP/Silva Genus	RDP/Silva %	Top Blastn Species Hit	Blast ID%	Seq Length	Well 3	Well 10	Pond
Proteobacteria	Sulfurospirillum	100	Sulfurospirillum deleyianum	97-98	681-722	5	2	0
		100	Sulfurospirillum multivorans	97-98	666-721	4	0	0
	Desulfovibrio	100	Desulfovibrio marrakechensis	99	697-742	3	0	0
	Rhizobium	88-100	Rhizobium petrolearium	99-100	678-716	0	10	0
	Polynucleobacter	100	Polynucleobacter difficilis	99	743-744	0	0	2
		100	Polynucleobacter cosmopolitanus	99	402-422	0	0	2
	Methylotenera	77	Methylotenera mobilis	98	745	0	0	1
Firmicutes	Desulfosporosinus	99-100	Desulfosporosinus orientis	97-98	712-746	7	0	0
	Oscillibacter	100	Oscillibacter valericigenes	96-97	721-722	7	0	0
	Desulfitibacter	99-100	Desulfitibacter alkalitolerans	88-92	436-731	3	0	0
	Acetobacterium	100	Acetobacterium carbinolicum	99	711-712	2	0	0
	Desulfitobacterium	94	Desulfitobacterium metallireducens	97	739	1	0	0
Bacteroidetes	Flavobacterium	100	Flavobacterium johnsoniae	99	499	0	0	1
	Sediminibacterium	99	Sediminibacterium salmoneum	92	736	0	0	1
	Arcicella	100	Pseudarcicella hirudinis	93	731	0	0	1
Verrucomicrobia	Prosthecobacter	100	Verrucomicrobia bacterium	97	745	0	0	1
	Unassigned	-	Cerasicoccus frondis	87	549	0	0	1
Actinobacteria	hgcI_clade	84	Sporichthya polymorpha	91	714	0	0	1
					Sum	32	12	11

Table S4. Distribution of Archived False Positive (AFP) OTUs discarded as contaminants from samples.

OTU #	Phylum	Genus	Blastn Top Hit	% AFP Reads	% Reads Discarded			*Known Lab Contaminant
					Well 3	Well 10	Pond	
1	Bacteroidetes	Cloacibacterium	Cloacibacterium normanense	20.58	0.49	3.67	0.14	
9	Proteobacteria	Acidovorax	Acidovorax radicis	12.90	0.22	1.52	0.09	X
20	Proteobacteria	Brevundimonas	Brevundimonas naejangsanensis	7.18	0.08	0.50	0.06	X
24	Proteobacteria	Halomonas	Halomonas pacifica	6.92	0.48	0.52	0.04	
10	Proteobacteria	Acinetobacter	Acinetobacter junii	6.56	0.22	0.69	0.09	X
11	Proteobacteria	Diaphorobacter	Acidovorax ebreus	3.40	0.15	1.27	0.02	
13	Proteobacteria	Aquabacterium	Aquabacterium parvum	3.21	0.14	1.53	0.06	X
23	Proteobacteria	Pseudomonas	Pseudomonas stutzeri	2.50	0.11	1.40	0.03	X
8	Firmicutes	Streptococcus	Streptococcus dentisani	2.26	0.10	0.58	0.02	X
3	Firmicutes	Bacillus	Bacillus toyonensis	2.06	1.46	0.27	0.05	X
335	Proteobacteria	Shewanella	Shewanella algae	1.32	0.17	0.06	0.00	
65	Proteobacteria	Dechloromonas	Dechloromonas agitata	0.93	0.02	3.37	0.00	
569	Proteobacteria	Limnohabitans	Limnohabitans parvus	0.47	0.00	0.04	2.83	
63	Actinobacteria	Candidatus_Aquiluna	Candidatus Aquiluna rubra	0.40	0.00	0.02	3.14	
248	Proteobacteria	Unassigned	Albidiferax ferrireducens	0.19	0.00	0.00	0.58	
171	Actinobacteria	Candidatus_Limnoluna	Candidatus Limnoluna rubra	0.18	0.00	0.01	1.76	
35	Bacteroidetes	Arcicella	Pseudarcicella hirudinis	0.18	0.01	0.03	5.63	
474	Actinobacteria	Leucobacter	Leucobacter chromiireducens	0.04	0.00	0.00	0.19	
193	Bacteroidetes	Unassigned	Owenweeksia hongkongensis	0.03	0.00	0.00	0.78	
896	Proteobacteria	Aquabacterium	Aquabacterium parvum	0.00	0.00	1.21	0.01	X
40	Cyanobacteria	Chloroplast	Aerosakkonema funiforme	0.00	0.12	0.02	0.00	
235	Actinobacteria	CL500-29_marine_group	Ilumatobacter fluminis	0.00	0.00	0.00	0.16	
270	Bacteroidetes	Unassigned	Owenweeksia hongkongensis	0.00	0.00	0.00	0.20	
314	Bacteroidetes	Unassigned	Owenweeksia hongkongensis	0.00	0.00	0.00	0.18	

*Based on Salter et al. (2014) list of contaminant genera detected in sequenced negative 'blank' controls

Note: OTUs 63, 65, and 569 may potentially represent genuine sample taxa based on their abundance in well samples, ecological precedent, and no prior identification as known lab contaminants, but were removed from this study based on AFP criteria.

Table S5A. RAST/IMG-annotated genes associated with inorganic metabolic pathways and cycling

METABOLIC / PHYSIOLOGICAL GENOMIC FUNCTIONAL ANNOTATION (RAST / IMG)					WELL 3										WELL 10			
					Ac34		Os35		Dv33		Ss31		Ss32		Ds*	Rz101		Ss102	
PATHWAY	GENE	DEFINITION	FUNCTION	KEGG Orthology (KO)	IMG	RAST	IMG	RAST	IMG	RAST	IMG	RAST	IMG	RAST	HMM	IMG	RAST	IMG	RAST
SULFUR METABOLISM																			
Sulfate Reduction	*Sat*	sulfate adenylyltransferase [EC:2.7.7.4]	Sulfate --> Adenylsulfate	KO:K00958															
	cysN	sulfate adenylyltransferase subunit 1 [EC:2.7.7.4]	Sulfate --> Adenylsulfate																
	cysD	sulfate adenylyltransferase subunit 2 [EC:2.7.7.4]	Sulfate --> Adenylsulfate	KO:K00957															
	aprA	adenylylsulfate reductase, subunit A [EC:1.8.99.2]	Adenylsulfate --> Sulfite	KO:K00394															
	aprB	adenylylsulfate reductase, subunit B [EC:1.8.99.2]	Adenylsulfate --> Sulfite	KO:K00395															
Sulfite reduction	*dsrA / asrA*	sulfite reductase alpha subunit [EC:1.8.99.3 1.8.99.5]	Sulfite --> Hydrogen Sulfide	KO:K11180; K16950															
	dsrB / asrB	sulfite reductase beta subunit [EC:1.8.99.3 1.8.99.5]	Sulfite --> Hydrogen Sulfide	KO:K11181; K16951															
	dsrD / asrC	dissimilatory sulfite reductase clustered protein	Sulfite --> Hydrogen Sulfide	KO:K00385															
	dsrMKJOP	sulfite reduction-associated complex	Sulfite --> Hydrogen Sulfide																
Sulfur oxidation	*soxA*	sulfur-oxidizing protein SoxA	Thiosulfate --> Sulfate	KO:K17222															
	soxB	sulfur-oxidizing protein SoxB	Thiosulfate --> Sulfate	KO:K17224															
	soxX	sulfur-oxidizing protein SoxX	Thiosulfate --> Sulfate	KO:K17223															
	soxY	sulfur-oxidizing protein SoxY	Thiosulfate --> Sulfate	KO:K17226															
	soxZ	sulfur-oxidizing protein SoxZ	Thiosulfate --> Sulfate	KO:K17227															
	soxC	sulfur oxidation molybdopterin C protein	Thiosulfate --> Sulfate	KO:K17225															
Sulfite oxidation	*sorAB / soxD*	sulfite dehydrogenase [EC:1.8.2.1]	Sulfite --> Sulfate																
	SUOX / yedY	sulfite oxidase [EC:1.8.3.1] / sulfoxide reductase catalytic subunit YedY [EC	Sulfite --> Sulfate	KO:K07147															
NITROGEN METABOLISM																			
Nitrate Reduction	*narG*	nitrate reductase alpha subunit [EC:1.7.99.4]	Nitrate --> Nitrite																
	narH	nitrate reductase beta subunit [EC:1.7.99.4]	Nitrate --> Nitrite	KO:K00374															
	narI	nitrate reductase gamma subunit [EC:1.7.99.4]	Nitrate --> Nitrite																
	narJ	nitrate reductase delta subunit	Nitrate --> Nitrite	KO:K02567															
	napA	periplasmic nitrate reductase NapA [EC:1.7.99.4]	Nitrate --> Nitrite	KO:K02568															
	napB	cytochrome c-type protein NapB	Nitrate --> Nitrite	KO:K02570															
	napD	periplasmic nitrate reductase NapD	Nitrate --> Nitrite	KO:K02572															
	napF	ferredoxin-type protein (periplasmic nitrate reductase)	Nitrate --> Nitrite	KO:K02573															
	napG	ferredoxin-type protein (periplasmic nitrate reductase)	Nitrate --> Nitrite	KO:K02574															
	napH	polyferredoxin quinol dehydrogenase	Nitrate --> Nitrite																
	napL	periplasmic nitrate reductase component	Nitrite --> Ammonia	KO:K00362															
Nitrite Reduction	*nirB*	nitrite reductase (NADH) large subunit [EC:1.7.1.15]	Nitrite --> Ammonia	KO:K00363															
	nirD	nitrite reductase (NADH) small subunit [EC:1.7.1.15]	Nitrite --> Ammonia	KO:K03385															
	nrfA	nitrite reductase (cytochrome c-552) [EC:1.7.2.2]	Nitrite --> Ammonia	KO:K15876															
	nrfH	cytochrome c nitrite reductase small subunit	Nitrite --> Ammonia	KO:K04014															
	nrfC	nitrite reductase component	Nitrite --> Ammonia	KO:K04015															
	nrfD	formate-dependent nitrite reductase, membrane subunit	Nitrite --> Nitric Oxide	KO:K00368															
Denitrification	*aniK*	nitrite reductase (NO-forming) [EC:1.7.2.1]	Nitrite --> Nitric Oxide	KO:K05601															
	nirS / hcp	nitrite reductase (NO-forming) / hydroxylamine reductase [EC:1.7.2.1 1.7.99	Nitric Oxide --> Nitrous Oxide	KO:K04561															
	norB	nitric oxide reductase subunit B [EC:1.7.2.5]	Nitric Oxide --> Nitrous Oxide	KO:K02305															
	norC	nitric oxide reductase subunit C	Nitrous Oxide --> N2	KO:K00376															
	nosZ	nitrous-oxide reductase [EC:1.7.2.4 1.7.99.6]	Nitric Oxide --> Nitrous Oxide	KO:K02448															
	norD	nitric oxide reductase subunit D	Nitric Oxide --> Nitrous Oxide	KO:K02164															
	norE	nitric oxide reductase NorE protein	Nitric Oxide --> Nitrous Oxide	KO:K04747															
	norF	nitric oxide reductase NorF protein	Nitric Oxide --> Nitrous Oxide	KO:K04748															
	norQ	nitric-oxide reductase [EC:1.7.2.5]	Nitric Oxide --> Nitrous Oxide																
	norV	anaerobic nitric oxide reductase flavorubredoxin	Nitrous Oxide --> N2	KO:K07218															
	nosD	nitrous oxide reductase maturation protein	Nitrous Oxide --> N2																
	nosF	nitrous oxide reductase maturation protein (ATPase)	Nitrous Oxide --> N2	KO:K19342															
	nosL	nitrous oxide reductase maturation protein, outer-membrane lipoprotein	Nitrous Oxide --> N2																
	nosX	nitrous oxide reductase maturation periplasmic protein	Nitrous Oxide --> N2																
	nosY	nitrous oxide reductase maturation transmembrane protein	Nitrous Oxide --> N2	KO:K02586															
Nitrogen Fixation	*nifD*	nitrogenase molybdenum-iron protein alpha chain [EC:1.18.6.1]	N2 --> Ammonia	KO:K02591															
	nifK	nitrogenase molybdenum-iron protein beta chain [EC:1.18.6.1]	N2 --> Ammonia	KO:K02588															
	nifH	nitrogenase iron protein NifH [EC:1.18.6.1]	N2 --> Ammonia	KO:K00531															
	anfG	nitrogenase delta subunit [EC:1.18.6.1]	N2 --> Ammonia																
	anfO	required for Mo- and V-independent nitrogenase	N2 --> Ammonia	KO:K02584															
	nifA	nitrogenase (molybdenum-iron)-specific transcriptional regulator	N2 --> Ammonia	KO:K02585															
	nifB	nitrogenase FeMo-cofactor synthesis FeS core scaffold and assembly prote	N2 --> Ammonia	KO:K02587															
	nifE	nitrogenase FeMo-cofactor scaffold and assembly protein	N2 --> Ammonia	KO:K02592															
	nifN	nitrogenase FeMo-cofactor scaffold and assembly protein	N2 --> Ammonia																
	nifO	nitrogenase-associated protein	N2 --> Ammonia	KO:K04488															
	nifU	iron-sulfur cluster assembly scaffold protein	N2 --> Ammonia	KO:K04487															
	nifS	cysteine desulfurase (EC:2.8.1.7)	N2 --> Ammonia																
	nifX	nitrogenase FeMo-cofactor carrier protein	N2 --> Ammonia	KO:K02595															
	nifW	nitrogenase stabilizing/protective protein	N2 --> Ammonia																
	fixJ	two-component nitrogen fixation transcriptional regulator	N2 --> Ammonia																
METALS / METALLOIDS METABOLISM																			
Fe(III) Reduction	*frd*	Ferric-chelate reductase [EC:1.16.1.10]	Fe(III) --> Fe(II)	KO:K18915															
	fhuF	ferric iron reductase protein FhuF	Fe(III) --> Fe(II)	KO:K13255															
Fe(II) Oxidation	*qcr*	Fe(II)-cytochrome c reductase [EC:1.9.98.1]	Fe(II) --> Fe(III)	KO:K00411; K00412															
As(V) Reduction	*arrA*	arsenate reductase (respiration) [EC:1.20.99.1]	As(V) --> As(III)																
	arrB	arsenate reductase (respiration) [EC:1.20.99.1]	As(V) --> As(III)																
	arsC	arsenate reductase (glutaredoxin) (detoxification) [EC:1.20.4.1]		KO:K00537; K03741															
As(III) Oxidation	*aoxA*	arsenite oxidase large subunit (EC:1.20.98.1 1.20.9.1)	As(III) --> As (V)	KO:K08355															
	aoxB	arsenite oxidase small subunit precursor (EC:1.20.98.1 1.20.9.1)	As(III) --> As (V)	KO:K08356															

LEGEND

Well	Bin	Strain
3	Ac34	*Acetobacterium* MD34
	Os35	*Oscillibacter* MD35
	Dv33	*Desulfovibrio* MD33
	Ss31	*Sulfurospirillum* MD31
	Ss32	*Sulfurospirillum* MD32
	Ds*	*Desulfosporosinus* MD36
10	Rz101	*Rhizobium* MD101
	Ss102	*Sulfurospirillum* MD102

Required Pathway Genes

Table S5B. RAST/IMG-annotated genes associated with organic metabolic pathways and hydrogenases

METABOLIC / PHYSIOLOGICAL GENOMIC FUNCTIONAL ANNOTATION (RAST / IMG)				WELL 3												WELL 10			
				Ac34		Os35		Dv33		Ss31		Ss32		Ds*		Rz101		Ss102	
PATHWAY	GENE	DEFINITION	KEGG Orthology (KO)	IMG	RAST	IMG	RAST	IMG	RAST	IMG	RAST	IMG	RAST	HMM		IMG	RAST	IMG	RAST
ANAEROBIC HYDROCARBON ACTIVATION																			
Benzene	*abcD**	anaerobic benzene carboxylase																	
Alkane	*assA**	alkylsuccinate synthase																	
Toluene	*bssA**	benzylsuccinate synthase																	
Ethylbenzene	*ebdA**	ethylbenzene dehydrogenase																	
Tyrosol	*HPAH*	4-hydroxyphenylacetic acid 3-hydroxylase	KO:K18089																
CARBON FIXATION																			
Calvin Cycle	rbcS	Ribulose bisphosphate carboxylase small chain (EC 4.1.1.39)	KO:K01602																
	rbcL	Ribulose bisphosphate carboxylase large chain (EC 4.1.1.39)	KO:K01601																
	cbbP	Phosphonbulokinase (EC 2.7.1.19)	KO:K00855																
	cbbF	Fructose-1,6-bisphosphatase (EC 3.1.3.11)	KO:K04041, K03841																
	tpiA	Triosephosphate isomerase (EC 5.3.1.1)	KO:K01803																
	cbbK	Phosphoglycerate kinase (EC 2.7.2.3)	KO:K00927																
	cbbG	Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	KO:K00134																
	cbbA	Fructose-bisphosphate aldolase (EC 4.1.2.13)	KO:K01623, K01624																
	cbbT	Transkelolase (EC 2.2.1.1)	KO:K00615																
	rpiAB	Ribose 5-phosphate isomerase (EC 5.3.1.6)	KO:K01808, K01807																
	cbbE	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)	KO:K01783																
Wood-Ljungdahl Pathway	cooS / acsA	carbon-monoxide dehydrogenase	KO:K00198																
	acsB	acetyl-CoA synthase	KO:K14138																
	cooF	CO dehydrogenase Fe-S protein	KO:K00196																
	cooC	CO dehydrogenase accessory protein	KO:K07321																
	coxS	CO dehydrogenase small subunit	KO:K03518																
Reductive TCA Cycle	aclA	ATP-citrate lyase alpha-subunit [EC 2.3.3.8]																	
	aclB	ATP-citrate lyase beta-subunit [EC 2.3.3.8]																	
	korA	2-oxoglutarate oxidoreductase, alpha subunit (EC 1.2.7.3)	KO:K00174																
	korB	2-oxoglutarate oxidoreductase, beta subunit (EC 1.2.7.3)	KO:K00175																
	korC	2-oxoglutarate oxidoreductase, gamma subunit (EC 1.2.7.3)	KO:K00177																
	korD	2-oxoglutarate oxidoreductase, delta subunit (EC 1.2.7.3)	KO:K00176																
	porA	Pyruvate ferredoxin oxidoreductase, alpha subunit (EC 1.2.7.1)	KO:K00169																
	porB	Pyruvate ferredoxin oxidoreductase, beta subunit (EC 1.2.7.1)	KO:K00170																
	porG	Pyruvate ferredoxin oxidoreductase, gamma subunit (EC 1.2.7.1)	KO:K00172																
	porD	Pyruvate ferredoxin oxidoreductase, delta subunit (EC 1.2.7.1)	KO:K00171																
CO_2 / Bicarbonate Fixation	cynT / can	Carbonic anhydrase (EC 4.2.1.1)	KO:K01673																
	ppc	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)	KO:K01595																
	carAB	Carbamoyl-phosphate synthase (EC 6.3.5.5)	KO:K01955, K01956																
HYDROGENASES																			
Hydrogenase 1	hyaA / hybO	hydrogenase small subunit [EC:1.12.99.6] / Uptake hydrogenase small subu	KO:K06282																
	hyaC	[NiFe] hydrogenase 1 B-type cytochrome subunit	KO:K03620																
Hydrogenase 2	hybD	hydrogenase maturation protease [EC:3.4.23.-]	KO:K03605																
	hybA	HybA, Fe-S-cluster-containing hydrogenase components 1 homolog																	
Hydrogenase 3	hycI	hydrogenase 3 maturation protease [EC:3.4.23.51]	KO:K08315																
Hydrogenase 4	hyfA	hydrogenase-4 component A [EC:1.-.-.-]	KO:K12136																
	hyfB	hydrogenase-4 component B	KO:K12137																
	hyfC	hydrogenase-4 component C [EC:1.-.-.-]	KO:K12138																
	hyfE	hydrogenase-4 component E [EC:1.-.-.-]	KO:K12140																
	hyfF	hydrogenase-4 component F [EC:1.-.-.-]	KO:K12141																
Fe Hydrogenase	hydA2	[NiFe] hydrogenase small subunit [EC:1.12.2.1]	KO:K18008																
	hydB2	[NiFe] hydrogenase large subunit [EC:1.12.2.1]	KO:K00437																
	hydA3	quinone-reactive Ni/Fe-hydrogenase small subunit [EC:1.12.5.1]	KO:K05927																
	hydB3	quinone-reactive Ni/Fe-hydrogenase large subunit [EC:1.12.5.1]	KO:K05922																
	hydE	[FeFe]-hydrogenase maturation protein HydE																	
	hydF	[FeFe]-hydrogenase maturation protein HydF																	
	hydC	[FeFe] hydrogenase, group A																	
NiFe Hydrogenase	hypA / hybF	hydrogenase nickel incorporation protein	KO:K04651																
	hypB	hydrogenase nickel incorporation protein HypB	KO:K04652																
	hypC	hydrogenase expression/formation protein HypC	KO:K04653																
	hypD	hydrogenase expression/formation protein HypD	KO:K04654																
	hypE	hydrogenase expression/formation protein HypE	KO:K04655																
	hypF	[NiFe] hydrogenase maturation protein HypF	KO:K04656																
Energy Conserving Hydrogenase	echA	ech hydrogenase subunit A	KO:K14086																
	echB	ech hydrogenase subunit B	KO:K14087																
	echC	ech hydrogenase subunit C	KO:K14088																
	echD	ech hydrogenase subunit D	KO:K14089																
	echE	ech hydrogenase subunit E	KO:K14090																
	echF	ech hydrogenase subunit F	KO:K14091																
NADP-Reducing Fe hydrogenase	hndB	NADP-reducing hydrogenase subunit HndB [EC:1.12.1.3]	KO:K17992																
	hndA	NADP-reducing hydrogenase subunit HndA [EC:1.12.1.3]	KO:K18330																
	hndC	NADP-reducing hydrogenase subunit HndC [EC:1.12.1.3]	KO:K18331																
	hndD	NADP-reducing hydrogenase subunit HndD [EC:1.12.1.3]	KO:K18332																
NAD-reducing hydrogenase	HoxF	NAD-reducing hydrogenase subunit HoxF (EC:1.12.1.2)																	
	HoxE	NAD-reducing hydrogenase subunit HoxE (EC:1.12.1.2)																	
	HoxA	NAD-reducing hydrogenase subunit HoxA (EC:1.12.1.2)																	
Fe Hydrogenase	HymA	[Fe] hydrogenase, HymA subunit, putative																	
	HymC	[Fe] hydrogenase, HymC subunit, putative																	
Membrane bound hydrogenase	MbhB	Membrane bound hydrogenase, MbhB subunit																	

LEGEND		
Well	Bin	Strain
3	Ac34	*Acetobacterium* MD34
	Os35	*Oscillibacter* MD35
	Dv33	*Desulfovibrio* MD33
	Ss31	*Sulfurospirillum* MD31
	Ss32	*Sulfurospirillum* MD32
	Ds*	*Desulfosporosinus* MD36
10	Rz101	*Rhizobium* MD101
	Ss102	*Sulfurospirillum* MD102

Required Pathway Genes

Table S5C. IMG-annotated genes associated with membrane transporters

Sulfurospirillum MD31	Sulfurospirillum MD32	Desulfovibrio MD33	Acetobacterium MD34
The Ammonia Transporter Channel (Amt) Family	The Ammonia Transporter Channel (Amt) Family	The Ammonia Transporter Channel (Amt) Family	The Ammonia Transporter Channel (Amt) Family
The CorA Metal Ion Transporter (MIT) Family	Sugar phosphate permease	The Lactate Permease (LctP) Family	Nitrate/nitrite transporter NarK
The C_4-Dicarboxylate Uptake (Dcu) Family	The C4-Dicarboxylate Uptake (Dcu) Family	The Glycerol Uptake (GUP) Family	MFS transporter, OFA family, oxalate/formate antiporter
The Lactate Permease (LctP) Family	The CorA Metal Ion Transporter (MIT) Family	The Sulfate Permease (SulP) Family	Sugar phosphate permease
The Formate-Nitrite Transporter (FNT) Family	The Lactate Permease (LctP) Family	The Arsenical Resistance-3 (ACR3) Family	The Lactate Permease (LctP) Family
The Sulfate Permease (SulP) Family	The Formate-Nitrite Transporter (FNT) Family	The Twin Arginine Targeting (Tat) Family	The Arsenical Resistance-3 (ACR3) Family
The Arsenical Resistance-3 (ACR3) Family	The Sulfate Permease (SulP) Family	The 10 TMS Putative Sulfate Exporter (PSE) Family	The Gluconate:H+ Symporter (GntP) Family
The Twin Arginine Targeting (Tat) Family	The Arsenical Resistance-3 (ACR3) Family	Simple sugar transport system substrate-binding protein	The Arsenite-Antimonite (ArsAB) Efflux Family
The Gluconate:H^+ Symporter (GntP) Family	The Twin Arginine Targeting (Tat) Family	Iron complex transport system substrate-binding protein	The PTS Glucose-Glucoside (Glc) Family
The Arsenite-Antimonite (ArsAB) Efflux Family	The Gluconate:H+ Symporter (GntP) Family	Iron complex transport system ATP-binding protein	The PTS Mannose-Fructose-Sorbose (Man) Family
The Ferrous Iron Uptake (FeoB) Family	The 10 TMS Putative Sulfate Exporter (PSE) Family	Iron(III) transport system permease protein	The Ferrous Iron Uptake (FeoB) Family
Aromatic amino acid transport protein AroP	The Arsenite-Antimonite (ArsAB) Efflux Family	Simple sugar transport system permease protein	The Ethanol Utilization/Transport (Eut) Protein Family
Alanine or glycine:cation symporter, AGCS family	The Putative 4-Toluene Sulfonate Uptake Permease (TSUP) Family	The PTS Mannose-Fructose-Sorbose (Man) Family	Iron complex transport system substrate-binding protein
Branched-chain amino acid transport system substrate-binding protein	The Ferrous Iron Uptake (FeoB) Family	The Ferrous Iron Uptake (FeoB) Family	Iron complex transport system permease protein
Branched-chain amino acid transport system permease protein	Iron complex transport system substrate-binding protein	Branched-chain amino acid transport system substrate-binding protein	Iron complex transport system ATP-binding protein
Polar amino acid transport system substrate-binding protein	Urea transport system substrate-binding protein	Branched-chain amino acid transport system ATP-binding protein	Simple sugar transport system permease protein
Glycine betaine/proline transport system ATP-binding protein	Urea transport system permease protein	Branched-chain amino acid transport system permease protein	Simple sugar transport system ATP-binding protein
ABC-type branched-chain amino acid transport system, substrate-binding protein	Nitrous oxidase accessory protein	Polar amino acid transport system substrate-binding protein	Putative glutamine transport system permease protein
Polar amino acid transport system permease protein	The Alanine or Glycine:Cation Symporter (AGCS) Family	Polar amino acid transport system permease protein	ABC-type nitrate/sulfonate/bicarbonate transport system, permease component
General L-amino acid transport system ATP-binding protein	Aromatic amino acid transport protein AroP	Polar amino acid transport system ATP-binding protein	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component
General L-amino acid transport system permease protein	The Aspartate:Alanine Exchanger (AAEx) Family	Peptide/nickel transport system substrate-binding protein	Multiple sugar transport system substrate-binding protein
General L-amino acid transport system substrate-binding protein	Carbon starvation protein CstA	Glycine betaine/proline transport system substrate-binding protein	Amino acid permease
Peptide/nickel transport system ATP-binding protein	Polar amino acid transport system substrate-binding protein	Glycine betaine/proline transport system permease protein	Alanine or glycine:cation symporter, AGCS family
Peptide/nickel transport system substrate-binding protein	Phosphate transport system substrate-binding protein	Glycine betaine/proline transport system ATP-binding protein	ABC-type dipeptide/oligopeptide/nickel transport system, ATPase component
Peptide/nickel transport system permease protein	Phosphate transport system permease protein	Peptide/nickel transport system ATP-binding protein	Peptide/nickel transport system substrate-binding protein
Carbon starvation protein CstA	Phosphate transport system ATP-binding protein	Peptide/nickel transport system permease protein	Peptide/nickel transport system permease protein
Phosphate transport system substrate-binding protein	Peptide/nickel transport system permease protein	Carbon starvation protein CstA	Oligopeptide transport system ATP-binding protein
Phosphate transport system permease protein	Branched-chain amino acid transport system substrate-binding protein	Phosphate transport system substrate-binding protein	Carbon starvation protein CstA
Phosphate transport system ATP-binding protein	Branched-chain amino acid transport system permease protein	Phosphate transport system permease protein	Phosphate/phosphite/phosphonate ABC transporter binding protein
Phosphonate transport system substrate-binding protein	Branched-chain amino acid transport system ATP-binding protein	Phosphate transport system ATP-binding protein	Phosphate transport system ATP-binding protein
	Peptide/nickel transport system substrate-binding protein		Phosphate transport system permease protein
	ABC-type amino acid transport substrate-binding protein		Phosphate transport system substrate-binding protein
	Phosphonate transport system ATP-binding protein		
	Phosphonate transport system substrate-binding protein		
	Phosphonate transport system permease protein		
	Putative phosphonate transport system ATP-binding protein		
	Polar amino acid transport system ATP-binding protein		
	Polar amino acid transport system permease protein		
	Glycine betaine/proline transport system substrate-binding protein		
	Glycine betaine/proline transport system permease protein		
	Glycine betaine/proline transport system ATP-binding protein		
	General L-amino acid transport system ATP-binding protein		
	General L-amino acid transport system permease protein		
	General L-amino acid transport system substrate-binding protein		
	D-methionine transport system substrate-binding protein		
	D-methionine transport system ATP-binding protein		
	D-methionine transport system permease protein		

Table S5C. IMG-annotated genes associated with membrane transporters (continued)

Oscillibacter MD35	Sulfurospirillum MD102	Rhizobium MD101
The Ammonia Transporter Channel (Amt) Family	The Ammonia Transporter Channel (Amt) Family	The Ammonia Transporter Channel (Amt) Family
Sugar phosphate permease	The CorA Metal Ion Transporter (MIT) Family	The Formate-Nitrite Transporter (FNT) Family
The Lactate Permease (LctP) Family	Nitrate/nitrite transporter NarK	The Benzoate:H+ Symporter (BenE) Family
The Glutamate:Na+ Symporter (ESS) Family	The C4-Dicarboxylate Uptake (Dcu) Family	The Sulfate Permease (SulP) Family
The Glycerol Uptake (GUP) Family	The Lactate Permease (LctP) Family	The Arsenical Resistance-3 (ACR3) Family
The Arsenical Resistance-3 (ACR3) Family	The Sulfate Permease (SulP) Family	The Twin Arginine Targeting (Tat) Family
The Gluconate:H+ Symporter (GntP) Family	The Arsenical Resistance-3 (ACR3) Family	The Aromatic Acid Exporter (ArAE) Family
The 10 TMS Putative Sulfate Exporter (PSE) Family	The Twin Arginine Targeting (Tat) Family	Multiple sugar transport system permease protein
The Dipicolinic Acid Transporter (DPA-T) Family	The Gluconate:H+ Symporter (GntP) Family	Multiple sugar transport system ATP-binding protein
The Putative 4-Toluene Sulfonate Uptake Permease (TSUP) Family	The 10 TMS Putative Sulfate Exporter (PSE) Family	Multiple sugar transport system substrate-binding protein
The Ferrous Iron Uptake (FeoB) Family	The Arsenite-Antimonite (ArsAB) Efflux Family	Iron(III) transport system ATP-binding protein
Simple sugar transport system permease protein	The Ferrous Iron Uptake (FeoB) Family	Iron(III) transport system permease protein
Multiple sugar transport system ATP-binding protein	The YedZ (YedZ) Family	Iron(III) transport system substrate-binding protein
Iron complex transport system substrate-binding protein	Iron complex transport system permease protein	Simple sugar transport system permease protein
Iron complex transport system permease protein	Nitrous oxidase accessory protein	Simple sugar transport system substrate-binding protein
ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component	Serine/threonine transporter	The Arsenite-Antimonite (ArsAB) Efflux Family
ABC-type nitrate/sulfonate/bicarbonate transport system, permease component	The Aspartate:Alanine Exchanger (AAEx) Family	Glucose/mannose transport system permease protein
The Alanine or Glycine:Cation Symporter (AGCS) Family	Branched-chain amino acid transport system permease protein	Glycerol transport system permease protein
L-asparagine transporter	Branched-chain amino acid transport system substrate-binding protein	Lactose/L-arabinose transport system ATP-binding protein
Oligopeptide transport system permease protein	Polar amino acid transport system substrate-binding protein	The Alanine or Glycine:Cation Symporter (AGCS) Family
Oligopeptide transport system ATP-binding protein	Glycine betaine/proline transport system substrate-binding protein	Peptide/bleomycin uptake transporter
Oligopeptide transport system substrate-binding protein	Glycine betaine/proline transport system permease protein	Phosphate transport system ATP-binding protein
Branched-chain amino acid transport system ATP-binding protein	General L-amino acid transport system permease protein	Phosphate transport system permease protein
Branched-chain amino acid transport system permease protein	General L-amino acid transport system substrate-binding protein	Phosphate transport system substrate-binding protein
Branched-chain amino acid transport system substrate-binding protein	Peptide/nickel transport system substrate-binding protein	Branched-chain amino acid transport system substrate-binding protein
Peptide/nickel transport system substrate-binding protein	Peptide/nickel transport system permease protein	Branched-chain amino acid transport system ATP-binding protein
Peptide/nickel transport system ATP-binding protein	Peptide/nickel transport system ATP-binding protein	Branched-chain amino acid transport system permease protein
Peptide/nickel transport system permease protein	Carbon starvation protein CstA	General L-amino acid transport system ATP-binding protein
Polar amino acid transport system substrate-binding protein	Phosphate transport system substrate-binding protein	General L-amino acid transport system permease protein
Polar amino acid transport system permease protein	Phosphate transport system permease protein	General L-amino acid transport system substrate-binding protein
Polar amino acid transport system ATP-binding protein	Phosphate transport system ATP-binding protein	Glycine betaine/proline transport system permease protein
Putative lysine transport system substrate-binding protein	Phosphonate transport system substrate-binding protein	Glycine betaine/proline transport system ATP-binding protein
Putative lysine transport system permease protein		Peptide/nickel transport system ATP-binding protein
Putative lysine transport system ATP-binding protein		Peptide/nickel transport system permease protein
Phosphonate transport system substrate-binding protein		Peptide/nickel transport system substrate-binding protein
Phosphate transport system substrate-binding protein		Oligopeptide transport system permease protein
D-methionine transport system ATP-binding protein		Polar amino acid transport system ATP-binding protein
D-methionine transport system permease protein		Polar amino acid transport system permease protein
D-methionine transport system substrate-binding protein		Polar amino acid transport system substrate-binding protein
Phosphate transport system ATP-binding protein		ABC transporter, substrate binding protein, PQQ-dependent alcohol dehydrogenase system
Phosphate transport system permease protein		D-xylose transport system permease protein
Carbon starvation protein CstA		D-xylose transport system substrate-binding protein
		Oligopeptide transport system ATP-binding protein
		Urea transport system ATP-binding protein
		Urea transport system permease protein
		Urea transport system substrate-binding protein
		Inositol-phosphate transport system substrate-binding protein
		Inositol-phosphate transport system permease protein
		Sorbitol/mannitol transport system substrate-binding protein
		Sorbitol/mannitol transport system permease protein
		Phosphonate transport system permease protein
		Phosphonate transport system substrate-binding protein
		Phosphonate transport system ATP-binding protein
		Manganese/iron transport system substrate-binding protein
		Manganese/iron transport system ATP-binding protein
		Manganese/iron transport system permease protein
		D-methionine transport system permease protein
		D-methionine transport system substrate-binding protein
		D-methionine transport system ATP-binding protein
		Cystine transport system ATP-binding protein
		Cystine transport system permease protein
		Cystine transport system substrate-binding protein
		Fructose transport system permease protein
		Fructose transport system substrate-binding protein

2.8 SUPPLEMENTARY METHODS

Site Characterization

476 billion m^3 of CO_2 in the 800 km^2 McElmo Dome field began to accumulate approximately 40 to 72 million years ago (Cappa and Rice, 1995; Gilfillan *et al.*, 2008). Trapped at depths of 1800 to 2600 m within the 100 m thick dolomite-rich Leadville Formation (Allis *et al.*, 2001; Gilfillan *et al.*, 2009), CO_2 exists as a supercritical fluid above the CO_2 critical point (>31°C, 71 atm) at the temperature (~65°C) and pressure (~135 atm) conditions of the formation. KinderMorgan operates 61 wells in the Leadville Formation that produce fluids with a wide range of CO_2/H_2O ratios, from pure CO_2 to nearly degassed samples. On site, an artesian mixture of supercritical CO_2 and saline formation water is produced and separated, after which the CO_2 is further dehydrated and compressed for pipeline delivery and commercial use, while the formation water is re-injected into the Leadville Formation (Stevens *et al.*, 2001).

Geochemical Analysis Methods

Chloride, Nitrate and Ammonium flow-injection analyses were conducted using a Thermo iCAP6300 ICP system with colorimetric determinations. The specific methods include, for chloride: Lachat QuikChem Method 10-117-07-1-C (FIA); for nitrate: Lachat QuickChem Method No. 12-107-04-1-B, using a cadmium-reduction column; and for ammonium: Latchat QuickChem Method No. 12-107-06-1-B.

DNA Extraction Methods

DNA-protective buffer was removed from Sterivex filters by a sterile 1 ml syringe. A hammer was then used to gently crack open Sterivex cartridges. Using a sterile razor blade, filter membranes were then removed, cut into halves, and added to screw cap tubes with ~0.25 g of sterile 0.1 mm zirconium beads, 0.25 mg/ml proteinase K, and 1 mg/ml lysozyme. Tubes were bead beat at 3000 rpm for 30 sec. Tubes were incubated at 55°C for 20 min, then 70°C for 5 min to inactivate enzymes. The solution was removed from tube, and added to 1 ml phenol:chloroform:isoamyl alcohol (25:25:1), inverted to mix, and incubated at

55°C for 5 minutes. The mixture was then centrifuged at 13,000 rpm for 10 min to separate phases, at which point the aqueous phase was transferred to a fresh tube. The mixture-incubation-phase separation steps were then repeated, and the aqueous phase decanted. The same steps were then repeated, but using 500 µl of pure chloroform. DNA was then precipitated with 10% 3M sodium acetate and 0.6 volumes of isopropanol. The mixture was then incubated overnight at -20°C. The next day, the mixture was centrifuged at 13,000 rpm for 10 min, supernatant removed, washed in 1 ml 70% ethanol, and centrifuged again at 13,000 rpm for 10 min. The supernatant was removed, and the pellet was resuspended in 50 µl of TE buffer, pH 8. The mixture was then allowed to sit overnight at 4°C, before being centrifuged down to trap any condensation. The extracted DNA solution was then stored at -20°C until use.

Preparation and sequencing of 16S rRNA gene clone libraries

Community genomic DNA extracted from each of the 10 wells, drilling mud source pond, and controls for potential laboratory contamination, consisting of DNA preservation buffer and MilliQ water, were PCR amplified with universal Bacterial small subunit 16S rRNA 357 forward (SSU_357_F) and 1100 reverse (SSU_1100_R) primers using *Taq* DNA Pol (New England Biolabs) to target the V3-V5 hypervariable regions of the 16S rRNA gene. All primer sequences are listed in Table S1. Final primer concentrations were 0.18 µM. Template, $MgCl_2$, and bovine serum albumin (BSA) concentrations were optimized to reduce impacts from PCR inhibition. As a result, the final PCR Master Mix included a final concentration of 1.88 µM $MgCl_2$ (including buffer $MgCl_2$ content) and 1.76 mg/ml BSA (New England Biolabs). 357F/1100R amplification was performed on the Veriti 96 Well Thermal Cycler (Applied Biosystems) using the following cycling conditions: initial denaturation at 95°C for 3 min, thirty cycles of denaturation at 95°C for 30 sec, annealing at 52°C for 30 sec and extension at 72°C for 1 min, before a final extension at 72°C for 7 min. PCR amplicons were ligated and transformed using the TOPO TA pCR4 cloning kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Transformed colonies were used as PCR template with M13 primers, generating amplicons

that were gel purified (QIAquick Gel Extraction Kit, Qiagen) and submitted for Sanger sequencing at Genewiz (Cambridge, MA).

Illumina MiSeq sample library preparation

All templates were amplified with primers PE_357_F and PE_806_R using Phusion High Fidelity Polymerase according to the following cycling conditions: initial denaturation at 96°C for 1:30 min, thirty cycles of denaturation at 96°C for 30 sec, annealing at 52°C for 30 sec, and extension at 72°C for 30 sec, followed by a final extension at 72°C for 5 min. After PCR amplification, amplicons were purified using Exo-SAP IT (Affymetrix) according to manufacturer's instructions. 1 μl of purified product was used as template in the subsequent Illumina adaptor addition and barcoding PCR cycle, using universal forward primer PE_SEQ_F and six nucleotide barcode-specific reverse primers PE_SEQ_R at final concentrations of 0.5 μM each, according to the following cycling conditions: initial denaturation at 96°C for 1:30 min, fifteen cycles of denaturation at 96°C for 30 sec, annealing at 65°C for 30 sec, and extension at 72°C for 30 sec, followed by a final extension at 72°C for 5 min. Barcodes were designed with three base differences between any two barcodes. The amplicons were loaded on a 2.0% agarose electrophoresis gel and run at 90V for 60 minutes. DNA bands of the appropriate size were excised from the gel using individual UV-treated scalpels and purified using the QIAquick Gel Extraction Kit (Qiagen). Purified barcoded samples were eluted in 30 μl DEPC water. The relative concentration of each barcoded sample was determined by qPCR. Each sample was used as template in triplicate 20 μl reactions using Illumina sequencing primers (0.5 μM final concentration) and 1X fluorescent Sybr. Samples were multiplexed in volumetric proportions based on minimum cycle number in order to enable equal reads per sample, following Preheim et al. (2013). Pooled samples were concentrated using the MinElute Reaction Cleanup Kit (Qiagen), eluted in 11 μl of DEPC water

16S rRNA clone library sequence processing

Clone library Sanger sequences were subjected to quality filtering and trimming using CLC Genomics Workbench 7. Sequences with quality scores >0.05 or fewer than 400 bases were excluded from analysis. Chimeras were removed using the

UPARSE command *uchime_ref* with the gold.db ChimeraSlayer database (Broad Microbiome Utilities). Sequences were annotated using the RDP classifier (Wang *et al.*, 2007) and Silva 16S database (Quast *et al.*, 2013).

16S rRNA Illumina sequence processing, OTU clustering and annotation

Paired-end sequences were merged (fastq_mergepairs) and filtered by designated overlap requirements (fastq_truncqual 3; fastq_minovlen 16). Merged pairs were then mapped using CLC Genomic Workbench to the PhiX virus genome (used by Illumina as an internal sequencing control). PhiX-mapped sequences were removed from downstream analysis. Merged pairs were then filtered according to the recommended expected error threshold (fastq_maxee 3.0; Robert Edgar, Personal Communication). Filtered merged pairs were then trimmed to a length of 385 bases (fastq_filter; fastq_trunclen 385). Sequences shorter than 385 bp were discarded. Trimmed sequences were dereplicated (derep_fulllength) and annotated by number of replicate reads per sequence (sortbysize). All singleton sequences were discarded (minsize 2), per recommended settings. Remaining sequences were then clustered to form OTUs (cluster_otus) at a 97% minimum identity threshold. After chimeric OTU were removed (uchime_ref), merged paired-end reads were then mapped to OTUs using the USEARCH algorithm (usearch_global) at a 97% minimum identity threshold. Readmaps for each sample were then converted into OTU tables using the python script *uc2otutab.py*. OTUs were taxonomically annotated by the python script *assign_taxonomy.py*, using the RDP classifier (Wang *et al.*, 2007) and Silva 16S database (Quast *et al.*, 2013) with a minimum 60% identity RDP confidence threshold. OTUs that displayed less than 60% confidence to any RDP assignment on the phylum level were discarded (n = 6). Remaining OTUs annotated as the PhiX virus (used as MiSeq internal standard) were removed from downstream analysis.

Binning of unique *Sulfurospirillum* (MD31, MD32, MD102) genomes

Binning of contigs in Metagenome 3 yielded a single genomic bin (63 contigs/5.9 Mbp) containing two strains of *Sulfurospirillum* (based on total number and best

Blastp hits of single copy genes (214 copies, assuming a single genome contains roughly 107 genes) and presence of two copies of non-overlapping 16S rRNA gene fragments). This finding is consistent with clustering of 16S rRNA amplicons (97%), which yielded two OTUs affiliated with *Sulfurospirillum* in which one OTU is more dominant than the other (~0.1%). Attempts to separate the two strains of *Sulfurospirillum* (63 contigs; 5.9 Mbp) using GC content (32-48%), tetranucleotide frequencies and contig coverage were unsuccessful. These contigs displayed comparable coverage (~150X) likely because high genome similarity made mapping of short reads indiscriminate. Blastn comparison between the 16S rRNA gene of *Sulfurospirillum* sp. MD102 (single copy of 1431 bp) and the two copies detected in the two strains of *Sulfurospirillum* in Metagenome 3 (fragment of 676 bp of 1088 bp contig (*Sulfurospirillum* MD32) and 523 bp of 1605 bp (*Sulfurospirillum* MD31) aligned partially with that in *Sulfurospirillum* MD102 revealed that one partial fragment (contig 442 in MD31) of the two 16S rRNA genes is 100% similar (overlapping region of 523 bp in MD31) with that in *Sulfurospirillum* MD102, while the second copy has 99% similarity with *Sulfurospirillum* MD32. Due to the similarity of 16S rRNA gene and results from clustering of 16S rRNA gene amplicon showing that MD10 and Metagenome 3 share a single dominant OTU affiliated with *Sulfurospirillum*, all 62 contigs from Metagenome 3 *Sulfurospirillum* bins were subjected to Blastn comparison against all 590 contigs in *Sulfurospirillum* MD102. Contigs in *Sulfurospirillum* MD31 having 100% nucleotide sequence similarity with at least 300 bp overlapping region with MD102 were preliminarily grouped in one genomic bin (*Sulfurospirillum* MD31; 34 contigs) with the remaining contigs assigned to *Sulfurospirillum* MD32 (29 contigs). Several instances suggest that this approach has separated the two strains of *Sulfurospirillum* into their proper respective bins. The number and category of single copy genes were equally distributed in both MD31 and MD32. Furthermore, two-way comparison using psi-Blast of ORFs among these three genomes and reference genome of *Sulfurospirillum deleyianum* DSM 6946 (NC_013512.1) shows that *Sulfurospirillum* MD31 and MD32 in comparison to *Sulfurospirillum* MD10 and S. *deleyianum* DSM 6946 have similar abundance and distribution of their sequence homolog (i.e. ORFS having >60% in amino acids; Figure S5). The close similarity between MD10 and MD31 is supported by 100% similarity of the partial 16S rRNA gene, while comparison of

Average Nucleotide Identity (ANI) between MD10 and MD31 showed that the two genomes displayed 99% ANI (9,110 fragments in 200 bp step size and 1,000 bp reading windows), while MD10 and MD32 had 81.4% AVI (4,200 fragments).

Autotrophic pathway "completeness"

Certain key genes must be present in binned genomes for canonical autotrophic pathways to be considered "complete." For example, according to RAST, two enzymes, RuBisCO and phosphoribulokinase, when present together, can confidently be used as indicators of the presence of the Calvin Cycle in a given organism. For the Wood-Ljungdahl (Reductive Acetyl-CoA) Pathway, the bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase enzyme (*acs*AB), which catalyzes the reactions from CO_2 to CO and from CO_2 to a methyl group, must be present for the pathway to be considered complete. Lastly, for the reductive TCA Cycle, four molecules of CO_2 are fixed by three crucial oxygen-sensitive enzymes, all of which must be present for the pathway to function: ATP-citrate lyase, 2-oxoglutarate oxidoreductase, and pyruvate:ferredoxin oxidoreductase.

Archived false positive (AFP)

The following criteria were used for OTU retention based on AFP sequence distribution: 1) Any OTU whose maximum relative abundance in a sample was at least ten times higher than its abundance in the AFP was retained. The factor of 10 cut off was implemented based on the assumption that abundances below this threshold may represent cross-contamination during the failed PCR run that originated in the AFP, and 2) Any OTUs that shared the same top Blastn hit as a removed OTU, and whose maximum percentage abundance in a sample was less than its abundance in the AFP, were also discarded. This second step targeted OTUs generated from sequencing errors that might represent false diversity from laboratory contaminants. The dominant taxa associated with the sequenced AFP sample control are presented in Table S4.

3

DEVELOPMENT AND CHARACTERIZATION OF BIOPROSPECTED STRAIN *BACILLUS MEGATERIUM* SR7 FOR ENHANCED GROWTH AND BIOPRODUCTION UNDER SUPERCRITICAL CO_2

ABSTRACT

A microbial bioproduction system that utilizes supercritical carbon dioxide ($scCO_2$) for non-polar product extraction would uniquely enable relief of end-product toxicity effects in a self-sterilizing environment. While previous studies have successfully demonstrated a broad diversity of biocatalytic reactions utilizing $scCO_2$ as a solvent and/or substrate, $scCO_2$ has previously been considered inaccessible to active microbial product biosynthesis due to its lethal effects on most microbes. Therefore, a bioprospecting approach was utilized in an attempt to isolate $scCO_2$-resistant microbial strains through enrichment culture and serial passaging of deep subsurface fluids from McElmo Dome $scCO_2$ reservoir. This approach enabled the isolation of six unique spore-forming *Bacillus* strains with the potential to serve as bioproduction host strains. After demonstrating superior growth under $scCO_2$ in pure culture when inoculated as endospores, strain *Bacillus megaterium* SR7 was selected for in depth characterization and development in order to enable eventual use in a dual-phase bioreactor scheme where bioproducts generated *in situ* would be extracted by the $scCO_2$ sustainable solvent phase. After sequencing *B. megaterium* SR7's 5.51 Mbp genome, natural metabolite profiles affirmed the strain's use of glycolytic pathways, the TCA Cycle and fermentation for anaerobic energy generation, as expected by functional genomic annotations. Process improvements, including optimized minimal medium formulation and altered mixing regimes, established consistent growth at 1 atm CO_2 as a higher throughput model system for $scCO_2$ conditions. Based on findings from 1 atm CO_2 cultures, including the ability to chemically-induce spore germination to improve growth frequency, L-alanine-amended semi-defined minimal media facilitated improved SR7 growth outcomes under $scCO_2$. The detection of extracellular natural fermentative products under $scCO_2$ serves as an endogenous proof of concept for heterologous production of engineered bioproducts in this unique bioreactor environment.

3.1 INTRODUCTION

Academic and industrial research has increasingly focused on supercritical carbon dioxide ($scCO_2$) as an environmentally benign and inexpensive substitute for conventional organic solvents that are typically hazardous, toxic, and/or flammable. As a result, $scCO_2$ has been extensively explored as a solvent for both *in vitro* and *in vivo* biocatalysis reactions that are difficult or expensive in aqueous phase reactors. Due to $scCO_2$'s non-polar chemistry, a specific focus of previous development has been on polar, hydrophobic compounds that readily partition into the $scCO_2$ solvent phase, enabling relief of end-product toxicity effects and purified product extraction. Broad classes of *in vitro* biocatalyzed reactions have been demonstrated in $scCO_2$, including amidation, esterification (Nakamura *et al.*, 1986; Marty *et al.*, 1992), and acetylation (Wimmer and Zarevúcka, 2010). Biocatalyzed reactions in $scCO_2$ are also known to generate enantiopure compounds with a single "handedness" or chirality, which are otherwise difficult to synthesize (Matsuda *et al.*, 2000; Matsuda *et al.*, 2008; Matsuda *et al.*, 2004; Salgin *et al.*, 2007). Biofuels are especially compelling with regard to $scCO_2$ harvesting systems because the moderately hydrophobic chemistry of alcohols like butanol would enable compound partitioning from the aqueous phase into $scCO_2$ (i.e. octanol-water partition coefficient (K_{ow} >4); Timko *et al.*, 2004).

Due to its potential to reduce dependence on carbohydrate-based feedstocks, CO_2-fixing carboxylation reactions are of particular interest with regard to the future of sustainable microbial-facilitated bioproduction. A wide range of *in vitro* studies have demonstrated enzymatic fixation of the $scCO_2$ solvent, including in the synthesis of urethane (Yoshida *et al.*, 2000) and styrene carbonates (Kawanami *et al.*, 2000), and pyrrole-2-carboxylate, which was catalyzed by purified decarboxylases from *Bacillus megaterium* (Wieser *et al.*, 1998; Yoshida *et al.*, 2000; Wieser *et al.*, 2001). A variety of central carbon metabolism enzymes have demonstrated *in vitro* $scCO_2$-fixation, including RuBisCO (Calvin Cycle; Hartman and Harpel, 1994), isocitrate dehydrogenase (Reverse TCA Cycle; Sugimura *et al.*, 1989), and malate dehydrogenase (Reverse TCA Cycle; Sugimura *et al.*, 1990; Matsuda, 2005). Beyond its use as a chemical

reagent, academic and industrial research has increasingly focused on $scCO_2$ as an environmentally benign and inexpensive substitute for conventional organic solvents that are typically hazardous, toxic, and/or flammable. In biological applications, $scCO_2$ is used for coffee decaffeination, product extraction, and select enzymatic reactions (Hammond *et al.*, 1985; Matsuda *et al.*, 2005; Matsuda *et al.*, 2008, Salgin *et al.*, 2007). As a result, while CO_2 is typically considered a waste product with a dangerous global impact potential, it also represents a directly useful, abundant resource that may be employed as a solvent and/or substrate with broad biotechnological applications.

The ability to conduct industrially relevant biosynthesis reactions by accessing $scCO_2$ as the solvent is limited by the availability of purified enzymes (which are progressively degraded by $scCO_2$ exposure) and the need to replenish small molecule reductants (i.e. NAD(P)H) to regenerate the active state of the enzymatic catalyst. Therefore, microbial growth under $scCO_2$ enables the cellular protection and regeneration of relevant biocatalytic compounds, relieving the need for supplementation. This principle has been demonstrated previously in limited fashion with reports of biocatalyzed ketone reduction by immobilized *Geotrichim candidum* cells and carboxylation of pyrrole using *Bacillus megaterium* cells (Matsuda *et al.*, 2000, 2001). While immobilized *G. candidum* cells were not metabolically active, the authors speculated that *B. megaterium* remained viable during the experiment and carried out active biocatalysis for $scCO_2$-reduction (Matsuda *et al.*, 2000, 2001). Cultures demonstrating robust growth under $scCO_2$ would therefore represent a significant improvement in the overall $scCO_2$ bioproduction development process.

Although $scCO_2$ has been identified as an attractive solvent for *in situ* extraction of fermentation products demonstrating chemical compatibility, its rapid sterilizing effects on nearly all bacteria have rendered this promising technology ineffective with living cells (Khosravi-Darani and Vasheghani-Farahani, 2005; Knutson *et al.*, 1999) due to cellular membrane disruption, desiccation, enzyme inactivation and cytosolic acidification (Spilimbergo and Bertucco, 2003). However, several recent studies have demonstrated the resilience of certain microbial populations that are able to withstand $scCO_2$ exposure in natural (Thesis Chapter 2; Mu *et al.*, 2015) and laboratory systems (Mitchell et al., 2008). Specifically, Peet *et al.* (2015) established that certain spore-forming

taxa can survive and grow in batch bioreactors in aqueous media with a $scCO_2$ headspace (Peet *et al.*, 2015). As a result, there appears to be a possibility of developing a $scCO_2$-tolerant bacterial strain for engineered product biosynthesis that may harness the utility of sustainable solvent $scCO_2$. Moreover, the largely lethal effect of $scCO_2$ on cells represents an engineering advantage by enabling the growth of a single $scCO_2$-resistant strain, while otherwise maintaining a sterile bioreactor environment free of contamination, a common challenge especially for scaled up bioproduction facilities.

In order to develop a microbial growth and bioproduction system, it is imperative to meet following goals: 1) identify strains that are biocompatible with $scCO_2$ as an extraction solvent, 2) optimize culturing conditions for these strains by process engineering, and 3) develop protocols for genetic modification enabling heterologous protein expression. Building on the $scCO_2$ enrichment and culturing protocols developed in Peet *et al.* (2015), which proved effective at isolating facultative aerobic/anaerobic spore-forming bacteria (genus *Bacillus*), this study sought to advance development of a system for bioproduction and *in situ* extraction by addressing the first two stated goals prior to genetic system development. It was hypothesized that a bioprospecting approach targeting microbes exposed to $scCO_2$ over long periods in the environment will enable the isolation of strains that have evolved natural mechanisms for enhanced $scCO_2$-resistance and growth potential. Analogously, the *Taq* polymerase variant isolated from *Thermus aquaticus* displays thermophilic properties due to its evolution upon long-term host emplacement in a high temperature geothermal system (Chien *et al.*, 1976). The isolation and genetic system development of environmental strains specifically adapted to $scCO_2$ may therefore hold a similar potential for exploiting naturally evolved phenotypes in establishing new solutions to current biotechnological challenges.

In the event that $scCO_2$-tolerant strains are isolated, a significant factor limiting the ability to initiate metabolic modifications for industrial applications is the challenge of genetic intractability (addressed in detail in Thesis Chapter 4). Previous work establishing genetic systems has enabled the investigation and exploitation of bioprospected strains with unique biochemical capacities with applications in wastewater treatment/bioremediation (Coppi *et al.*, 2001), and the production of pharmaceutical and agricultural agents (Xiong *et al.*, 2013).

Engineering of *Bacillus* strains has been broadly successful, including in the production of biofuels and industrially relevant compounds (Nielsen *et al.*, 2009; Hu and Lidstrom, 2014). *B. megaterium* in particular has garnered popularity as a host for genetic engineering due to its high secretion capacity, lack of exotoxins, and ability to maintain plasmids (Korneli *et al.*, 2013). Promisingly, strains QM B1551 and DSM319, and their derivatives, have been used as hosts for protein and bioproduct expression for over 30 years (Vary *et al.*, 2007).

To isolate strains able to grow in biphasic $scCO_2$-water reactors, fluid samples were collected from the deep subsurface McElmo Dome CO_2 field in Colorado, where $scCO_2$ had accumulated over 40-70 million years. Metagenomic analysis of formation fluids at this site suggested existence of an anaerobic microbial ecosystem (Thesis Chapter 2). It was hypothesized that by using fluid samples from McElmo Dome as inocula for enrichment passaging under $scCO_2$ conditions, we should be able to isolate strains that consistently exhibit robust growth under $scCO_2$. Following isolation of strain *Bacillus megaterium* SR7, improvement in growth was achieved through optimization of 1) a semi-defined minimal growth medium supplemented with nutrient and metals amendments, 2) culture conditions through phenotypic fingerprinting and mixing regime growth assays, and 3) spore germination frequency under $scCO_2$ by exposure to chemical inducer L-alanine, as confirmed by flow cytometry (FCM). Isolate genome sequencing enabled annotation of relevant pathway genes and an endogenous promoter system. This work contributes to the establishment of a new technology for microbial bioproduction by enabling a bacterial strain capable of bioproduct generation to access the unique properties of sustainable solvent supercritical carbon dioxide and paves the way for future bioengineering for enhanced generation of heterologous bioproducts under $scCO_2$ (Thesis Chapter 4).

3.2 Methods

Subsurface fluid sample collection and storage

Formation water samples from the McElmo Dome CO_2 field were used as inocula for microbial strain isolation through $scCO_2$-exposed enrichment culture and passaging. Sample fluids were sourced from the deep subsurface, where CO_2

is trapped at depths of 1800 to 2600 m within the 100 m thick dolomitic Leadville Formation (Allis *et al.*, 2001; Gilfillan *et al.*, 2009) and exists as a supercritical fluid at a temperature and pressure of approximately 65°C and 135 atm (Allis *et al.*, 2001). Sampled fluids from each of ten CO_2 production wells (operated by KinderMorgan CO_2) were collected from fluid-gas separators that were decanted and filled 15 hours prior to sample collection. At each separator, one liter of degassed fluid was collected in an acid-washed bottle (Nalgene) and placed immediately on ice for use as enrichment culture inocula. Fluids were shipped on ice and stored at 4°C.

Supercritical CO_2-exposed enrichment passaging

Culturing media and vessels

Media for enrichment culture and passaging of McElmo Dome samples was a modified version of MS media (Colwell *et al.*, 1997) with supplements targeting diverse metabolic groups as described in (Peet *et al.*, 2015). Media consisted of (in g/l) 0.5 yeast extract, 0.5 tryptic peptone, 10.0 NaCl, 1.0 NH_4Cl, 1.0 $MgCl_2 \cdot 6H_2O$, 0.4 K_2HPO_4, 0.4 $CaCl_2$, 0.0025 EDTA, 0.00025 $CoCl_2 \cdot 6H_2O$, 0.0005 $MnCl_2 \cdot 4H_2O$, 0.0005 $FeSO_4 \cdot 7H_2O$, 0.0005 $ZnCl_2$, 0.0002 $AlCl_3 \cdot 6H_2O$, 0.00015 $Na_2WoO_4 \cdot 2H_2O$, 0.0001 $NiSO_4 \cdot 6H_2O$, 0.00005 H_2SeO_3, 0.00005 H_3BO_3, and 0.00005 $NaMoO_4 \cdot 2H_2O$. MS medium supplements (g/l) consisted of 0.5 glucose (MS-FM); or 1.3 MnO_2, 2.14 $Fe(OH)_3$, and 1.64 sodium acetate (MS-MR); or 0.87 K_2SO_4, 0.83 $FeSO_4$, 0.82 sodium acetate (MS-SR). Following enrichment culturing and three passages using MS medium, Luria-Bertani Broth (LB) (Difco) was included as an additional growth media for $scCO_2$ culturing. Phosphate buffered LB (P-LB) is amended with 50 mM K_2HPO_4. During all rounds of culturing, CO_2-incubated media was amended with 0.25 g/l of reducing agent Na_2S (at 0.25 g/l) and 0.001 g/l of the redox indicator resazurin. A summary of all media utilized during enrichment passaging and subsequent culturing is presented in Table 1.

High-pressure culturing vessels were constructed of $\frac{3}{4}$ inch 316 stainless steel tubing for a 10 ml total capacity, and fitted with quarter turn plug valves (Swagelok or Hylok). Between uses, all vessel components were cleaned and soaked for at least two hours with 10% bleach and detergent, then autoclaved

Table 1. Summary of microbial culturing media used in study.

Use	Name	Base	Supplements	Base Reference
Enrichment Passaging	**MS-MR**	**MS:** Yeast Extract, Trypticase Peptone, Salts	[a]Metals (Mn, Fe) + Acetate	Colwell et al., 1997
	MS-SR		[b]Sulfates + Acetate	
	MS-FM		0.5 g/ L Glucose	
Pure Culture	**P-LB**	**LB:** Yeast Extract Tryptone, NaCl	50 mM Phosphate (Dibasic)	BD Difco™
	P-LBA		50 mM Phosphate (Dibasic) + 100 mM L-alanine	
	P-LBL		50 mM Phosphate (Dibasic) + 10 mM L-leucine	
	P-LBAL		50 mM Phosphate (Dibasic) +100 mM L-alanine + 10 mM L-leucine	
	M9+	**M9:** Phosphates, Salts	M9 + [c]0.1X trace metals + 50 mM YE + 0.4% Glucose	thelabrat.com*
	M9A+		M9 + [c]0.1X trace metals + 50 mM YE + 0.4% Glucose + 100 mM L-alanine	

*www.thelabrat.com/protocols/m9minimal.shtml
[a]1.30 g/l MnO_2, 2.14 g/l $Fe(OH)_3$, 1.64 g/l Na-Acetate
[b]0.87 g/l K_2SO_4, 0.83 g/l $FeSO_4 \cdot 7H_2O$, 0.82 g/l Na-Acetate
[c]0.1X trace metals solution (Boone, 1989); see methods

prior to assembly. All tubing in the pressurization manifold was filled before use with 10% bleach for 30 minutes, flushed with MilliQ-H_2O, rinsed with 70% ethanol for 20 minutes, and dried with CO_2 gas. Prior to reactor loading, culture media was added to 100 ml serum bottles and degassed with a stream of 100% CO_2 for 30 minutes. Vessels were then filled to $\frac{1}{2}$ capacity (5 ml) with inocula and degassed media, after which the headspace was pressurized with extraction grade CO_2 gas at a rate of 2-3 atm min^{-1} until reaching a final pressure of 100 atm. Since the CO_2 tank used for reactor pressurization contained a helium (He) cushion (in order to reach elevated pressures) the gas tank mixture contained 97-99% CO_2. Unless stated otherwise, after pressurization, reactors were incubated in a 37°C warm room (to reach supercritical conditions) with shaking at 100 rpm to increase the extent of inocula and media mixing.

As described previously (Peet *et al.*, 2015) prior to depressurization, culturing vessels were connected via 316 stainless steel tubing and fittings to a pressure gauge (Hunter) to measure the final vessel pressure. All reported vessel incubation data maintained pressures above the CO_2 critical point (>72.9 atm) when mixed with ≤3% inert Helium at 37°C (Roth, 1998). Reactors were depressurized at a rate of 3-5 atm min^{-1} over approximately 30 min, at which point the vessels were transferred to the anaerobic chamber for sub-sampling, glycerol stock preparation and passaging.

Enumeration of cell density

In order to quantify biomass of CO_2 cultures, 0.5-1.0 ml samples were treated with Syto 9 stain (Life Technologies), left in a dark room for 15 minutes

to allow the stain to adhere to nucleic acids, collected on 0.22 um polycarbonate filters (Whatman Nucleopore) by vacuum pump, and washed twice with phosphate buffered saline (PBS) to remove excess stain. Each filter was mounted on glass slides for visualization by epifluorescent microscopy (Zeis Axioscope) with immersion oil dropped below and above the filters, after which a cover slip was applied. Filters were stored at 4°C in the dark until use. Cell densities were extrapolated by multiplying individual cell counts in a 10x10 microscope eyepiece grid by a dilution factor (if <1.0 ml of sample was filtered), and then multiplied by $3.46x10^4$, because a 10x10 grid at 1000X magnification corresponds to $1/(3.46x10^4)$ of a 25 mm filter. Final cell densities represented the mean values of cell counts in 15-20 separate 10x10 grids/sample. The limit of detection was considered to be one half of a cell per 15 grids, which corresponds to 1150 cells/ml. Fluorescent images were captured on a Nikon D100 camera using the NKRemote live-imaging software. Cell density calculations and morphological observations were conducted for inocula prior to pressurization as well as after incubation in order to determine the extent to which growth had occurred. CFU plating was performed using LB Agar with order of magnitude dilutions in autoclaved PBS buffer prior to plating with cell-free negative controls. Plates were dried, inverted and incubated overnight at 37°C prior to colony counting.

Enrichment cultures and serial passaging

Fluids from four of the sampled wells (HB-5/Well 2, HE-1/Well 4, HF-3/Well 5, YB-4/Well 7) were selected as inocula for enrichment culturing under $scCO_2$ because they appeared to harbor elevated cell density by fluorescent microscopy (Thesis Chapter 2). 100 ml of fluids from the four respective wells were filtered onto 5 mm, 0.2 μm pore size, polycarbonate filters (Nucleopore) in order to concentrate microbial content. The filters were then placed inside an anaerobic chamber (Coy Lab Products) containing 95% CO_2 / 5% H_2. Using sterilized tweezers, filters were then placed inside 10 ml 316 stainless steel pressure reactors with 1 ml of formation fluid from the same well used to concentrate biomass on the filter. The filters and formation fluids were then amended with 4 ml of growth media. After the initial round of growth using filter-concentrated biomass inoculum, cultures were inspected by epifluorescence microscopy to identify biomass accumulation. Cultures that showed growth

relative to inocula based on cell counts were serially passaged by dilution in freshly degassed growth media to achieve initial concentrations of approximately 10^4 cells/ml. The McElmo Dome enrichment culture (M1) was incubated for 46 days, while subsequent passages were incubated for 19 days (M2), 33 days (M3), and 35 days (M4).

Strain isolation and $scCO_2$ growth verification

Samples from the final round of passaging (M4) were plated on LB agar and incubated overnight at 37°C under aerobic conditions at ambient pressure. Single colonies with unique morphologies were used to inoculate liquid LB. DNA extraction from overnight grown cultures was performed using the Qiagen Blood and Tissue DNA extraction kit protocol for gram-positive cells (Qiagen). Resulting DNA was used as template for 16S rRNA PCR using universal Bacterial primers 515F (3'–GTGCCAGCMGCCGCGGTAA–5') and 1492R (3'–GGTTACCTTGTTACGACTT–5'; Turner et al., 1999). PCR mixtures (20 µl per reaction) included 1X Phusion High Fidelity Polymerase buffer, 0.4 uM of each primer (IDT), 0.4 uM deoxynucleotide mixture and 1 U Phusion Polymerase (New England Biolabs). Thermal cycling conditions consisted of an initial 5 minutes at 95°C followed by 30 cycles of denaturation at 95°C for 30 sec, annealing at 55°C for 30 sec, and extension at 72°C for 90 sec; followed by a final extension time of 7 min. Every PCR reaction included negative and positive controls (Peet *et al.*, 2015). PCR products were then purified using Exo-SAP IT (Affymetrix) and submitted for Sanger sequencing (Genewiz, Cambridge, MA). Returned sequences were processed in CLC Genomics Workbench (Version 7), including primer removal and universal sequence trimming to 918 bp for all isolates. Sequence alignment and tree building of isolates and reference sequences consisting of *Bacillus*, closely related taxa, and an *E. coli* outgroup using a 914 bp alignment was also conducted in CLC Genomics Workbench. Tree building used a bootstrapped (100X) neighbor-joining method, which was visualized in FigTree v1.4.2. 16S rRNA reference sequences were downloaded from Genbank (NCBI) or generated in Peet et al., (2015; i.e. *Bacillis* sp. OT1, *Bacillus* sp MIT0214).

Because *Bacillus spp.* spores were previously demonstrated to be able to germinate and grow under 1 atm CO_2 and $scCO_2$ headspace conditions (Peet *et al.*, 2015), spores of all *Bacillus spp.* strains isolated from McElmo Dome fluids were prepared using the protocol described in Kim and Goepfert (1974). Briefly, colonies streaked from glycerol stocks were used to inoculate overnight cultures in LB medium that were incubated under aerobic conditions at 37°C while shaking at 100 rpm. Dense, stationary-phase cultures were then diluted 1:50 into 100 ml of Modified G Medium, which is composed of (in g/l): yeast extract 2.0, $CaCl_2 \bullet 2H_2O$, 0.025, K_2HPO_4 0.5, $MgSO_4 \bullet 7H_2O$ 0.2, $MnSO_4 \bullet 4H_2O$ 0.05, $ZnSO_4 \bullet 7H_2O$ 0.005, $CuSO_4 \bullet 5H_2O$ 0.005, $FeSO_4 \bullet 7H_2O$ 0.0005, $(NH_4)_2SO_4$ 2.0, adjusted to pH 7.1 after autoclaving. Modified G Medium-inoculated cells were incubated at 37°C for 72 hours to induce sporulation, then centrifuged for 10 minutes at 10,000 x g. The pellet was resuspended and centrifuged 5 times in autoclaved wash buffer containing 0.058 g/l $NaH_2PO_4 \bullet H_2O$ and 0.155 g/l $Na_2HPO_4 \bullet 7H_2O$ with 0.01% (v/v) Tween20 to prevent clumping. Spores were stored in wash buffer at 4°C until use and periodically assayed for continued viability after extended storage by LB agar colony plating.

Growth of *B. megaterium* SR7 (and other isolates) under $scCO_2$ was validated by triplicate incubation for 28-42 days inoculated in pure culture from spores loaded at $\sim 1x10^4$ spores/ml using multiple media (Table 1). Cultures were scored for growth by filter counts, as previously described.

Isolate strain SR7 genomics

Understanding the genomic landscape of strain SR7 provides useful insight into endogenous physiological and metabolic capacities and will aid future development of SR7 as a strain for bioengineered product generation for *in situ* $scCO_2$ extraction. SR7 genomic DNA was extracted from a 10 ml overnight aerobic LB culture using the Qiagen Blood & Tissue Kit, following the Gram-positive bacteria protocol. Eluted DNA was submitted to the MIT BioMicro Center for sequencing using PacBio SMRT technology. Following sequencing, the PacBio assembler software was used to assemble SR7 contigs, which were then compared to the genome of closely related strain *B. megaterium* QM B1551 (Eppinger *et al.*, 2011) using the online tools nucmer and "double act" (www.hpa-

bioinfotools.org.uk/pise/double_act.html), the latter of which cuts the query and reference DNA into smaller pieces to create an inter-genome Blastn comparison file that can be viewed in the Artemis Comparison Tool (ACT; Carver *et al.*, 2005). Based on the ACT comparison, the putative SR7 chromosome (longest contig) was adjusted to start at the beginning of gene *dnaA* in agreement with the reference genome. The closed chromosome was then plotted by DNA Plotter (sanger.ac.uk/resources/software/dnaplotter) and submitted to RAST (Aziz *et al.*, 2008) for gene prediction and functional annotation. Remaining contigs, potentially indicative of endogenous plasmid based on sequenced *B. megaterium* strains, were also submitted to RAST for annotation. Shared and unique RAST-annotated genes between SR7 and *B. megaterium* reference genomes QM B1551, DSM319 (Eppinger *et al.*, 2011), and WSH-002 (Liu *et al.*, 2011) were determined using online tool Venny 2.1. Inter-strain sequence comparisons were conducted using the Average Nucleotide Identity (ANI) calculator (http://enve-omics.ce.gatech.edu/ani).

Physiological characterization of strain SR7

To help guide optimization of growth conditions for strain SR7, physiological tests were conducted under aerobic conditions. To determine tolerance for pH, salinity and bicarbonate, high throughput culturing was done in 96 well plates and scored for growth by OD_{600} using a microplate reader (BioTek Synergy 2). 200 uL LB solutions/well were inoculated in triplicate with 10^4 spores/ml (based on SR7 spore stock filter counts) and incubated on a plate rocker at 37°C with unamended positive and cell-free negative LB controls. Tests for pH tolerance (pH 2-10) were conducted in LB medium amended with HCl or NaOH. The effect of salinity and bicarbonate on growth was determined by adding NaCl (1-10%) and $NaHCO_3$ (0.1-0.5M), respectively, to LB media. Optimal SR7 growth temperature was tested by inoculating 10^4 spores/ml in 5 ml of LB in triplicate at temperatures of 9-55°C. Cultures and cell-free negative controls were incubated without shaking, with subsamples taken for periodic OD_{600} measurements. SR7 antibiotic sensitivity was determined by supplementing 5 ml LB with ampicillin (5-50 ug/ml), chloramphenicol (3.5-35 ug/ml), kanamycin (5-50 ug/ml), spectinomycin (10-100 ug/ml), streptomycin

(10-100 ug/ml), or tetracycline (1.5-15 ug/ml). 5 mL cultures inoculated with 10^4 spores/ml were incubated in a spinning rack at 100 rpm for 24 hours at 37°C and assayed for growth by comparing OD_{600} measurements to unamended positive and cell-free negative LB controls.

Biolog GenIII Microplates 96 well plates (unamended and with a trace metals solution amendment (Boone *et al.*, 1989)) were used to determine SR7 growth substrates and to test growth sensitivities relative to a positive control. Plates were inoculated with 2-4 SR7 colonies grown overnight on solid BUG media (Biolog) such that starting OD_{490} transmission was 90-94%. Plates were incubated at 37°C on a plate shaker at 200 rpm and assayed for growth using NADH-dependent colorimetric changes measured by OD_{490} on a microplate reader (BioTek Synergy 2). Total growth was quantified by integrating the area under the curve of OD_{490} values over the course of the incubation, and categorized as: "-" displays an area less than the negative control, "+" is greater than the negative control, but less than half of the maximum value, and "+++" is between "+" and the maximum value.

Process improvements for SR7 growth under 1 atm CO_2 and $scCO_2$

After initial physiological characterization assays, subsequent culturing improvements sought to establish consistent, replicable growth of SR7 under $scCO_2$ by conducting experiments under 1 atm CO_2 as a proxy for pressurized conditions (e.g. Peet *et al.*, 2015). In order to improve spore germination frequencies, the effects of chemical inducers and mixing regimes (i.e. culture volume and shake speed) were examined, as the literature has shown certain compounds (i.e. amino acids, KNO_3, peptidoglycan, Ca-dipicolinic acid, and others; Ghosh and Setlow, 2009) and conditional treatments (temperature, pressure; Wei *et al.*, 2010) increase *Bacillus* spore germination rates. Experiments testing the role of mixing speed and modified culture media on rates of vegetative outgrowth were conducted under 1 atm CO_2 with CO_2-degassed media or buffer in 100 ml serum vials with clamped rubber stoppers.

Evaluating the effect of spore germination inducers

The effect of shake speed on spore germination was assayed by inoculating 5 ml LB medium with SR7 spores at a starting concentration of 10^5 spores/ml. Singleton cultures were subjected to shake speeds of 150, 250, and 350 rpm and scored for growth by OD_{600} and LB agar colony plating.

The ability to induce spore germination based on literature precedent was tested by inoculating triplicate 10 ml cultures of SR7 spores at OD_{600} 0.01 in LB amended with 100 mM L-alanine, LB subjected to a heat activation (65°C for 15 minutes) upon inoculation, or unamended LB as a control. Growth was scored by OD_{600} and LB agar CFU plate counts. In addition, sub-samples were heat-killed by exposure to 80°C for 10 minutes (Setlow, 2006) prior to plating, to ascertain remaining spore concentrations, as heat exposure is lethal to vegetative cells.

The role of candidate germination inducers was subsequently investigated in PBS buffer rather than growth medium to decouple the germination process from outgrowth. SR7 spores were loaded in triplicate 10 ml PBS amended with 100-250 mM L-alanine, 100 mM L-alanine with heat treatment, 25 mM L-leucine, or an unamended PBS control. The extent of germination was measured by fluorescence microscopy staining patterns (i.e. the degree of cell membrane penetration by DNA stain), bulk fluorescence, OD_{600}, and flow cytometry (FCM). A total of 100-300 cells per filter were counted and categorized as either "dormant" or "germinated" if the spore stain was localized to the cell membrane or diffused within the interior of the cell, respectively (Cronin and Wilkinson, 2007). Cells displaying an intermediate degree of stain dispersal ("activated") were categorized as germinated (Figure 1).

Sub-sample bulk fluorescence (Syto9) was measured by microplate reader (BioTek Synergy 2; 485/20 excitation, 528/20 emission) and OD_{600} was measured by microplate reader. OD should decrease in germinated cells (the index of

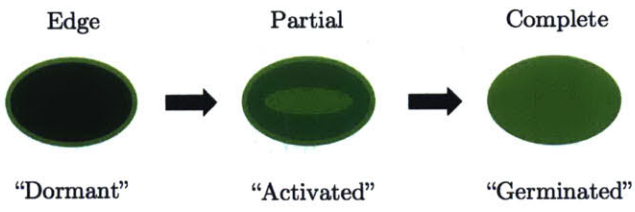


Figure 1. Expected (and observed) DNA staining patterns of differentially germinated SR7 spores.

refraction decreases due to hydration upon spore coat degradation) while bulk fluorescence should increase as the nucleic acid stain progressively penetrates and permeates the cell (Magge *et al.*, 2009). To test for germination after a delayed inducer spike rather than at the moment of inoculation, SR7 spores loaded at OD_{600} 0.01 were incubated overnight in 30 ml of PBS, passaged into PBS amended with L-alanine (25-250 mM) or L-leucine (10-25 mM) and then assayed for germination by bulk fluorescence and OD_{600} during incubation.

FCM was employed as a high-throughput germination assay based on Baier *et al.*, (2010). Triplicate cultures of SR7 spores loaded at OD_{600} 0.01 were incubated overnight in PBS and PBS amended with 2.5-250 mM L-alanine, along with cell-free PBS controls. Prior to loading on the flow cytometer (BD FACS Canto II HTS-1) cultures were diluted 1/50 in PBS and stained with Syto16 and propidium iodide (PI) in the dark for at least 30 minutes prior to analysis. After spore and media-only samples were used to set forward scatter, side-scatter, Syto16 and PI gates, sample data was collected and analyzed using FACSDIVA software.

Testing the effect of mixing on vegetative growth

After testing for the potential to induce germination in SR7, the next priority was to accelerate growth rates in order to increase metabolic activity for eventual product pathway expression. Experiments testing the role of shake speed on vegetative growth rate were inoculated with passaged cells of spore-loaded overnight cultures grown under 1 atm CO_2. Triplicate 25 ml LB cultures of vegetative cells loaded at OD_{600} 0.01 were subjected to shake speeds of 150, 250, and 350 rpm, with growth assayed by OD_{600} and LB agar colony plating.

Minimal media development to improve growth

Development of a minimal medium enables individual chemical components to be tuned in order to establish optimized growth from a sole carbon source under 1 atm CO_2. Initial attempts to generate SR7 growth in triplicate cultures tested various amendments to M9 base medium (thelabrat.com; Table 1), including 0.4% glucose or 0.4% xylose amendments as sole carbon sources, with or without trace metals solution (Boone *et al.*, 1989). The 1X concentration trace metals solution consisted of (in g/l): 0.005

Na_2(EDTA), 0.0002 $NiSO_4$•$6H_2O$, 0.0005 $CoCl_2$•$6H_2O$, 0.0001 H_2FeO_3, 0.001 $FeSO_4$•$7H_2O$, 0.0001 H_3BO_3, 0.001 $ZnCl_2$, 0.0001 $NaMoO_4$•$2H_2O$, 0.0004 $AlCl_3$•$6H_2O$, 0.001 $MnCl_2$•$4H_2O$, 0.0003 Na_2NO_4•$2H_2O$, 0.0002 $CaCl_2$. To further boost growth, triplicate cultures of M9 + 0.4% glucose were amended with dilute LB (0.001-0.01X) or yeast extract (YE; 0.001-0.01X) as de facto vitamin and co-factor solutions, and/or $NaNO_3$ (5 mM) as an alternative electron acceptor. All M9 incubations were scored for growth by OD_{600} and designated as robust above OD_{600} >0.600, low level between 0.2-0.6, and no growth below 0.2. Passaged vegetative cultures were also assayed in duplicate for growth (by OD_{600}) amended with a range (0.1X, 0.25X, 1X) of trace metals solutions in M9 + 0.4% glucose + 0.01X YE media, including in the presence and absence of 5 mM $NaNO_3$.

Growth curves under optimized shaking conditions were generated to establish baseline metabolic characteristics of strain SR7. Vegetative SR7 cells were passaged in quadruplicate at OD_{600} 0.01 in minimal or LB media and assayed for growth by OD_{600}, LB agar colony plating, and glucose consumption (for minimal medium cultures only) measured on the YSI 2900 with 2814 glucose starter kit. Doubling times were calculated using a log-linear fit of CFU and OD_{600} data points during exponential growth.

Analysis of SR7 fermentation products under 1 atm CO_2 and scCO_2

Following optimization of growth conditions under 1 atm CO_2 and scCO_2, identification of fermentation products would establish potential target pathways for redirecting carbon flux and would demonstrate the ability to generate extracellular natural products. Metabolite identification and quantification was conducted by high performance liquid chromatography (HPLC). Triplicate cultures of SR7 vegetative cells inoculated in M9+ or LB at OD_{600} 0.01 were scored for growth by OD_{600}. 500 ul of supernatant from each spun down sample (5 mins at 21,000 x g) was loaded on the HPLC (Agilent 1100 series) for analysis. Compound separation was achieved using an Aminex HPX-87H anion exchange column; Bio-Rad Laboratories, Hercules, CA) according to the protocol established by Buday *et al.* (1990) using 5 mM H_2SO_4 as the mobile phase. Analyte concentrations were established using standard curves for fermentative substrates and products, including glucose, succinate, lactate, formate, acetate,

and ethanol. Though retention times were determined for pyruvate, malic acid, propionate, 2-3 butanediol, butyrate, propanol, crotonate, butyraldehyde, valerate, butanol, and pentanol, standard curves were not generated because no apparent peaks were detected for these compounds.

Evaluating SR7 $scCO_2$ growth using 1 atm CO_2-optimized conditions

SR7 growth outcomes were investigated under $scCO_2$ headspace (90-100 atm; 37°C) while shaking at 250 rpm. SR7 spores were inoculated at starting concentrations of ~$3x10^4$ spores/ml (unless otherwise specified) in either 50 mM K_2HPO_4-buffered LB (P-LB) or M9+ media (Table 1). Experiments assaying the effect of germination induction included 100 mM L-alanine and 10 mM L-leucine media amendments and heat treatment upon reactor pressurization (70°C for 10 minutes). Incubations were conducted in 316 stainless steel vessels and gradually pressurized to supercritical conditions using a CO_2/He cylinder, as previously described. SR7 germination was verified by the identification of vegetative cell morphologies using fluorescence microscopy of Syto9-stained cultures. Growth was defined by an increase of at least 10-fold growth in cell counts relative to t_0.

In order to ascertain whether L-alanine, L-leucine, and heat treatment induce germination under $scCO_2$ headspace, three replicate experiments were conducted comparing growth for SR7 spores when loaded in media P-LB, P-LBL,

Table 2. Incubation conditions assaying chemical germination induction and heat treatment effects on SR7 growth under $scCO_2$

Incubation	Duration	Media	# Columns
A	18 days	P-LB	7
		P-LBA	7
		P-LBA (+ Heat)	6
		P-LBL	7
		P-LBAL	7
		Neg Ctrl	4
B	20 days	P-LB	7
		P-LBA	7
		P-LBL	6
		P-LBAL	6
		Neg Ctrl	4
C	18 days	P-LB	6
		P-LBA	5
		Neg Ctrl	4

P-LBAL, or P-LBA ± heat treatment (Table 1; Table 2). Reactors were depressurized and scored for germination and growth by fluorescence microscopy, as previously described.

Cell densities of P-LB and L-PBA incubations from the three experiments (Table 3; Incubations A-C) were subjected to statistical analysis to establish the significance of 100 mM L-alanine on spore growth outcomes. A non-parametric Wilcoxon/Kruskal-Wallis Test was performed on the dataset (JMP Pro v.12) where growth outcome (growth/no growth) and cell density fold change (relative to t_0) were dependent variables and incubation time and inducer presence/absence (± 100 mM L-alanine) were independent variables.

Table 3. L-alanine-amended $scCO_2$ incubations in P-LBA and M9A+

Media	Duration	# Cultures	# Neg Ctrl
P-LBA	18 days	7	4
	20 days	7	4
	18 days	5	4
	Total	**19**	**12**
M9A+	18 days	18	6
	20 days	7	4
	Total	**25**	**10**

Growth was compared for spore-inoculated cultures in L-alanine-amended M9+ (M9A+; Table 1) and P-LBA to determine whether either medium facilitates superior growth under $scCO_2$ when controlling for the presence of L-alanine. Buffering capacity was comparable for both media based on similar phosphate content. A summary of the M9A+ vs. P-LBA incubations is provided in Table 4. A non-parametric Wilcoxon/Kruskal-Wallis Test (JMP Pro v. 12) was run on the P-LBA and M9A+ datasets, where growth outcome (growth/no growth) and cell density fold change (relative to starting concentrations) were the dependent variables and incubation time and inducer presence/absence (±100 mM L-alanine) were the independent variables.

To establish whether increasing starting spore concentrations and incubation time improves the likelihood of growth, replicate cultures in M9A+ loaded with four starting spore concentrations ($5x10^5$, $5x10^3$, $5x10^1$, $5x10^{-1}$ cells/ml) were run over an 18-day time course. Samples were prepared for cell counts by fluorescence microscopy according to previously described protocols. Because reactors inoculated with $5x10^1$ and $5x10^{-1}$ cells/ml are below the limit of

detection by direct counts, their concentrations are recorded as one half the detection limit ($1.15x10^3$ cells/ml, as previously discussed). M9A+ time course data was combined with prior M9A+ $scCO_2$ results to develop a logistic regression model (JMP Pro v. 12) for growth frequency where outcome (growth/no growth) was the dependent variable, and inocula concentration and incubation time were independent variables.

3.3 Results

3.3.1 Isolation of $scCO_2$-tolerant strains from McElmo Dome fluids

Enrichment cultivation and serial passaging of McElmo Dome formation fluids with microbial growth media in high-pressure reactors under supercritical CO_2 headspace enabled the isolation of six different microbial strains, all of which are taxonomically classified within the *Bacillus* genus. Cultures were assayed for growth after the enrichment (M1 = 45 days) and each of three subsequent passages (M2 = 19 days, M3 = 33 days, M4 = 35 days) by epifluorescence microscopy methods (Table 4). Cell density from enrichment cultures was regularly observed to be greater than 10^5 cells/ml. The second (M2) and third (M3) round of culturing winnowed down the number of reactors demonstrating growth, with passaging of most media-inocula combinations discontinued due to lack of growth (or in some cases loss of pressure in reactors). The media-inocula combinations that were incubated during the fourth round (M4) of culturing

Table 4. Enrichment passaging diversity, biomass, and isolate strain summary.

Passage	M1	M2	M3	M4	Isolate Strain(s)
Duration (d)	45	19	33	35	
WELL 2 / HB-5					
MS-FM	+++	++	n.d.	n.d.	No Isolate
MS-MR	+++	++	+++	++	MR2
MS-SR	+++	++	+	n.d.	No Isolate
WELL 4 / HE-1					
MS-FM	++	n.d.	n.d.	n.d.	FM4
MS-MR	+++	+	+++	+++	MR4
MS-SR	+++	+	-	n.d.	No Isolate
WELL 5 / HF-3					
MS-MR	+++	+	+	n.d.	No Isolate
WELL 7 / YB-4					
MS-FM	+++	+	n.d.	n.d.	No Isolate
MS-MR	+++	+	+++	+++	MR7C / MR7R
MS-SR	+++	+++	+++	+++	SR7

Biomass Concentration (direct filter counts)	
-	cells below detection limit (<1.2E3 cells/ml)
+	biomass observed at <5E4 cells/ml
++	biomass observed at 5E4 to 1E6 cells/ml
+++	biomass observed at >1E6 cells/ml
n.d.	no data

showed maximum biomass accumulations of at least $7x10^5$ cells/ml (Table 6), including Well 2 + MS-MR media ($7.4x10^5$ cells/ml), Well 4 + MS-MR ($1.2x10^8$ cells/ml), Well 7 + MS-MR ($3.1x10^7$ cells/ml), and Well 7 + MS-SR ($6.9x10^6$ cells/ml).

After the fourth passage (M4), individual strains were isolated by plating on LB agar. Colonies with unique morphologies were identified by 16S rRNA gene sequencing and taxonomic annotation (Table 5). In most cases (except Well 7 + MS-MR, which enabled isolation of two strains of the same species), a single dominant strain was able to be isolated from specific combinations of media and inocula. One additional strain was isolated by LB agar colony culturing after M1 in MS + FM media with Well 4 fluids. 16S rRNA Blastn annotations of isolated strains are presented in Table 5. A 16S rRNA phylogenetic tree of McElmo Dome CO_2-passaged isolates and several closely related Bacilli is presented in Figure 2.

Table 5. Summary of passaged isolate morphologies and taxonomic annotations.

Well Inocula	Passage Media	Colony Morphology	16S rRNA Blastn Top Hit	Blastn ID%	Designated Strain Name
2	MS-MR	Circular, entire, umbonate, dull, cream, opaque	*Bacillus safensis*	99	*B. safensis* MR2
4	MS-MR	Circular, filamentous, flat, dull, nonpigemented, translucent	*Bacillus licheniformis*	99	*B. licheniformis* MR4
4	MS-FM	Circular, entire, umbonate, dull, cream, opaque	*Bacillus safensis*	99	*B. safensis FM4*
7	MS-MR	Circular, entire, umbonate, dull, cream, opaque	*Bacillus safensis*	100	*B. safensis* MR7C
7	MS-MR	Circular, undulate, umbonate, dull, cream, opaque	*Bacillus safensis*	99	*B. safensis* MR7R
7	MS-SR	Circular, entire, convex, dull, white, opaque	*Bacillus megaterium*	100	*B. megaterium* SR7

Because enrichment passaging led to the isolation of several strains demonstrating spore-like morphologies and annotated as spore-forming taxa, isolated strains were prepared as spores for long-term storage. Previous work by Peet *et al.*, (2015) demonstrated that spores loaded into replicate reactors under an $scCO_2$ headspace (i.e. *Bacillus* sp. OT1, *Bacillus* sp. MIT0214; Figure 2) grew with frequencies dependent on incubation time and starting spore concentrations, while vegetative cells were unable to survive $scCO_2$ exposure. Spore preparations of *B. megaterium* SR7 and *B. licheniformis* MR4 maintained consistent viability over long periods (>2 years) in spore prep wash buffer at 4°C, though all *B. safensis* strains demonstrated markedly lower survival (decrease of CFUs/ml by at least four orders of magnitude in <6 months). Growth of *B. megaterium* SR7 and other strains was validated by triplicate incubation of spore stocks for 28-42 days using multiple media (Table 1). Growth was defined as demonstrating at least one order of magnitude increase in cell density relative to starting concentration (~10^4 spores/ml; Table 6).

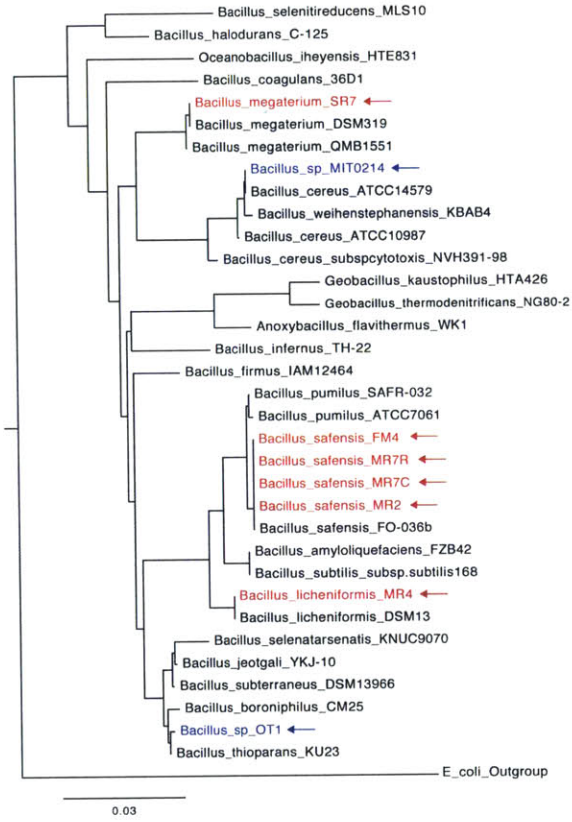


Figure 2. 16S rRNA phylogenetic tree of McElmo Dome isolates (red), isolates studied in Peet et al. (2015; blue) and additional closely related Bacilli.

Table 6. Summary of results from strain isolate $scCO_2$ incubations in pure culture

Incubation	Duration	Strain	Media	Growth	Max cells/mL
P1	33 days	B. megaterium SR7	SR	1/2	$1.0x10^8$
		B. licheniformis MR4	MR	3/3	$1.2x10^8$
		B. safensis MR7C	MR	1/3	$8.1x10^7$
		B. safensis MR7R	MR	1/3	$4.9x10^7$
		B. safensis MR2	MR	1/2	$8.8x10^6$
		B. safensis FM4	FM	2/3	$3.5x10^7$
P2	28 days	B. megaterium SR7	SR	1/3	$2.0x10^7$
			LB	3/3	$6.8x10^7$
		B. licheniformis MR4	MR	1/3	$1.8x10^6$
			LB	1/3	$2.3x10^7$
		B. safensis MR7C	MR	1/3	$1.5x10^8$
			LB	1/3	$4.4x10^6$
P3	42 days	B. megaterium SR7	LB	3/3	$3.5x10^7$
		B. safensis MR2	MR	0/2	$3.2x10^4$
			LB	0/2	$1.8x10^4$
		B. safensis MR7R	MR	1/3	$1.9x10^6$
			LB	1/3	$6.3x10^6$
		B. safensis FM4	MR	3/3	$6.1x10^7$
			LB	3/3	$1.2x10^7$

Based on the results from the original four rounds of enrichment passaging (M1-M4) and subsequent pure culture $scCO_2$ incubation trials from spore stocks (P1-P3), strain *B. megaterium* SR7 generated the most consistently robust growth, especially in LB media (6/6 combined growth in P2-P3). Thus, strain SR7 was selected for physiological, metabolic and genomic investigation with the intent of optimizing growth under $scCO_2$. Isolates *B. licheniformis* MR4 and *B. safensis* FM4 also demonstrated strong growth during enrichment passaging and in pure culture, and thus may be useful strains for future development.

3.3.2 Isolate SR7 genomics

The genome of *B. megaterium* SR7 was sequenced to determine its metabolic capacity and enable the development of genetic manipulation tools for bioproduct pathway engineering. PacBio sequencing and assembly resulted in six contigs (DNA fragments) from *B. megaterium* SR7 (Table 7).

Table 7. SR7 Summary of PacBio genome sequencing/assembly and RAST annotation statistics

Contig	DNA Type	Length	% Bases Called	Coverage	ORFs	Plasmid-Associated	Sporulation/ Germination
1	Chromosome	5449642	100.0	40.7	5,696	3	194
2	Plasmid p1	21958	99.9	57.0	35	11	6
3	Plasmid p2	17283	100.0	50.5	19	4	2
4	Plasmid p3	9202	79.2	20.8	13	3	1
5	Plasmid p4	7873	92.5	6.7	8	3	1
6	Plasmid p5	2921	52.2	0.5	4	2	1

The largest contig is 5,449,642 bp with 40.7X coverage and 39% GC content, while the other five contigs are between 2.9 kb and 22.0 kb (Table 7). Comparison of SR7 contigs with reference *B. megaterium* strain QM_B1551 showed nearly 1:1 synteny of the largest SR7 contig and the main chromosome of QM B1551, as well as similarity between the smaller SR7 contigs and QM B1551 plasmids. After synteny-based adjustments enabled the SR7 chromosome to be closed (Figure 3), it was submitted to RAST for functional annotation along with the five smaller contigs. RAST chromosome analysis called 5,696 coding ORFs, with 13 complete rRNA operons with 5S, 16S and 23S rRNA genes and one extra 5S rRNA gene.

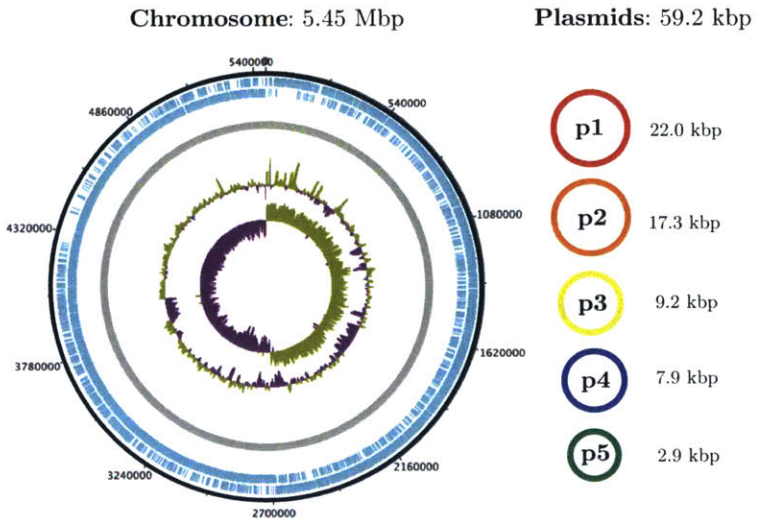


Figure 3. Schematic of the *B. megaterium SR7* 5.51 MBp genome, including the closed 5.45 MBp chromosome. Concentric circles (outside in) are RAST ORFs (blue), rRNA and tRNA (green lines on grey circle), GC content, and GC skew. Asymmetry in GC skew indicates proper chromosome assembly (Grigoriev, 1998). Circles at right represent five putative plasmids native to SR7.

Genomic annotations of carbon metabolism in SR7 include genes associated with glycolysis, the Entner-Doudoroff Pathway, TCA Cycle, Pentose Phosphate Pathway, Glyoxylate Bypass, and acetogenesis from pyruvate. Annotation of the SR7 chromosome also reveals the genomic potential for broad fermentative reactions, including utilization of glucose, fructose, mannose, and xylose, and the production of butyrate, lactate, butanol, acetate, 2,3-butanediol, and ethanol.

No genes associated with direct carbon fixation pathways were detected in the genome (i.e. Calvin Cycle, Wood-Ljungdahl Pathway, rTCA cycle, etc.). However, the annotation of carbonic anhydrase, which facilitates conversion of CO_2 to bicarbonate (Smith and Ferry, 2000), carbamoyl-phosphate synthase, which incorporates bicarbonate for pyrimidine and arginine biosynthesis (Arioli *et al.*, 2009), and phosphoenolpyruvate carboxylase, which catalyzes the irreversible addition of bicarbonate to phosphoenolpyruvate, indicates the capacity for SR7 to utilize and assimilate CO_2 species, potentially as a mechanism for aiding in high pCO_2 exposure survival (Santillan *et al.*, 2015, Arioli *et al.*, 2009). The presence of carboxylase may prove useful for future engineering of CO_2-consuming metabolic pathways as a sustainable substrate in addition to solvent under $scCO_2$ conditions, especially in light of the previous demonstration of *B. megaterium* carboxylase activity under $scCO_2$ (Matsuda *et al.*, 2001).

Annotated inorganic redox metabolism-associated genes may ultimately prove useful by informing growth media amendments or elucidating the capacity for SR7 to grow on alternative substrates, including treated wastewater, e.g. by denitrification (Yang *et al.*, 2012), reducing the need for expensive carbohydrate substrates. SR7 genes of this nature include assimilatory sulfite reductase (NAPDH-dependent), sulfite oxidase, assimilatory nitrate reductase, dissimilatory nitrite reductase (*nir*BD), nitric oxide reductase denitrification genes (*nor*QD), and an arsenate reductase detoxification gene (*ars*C). Physiological annotations of the SR7 chromosome that hold potential utility as components of a microbial bioproduction system include a full suite of sporulation genes, siderophore assembly and uptake, flagellar motility, the twin-arginine translocation (TAT) system, and PHB metabolism, the last of which indicates a capacity for redirecting flux toward concentrated storage of excess carbon. The endogenous TAT secretion system, may be useful for developing the ability to secrete specific products in the event that bioproduction focuses on the generation of proteins or enzymes.

Because the five smaller contigs failed to thoroughly annotate via RAST (i.e. a majority of hypothetical genes), RAST-called ORFs were submitted to Blastx for amino acid level annotation. All five contigs are annotated as containing plasmid replication, recombination, and mobility genes, as well as genes previously identified on other *Bacillus spp.* plasmids, and sporulation-related genes, content consistent with previously characterized *B. megaterium* plasmids (Eppinger *et al.*, 2011). As a result, the five putative plasmids native to SR7 are designated (in order of decreasing size) plasmids p1 through p5, the RAST statistics and Blastx annotations for which are listed in Table 7. In comparison to the five putative plasmids in strain SR7 (59.2 kb total), seven (426 kb total) and three plasmids (91.3 kb total) were previously detected in strains QM B1551 and WSH-002, respectively, providing precedent for extra-chromosomal gene content in *B. megaterium.*

The *B. megaterium* SR7 genome size (5.51 Mbp) is slightly larger than several previously sequenced *B. megaterium* strains, including QM B1551 (5.1 Mbp) and DSM319 (5.1 Mbp), and approximately 33% larger than strain WSH-002 (4.14 Mbp). *B. megaterium* isolate SR7 and industrial strains QMB1551 and DSM319 share 96-97% average nucleotide identity (ANI). A comparison of shared

gene content based on RAST annotations of SR7 and the three *B. megaterium* type strains reveal that approximately 12% of the SR7 genome consists of gene content not observed in three fully sequenced *B. megaterium* strains. However, the number of ORFs called by RAST appears to underestimate the number of gene calls in the original sequencing studies associated with each strain (i.e. DSM319 RAST = 2,898 ORFs, Eppinger *et al.* (2011) = 5,272 ORFs; QM B1551 RAST = 2,915 ORFs, Eppinger et al., (2011) = 5,284 ORFs; WSH-002 RAST = 2,872 ORFs, Liu et al. (2011) = 5,269 ORFs). According to the RAST re-annotation of these submitted genomes, genes unique to SR7 include a gas vesicle structural protein (*gvp*A), genes associated with biotin synthesis/regulation (*bio*HR), a carboxysome structural gene (*ccm*M), a cell wall teichoic acid glycosylation gene (*gtc*A), several phage annotations, and chromosome/plasmid partitioning genes (*par*AB).

3.3.3 Physiological characterization of SR7 under ambient conditions

Strain *B. megaterium* SR7 was subjected to chemical and temperature characterization experiments under an ambient atmosphere to establish conditional growth ranges and optima of facultative aerobic growth. The results of these assays are presented in Table 8. pH growth experiments revealed the

Table 8. Summary of viable SR7 growth in LB media over chemical and temperature ranges under aerobic conditions

Condition	Range	Optimal
pH[a]	4-10	6-7
[NaCl] (g/L)[b]	0-100	0-10
$NaHCO_3$ (mM)[c]	0-300	0-100
Temperature (°C)[d]	23-45	37
Assayed ranges and durations: [a]pH: 2, 4, 6, 7, 8, 10, 12 over 123 h [b]Salinity: 0, 5, 10, 50, 80, 100 g/L over 36 h [c]Bicarbonate: 0, 100, 300, 500 mM over 36 h [d]Temperature: 9, 23, 27, 30, 37, 45, 55°C over 73 h		

Table 9. SR7 Antibiotic sensitivity assay summary

Antibiotic	[a]ug/mL	[b]% Control	Assay	[c]Sensitivity
Spectinomycin	100	47%	LB	S/R
Nalidixic Acid	-	14%	Biolog	S
Tetracycline	1.5	11%	LB	S
Minocycline	-	11%	Biolog	S
Lincomycin	-	11%	Biolog	S
Rifamycin SV	-	11%	Biolog	S
Aztreonam	-	10%	Biolog	S
Vancomycin	-	10%	Biolog	S
Streptomycin	10	10%	LB	S
[d]D-Serine	-	10%	Biolog	S
Fusidic Acid	-	9%	Biolog	S
[d]Niaproof 4	-	9%	Biolog	S
Chloramphenicol	35	9%	LB	S
Kanamycin	5	9%	LB	S
Troleandomycin	-	9%	Biolog	S
Ampicilin	50	8%	LB	S
[a]Biolog does not publish antibiotic concentrations [b](OD_{600} AB/OD_{600} Control)*100 in LB, average n = 2 (OD_{490} AB/OD_{490} Control)*100 for Biolog, average n = 2 [c]S = sensitive, R = resistant [d]Non-antibiotic treatment				

fastest growth between pH 6-7 with an extended lag phase of 76 hours for pH 4 and 10, and no growth after 123 hours at pH 2 and pH 12. LB and Biolog salinity assays revealed diminished growth of SR7 above 10 g/l NaCl. Increasing bicarbonate above 100 mM also led to decreased growth. SR7 growth is supported between 23°C and 45°C, with growth not observed after 73 hours at 9°C and 55°C. Sensitivity to all tested antibiotics (with intermediate sensitivity to spectinomycin; Table 9) may be exploited for aspects of biotechnology development methods, including selective markers for transformations. Biolog assays revealed SR7 growth was also inhibited by D-serine and Niaproof 4, which are known to inhibit cell wall synthesis and emulsify lipid membranes, respectively.

Biolog assays also established which potential sole carbon sources may be useful in future SR7 culturing and allowed comparison between SR7 and closely related *B. megaterium* strains DSM319 and QM B1551. While all three strains demonstrated robust growth on TCA Cycle intermediates citric acid and L-malic acid, DSM319 and QM B1551 both grew on L-lactic acid and L-glutamic acid, while SR7 did not (Table 10).

Table 10. SR7 and alternative *B. megaterium* strains categorized by robust (+++), marginal (+) or no (-) growth on Biolog sole carbon sources (no metals added). Only carbon sources enabling at least one strain to demonstrate robust growth are listed.

Carbon Substrate	SR7	DSM319	QM B1551
Citric Acid	+++	+++	+++
L-Malic Acid	+++	+++	+++
L-Lactic Acid	+	+++	+++
L-Glutamic Acid	+	+++	+++
α-D-Glucose	+	+++	+
Dextrin	+	+++	+
D-Mannitol	+	+++	+
D-Gluconic Acid	+	+++	+
L-Aspartic Acid	+	+++	+
N-Acetyl-D- Glucosamine	+	+++	+
L-Histidine	+	+++	+
Bromo-Succinic Acid	+	+++	+
D-Maltose	-	+++	+
Sucrose	-	+++	+
β-Hydroxy-D,L-Butyric Acid	+	+	+++
D-Saccharic Acid	-	-	+++

SR7 growth was markedly increased upon addition of trace metals solution to Biolog media (Table 11), including on substrates D-raffinose, α-D-glucose, γ-amino-butryric acid, myo-inositol, L-arginine, D-gluconic Acid, citric acid, N-

acetyl-D-glucosamine, L-glutamic acid, D-turanose, and L-pyroglutamic acid. Malic acid appears to have facilitated robust growth only in the absence of metals. SR7 was able to grow on several carbon sources in the presence of metals that strains DSM319 and QM B1551 grew on without amendment (e.g. L-glutamic acid, α-D-glucose, sucrose, N-acetyl-D-glucosamine, etc.), which suggests that metal-bearing co-factors specific to SR7 catabolism may require elevated metals concentrations to properly function. Initially, SR7 demonstrated robust growth on 2/71 Biolog substrates, improving to 12/71 upon addition of metals. These 12 substrates have thus been identified as potential sole carbon sources for metals-amended defined media.

Table 11. SR7 robust (+++), marginal (+) and no (-) growth in unamended (I & II) and trace metals-amended carbon source Biolog plates. Maximum growth for each plate trial is noted by an asterisk. All substrates (and negative control) listed.

Carbon Substrate	I	II	+Metals
Citric Acid	+++	+++*	+++
α-D-Glucose	+	+	+++
L-Arginine	+	+	+++
D-Gluconic Acid	+	+	+++
L-Aspartic Acid	+	+	+++
N-Acetyl-D- Glucosamine	+	+	+++*
L-Glutamic Acid	+	+	+++
D-Turanose	+	+	+++
L-Pyroglutamic Acid	+	+	+++
D-Raffinose	-	+	+++
γ-Amino-Butryric Acid	-	+	+++
myo-Inositol	-	+	+++
L-Malic Acid	+++*	+++	+
Gelatin	+	+	+
Pectin	+	+	+
Dextrin	+	+	+
α-D-Lactose	+	+	+
D-Mannitol	+	+	+
Methyl Pyruvate	+	+	+
D-Melibiose	+	+	+
D-Fructose	+	+	+
D-Arabitol	+	+	+
L-Alanine	+	+	+
D-Lactic Acid Methyl Ester	+	+	+
D-Galactose	+	+	+
L-Lactic Acid	+	+	+
β-Hydroxy-D,L- Butyric Acid	+	+	+
D-Cellobiose	+	+	+
D-Salicin	+	+	+
Glycerol	+	+	+
Gentiobiose	+	+	+
α-Keto-Glutaric Acid	+	+	+
L-Histidine	+	+	+
Stachyose	+	+	+
Bromo-Succinic Acid	+	+	+
D-Maltose	-	+	+

Carbon Substrate	I	II	+Metals
D-Trehalose	-	+	+
β-Methyl-D-Glucoside	-	+	+
Sucrose	-	+	+
Inosine	-	+	+
D-Sorbitol	+	-	+
3-Methyl Glucose	-	-	+
D-Glucuronic Acid	+	+	-
Acetic Acid	+	+	-
L-Serine	+	+	-
Tween 40	-	+	-
D-Galacturonic Acid	-	+	-
L-Galactonic Acid Lactone	-	+	-
Acetoacetic Acid	-	+	-
Mucic Acid	-	+	-
Propionic Acid	-	+	-
Quinic Acid	-	+	-
D-Saccharic Acid	-	+	-
D-Fructose- 6-PO4	+	-	-
N-Acetyl-D- Galactosamine	+	-	-
Formic Acid	+	-	-
Negative Control	-	-	-
p-Hydroxy- Phenylacetic Acid	-	-	-
D-Mannose	-	-	-
Glycyl-L-Proline	-	-	-
α-Hydroxy- Butyric Acid	-	-	-
α-Keto-Butyric Acid	-	-	-
D-Fucose	-	-	-
D-Glucose- 6-PO4	-	-	-
Glucuronamide	-	-	-
N-Acetyl-β-D- Mannosamine	-	-	-
L-Fucose	-	-	-
D-Malic Acid	-	-	-
L-Rhamnose	-	-	-
D-Aspartic Acid	-	-	-
N-Acetyl Neuraminic Acid	-	-	-
D-Serine	-	-	-

3.3.4 SR7 activity under 1 atm CO_2

As described in the methods, culturing experiments under 1 atm CO_2 were used as a proxy for scCO_2 conditions. Modeling using the ideal gas law indicates

that for rich media, predicted dissolved CO_2 concentrations for ambient air, 1 atm CO_2, and $scCO_2$ are $1.2x10^{-5}$ M, $2.6x10^{-2}$ M and 2.7 M, respectively (Peet *et al.*, 2015). Therefore, exposure of SR7 cultures to intermediate dissolved CO_2 content and pH conditions at 1 atm CO_2 may inform beneficial process improvements for enhanced growth under $scCO_2$.

Growth dynamics and process engineering

Assays conducted at 1 atm CO_2 showed that increased shake speed led to faster cell growth in spore-inoculated cultures (Figure 4) and also facilitated more rapid growth of passaged vegetative cells (Figure 5).

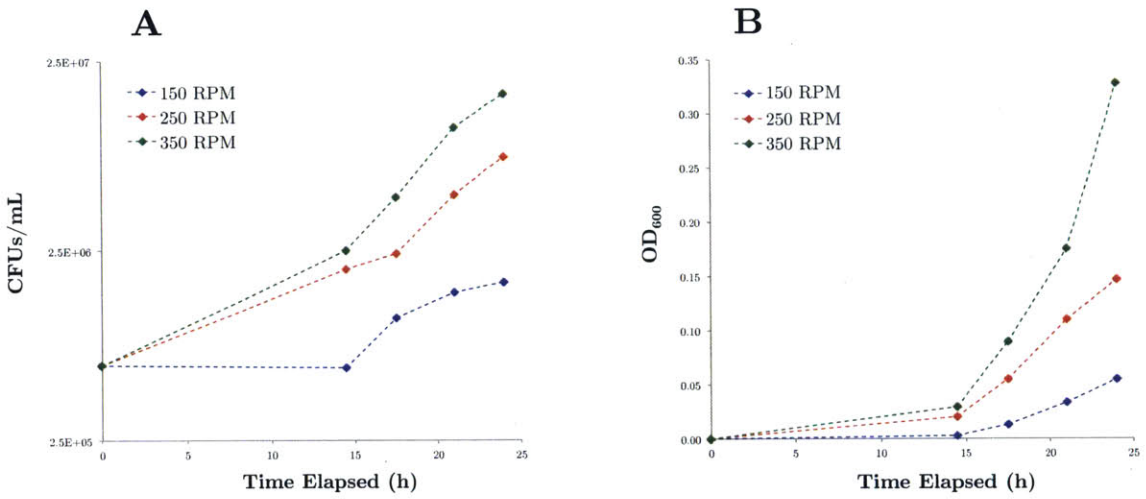


Figure 4. Effect of mixing rates on SR7 spore germination in LB under 1 atm CO_2 as measured by **A)** CFU/mL and **B)** OD_{600}

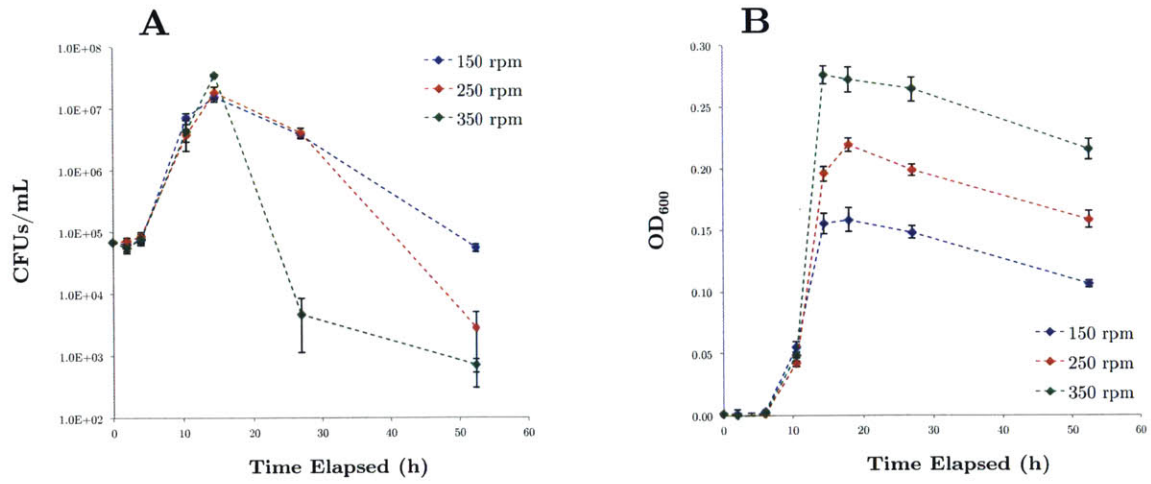


Figure 5. Effect of mixing rates on passaged SR7 vegetative growth in LB under 1 atm CO_2 as measured by **A)** CFU/mL and **B)** OD_{600}

Increased shake speeds also enabled higher biomass accumulation, as the maximum OD_{600} reached by 150 and 250 rpm samples were 57% and 79% the OD_{600} maximum for 350 rpm, while maximum CFU counts reached by 150 rpm ($1.5x10^7$ CFUs/ml) and 250 rpm ($1.8x10^7$ CFUs/ml) samples were 43% and 51% of the maximum count for 350 rpm ($3.5x10^7$ CFUs/ml), respectively (Figure 5). However, it appears that cultures that reach maximum biomass accumulation due to increased mixing rates also may reach stationary phase and crash more quickly, a result often associated with end product toxicity in fermenting cultures (Figure 5A). Therefore, due to the accelerated growth rate of SR7 at 250 RPM and the ability to sustain high biomass without experiencing a precipitous drop in CFU counts (as with 350 rpm), a shake speed of 250 RPM was utilized for all subsequent incubation experiments.

Minimal medium formulation (M9+)

Development of a minimal growth medium enables examination of microbial physiology, determination of nutritional growth requirements, and holds potential to reveal the metabolic pathways through which carbon flux occurs during growth under various conditions. Initial attempts to grow SR7 in

Table 12. M9 supplemented growth outcomes under 1 atm CO_2

M9 Amendments	Growth
0.4% Glucose	-
0.4% Xylose	-
0.4% Glucose + 1X Metals	-
0.4% Xylose + 1X Metals	-
0.001X LB	-
0.01X LB	-
0.001X YE	-
0.01X YE	-
0.001X LB + 0.4% Glucose	+
0.01X LB + 0.4% Glucose	+
0.01X YE + 0.4% Glucose	+
0.01X YE + 0.4% Glucose + 5 mM $NaNO_3$	+
0.001X LB + 0.4% Glucose + 5 mM $NaNO_3$	+
0.01X LB + 0.4% Glucose + 5 mM $NaNO_3$	+
0.001X LB + 0.4% Glucose + 0.1X Metals	+++
0.01X LB + 0.4% Glucose + 0.1X Metals	+++
0.01X YE + 0.4% Glucose + 0.1X Metals	+++
0.01X YE + 0.4% Glucose + 0.25X Metals	+++
0.01X YE + 0.4% Glucose + 1X Metals	+++
0.01X LB + 0.4% Glucose + 5 mM $NaNO_3$ + 0.1X Metals	+++
0.01X YE + 0.4% Glucose + 5 mM NaNO3 + 0.1X Metals	+++
LB/Yeast Extract (YE) Dilutions	
0.01X LB = 100 mg/L tryptone, 50 mg/L YE, 100 mg/L NaCl	
0.001X LB = 10 mg/L tryptone, 5 mg/L YE, 10 mg/L NaCl	
0.01X YE = 50 mg/L	
0.001X YE = 5 mg/L	

M9 base medium under 1 atm CO_2 with 0.4% glucose or 0.4% xylose as sole carbon source in the presence and absence of a trace metals solution were unsuccessful (Table 12). Subsequent growth assays revealed the necessity for both a de facto vitamin/co-factor supplement (i.e. dilute LB/YE at concentrations insufficient to independently support observable growth) and trace metals solution to be present in glucose-amended media to enable robust growth (Table 12; Figure 6). The use of NO_3^- as a potential alternative electron acceptor did not demonstrate any pronounced effects on growth rates or biomass accumulation, despite genomic evidence for potential nitrate/nitrite reduction capacity. Due to potential conflicts between xylose-induced biomass accumulation and heterologous gene expression, media development proceeded with glucose as sole carbon source. Since substituting out 0.01X LB for 0.01X YE (i.e. 1X is the concentration of YE present in LB, 5 g/l; 0.01X YE = 50 mg/l) generated similar outcomes, media development proceeded with YE due to its more defined nature.

0.1X trace metals solution proved the most effective concentration for enabling rapid growth of passaged vegetative cultures. Although 1 atm CO_2 passaged cultures in M9 + 0.4% glucose + 0.01X YE amended with 0.25X and 1.0X trace metals achieved the same maximum OD_{600} as 0.1X metals-amended cultures, lower OD_{600} values at intermediate time points suggested diminished growth rates relative to 0.1X metals (Figure 6A). Further investigation revealed

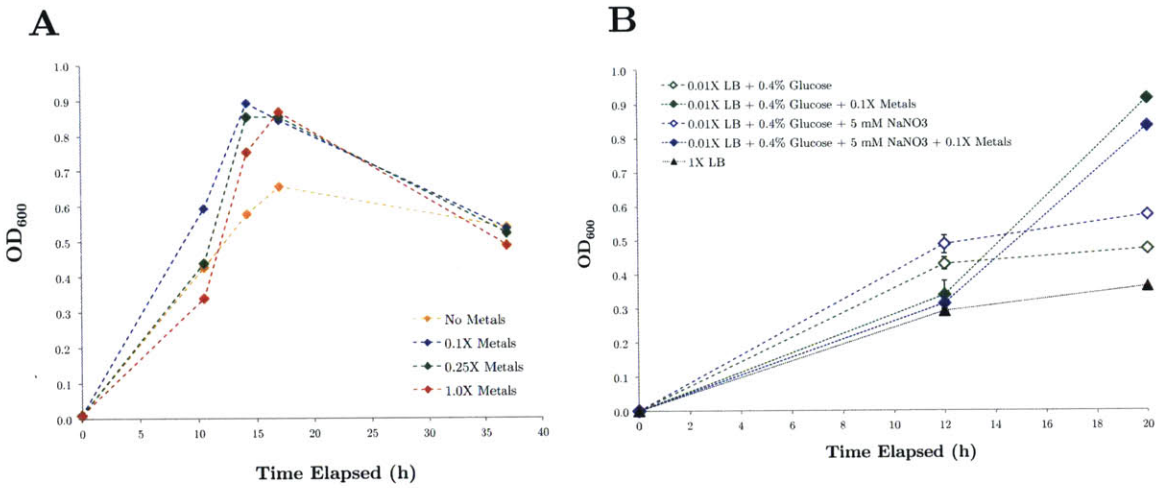


Figure 6. SR7 growth under 1 atm CO_2 at 37°C **A)** as a function of metals concentration and **B)** in the presence (filled diamonds) and absence (open diamonds) of trace metals in M9 media types and unamended LB (black filled triangles).

that while cultures in the presence and absence of 0.1X trace metals reach intermediate OD_{600} values at approximately the same rate, metals-amended cultures continue to grow while non-amended cultures appear to enter stationary phase (Figure 6B). The effect of trace metals on accelerated anaerobic growth has previously been observed in (David *et al.*, 2010), who suggested that bacteria require metal co-factors to improve growth outcomes.

The final combination of M9 + 0.01X YE + 0.1X metals + 0.4% glucose is designated "M9+" medium, and was used as the base semi-defined minimal medium for all subsequent sole carbon source experiments. After establishing M9+ medium components, 1 atm CO_2 growth curves conducted in both M9+ and LB media revealed SR7 anaerobic doubling times based on OD_{600} of 1.93 ± 0.1 h and 2.07 ± 0.1 h, respectively (Figure 7). OD_{600} values and glucose consumption for M9+ media incubations appear to indicate log phase growth, then a brief stationary phase, followed by steady increases in OD and glucose consumption (though a decrease in CFUs). It is thus possible that M9+ media leads to diauxic growth. Though the source of this behavior is uncertain, 1 atm CO_2 fermentation products analysis (discussed below) shows that M9+ cultures rapidly accumulate ethanol, followed by the production of several acids (acetate, lactate, succinate), which may indicate a shift in metabolic pathway utilization due to substrate exhaustion or a response to end-product toxicity.

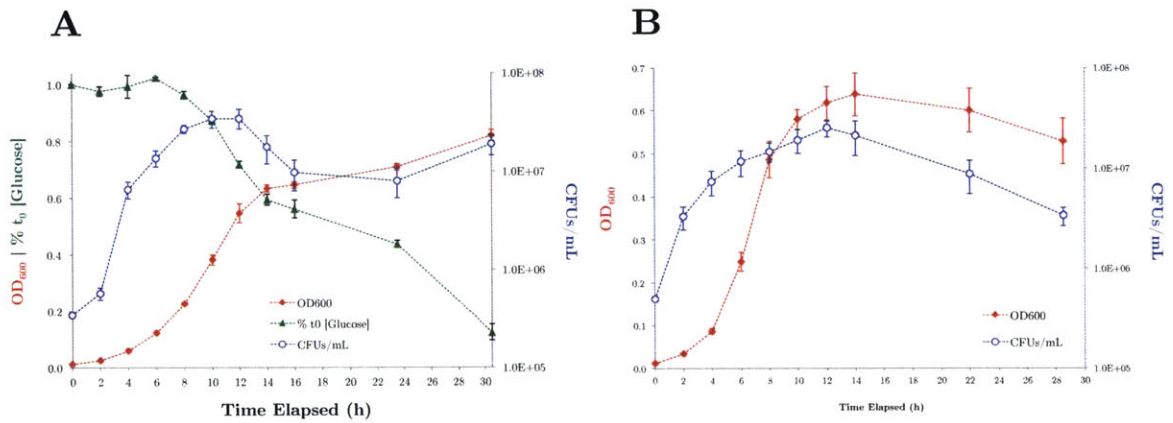


Figure 7. SR7 growth dynamics under 1 atm CO_2, 37°C, 250 rpm (OD_{600}: red; CFU/mL: blue) and glucose consumption (green; M9+ only) in **A)** M9+ media and **B)** LB

Despite positive growth on glucose as a sole carbon source in M9+ defined medium, SR7 showed a reduced growth rate under the same culturing conditions

in LB supplemented with 0.4% glucose, including both with and without metals solution, as demonstrated by an extended lag (Figure 8A). Similar results were generated under aerobic conditions (data not shown). Incomplete glucose

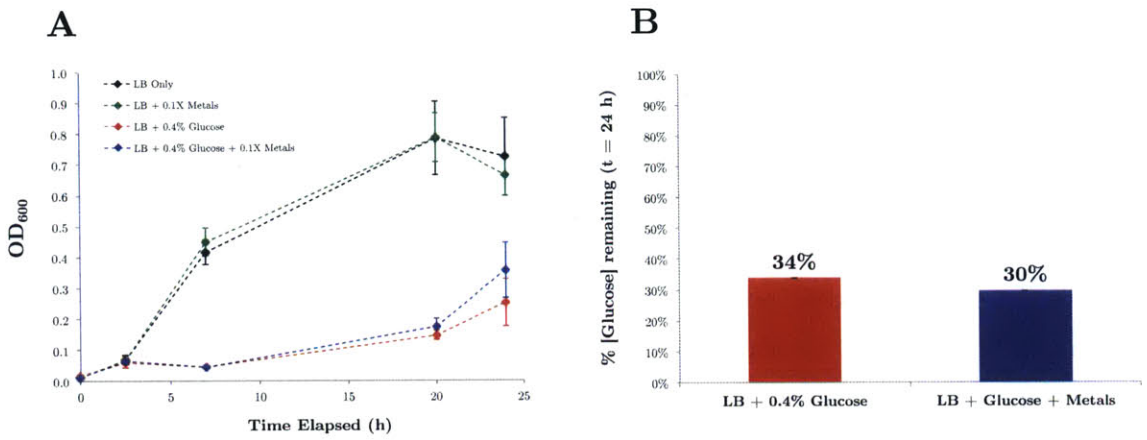


Figure 8. SR7 cultures in LB under 1 atm CO_2, 37°C, 250 rpm show **A)** a lag phase when glucose-amended, as measured by OD_{600} and **B)** incomplete glucose consumption after 24 hours

consumption after 24 hours (30-34% glucose remaining; Figure 8B) further indicates that LB-amended glucose is not fully utilized either due to growth on an alternative substrate (i.e. dilute YE or tryptone), or because glucose consumption during growth in LB medium is generating a toxic concentration of metabolites. Evidence from aerobic cultures demonstrate that after 2-3 hours, SR7 accumulates and maintains ~0.6 g/l acetate, while LB only cultures accumulate 0.4-0.5 g/l acetate after 5 hours, at which point it is consumed as a

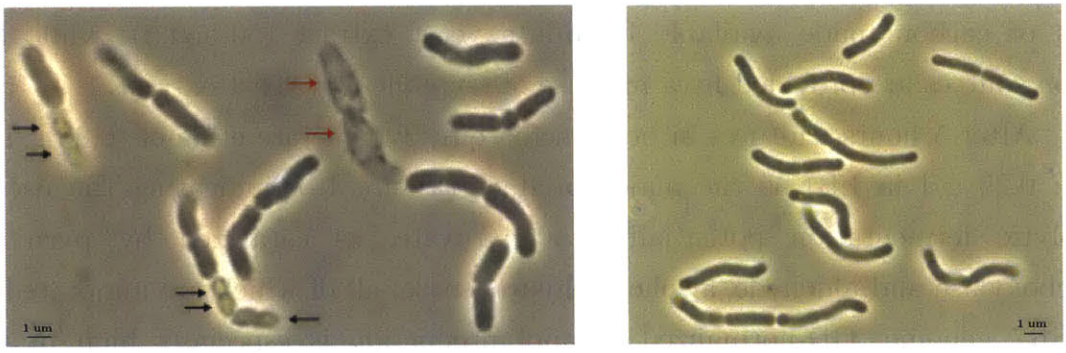


Figure 9. Phase contrast light microscopy of SR7 vegetative cultures. (Left) PHB-filled (bright cells, black arrow) and membrane-degraded SR7 cells (transparent, red arrows) in glucose-amended LB grown under 1 atm CO_2. (Right) Cells grown in LB under 1 atm CO_2 without glucose accumulate observable, but smaller amounts of PHB granules, often distributed at the cellular poles.

substrate (data not shown). These results suggest a potential mechanism for growth inhibition by glucose-associated end product toxicity as well as glucose repression of acetate re-assimilation. Inspection by phase contrast light microscopy of SR7 grown in LB + 0.4% glucose under 1 atm CO_2 appears to show an increase in PHB granules (Figure 9; polyhydroxybutyrates (PHBs) confirmed by GC-MS, personal communication, Yekaterina Tarasova). The membranes of many other cells appear to be nearly completely degraded, causing cells to look completely transparent. In carbon-rich media like LB, it is possible that the addition of glucose causes SR7 to become nutrient (P, N, etc.) limited, but this is doubtful due to the proteinaceous nature of LB medium. As an alternative hypothesis, as the consumption of glucose causes accumulation of a toxic product (e.g. lactic acid, ethanol; Figure 10) SR7 begins storing glucose as PHBs rather than increasing biomass. Therefore, only after the product has been re-assimilated or conditions become favorable (e.g. media buffering) will cells begin metabolizing stored PHBs and re-commence growth. Because previous studies have shown high production of PHBs do not have a toxic affect on *B. megaterium* (Rodríguez-Contreras *et al.*, 2013) it is considered unlikely that PHB accumulation itself is disrupting metabolism or cellular integrity.

SR7 fermentation products under 1 atm CO_2

SR7 cultures incubated under 1 atm CO_2 in M9+ and LB media generated a variety of fermentative products detected by HPLC. M9+ products are expected to be derived from sole carbon source glucose, with otherwise very low levels of carbon made available by dilute yeast extract (50 mg/l), while LB carbon sources are provided by a mixture of tryptone and yeast extract.

After 5 hours, cultures in both media quickly generate ethanol (0.25 g/l in M9+; 0.29 g/l in LB) as the major product (Figure 10), indicating the use of glycolytic fermentation potentially via pyruvate, as catalyzed by pyruvate decarboxylase and aldehyde/alcohol dehydrogenase, all of which are annotated in the SR7 genome. OD-normalized ethanol production is especially high in the M9+ incubations (3.54 g/l per OD_{600}), suggesting that glucose is first metabolized to ethanol before generating alternative products. In LB, however, due to the less defined nature of available carbon, growth at 5 hours also generates acetate (0.18 g/l) and succinate (0.13 g/l), possibly due to the more

complex proteinaceous available substrates (Figure 10). Acetate may also be generated through glycolytic fermentation, including via phosphate

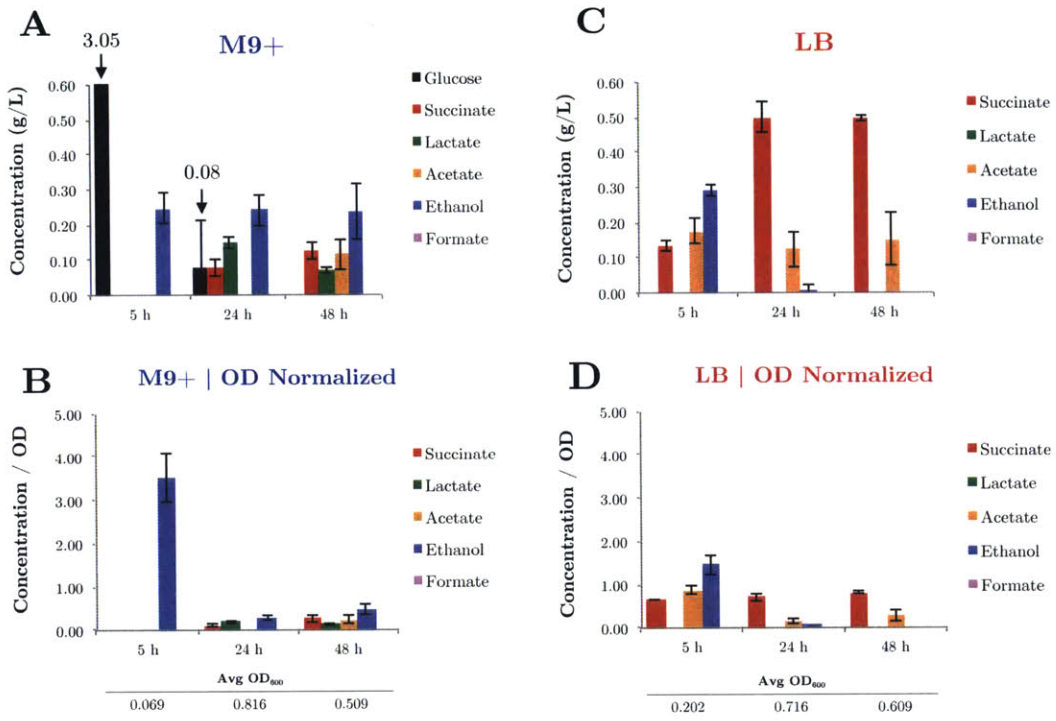


Figure 10. SR7 fermentation products under 1 atm CO_2 in **A)** M9+, raw values **B)** M9+, OD_{600}-normalized, **C)** LB, raw values, and **D)** LB, OD_{600}-normalized as detected and measured by HPLC. Average OD_{600} values (normalization factors) of producing cultures are listed below each time point.

acetyltransferase and acetate kinase, both of which are annotated in the SR7 genome. The production of succinate in cultures grown in LB may indicate TCA Cycle activation, as succinate is one of its major intermediate products. OD-normalized concentrations are comparable for all three products, which suggests that multiple pathways may be active, unlike the concentrated ethanol production in M9+, which appears to indicate a single dominant pathway for metabolic carbon flux.

After 24 and 48-hour incubations in M9+ medium, alternative products begin to appear, including lactate (24 hrs: 0.15 g/l; 48 hrs: 0.07 g/l), succinate (24 hrs: 0.08 g/l; 48 hrs: 0.12 g/l), and acetate (0.11 g/l; 48 hrs only), while measured ethanol concentrations remain steady. The detection of succinate is a potential indicator of TCA cycle activity and the presence of lactate suggests active lactic acid fermentation via SR7 genome-annotated enzyme lactate

dehydrogenase. OD-normalized results indicate that metabolite production is significantly reduced on a per-cell basis after the 5-hour time point. Metabolite profiles in LB incubations at 24 and 48 hours, in contrast, appear streamlined, with only two major products detected: succinate (24 hrs: 0.50 g/l; 48 hrs: 0.49 g/l) and acetate (24 hrs: 0.12 g/l; 48 hrs: 0.15 g/l). OD-normalized values indicate that metabolites continue to be generated at nearly consistent levels on a per-cell basis through 48 hours. No additional volatile products (e.g. isobutanol, isopentanol, phenethyl alcohol) were detected by gas chromatography at any sub-sampled time points.

Observation of ethanol in LB cultures after 5 hours, but not at 24 and 48 hours appears to indicate that ethanol is being re-assimilated or utilized for alternative product generation. For example, the decrease in ethanol concentration between 24 and 48 hours is nearly equivalent to the increase in succinate concentration, suggesting that ethanol may be converted to acetaldehyde, acetate, and then acetyl-CoA (via the reversible enzymatic activity of alcohol dehydrogenase, aldehyde dehydrogenase, and acetate kinase, respectively; e.g. Camarena *et al.*, 2010) prior to entry into the TCA cycle. Since OD-normalized metabolite concentrations in M9+ also show a pattern of decreasing ethanol concentration with increasing alternative metabolite concentrations, the re-assimilation of ethanol may be taking place during growth in both media types. Alternatively, if end products are accumulating to toxic concentrations, it is possible that cultures may perpetuate growth by metabolizing components of dead cells.

SR7 germination induction

Because germination and growth of spores under $scCO_2$ conditions has previously been shown to be a stochastic process (Peet *et al.*, 2015), an effort was made to improve germination rates during $scCO_2$ incubations in order to be able to express heterologous enzymes more quickly and consistently. The literature has shown that a broad array of compounds, including several L-amino acids, and peptidoglycan are able to induce metabolically dormant endospores to germinate (Wei *et al.*, 2010). These inducers have to be shown to activate germination in *Bacillus* through several independent pathways (Hyatt and Levinson, 1962). Two amino acids (L-alanine and L-leucine) previously shown to induce germination

through different pathways were chosen for investigation with SR7 to increase the likelihood of success in an uncharacterized strain. Initial assays in LB under 1 atm CO_2 as a proxy for $scCO_2$ conditions demonstrated that L-alanine-amended cultures germinated by 4.5 hours, while unamended cultures grew between 4.5 and 6 hours after inoculation (Figure 11).

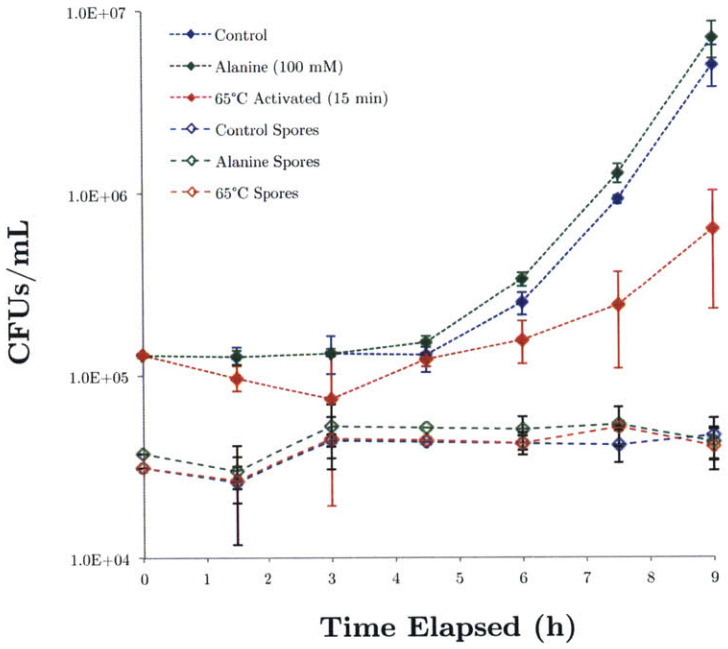


Figure 11. Germination of SR7 spores under 1 atm CO_2 in LB media, 100 mM L-alanine-amended LB, and LB heated at 65°C for 10 minutes. Open triangles representing viable spore counts based on heat-killed (80°C for 15 minutes) CFU counts.

After germination, growth occurs at nearly identical doubling times by OD_{600} (M9A+: 0.86 h; M9+: 0.89 h), suggesting that the effect of alanine is specific to the germination process rather than improved growth rates. Heat treatment reduced SR7 germination (marginal growth at 6 hours) rates and increased doubling times (1.11 h), despite previously being shown to induce spore germination spores for certain *Bacillus* species (Hyatt and Levinson, 1962). It is possible though that in the case of SR7, rather than inducing germination, the heat treatment is lethal to a sub-population of spores, decreasing the number of viable spores available to germinate and grow. During the initial period of vegetative outgrowth spore concentrations remain nearly constant (Figure 11). As a result, it appears that individual spores or sub-populations will germinate

and commence vegetative growth while remaining spores stay dormant, at least initially. Therefore, adding an inducer such as L-alanine provides a consistent source of growth potential to a pool of dormant cells.

Because phosphate buffered saline (PBS) solution does not have an available carbon source, spore-inoculated PBS cultures with inducer amendments under 1 atm CO_2 headspace enabled investigation of the anaerobic germination process independent of growth. Results assayed by fluorescence microscopy, bulk fluorescence and OD_{600} demonstrate that germination is induced by 3 hours, including with 100 and 250 mM L-alanine, 25 mM L-leucine and heat-treated 100 mM L-alanine (Figures 1 and 12). All incubated cultures showed an approximate 2-fold increase (1.9-2.1-fold) in fluorescence magnitude after 3 hours, eventually reaching a maximum at the 8.5 hour endpoint of 2.4-fold the bulk fluorescence (100 mM L-alanine) of PBS only incubated spores. By 24 hours, every inducer-amended sample also had a lower OD_{600} than unamended PBS samples, indicative of the flooding of the spore interior after spore coat degradation, decreasing the cell's index of refraction.

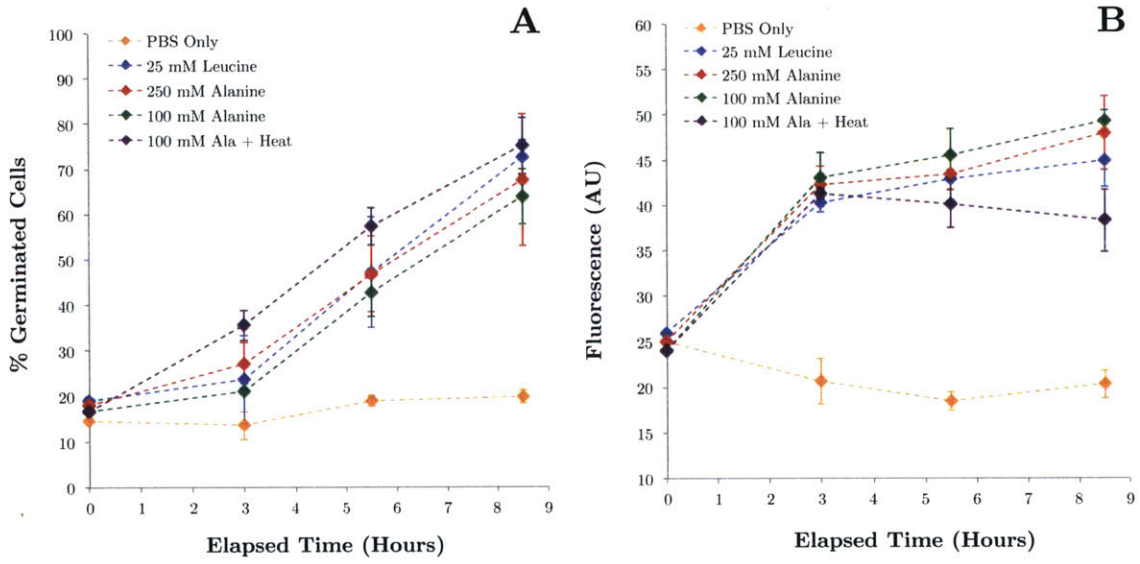


Figure 12. Assays tracking the extent of population-level germination progress in PBS buffer by **A)** cell stain pattern by fluorescence microscopy and **B)** culture bulk fluorescence.

A similarly pronounced inducer effect was observed by fluorescence microscopy direct filter counts based on spore staining patterns indicative of dormancy and germination. All treatments increased the percentage of

germinated cells by 3 hours relative to PBS incubated spores (Figure 12). According to filter counts, unamended PBS incubated spores maintained a constant, low-level abundance of germinated cells, increasing from 14.7% germinated at t_0, to 19.8% at 8.5 hours. Inducer-treated cultures increase more substantially, from 16.7-18.9% at t_0 to 63.8-75.1% at 8.5 hours. Since all cells that showed stain membrane penetration, including whole cell and center-localized (Figure 1), were considered "germinated," it is possible that "percentage germination" values in Figure 12A may be overestimates. Microscopic inspection of inducer-amended PBS incubations did not reveal any vegetative cell morphologies, suggesting that L-alanine and L-leucine are not being utilized as a carbon source for growth in PBS.

The effect of heat treatment on PBS cultured spores (also amended with 100 mM L-alanine) generated mixed results. After initially increasing in bulk fluorescence, heat-treated cultures steadily decreased in bulk fluorescence, while microscopy indicated that heat treatment increased germination to the highest observed frequencies (Figure 12A). However, this result may be due to spore coat damage during heat treatment at 70°C that caused more cells to become susceptible to membrane penetration by Syto9 cell stain. Therefore, without further physical evidence it is difficult to conclude whether the apparent germination inducing effect by heat treatment is a genuine result or false positive.

Delayed germination induction experiments in which spores incubated under 1 atm CO_2 for 14.5 hours in PBS were amended with L-alanine or L-leucine and assayed for germination after 9.5 hours demonstrated the capacity to actively germinate in the presence of inducers mid-culture. Higher concentrations of L-alanine and L-leucine did not appear to improve the extent of germination relative to lower concentrations, suggesting that the capacity for SR7 spores to be germinated saturates at or below 100 mM L-alanine and 10 mM L-leucine. The observed effects caused by both amino acid inducers were comparable in magnitude (Table 13). Follow-up investigation of a physiological state change in endospores caused by alanine amendment to carbon-free PBS cultures was ultimately verified by use of FCM, as explained below in section "Physiological signatures of induced germination."

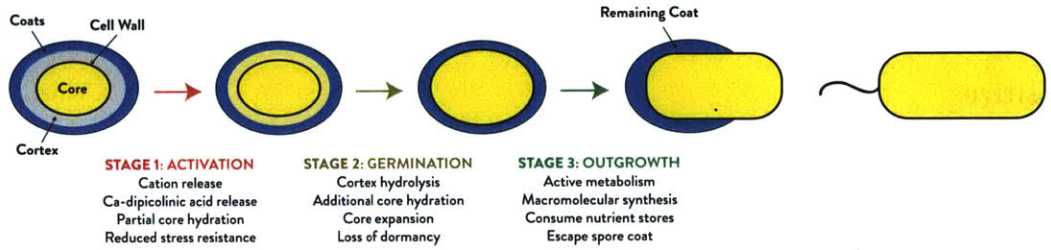
Table 13. Germination assays 9.5 hours after delayed induction in PBS under 1 atm CO_2 (OD decrease indicative of germination)

Inducer	mM	Fluorescence Fold Increase	OD_{600} Fold Decrease
L-alanine	250	1.4	1.2
	100	1.4	1.2
	25	2.1	1.5
L-leucine	25	2.0	1.2
	10	2.0	1.2

3.3.5 Physiological signatures of induced spore germination

Additional investigation using flow cytometry (FCM) sought to build upon preliminary LB-based evidence (Figures 11-12; Table 13) to verify a physiological effect of alanine on spores during the transition from dormant to germinated cell. FCM data collected on Syto16-stained SR7 cells from unamended and L-alanine-amended PBS cultures revealed two populations capable of gating on side and forward scatter (Figure 14): 1) PBS only incubated cells (Population 1) and 2) L-alanine-amended PBS cells (Population 2). Based

A



B

Classification	FCM Fluorescence	DNA Stain Distribution	Physiological State
"Dormant"	Low	Edge-localized	Intact Spore
"Activated"	Medium	Center-localized	Stage 1-2
"Germinated"	High	Whole cell	Stage 2-3

Figure 13. A) Schematic illustrating the physiological process of endospore germination (Adapted from Reineke, 2013; Setlow, 2003), and **B)** The defining traits and hypothesized corresponding physiological state of three detected spore populations.

on Syto16 and propidium iodide (PI) fluorescence intensities, three additional populations could subsequently be gated based on unique Syto16 and propidium iodide (PI) fluorescence signatures. These individual populations appear well correlated with visual evidence by fluorescence microscopy of three staining patterns of varying intensity (whole cell, center localized, edge localized; Figure 1). These staining patterns and germination stage categories can be thus be mapped onto each other schematically (Figure 13).

Population 1 (PBS) and 2 (PBS + L-alanine) display marked differences in terms of fluorescence magnitudes and distributions (Figure 14).

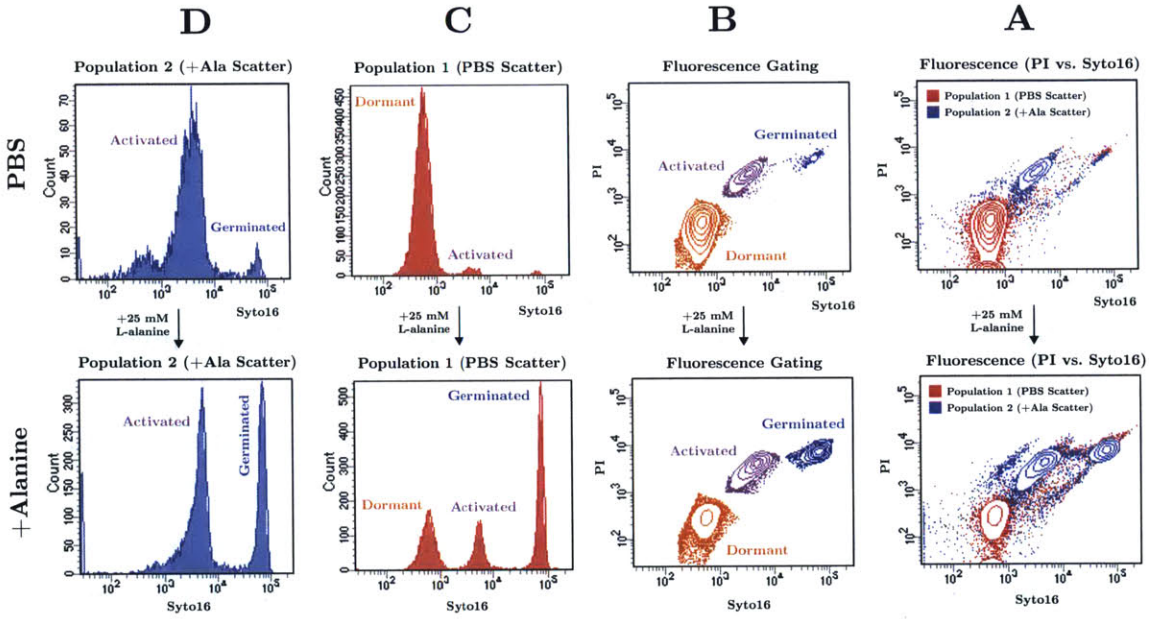
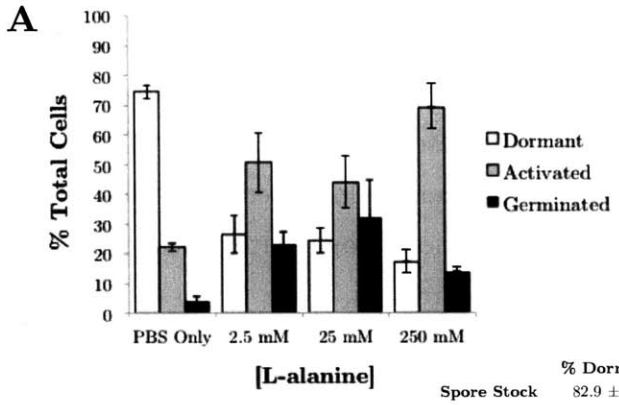


Figure 14. Flow cytometry results. **A)** Fluorescence distribution of all counts previously gated by side/forward scatter gates (red: Population 1, blue: Population 2). **B)** A subset of those counts were then gated by GFP and PI fluorescence intensity (orange: low/dormant; purple: intermediate/activated; blue: high/germinated. **C)** The distribution and intensity of Syto16 fluorescence in Population 1 and **D)** Population 2. The top panel is for SR7 spores incubated in PBS and the bottom panel is for SR7 spores incubated in PBS amended with 25 mM L-alanine.

A wide majority of spores incubated in PBS are gated as dormant cells (74.3%), with 22.0% and 3.7% gated as activated and germinated, respectively (Figures 14 and 15). When L-alanine is amended to PBS cultures, fluorescence distributions shift towards activated (44-69%) and germinated (1.38-31.8%) fluorescence signatures (Figures 14 and 15). These results reinforce the implication that L-alanine acts to induce physiological changes involved in the progression from dormant endospore to germinated cell.



B

	% Dormant	% Activated	% Germinated
Spore Stock	82.9 ± 3.2	1.0 ± 0.2	16.1 ± 2.1
PBS Only	74.3 ± 2.2	22 ± 1.2	3.7 ± 1.7
2.5 mM	26.5 ± 6.4	50.5 ± 9.9	23.1 ± 4.1
25 mM	24.2 ± 4.1	44 ± 8.9	31.8 ± 13
250 mM	17.2 ± 4.5	69 ± 7.0	13.8 ± 3.2

Figure 15. Summary of flow cytometry signatures of SR7 spores incubated in PBS or L-alanine-amended (2.5, 25, 250 mM) media. **A)** Summed distribution of Population 1 and Population 2 counts within each of the three Syto16/PI fluorescence gates. **B)** Values presented in plot, as well as spore stock distributions.

3.3.6 SR7 growth and activity under supercritical CO_2

Results generated under aerobic and 1 atm CO_2 conditions investigating the physiology, growth dynamics and germination induction of *B. megaterium* SR7 were integrated in an effort to generate robust growth and production of natural products under $scCO_2$. Chemical induction experiments in P-LB medium spore-loaded $scCO_2$ incubations (Table 14) revealed that L-alanine confers a statistically significant improvement in germination rates and growth outcomes relative to inducer-free cultures (Figure 16), while L-leucine reduced growth frequency under $scCO_2$ conditions relative to controls. Growth was defined as at least one order of magnitude increase in biomass according to epifluorescence cell counts:

1) high growth: ≥40-fold increase in direct cell counts relative to t_0 cell density
2) low growth: >10-fold increase in direct cell counts relative to t_0 cell density
3) germinated: <10-fold increase, mixture of vegetative cells and spores
4) dormant: <10-fold increase, only spore morphologies observed

Table 14. Growth outcomes for unamended and induced $scCO_2$ cultures

Incubation	Duration	Media	Growth
A	18 days	P-LB	3/7
		P-LBA	5/7
		P-LBA (+Heat)	3/6
		P-LBL	1/7
		P-LBAL	2/7
		Neg Ctrl	0/4
B	20 days	P-LB	1/7
		P-LBA	5/7
		P-LBL	0/6
		P-LBAL	2/6
		Neg Ctrl	0/4
C	18 days	P-LB	1/6
		P-LBA	3/5
		Neg Ctrl	0/4

Overall, growth was observed in 63% of all cultures amended with L-alanine, while only 36% of unamended reactors showed growth (Table 14). Median fold increase in cell concentration for P-LBA cultures was 37.5 and for unamended phosphate-buffered LB (P-LB; Table 1) was 22.8. Using growth frequency and fold change as inputs for non-parametric modeling of $scCO_2$ growth outcomes established that L-alanine conferred a statistically significant improvement on growth ($p = 0.0036$) relative to P-LB cultures by a Wilcoxon/Kruskal-Wallis Test. L-leucine (P-LBL media) only generated growth in 7.7% of reactors, while the combined treatment of L-alanine and L-leucine (P-LBAL) resulted in 31% growth frequency. Diminished growth in L-leucine reactors suggests a neutral to inhibitory effect on SR7 $scCO_2$ germination and growth, which is unexpected based on 1 atm CO_2 results (Table 14). As the L-alanine + heat treatment reactors (50%) also did not grow as well as non-heated L-alanine reactors, L-leucine and heat treatment were discarded as potential growth enhancing components of the microbial bioproduction system.

After verifying the positive growth effect of L-alanine on spore-loaded P-LBA $scCO_2$ cultures, two rounds of 18-20 day $scCO_2$ incubations of SR7 spores in M9A+ displayed growth in 11/18 reactors and 5/7 reactors, respectively. The total frequency of growth in M9A+ (64%) is thus comparable to P-LBA (63%), though M9A+ appears to increase the frequency of high level (>40 fold) growth (56% in M9A+ vs. 26% in P-LBA; Figure 16). Despite similar overall frequencies, median biomass accumulation was improved for cultures grown in M9A+ (64.3 fold increase) relative to P-LBA (37.5 fold).

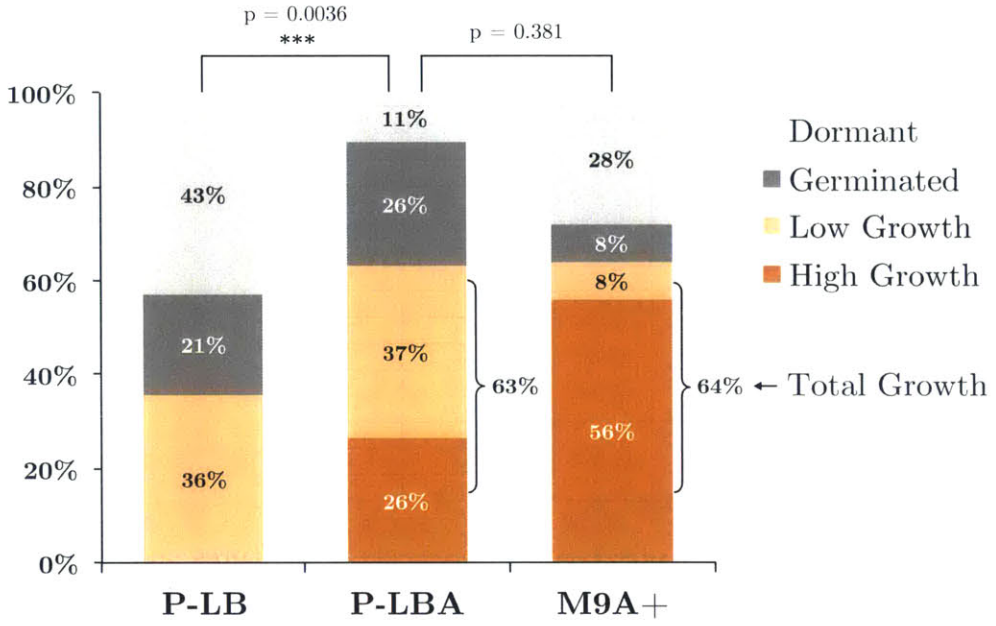


Figure 16. SR7 germination and growth frequencies under $scCO_2$ in respective media based on cumulative results pooled from multiple individual experiments, as summarized in Table 14.

As statistical tests did not establish significance ($p = 0.381$) in differential growth outcomes for P-LB and M9+, subsequent system development proceeded with semi-defined M9A+ minimal media, which simplifies pathway engineering architecture and metabolic flux analysis due to growth on a single carbon source. In order to more fully understand the relationship between starting SR7 spore concentration and likelihood of growth in M9A+ media, a logistic regression model for growth frequency was generated in part using data from an 18 day $scCO_2$ time course experiment (sampled at 6, 12, and 18 days) with starting spore concentrations varied over six orders of magnitude. The results of the incubation are summarized in Table 15, using a 10-fold increase in filter cell counts as the threshold for growth. After merging the time course growth data

Table 15. Growth summary of $scCO_2$ incubations in M9A+ media inoculated with a range of spore concentrations

t_0 [spores/mL]	6 days	12 days	18 days
$5.6x10^5$	1/4	2/4	3/4
$4.6x10^3$	0/4	0/4	2/4
[c] $5x10^1$	0/4	0/4	0/4
[c] $5x10^{-1}$	0/4	0/4	0/4
Media Only	n.d.	n.d.	0/6
[c] Below detection: recorded as 1.15×10^2			

(Table 15) with previously generated results from M9A+ incubated spores (Table 14) a total of 91 experimental samples and 24 negative controls subjected to logistic regression analysis demonstrated that both loaded spore density ($p = 0.0057$) and incubation time ($p = 0.003$) have statistically significant impacts on growth frequency, while the interaction of their effects was not significant ($p = 0.89$). The overall regression model generated the following equation (plotted in Figure 17) to describe growth frequency (Z) as a function of incubation time (X) and starting inocula concentration (Y).

$$Z = \frac{1}{1 + e^{4.7849 - 0.228 * X - (3.8207 * 10^{-6}) * Y}}$$

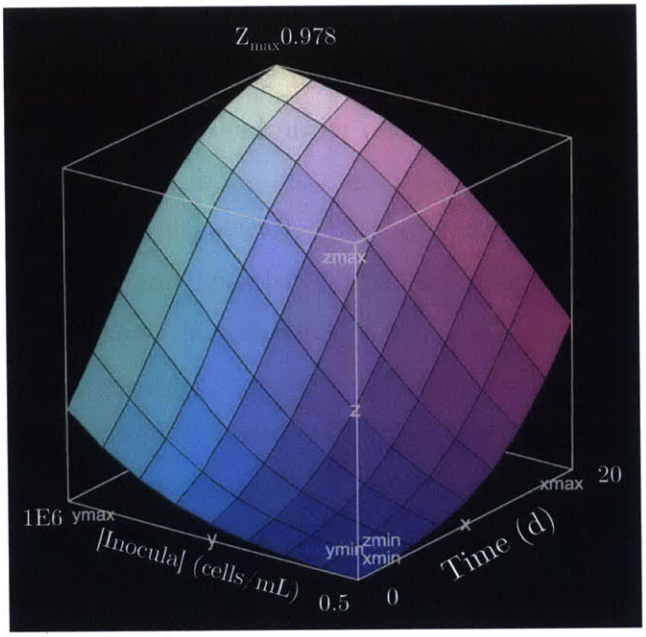


Figure 17. Nominal logistic regression of SR7 $scCO_2$ growth frequency (Z axis) as a function of inocula spore density ($p = 0.0057$; Y axis) and incubation time ($p = 0.003$; X axis).

SR7 fermentation products under $scCO_2$

Cultures from 6, 12, and 18-day SR7 time course incubations in M9A+ media demonstrating growth (>10-fold increase in cell counts) under $scCO_2$ were analyzed for natural fermentation products by HPLC. Cultures generated several detectable metabolites typically on the order of 0.1-10 mg/l, including for succinate (up to 2.5 mg/l), lactate (up to 13.3 mg/l), and acetate (up to 9.5

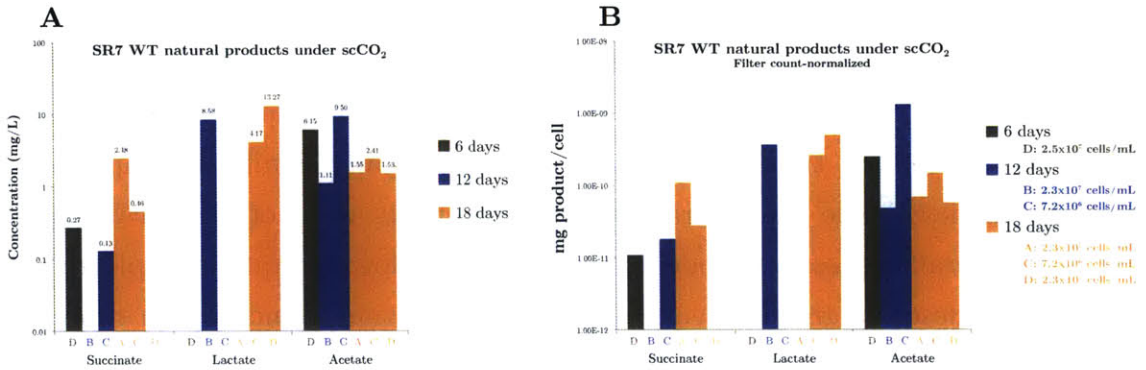


Figure 18. Natural fermentative products generated by SR7 cultures under $scCO_2$ showing growth, as detected by HPLC. **A)** total final titers and **B)** filter count-normalized per cell metabolite productivity. Final cell concentrations for each sample listed in legend. No metabolites were detected in media-only reactors and reactors without cell growth (data not shown).

mg/l) (Figure 18). All metabolites were also detected in culture under 1 atm CO_2 conditions grown in similar M9+ media (Figure 10), suggesting shared features of active fermentative pathways under both conditions. The presence of succinate suggests active use of the TCA cycle, while lactate (via genome-annotated lactate dehydrogenase) indicates utilization of lactic acid fermentation. Acetate production under $scCO_2$ was expected after its prior detection under all SR7 culturing conditions including in both M9-based and LB media. The absence of detectable ethanol in $scCO_2$ culture supernatant suggests a potential loss of volatile product during degassing, the use of alternative fermentative pathways under $scCO_2$, or re-assimilation of produced ethanol, as suggested to have potentially occurred under 1 atm CO_2. Overall, these metabolites are the first reported bioproducts generated by complex central carbon metabolism (rather than single enzyme reactions) under $scCO_2$. Normalization of product concentrations by total cell counts enables calculation of metabolite productivities on per cell basis. Maximum per cell productivity values (mg product $cell^{-1}$) are $1.1x10^{-10}$, $5.0x10^{-10}$, and $1.3x10^{-9}$ for succinate, lactate and acetate, respectively. These productivities are comparable to results observed under 1 atm CO_2 (assuming OD_{600} = 1.0 corresponds to 10^8 cells/ml, based on filter counts), which displayed maximum productivities (mg product $cell^{-1}$) in M9+ media of $5.5x10^{-10}$, $3.4x10^{-10}$, $5.8x10^{-10}$ and $7.0x10^{-9}$ for succinate, lactate, acetate and ethanol, respectively. Therefore, a relationship within roughly an order of magnitude appears to exist between concentration of per cell metabolite production and total cell numbers per culture.

3.4 Discussion

Bioprospecting supercritical carbon dioxide-resistant microbial strains through enrichment culture and serial passaging of deep subsurface McElmo Dome $scCO_2$ reservoir formation fluids enabled the isolation of six unique spore-forming *Bacillus* strains with the potential for bioplatform technology development. Bacilli have been detected at low levels in natural systems containing high pCO_2 and have shown the ability to grow in biphasic cultures exposed to $scCO_2$ (Peet *et al.*, 2015). In addition, because *Bacillus* species commonly display the capacity to grow on a broad substrate spectrum and demonstrate facultative anaerobic metabolism (Bunk *et al.*, 2010), the taxonomic identity of isolated strains is unsurprising, especially after passaging methods developed in Peet *et al.* (2015) proved effective at isolating facultative anaerobes. The shared genus of McElmo Dome environmental strains and previous $scCO_2$-tolerant isolates (Figure 2) may indicate that the capacity for sporulation, membrane structure (Peet *et al.*, *in review*) or another unifying trait of Firmicutes bacteria provides a fitness advantage enabling survival and growth in high pCO_2 environments. While a taxonomic survey of McElmo Dome conducted in Chapter 2 indicated the presence of several dominant taxa in formation fluids, including *Sulfurospirillum*, *Rhizobium* and several Clostridiales, several factors may have contributed to their inability to be isolated. First, the sampled fluids were subjected to depressurization over 15 hours (from >100 atm to 1 atm) shipped on ice and stored at 4°C until use for ~4 months, which may have proven stressful and lethal to thermophilic and even mesophilic species. Furthermore, the initial enrichment passage included aerobic filtering of fluids to concentrate cell biomass, which in addition to sample transport and handling may have caused oxygen poisoning of any exposed obligate anaerobes. As a result, the isolation of strains specifically capable of O_2 tolerance and sporulation makes sense due to the stresses associated with variable temperature, pressure, oxygen exposure, and culture media.

Follow-up incubation experiments with isolated strains in pure culture under $scCO_2$ established *Bacillus megaterium* SR7 as the best growing McElmo Dome isolate strain under pressure in terms of biomass accumulation and

frequency of spore-inoculated culture growth (Table 6). Most other strains (and to a lesser extent *B. licheniformis* MR4 and *B. safensis* FM4) demonstrated low-level stochastic growth frequencies in $scCO_2$ pure culture (Tables 4 and 6). As a result, downstream process engineering and strain development optimizations focused on improving growth outcomes in environmental isolate strain *B. megaterium* SR7. Additional motivation for proceeding with *B. megaterium* SR7-based system development is that *B. megaterium* is an extremely well-described species due to its biotechnological and industrial utility in generating chemical products through natural and heterologous pathways (Korneli *et al.*, 2013). Vegetative cells of *B. megaterium* are physically quite large (4 x 1.5 ug; Bunk *et al.*, 2010) and because as gram-positive bacteria they lack an outer cell membrane are thus able to secrete significant amounts of enzymes or protein products (Korneli *et al.*, 2013). In addition to the potential industrial utility afforded by exploitable secretion systems, enzymatic secretion enables *B. megaterium* cells to degrade complex polymeric nutrients like sugars, peptides and lipids into simpler and smaller substrates able to be taken up into the cell for consumption (Bunk et al., 2010).

Physiological characterization of SR7 established viable growth ranges in terms of pH and salinity (Table 8) that are consistent with the *in situ* conditions in deep subsurface McElmo Dome formations (Thesis Chapter 2). However, assays demonstrating 45°C as an upper temperature boundary for viable growth indicates that SR7 may not be metabolically active *in situ* at McElmo Dome where temperatures in the fluid-sourced geologic formation are typically above 60°C (Thesis Chapter 2). Instead, SR7 (and possibly all isolated strains) appears more prone to persist as dormant endospores subject to low frequency condition-independent stochastic germination than vegetative growth, in line with the "microbial scout hypothesis," which articulates that sub-populations attempt growth rather than putting complete dormant communities at risk (Buerger *et al.*, 2012; Peet *et al.*, 2015). The detection of 16S rRNA genetic signatures of *Bacillus* in the McElmo Dome Well 3 community survey (<2% abundance; Thesis Chapter 2) and other deep subsurface formations (Nicholson, 2002; Vary, 1994) reinforces the notion of *in situ* habitation, possibly emplaced through saline aquifer fluid migration, including with access to the $scCO_2$ formation by localized geologic faulting structures (Holloway *et al.*, 2005). Therefore, despite not

growing at *in situ* elevated temperatures, superior growth of SR7 under $scCO_2$ than other tested strains (this study; Peet *et al.*, 2015) suggests potential acclimation to $scCO_2$ exposure. In regard to the $scCO_2$ bioproduction scheme, however, the ability to reliably grow at *in situ* temperatures is less relevant than the strain's demonstrated ability to grow at temperatures near and above the CO_2 critical point of 31.1°C, which due to the tunability of $scCO_2$'s chemical properties near its critical point (Matsuda *et al*, 2005) will enable downstream culturing and product extraction optimization.

Characterization of SR7 metabolic flexibility and preferred growth conditions will in the future enable culturing optimization that enhances reproducible growth as a means for high bioproduct yields. Consistent with the general capacity for *B. megaterium* to grow on a wide variety of compounds (Vary, 1994), strain SR7 demonstrated a broad substrate spectrum, including organic acids, sugars and amino acids according to Biolog phenotypic fingerprinting (Table 11). In the development of semi-defined minimal media, SR7's inability to grow robustly without a trace metals amendment (Tables 12; Figure 6) is consistent with a previous study projecting that *B. megaterium* is dependent on calcium, manganese, cobalt and especially magnesium for biomass formation and product generation (David *et al.*, 2010; Korneli *et al.*, 2013). As M9 base does not contain manganese or cobalt, these two elements may be specifically responsible for the observed phenomenon. The iterative development of a semi-defined minimal medium (M9+) optimized for SR7 enabled consistently robust growth under 1 atm CO_2, a crucial microbial physiology breakthrough for an environmental strain whose active metabolism under $scCO_2$ is foundational to the proposed microbial bioproduction system.

Functional annotation of the bioplatform strain SR7 genome provides actionable insights into metabolic and physiological capacities that may be exploited for pathway engineering and development of the proposed microbial $scCO_2$ bioproduction system. Specifically, upon developing the capacity for chromosomal integration of exogenous genes, the ability to knock out or modulate expression (Brockman *et al.*, 2015) of growth pathways to limit carbon flux away from compound production is predicated on knowledge of central carbon metabolism. Natural fermentative products detected from SR7 under 1 atm CO_2 and $scCO_2$ provide insight into potential competing pathways and

active fermentation capacity. Furthermore, the detection of metabolites under $scCO_2$ represents the first ever endogenous products generated by active central carbon pathway metabolism under $scCO_2$. *B. megaterium* strains typically grow using the glycolytic Embden-Meyerhof-Parnas (EMP) pathway and pentose phosphate (PP) pathway upstream of the TCA Cycle (Korneli *et al.*, 2013). Furthermore, *B. megaterium* strains are known to employ mixed acid fermentation via alcohol/aldehyde dehydrogenase (Adh; Slostowski *et al.*, 2012) and are unique in utilizing the glyoxylate pathway (Korneli *et al.*, 2013), which enables re-assimilation of acetate generated as a fermentative product. Based on genomic annotations, culturing experiments and natural products surveys, strain SR7 metabolism appears consistent with previously observed *B. megaterium* traits, including full EMP, PP and TCA cycle pathways and the presence of fermentative genes lactate dehydrogenase (*ldh*), acetolactate synthase (*als*S), acetolactate dsecarboxylase (*als*D), phosphate acetyltranferase (*pta*) and acetate kinase (*ack*A). Glyoxylate pathway genes isocitrate lyase (*ace*A) and malate synthase (*glc*B) are also present in SR7, indicating the genomic capacity to recycle acetate for product generation, improving substrate utilization efficiency by recouping initial byproduct losses to acetate. The suite of detected fermentative products in both M9+ and LB media (Figures 10 and 18) appear to confirm active use of annotated central carbon pathways, especially fermentation to generate ethanol, lactate and acetate, as well as the production of major TCA Cycle intermediate succinate. In OD-normalized LB incubated cultures, acetate concentration appears to decrease with time, potentially indicating the use of the glyoxylate pathway to re-assimilate acetate for alternative compound production. Transcriptomic and/or proteomic investigation may further elucidate the active utilization of this pathway.

A critical development established by this study towards the realization of a $scCO_2$ bioproduction system was the optimization of growth under 1 atm CO_2 as an effective proxy for $scCO_2$ culturing conditions (as exhibited in Peet *et al.*, 2015). As *B. megaterium* was considered a model system for studying gram-positive organisms for several decades, especially in regards to studies on sporulation and germination, a significant amount of literature describing the physiology, metabolism, genomics, and genetics of *B. megaterium* vegetative cells and endospores aided in the process of developing *B. megaterium* SR7 as a

bioproduction host. One of the key insights provided by the literature was the correlation between robust *B. megaterium* growth and disruptive mixing (Santos *et al.*, 2014). 1 atm CO_2 culturing experiments reinforced this finding, as improved doubling times (Figure 5) and possibly spore germination frequencies (Figure 4) were induced by increased mixing rates.

Additional literature on the physiology of *Bacillus* spore germination (Wei *et al.*, 2010; Cronin and Wilkinson, 2007), including studies specific to *Bacillus megaterium* (Levinson and Hyatt, 1970; Hyatt and Levinson, 1962; Roth and Lively, 1956; Vary, 1973), elucidated that several independent pathways result in germination activation, including through conditional (e.g. pressure; temperature; Wei *et al*, 2010) or chemical (Ghosh and Setlow, 2009) induction. As a result, in confronting previous challenges associated with stochastic growth frequencies under $scCO_2$ conditions (Peet *et al.*, 2015), an attempt was made to deliberately induce spore germination through chemical induction by L-alanine and L-leucine. The working hypothesis was that while most spores remain dormant during incubation, if sub-populations can be caused to germinate, subsequent vegetative outgrowth will enable active metabolic pathway expression, including that of endogenous (and ultimately heterologous) proteins. Investigation of this hypothesis in PBS suspension experiments assayed by fluorescence microscopy (Figure 12A), bulk fluorescence (Figure 12; Table 13) and OD_{600} (Table 13) demonstrated that L-alanine and L-leucine exposure induced signatures of transitional states (dormant, activated, germinated) of germinated spore physiology (Figure 13). Results from initial heat induction experiments were consistent with chemical inducers (Figure 12A), though may reduce population-wide growth viability through spore-coat damage (Figures 11 and 12B; Setlow, 2006). FCM methods developed in this study hold potential for significant future use with $scCO_2$-grown cultures in characterizing the distribution of spore states in induced and non-induced incubations. These results would elucidate details on the timescale and frequency of spore germination that may inform additional growth improvements. As initial tests with $scCO_2$-incubated cultures revealed significant detritus that confounded gating, further methods development is necessary to develop a FCM assay specific to $scCO_2$-incubated cultures.

Chemical induction results were further investigated in depth, and can thus also be understood mechanistically in terms of physiological cell

modifications during germination. Cells form multiple layers of protection as they sporulate, including the growth of a spore coat and cortex around the central core (Setlow *et al.*, 2006). Based on the results of the PBS incubation FCM analysis, it appears that L-alanine accelerates the process of cortex hydrolysis and spore coat escape by activating a sub-population of cells into an intermediate and then final stage of germination. Both stages (as demonstrated by mid-range and high-level fluorescence intensities; Figures 12-15) are hypothesized to confer a physiological plasticity that increases the likelihood of germination and outgrowth relative to dormant endospores. As a result, germination inducers may be considered a new class of bioprocess engineering tools that may be exploited in allowing sporulated inocula to survive initial exposure to stressful conditions (that would otherwise be lethal to vegetative cells) before germinating and commencing outgrowth, as with SR7 in $scCO_2$.

Successful transition of L-alanine-induced germination capacity from 1 atm CO_2 to $scCO_2$ conditions was a fundamentally enabling improvement of high pressure culturing outcomes in both P-LBA and M9A+ media (Figure 16). Relative to Peet *et al.* (2015), which observed overall $scCO_2$ growth frequencies of between 33% (MIT0214) and 55% (MITOT1) after at least 20 days, SR7 outperformed these strains by demonstrating 64% growth frequency after 18-20 days in M9A+ when inoculated with $\sim 10^4$ spores/ml (Figure 16), and 75% when inoculated with $\sim 10^5$ spores/ml (Table 15). The generation of a nominal logistic regression (Figure 17) provides mathematical predictive power regarding expected growth outcomes, which informs the design of high-throughput $scCO_2$ culturing experiments (including inocula concentrations and incubation duration) and provides a benchmark against which to measure future system improvements.

The novel demonstration of natural microbial product generation under $scCO_2$ (Figure 18) represents proof of concept for extracellular chemical production. Building upon this finding, heterologous compound generation under $scCO_2$ by SR7 via natural or engineered metabolic pathways appears technically feasible. The logical next step in the development of this microbial bioproduction scheme is thus the development of a genetic system for SR7 to enable heterologous enzyme production of compounds demonstrating chemical compatibility with $scCO_2$ for direct *in situ* extraction. Since a vast range of

recombinant proteins have been produced in *B. megaterium* (Bunk *et al.*, 2010, Martens *et al.*, 2002) for pharmaceutical, industrial and energy-related applications, there are myriad prospective compounds that represent chemically viable products for $scCO_2$-exposed production and harvesting. Specifically, genomic sequencing of SR7 reveals that isobutanol production by the valine-specific amino acid biosynthesis pathway should be possible with the addition of only two exogenous genes. As short-chain alcohols also readily partition into the $scCO_2$ phase at a rate of approximately 4:1 (Timko *et al.*, 2004) and biofuel production has been demonstrated extensively in *Bacillus* species, including 1-butanol in *B. subtilis* (Nielsen *et al.*, 2009), short-to-medium chain alcohol production would be an exciting potential proof of principle demonstrating the industrial utility of a $scCO_2$ bioreactor system using bioproduction host *B. megaterium* SR7.

3.5 Conclusions

This study used a bioprospecting approach to identify and isolate a $scCO_2$-compatible strain, *B. megaterium* SR7, which was developed by process engineering and culturing modifications in order demonstrate the capacity for enhanced growth and natural product generation under $scCO_2$. Specifically, the development of M9+ minimal media with trace metals amendments and optimized mixing regimes promoted improved growth rates, while L-alanine demonstrated the capacity to increase endospore germination frequency under 1 atm CO_2 and $scCO_2$ conditions. Genome sequencing provided insights into potentially useful and competing pathways with regard to bioproduct generation and may enable genetic tool development using endogenous features. The detection of natural fermentative products under 1 atm CO_2 and $scCO_2$ were consistent with genomically annotated pathways and demonstrate proof of concept that extracellular product generation is possible under $scCO_2$. Altogether, this work represents a major step towards the use of $scCO_2$ as a solvent for *in situ* extraction of endogenous or heterologous bioproducts. Because *B. megaterium* strains have previously been used to produce compounds for food, energy, pharmaceuticals and household goods industries both by natural and

engineered metabolic pathways, future development of bioproduction strain SR7 may be able to target the production and secretion of valuable industrial products while utilizing sustainable solvent $scCO_2$.

3.6 Acknowledgements

I thank Dr. Kevin Penn for his significant contributions to the work presented in this chapter. Specifically, Dr. Penn prepared SR7 genomic DNA for PacBio sequencing and through analysis of returned contigs assembled and closed the SR7 genome.

4

METABOLIC ENGINEERING OF *BACILLUS MEGATERIUM* SR7 FOR HETEROLOGOUS GENE EXPRESSION AND ADVANCED BIOFUEL SYNTHESIS AND RECOVERY UNDER BIPHASIC AQUEOUS-SUPERCRITICAL CARBON DIOXIDE CONDITIONS

ABSTRACT

Bacillus megaterium SR7 is biocompatible with the solvent supercritical (sc) CO_2 and has been shown to exhibit reproducible growth and natural product biosynthesis under scCO2 in an optimized growth medium. In this study development of a genetic system for strain SR7 was pursued to enable heterologous enzyme expression and synthesis of engineered products that partition into a $scCO_2$ headspace, enabling energy-efficient in situ compound extraction. Genetic transformation of SR7 was achieved using a protoplast-based method that introduced host-compatible plasmid pRBBm34. Plasmid maintenance in SR7 was demonstrated under aerobic, 1 atm CO_2, and $scCO_2$ conditions. Xylose-inducible (P_{Xyl}) and IPTG-inducible ($P_{Hyper\text{-}spank}$) promoters cloned into plasmids pJBxL and pJBhL enabled heterologous expression of the reporter LacZ under 1 atm CO_2 and $scCO_2$ conditions. Next, the two-gene pathway for biosynthesis of the C4 advanced biofuel isobutanol from the precursor α-ketoisovalerate (α-KIV) was cloned into pJBxKA6 under the xylose-inducible promoter. Variants of the pathway gene alcohol dehydrogenase were tested under aerobic and 1 atm CO_2 conditions to optimize isobutanol yield and decrease accumulation of pathway intermediate isobutyraldehyde. Initial tests of the optimized two-gene pathway introduced into strain SR7 and incubated under $scCO_2$ resulted in generation of up to 93.5 mg/l isobutanol and 29.7 mg/l isopentanol from 5 mM (580.6 mg/l) of α-KIV indicating a yield of 21.2% with partitioning of 5.2% of the isobutanol product into the $scCO_2$ headspace. This result represents the first demonstration of heterologous product synthesis and

extraction in a single $scCO_2$-exposed bioreactor. Use of $scCO_2$ as a solvent in a dual phase harvesting system would reduce the likelihood of bacterial contamination, help to relieve end product toxicity effects, enable energy-efficient extraction of products and alleviate the need for additional product dehydration due to the desiccating properties of $scCO_2$. Establishment of a genetic system for heterologous protein production in $scCO_2$ biocompatible strain Bacillus SR7 now provides a conduit to exploiting the $scCO_2$ solvent phase, which was previously thought inaccessible for microbial-mediated product generation and solvent extraction. Now that this specialized growth system has proven feasible, attention may turn to increased pathway complexity, genetic tool development, optimization of reactors to promote increased $scCO_2$ extraction efficiency, and economically viable applications of this novel technology.

4.1 INTRODUCTION

This study aimed to develop environmental strain *Bacillus megaterium* SR7, isolated from the McElmo Dome supercritical carbon dioxide ($scCO_2$) system (Thesis Chapter 3), into a host for heterologous biochemical production coupled with *in situ* product extraction by $scCO_2$ stripping. The exposure of growing cultures to $scCO_2$ provides several advantages over ambient or anaerobic growth reactor systems because $scCO_2$ rapidly sterilizes nearly all bacteria (Spilimbergo and Bertucco, 2003), while affording unique solvent properties for purified product extraction and relief of concentrated end product toxicity. After improving growth capacity of strain SR7 under $scCO_2$ through media and culturing optimization, the development of a genetic system would hold the potential to enable heterologous enzyme expression and product generation under $scCO_2$. Microbial bioproduction under $scCO_2$ solvent would thus represent a feasible new method for compound production and extraction in a single bioreactor system with living cells. Though $scCO_2$ has previously been used as a solvent for biocatalysis and chemical extraction from *in vitro* systems, and as a substrate for carboxylation using cells of *B. megaterium* PYR 2910 (Matsuda *et al.*, 2005, 2001), $scCO_2$ extraction of products generated by robust growth and heterologous enzyme production has not previously been demonstrated.

A significant factor limiting the capacity for environmental isolates to be utilized for industrial applications is the challenge of genetic intractability. Previous work establishing genetic systems has enabled investigation and exploitation of bioprospected strains with unique metabolic properties for applications in wastewater treatment/bioremediation (Coppi *et al.*, 2001), and the production of pharmaceutical and agricultural agents (Xiong *et al.*, 2013). *Bacillus* strains have been successfully engineered for a range of applications, including the production of biofuels and industrially relevant compounds (Nielsen *et al.*, 2009; Hu and Lidstrom, 2014). *B. megaterium* in particular has long been considered an attractive bioproduction host due to its capacity to stably maintain plasmids, lack of endogenous endotoxins and alkaline proteases, high protein secretion, facultative anaerobic metabolism, and an ability to grow directly on inexpensive sole carbon sources (Vary *et al.*, 2007; Korneli *et al.*,

2013). Promisingly, *B. megaterium* strains QM B1551 and DSM319, and their derivatives, have been used as hosts for protein and bio product expression for over 30 years (Vary *et al.*, 2007).

The production of biofuels is compelling with regard to $scCO_2$ harvesting systems due to the semi-hydrophobic chemistry of alcohols like isobutanol and butanol, which readily causes compound partitioning from the aqueous phase into $scCO_2$ (i.e. K_{ow} >4; Timko *et al.*, 2004). Product partitioning could be harnessed for *in situ* compound recovery concomitant with alcohol biosynthesis in aqueous growth media, i.e. single-step continuous flow $scCO_2$ extraction of 1-butanol has demonstrated nearly complete recovery of 1-butanol (up to 99.7 wt%) from aqueous solutions (Laitinen and Kaunisto, 1999). Advanced biofuels also represent attractive products for $scCO_2$ extraction because current harvesting technologies (e.g. steam stripping, gas stripping, adsorption, pervaporation) often require energy-intensive dehydration steps, which may be mitigated due to the low solubility of water in $scCO_2$ (Sabirzyanov *et al.*, 2002).

Advanced biofuel production is further motivated by performance improvements over common gasoline additive ethanol, which displays low energy density (~70% of gasoline), high hygroscopicity (ability to hold water), and elevated corrosiveness relative to longer chain hydrocarbons (Connor and Liao, 2009). As a result, C3-C5 alcohols are better suited for integration with current transportation infrastructure (Nigam and Singh, 2011). Higher chain alcohols may also be biocatalytically dehydrated to alkenes as a feedstock for sustainable generation of commodity chemicals including paints, surface coatings, solvents, plastics, and resins (Connor and Liao, 2009). As a result, while short-to-medium chain alcohols (e.g. isobutanol, isopentanol) were the first targets of interest for this study, $scCO_2$ holds high potential for extraction of a broad range of high-value bioproducts.

Biofuel pathway construction in heterologous hosts has been broadly successful (Liu and Qureshi 2009), including in *Escherichia coli* (Atsumi *et al.* 2008; Inui *et al.* 2008; Nielsen *et al.* 2009), *Saccharomyces cerevisiae* (Steen *et al.* 2008), *Clostridium ljungdahlii* (Kopke *et al.* 2010) and organic solvent-tolerant strains *Pseudomonas putida* S12 and *Bacillus subtilis* KS438 (Nielsen *et al.*, 2009). As end-product inhibition often limits the industrial utility of bioproduction strains (Nielsen *et al.*, 2009; Liu and Qureshi, 2009; Ezeji *et al.*, 2010),

acclimatization and media modifications have typically been used to improve product tolerance, including for 1-butanol with *B. subtilis* (Kataoka *et al.*, 2011) and *S. cerevisiae* (Lam *et al.*, 2014). Further work using metabolic engineering and directed evolution may help guide researchers toward more efficient biochemical tolerance, especially in conventional batch bioreactor systems that lack the capacity for *in situ* product extraction.

Isobutanol production requires modification of the amino acid valine biosynthesis pathway by directing flux of the intermediate α-ketoisovalerate (α-KIV) away from L-valine production and instead towards isobutyraldehyde and finally isobutanol (Atsumi *et al.*, 2008; Figure 1). α-KIV itself is generated from the condensation of two pyruvates (via pyruvate kinase, Pyk), which is decarboxylated (via acetolactate synthase, IlvIH) to form 2-acetolactate, then reduced (via acetohydroxy acid isomeroreductase, IlvC) and dehydrated (via dihydroxy acid dehyratase, IlvD) to α-KIV. Insertion of exogenous pathway genes for keto-acid decarboxylase (*kiv*D) and alcohol dehydrogenase (*adh*) then facilitates isobutanol production from α-KIV.

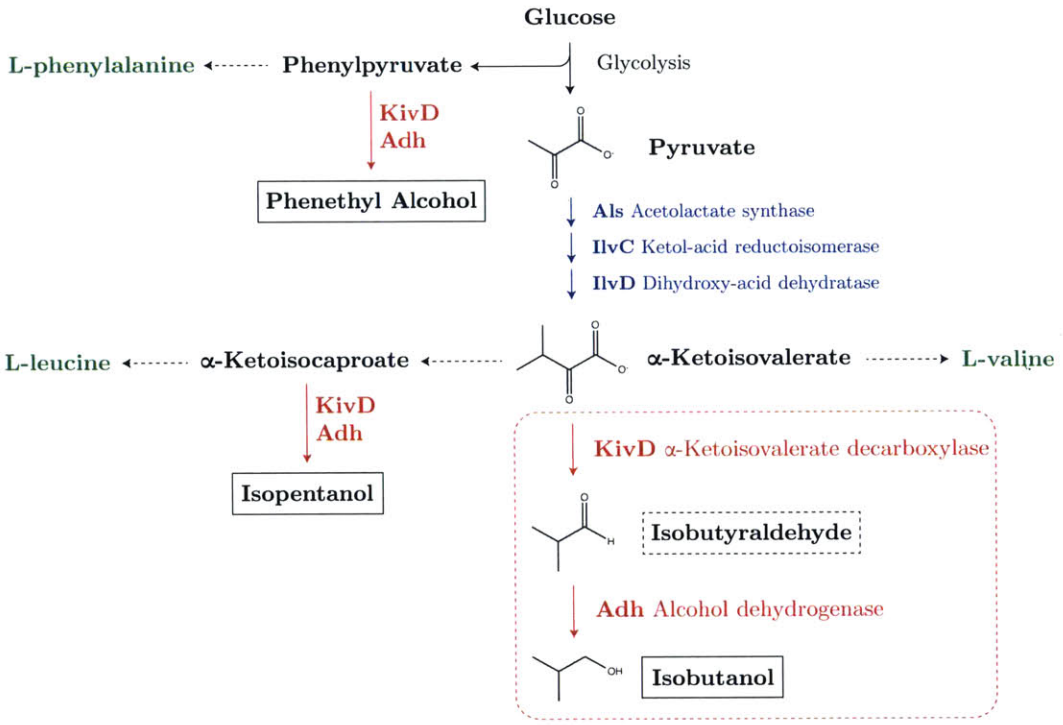


Figure 1. Isobutanol biosynthesis pathway from α-KIV encompassed by red dashed line. Amino acids highlighted in green. Heterologous enzymes highlighted in red. Enzymes highlighted in blue hold potential for upstream optimization of isobutanol pathway, but were not considered in this study.

As a result, biosynthesis of isobutanol should be possible directly from glycolytic processing of glucose to pyruvate, which may then enter the isobutanol pathway. Metabolic engineering of this nature has already been demonstrated in several organisms, including *Bacillus subtilis* (Li *et al.*, 2011), *S. cerevisiae* (Lee *et al.*, 2012), and *E. coli* (Atsumi *et al.*, 2008).

Overall, this study provides proof-of-concept for a two-phase harvesting method for stripping of microbially produced chemicals using *in situ* $scCO_2$ extraction. This study marks the first development of a genetic system for expression of single and multi-gene pathways under $scCO_2$ and the first demonstration of *in situ* bioproduct recovery by partitioning into $scCO_2$. This breakthrough establishes a new branch of microbial bioproduction by enabling access of an engineered bacterial strain to the unique properties of sustainable solvent supercritical carbon dioxide.

4.2 Methods

Strain, media and culture conditions

Environmental strain *Bacillus megaterium* SR7 was isolated through enrichment culture and serial passaging of fluids sourced from the deep subsurface McElmo Dome supercritical CO_2 formation (Thesis Chapter 3). Previous SR7 development under 1 atm CO_2 included the formulation of semi-defined minimal medium M9+, which consists of M9 base medium amended with 0.4% D-glucose, 50 mM yeast extract, 0.1X trace metals solution (Boone *et al.*, 1989). The addition of 100 mM L-alanine to M9+ (resulting in medium "M9A+") was previously shown to increase rates of SR7 spore germination and growth rate under $scCO_2$ conditions. Therefore, all culturing experiments conducted under 1 atm CO_2 occur in M9+ medium, and under $scCO_2$ in M9A+ medium. All cultures were incubated at 37°C and 250 rpm based on previous results showing enhanced growth rates and population longevity under these conditions. All 1 atm CO_2 and $scCO_2$ experiments were prepared within an anaerobic chamber (Coy Products) containing an atmosphere of 95% CO_2 and 5% H_2. Experiments conducted under 1 atm CO_2 used 10 ml of CO_2-degassed culture media in 100 ml serum vials with clamped rubber stoppers. Incubations under $scCO_2$ used $\frac{3}{4}$ inch

316 stainless steel tubing fitted with quarter turn plug valves (Swagelok or Hylok) for 10 ml total capacity. As previously described (Thesis Chapter 3), reactors were filled to ½ capacity (5 ml) with inocula and degassed media, after which the headspace was pressurized with extraction grade CO_2 gas at a rate of 2-3 atm min^{-1} until reaching a final pressure of 100 atm. After pressurization, reactors were incubated in a 37°C warm room and mixed at 250 rpm until unloading.

Development for SR7 genetic manipulation and expression

Vector construction

All primers used in plasmid construction, final vector constructs, transformed strains and associated references are presented in Tables 1A and 1B. The *lac*Z gene was PCR amplified from plasmid pKVS45 LacZ_LVA with primers LacZ_F and LacZ_R. Shuttle vector pRBBm34 (Amp^R (*E. coli*), Tet^R (*B. megaterium*); pBR322 Ori (*E. coli*), RepU (*B. megaterium*)) was used as a scaffold for pathway genes. PCR products and pRBBm34 were digested with SpeI and SphI prior to ligation to create the pJBxL plasmid (Figure 2A). The xylose repressor and promoter of pRBBm34 were replaced with a hyper-spank promoter ($P_{Hyper\text{-}spank}$) and *lac*I using circular polymerase extension cloning (CPEC). The pRBBm34 plasmid was PCR linearized with two sets of primers to remove xylR and P_{Xyl}: pRBBm34_F / Bla_R and Bla_F / pMM1520R. $P_{Hyper\text{-}spank}$ and *lac*I were PCR amplified from pDR111 using pMM1520-P_{Hysp}_F and LacI-pRBBm34_R. Standard CPEC cloning was used to assemble the three PCR products into the $P_{Hyper\text{-}spank}$ plasmid. The *lac*Z gene with a ribosome-binding site was PCR amplified from the plasmid pKVS45 LacZ_LVA using: RBS-LacZ_F and LacZ_R. PCR products and the $P_{Hyper\text{-}spank}$ plasmid were digested with SalI and SphI prior to ligation to create the pJBhL plasmid.

Table 1A. Primers used for vector construction

Name	Sequence (5' > 3')	Target	Reference
LacZ_F	GTCCAAACTAGTACCATGATTACGGATTCACTGGC	pKVS45 LacZ_LVA	Solomon et al., 2012
LacZ_R	CCGCCGGCATGCTCATTATTTTTGACACCAGACCAACTGG		
pRBBm34_F	CGGCGGCACCTCGCTAAC	pRBBm34	Biedendieck et al., 2007
Bla_R	GGTGCCTCACTGATTAAGCATTGG		
Bla_F	CCAATGCTTAATCAGTGAGGCACC	pMM1520	Malten et al., 2005
pMM1520_R	AGATCCACAGGACGGGTGTG		
pMM1520-P_{Hyp}_F	CACACCCGTCCTGTGGATCTGACTCTCTAGCTTGAGGCATC	pDR111	Guerout-Fleury et al., 1996
LacI-pRBBm34_R	GTTAGCGAGGTGCCGCCGGGATCCTAACTCACATTAATTGCG		
RBS-LacZ_F	AGCTTAGTCGACAGGGGGAAATGTACAATGACCATGATTACGGATTCACTGGC	pKVS45 LacZ_LVA	Solomon et al., 2012
KivD_F	GTCCAAACTAGTATGTATACAGTAGGAGATTACCTATTAGACCG	pCOLA KivD_Fjoh_2967	Sheppard et al., 2014
KivD_R	GAGGAGCATGCGAGCTCGGATCCTCATTATGATTTATTTTGTTCAGCAAATAGTTTACCC		
RBS-ADH6_F	GAGGAGGGATCCTCGACAGGGGGAAATGTACAATGAGCTACCCGGAAAAGTTCG	pACYC (car,sfp)	
ADH6_R	CCGCCGGCATGCAATGCGGCCGCTCATTAGTCGCTGAATTCTTTATCGTAACCAACC		
RBS-YqhD_F	TAATGAGGATCCTCGACAGGGGGAAATGTACAAATGAACAACTTTAATCTGCACACCC	*E. coli* MG1655 gDNA	Common lab strain
YqhD_R	GCATGCAATGCGGCCGCTCATTAGCGGGCGGCTTCGTATATAC		

Table 1B. Vector constructs and strains used in this study

Strain	Plasmid	Description / Genotype	Reference
B. megaterium SR7	Endogenous only	Wild-type isolate from $scCO_2$ subsurface formation	Thesis Ch. 3
SR7JR1	pJR1	CmR; mob, oriT , rep (*E. coli*), pUCTV2 ori^{ts} *(Bacillus)*, sacB (*B. subtilis*)	Richhardt et al., 2010
SR7x	*pJBx	P_{Xyl}-empty construct	This study
SR7xL	*pJBxL	P_{Xyl} lacZ; Tet^R	This study
SR7h	*pJBh	$P_{Hyper\text{-}spank}$-empty construct; Tet^R	This study
SR7hL	*pJBhL	$P_{Hyper\text{-}spank}$ lacZ; Tet^R	This study
SR7xK	*pJBxK	P_{Xyl} $kivD_{Ll}$; Tet^R	This study
SR7xKA6	*pJBxKA6	P_{Xyl} $kivD_{Ll}$, $adh6_{Sc}$; Tet^R	This study
SR7xKY	*pJBxKY	P_{Xyl} $kivD_{Ll}$, $yqhD_{Ec}$; Tet^R	This study
SR7xKAB	*pJBxKAB	P_{Xyl} $kivD_{Ll}$, $adhA_{Bm}$; Tet^R	This study
SR7xKAL	*pJBxKAL	P_{Xyl} $kivD_{Ll}$, $adhA_{Ll}$; Tet^R	This study
SR7xKAP	*pJBxKAP	P_{Xyl} $kivD_{Ll}$, $adhA_{Ec}$; Tet^R	This study
SR7xGFP	*pJBxGFP	P_{Xyl} sfGFP; Tet^R	This study
SR7hGFP	*pJBhGFP	$P_{Hyper\text{-}spank}$ sfGFP; Tet^R	This study
*pRBBm34 derivative (Biedendieck et al., 2007)			

For solventogenesis strain engineering, *kiv*D_{Ll} sourced from *Lactococcus lactis* and *adh*6_{Sc} from *Saccharomyces cerevisiae* were placed downstream of xylose-inducible promoter P_{Xyl} on pRBBm34. Vector construction began by PCR amplifying *kiv*D_{Ll} from pCOLA KivD, Fjoh_2967 using primers KivD_F and KivD_R. PCR products and the pRBBm34 plasmid were digested with *SpeI* and *SphI* prior to ligation to create the pJBxK plasmid. *Adh*6_{Sc} from *S. cerevisiae* was PCR amplified from pACYC (car,sfp; adh6) with the same ribosome binding site as was used for *kiv*D_{Ll} using primers RBS-ADH6_F and ADH6_R. *Adh*6_{Sc} was added between the *BamHI* and *SphI* restriction sites in P_{Xyl} $KivD_{Ll}$ to create pJBxKA6. *Yqh*D_{Ec} from *E. coli* was PCR amplified from *E. coli* MG1655 genomic DNA with the same ribosome binding site as was used for *kiv*D_{Ll} using primers RBS-YqhD_F and YqhD_R. *Yqh*D_{Ec} was added between the *BamHI* and *SphI* restriction sites in P_{Xyl} $KivD_{Ll}$ to create pJBxKY (Figure 2B). All constructs were verified by DNA sequencing.

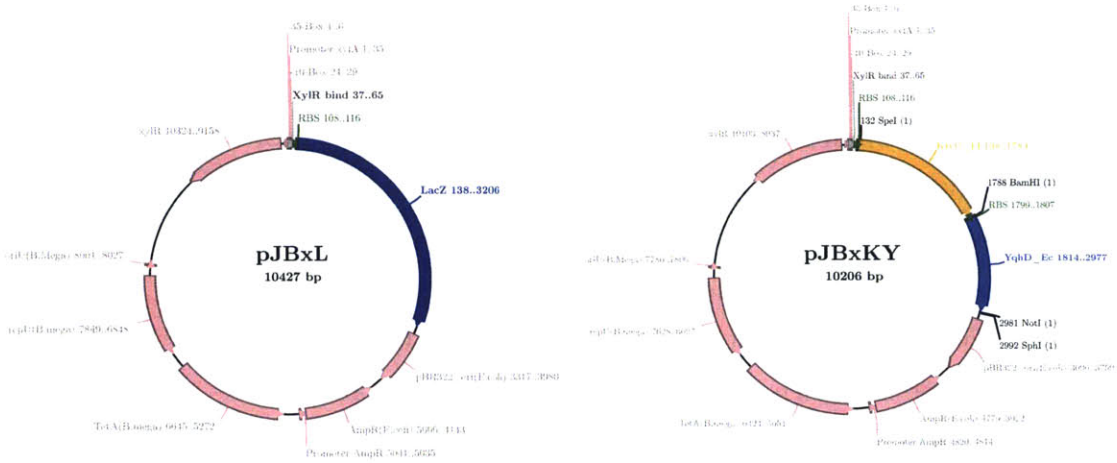


Figure 2. Vector maps for **A)** pJBxL (left) and **B)** pJBxKY (right) using scaffold pRBBm34

Transformation methods

Initial attempts to genetically transform strain SR7 with shuttle vector pRBBm34 (Addgene) used an established *Bacillus* electroporation protocol (Zhang *et al.* (2011; Analytical Biochemistry). Modifications to the method included the addition of cell wall weakeners (3.9% glycine, 80 mM DL-threonine) one hour prior to electroporation, and testing a wide range of plasmid concentrations (10-200 ng/µl) and cell densities (OD_{600} 0.6-1.2). Conjugation-based transformation was attempted with SR7 using mating strain *E. coli* S-17 and plasmid pJR1 (provided courtesy of the Meinhardt Lab, University of Muenster, Germany; Table 1B), following the protocol of Richhardt *et al.* (2010). To optimize the protocol, a range of donor to recipient strain volumes were tested (i.e. 10:1 to 1:1000) after reaching protocol-prescribed OD_{600} values. Post-transformation counter-selection included pasteurization and the *sacB* suicide system. The final transformation method attempted was protoplast fusion based on von Tersch and Robbins (1990) and Biedendieck *et al.* (2011) using shuttle vector pRBBm34. The cell wall removal step was optimized to increase viable protoplasts by modifying lysozyme concentrations and transformed protoplast incubation times. Counter-selection occurred by plating protoplasts on a soft agar overlay above LB agar containing 5 µg/ml tetracycline.

Plasmid maintenance

Several assays were utilized to verify exogenous plasmid stability in SR7 during growth under 1 atm CO_2. To assay for maintenance of pRBBM34 in SR7 under 1 atm CO_2, singleton incubations of SR7 empty vector control strain (SR7x), which constitutively expresses tetracycline resistance, were inoculated at a concentration of 10^5 spores/ml and passaged three times in LB for 24 hours with and without supplementation of 0.5 μg/ml tetracycline. After each passage, cultures were plated on LB agar with or without 0.5 μg/ml tetracycline to determine if cultures grown without antibiotics maintained the transformed vector over multiple growth cycles in the absence of a selective pressure. SR7 wild-type and SR7x strains were also assayed to determine minimum required tetracycline concentration to select for transformed strains containing the vector. Cultures inoculated with 10^5 spores/ml were incubated in LB amended with a range of tetracycline concentrations under both aerobic (Tet 0.05-10.0 μg/ml) and 1 atm CO_2 conditions (Tet 0.1-10.0 μg/ml) and scored for growth by OD_{600} relative to cultures that were not amended with Tet.

Heterologous single gene expression under 1 atm CO_2 and scCO_2

SR7 strains SR7xL and SR7hL (bearing genetic constructs pJBxL and pJBhL, respectively) and empty vector control strains SR7x and SR7h were assayed for protein expression in SR7 under 1 atm CO_2 and scCO_2 conditions. 1 atm CO_2 cultures grown overnight were diluted in fresh media to OD_{600} 0.01, cultured for 2 hours, then amended with 0.4% D-xylose (P_{Xyl}; SR7x, SR7xL) or 5 mM ITPG ($P_{Hyper\text{-}spank}$; SR7h, SR7hL) to induce expression. After 24 hours, 1 ml of culture volume was spun down for 5 min x 21,000g and the remaining pellet was stored at -20°C until analysis.

Supercritical CO_2 cultures were loaded with $3x10^5$ spores/ml (prepared as described in Thesis Chapter 3) of strain SR7xL. A subset of reactors was amended with 0.5% xylose inducer. Reactor cultures were incubated for 21 days then depressurized and prepared for fluorescence microscopy as previously described. 2 ml of culture volume was spun down for 5 min x 21,000g, after which the supernatant and pellet were separately stored at -20°C until analysis.

Supernatant was prepared for GC-MS analysis by methods described below and for HPLC analysis by previously described methods (Thesis Chapter 3).

Pellets from 1 atm CO_2 and $scCO_2$ cultures were lysed by addition of 100 µl of Bacterial Protein Extraction Reagent (B-PER; Thermo Scientific) and vortexed for 30 minutes. After lysed pellets were centrifuged for 20 min x 18,500g at 4°C, supernatants were prepared for total protein analysis using the Pierce™ BCA Protein Assay Kit (Thermo Scientific) according to manufacturer's instructions. Colorimetric signatures proportional to total sample protein were measured by OD_{562}, including for cell-free B-PER negative controls. Total protein standard curves were generated using 0.05-1.0 mg/ml of bovine serum albumin (BSA) according to the same protocol. Samples and B-PER negative control were prepared for LacZ activity assays by adding 70 µl of lysed culture supernatant to 730 µl of assay buffer (0.1 M sodium phosphate buffer, 10 mM KCl, 1 mM MgSO4) and 200 µl of β-galactosidase substrate (4 mg/ml o-nitrophenyl-β-D-galactoside, ONPG). LacZ activity was quantified as the rate of OD_{420} absorbance per minute, as the product of ONPG cleavage by B-galactosidase absorbs at 420 nm. Absorbance rate was normalized by total protein per culture using BSA standard curves. Protein-normalized rates were converted to specific activity using the assay extinction coefficient and volume to generate units of µmol min^{-1} mg^{-1}.

Heterologous biofuel production under 1 atm CO_2

Vegetative cultures of SR7x and SR7xKA6 were prepared by growing 10^5 spores/ml of each strain in CO_2-degassed LB $tet_{0.5}$ for overnight growth. Stationary phase cultures were then diluted in 10 ml of fresh LB + $tet_{0.5}$ to OD_{600} 0.01. After 2 hours, passaged cultures were amended with 5 mM α-KIV substrate and 0.4% D-xylose to induce gene expression. Passaged cultures were grown for 24 hours post-induction, with sub-sampling at 4 and 24 hours by aseptic needle extraction. After 1 ml samples were centrifuged for 5 minutes x 21,000g, 500 µl supernatant was pipetted into separate tubes with 500 µl of ethyl acetate solvent (≥99.9% pure GC-grade, Sigma Aldrich) and vortexed for 5 minutes. The ethyl acetate fraction was pipetted into analysis vials (Agilent) and loaded on the Agilent Technologies 7890B GC system (using Agilent J&W VF-

WAXms GC Column) and 5977A MSD for gas chromatography-mass spectrometry (GC-MS) analysis using MassHunter Qualitative Analysis (Agilent) software to measure compound concentrations. Peaks in the resulting total ion current (TIC) chromatogram were input into the NIST MS Search 2.2 database for compound prediction. Prior to running incubated samples, standard curves were generated using a range of concentrations (0.2-5.0 g/l) of expected products (isobutyraldehyde, isobutanol, isopentanol, phenethyl alcohol, acetate) using flame ionization detector (FID) spectra. Integrated total ion current (TIC) chromatogram peaks for differentially produced compounds were measured and converted to g/l concentrations according to standard curve conversion factors.

Alcohol dehydrogenase screening

To assay for differential alcohol dehydrogenase activity under aerobic and 1 atm CO_2 conditions, pRBBm34 vectors were constructed using previously described methods with xylose-inducible promoter P_{Xyl} upstream of *kiv*D_{Ll} and one of five alcohol dehydrogenase variants: *adh*6 / pJBxA6 (*S. cerevisiae*), *adh*A_{Bm} / pJBxKAB (*B. megaterium* SR7), *adh*A_{LL} / pJBxKAL (*L. lactis*), *adh*P_{Ec} / pJBxKAP (*E. coli*), and *yqh*D / pJBxKY (*E. coli*) (constructs and strains summarized in Tables 1A-B).

For aerobic screens, freezer stocks of each strain were streaked onto LB agar plates supplemented with tetracycline (5 μg/ml) and grown at 37°C overnight. For each alcohol dehydrogenase, three colonies were added separately to 5 ml of LB with tetracycline and grown at 37°C overnight. Each sample was sub-cultured to an OD_{600} of 0.05 in 3 ml of LB with tetracycline in a 50 ml screw-capped glass tube. Cultures were grown at 37°C and 250 RPM until an OD_{600} of ~1.0 was reached, at which point 5mM α-ketoisovalerate (α-KIV) precursor was added and protein production induced by supplementing with 0.5% D-xylose. Cultures were grown at 37°C and 250 RPM. Time points were taken at 4 hours, 24 hours and 48 hours by removing 1 ml of culture volume. Samples were centrifuged at 20,000xg for 5 minutes and the supernatant removed. Alcohols were extracted from the supernatant using a 1:1 ratio of supernatant to ethyl acetate and vortexed at maximum speed for 5 minutes. The ethyl acetate fraction was recovered by centrifugation at 20,000xg for 5 minutes and removal of the upper solvent fraction. Sample analysis by GC-FID and concentrations of

produced alcohols by standard curve calculations used previously described methods.

1 atm CO_2 cultures inoculated with 10^5 spores/ml of each strain were grown overnight and passaged by syringe needle into fresh CO_2-degassed LB + $tet_{0.5}$. Two hours after passaging, cultures were amended with 5 mM α-KIV substrate and 0.5% D-xylose for gene induction. Sampling of strain variant cultures took place at 24 and 48 hours by syringe needle. Samples were then prepared for GC-MS analysis and post-run data processing as previously described.

Assay for quantification of isobutanol production under $scCO_2$

A headspace extraction setup was constructed to collect the $scCO_2$ phase with dissolved species from high-pressure growth reactors. 316 stainless steel lines and valves connected to the reactor pressurization manifold enabled direct depressurization of the $scCO_2$ headspace into ethyl acetate solvent for subsequent GC-MS analysis. To generate standard curves for isobutanol $scCO_2$ phase extraction, duplicate 10 ml incubation reactors were filled with cell-free LB medium amended with 5% isobutanol, 0.5% isobutanol and unamended LB. Reactors were pressurized with CO_2 to 100 atm, heated to 37°C and shaken for 3 days to allow equilibration. Individual reactors were then slowly depressurized into 10 ml of chilled ethyl acetate at a rate of ~1 atm/min. This process was repeated a second time with several modifications, including submerging reactors in a heat bath at 55°C, increased depressurization rates (1.5-2 atm/min) and extraction into larger solvent volume (20 ml). The extraction setup is shown in Figure 3. A standard curve generated by GC-MS analysis of the initial extraction run enabled conversion of FID isobutanol peak areas to mg/l concentrations.

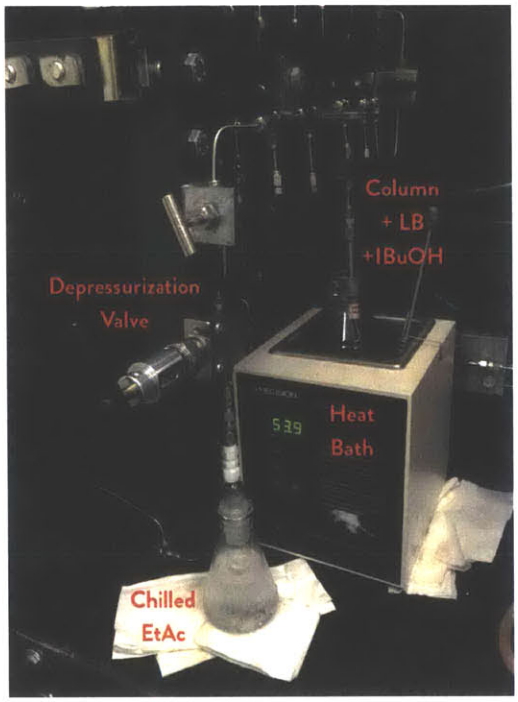


Figure 3. $ScCO_2$ phase extraction setup used for harvesting biofuel from $scCO_2$ SR7xKY cultures and abiotic isobutanol suspensions.

Heterologous biofuel production under $scCO_2$

To determine whether SR7xKY produces isobutanol during growth under $scCO_2$ headspace, high-pressure reactors were loaded with $3x10^5$ spores/ml of SR7xKY, control strain SR7xL, or cell-free media controls. A subset of reactors was amended with 0.5% xylose to induce gene expression. Reactors were pressurized to 100 atm CO_2, heated to 37°C and incubated under $scCO_2$ for 21-22 days while shaking at 250 rpm. Reactors with identical inocula/media conditions (i.e. strain ± xylose) or cell-free controls were simultaneously depressurized into chilled ethyl acetate at a rate of ~1 atm/min. Between each round of unloading, manifold lines and valves were flushed with ethyl acetate. Samples were prepared for GC-MS analysis as previously described. Quadruplicate technical replicates were run for all $scCO_2$ bulk phase-collected samples. Culture supernatant glucose concentrations were measured using the YSI 2900 with YSI 2814 glucose starter kit after spinning down 1 ml culture volume for 5 minutes x 21,000g. Depressurized cultures were prepared for epifluorescence microscopy by methods described in Thesis Chapter 3. Cultures were considered to have grown when demonstrating at least 10-fold increase in cell counts relative to loaded spore

concentrations. The limit of detection was considered to be one half of a cell per 15 grids, which corresponds to $1.15x10^3$ cells/ml.

4.3 RESULTS

4.3.1 Development of a genetic system for *B. megaterium* SR7

Three methods for transforming strain *B. megaterium* SR7 were tested: electroporation, conjugation and protoplast fusion. Genetic transformation of SR7 by electroporation using plasmid pRBBm34 was unsuccessful despite multiple attempts to modify protocol parameters based on published studies in other *B. megaterium* strains (Moro *et al.*, 1995). Transformation by conjugation using the *E. coli* S-17 mating strain and vector pJR1 (Richhardt *et al.*, 2010; Table 1B) gave mixed results with conferral of chloramphenicol resistance up to 10 μg/ml and positive PCR amplification of the plasmid-specific marker (*sac*B) in the resistant strains confirming transformation, albeit at a low frequency (i.e. 1 transformant per 10^7 SR7 cells). However, subsequent attempts to transform SR7 by the described conjugation protocol were not successful and thus prevented its further use in this study.

Protoplast fusion transformation, previously demonstrated in several *B. megaterium* strains (Bunk *et al.*, 2010) proved successful at introducing all constructs used in this study via shuttle vector pRBBm34. Despite protocol modifications that increased viable protoplasts by fifty-fold, transformation efficiencies remained low (~1 transformed cell/10^7 viable protoplasts), frequently generating 1-3 successfully transformed colonies per ug of plasmid DNA. Protoplast transformation first enabled conferral of constitutive tetracycline resistance (10 μg/ml aerobic; 1.0 μg/ml under 1 atm CO_2). Maintenance of tetracycline resistance under 1 atm CO_2 was verified by nearly identical growth of cultures passaged three times in either LB or tetracycline-amended LB on unamended and tetracycline-amended plates. All subsequent genetic manipulation of strain SR7 was conducted using the modified protocol for protoplast fusion transformation.

4.3.2 Inducible heterologous enzyme production under 1 atm CO_2 and $scCO_2$

After demonstrating constitutive antibiotic expression, two promoters (P_{Xyl} and $P_{Hyper\text{-}spank}$) were investigated for inducible protein expression in SR7. The D-xylose-inducible P_{Xyl} promoter (Figure 4) is endogenous to all sequenced *B. megaterium* strains, including SR7:

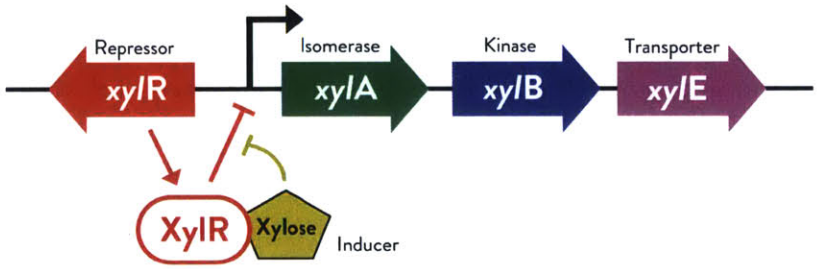


Figure 4. Xylose promoter system endogenous to *B. megaterium*

Rather than using the xylose promoter native to SR7, a previously optimized P_{Xyl} system (Biedendieck *et al.*, 2007) was used to avoid uncharacterized endogenous promoter regulation specific to SR7. The IPTG-inducible hyper-spank promoter ($P_{Hyper\text{-}spank}$) had previously been transformed into and expressed in *B. subtilis* (van Ooij and Losick, 2003), but had never been utilized in *B. megaterium.*

Plasmids pJBxGFP and pJBhGFP, where reporter superfolder GFP was placed under the control of the P_{Xyl} and $P_{Hyper\text{-}spank}$ promoters, respectively, demonstrated induced expression in SR7 at nearly equal strengths under aerobic conditions assayed by GFP fluorescence intensity (data not shown). Low-level fluorescence in uninduced cultures appeared to show minor leakiness by P_{Xyl}. However, since fluorescent protein reporters including GFP are typically active only under aerobic conditions, an anaerobically functional reporter was necessary to verify promoter activity under 1 atm CO_2. Therefore, both promoters were placed upstream of the *lacZ* reporter, which is O_2-indepenent. Cultures of transformed strains SR7xL and SR7hL passaged under 1 atm CO_2 and induced with D-xylose and IPTG, respectively, were analyzed for LacZ production 24 hours after induction. Total protein-normalized LacZ specific activities (i.e. activity of enzyme per mg total protein; μmol min^{-1} mg^{-1}; U/min) for duplicate

lacZ strain cultures, empty vector controls, and a LacZ assay reagent (B-PER) control are displayed in Figure 5.

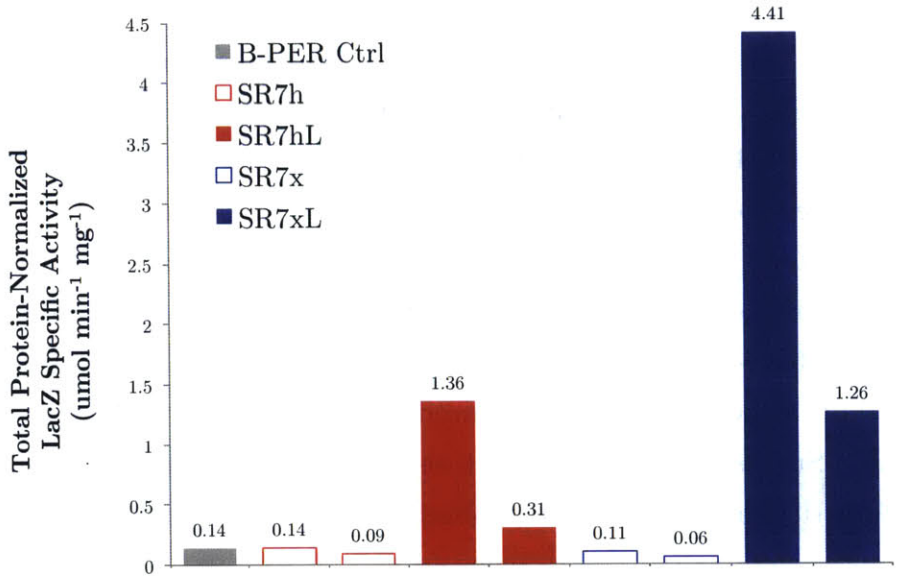


Figure 5. LacZ specific activity of duplicate cultures using IPTG and xylose-inducible promoters under 1 atm CO_2, with B-PER negative control.

LacZ specific activity values from duplicate incubations of xylose-amended cultures of SR7xL (1.26-4.41 U/min) and SR7x (0.06-0.11 U/min) demonstrate that LacZ activity is increased by induction relative to empty vector controls. Relative to empty vector samples and the B-PER assay control (≤0.14 U/min), LacZ activity increased 9-32 fold. Differential expression of LacZ by IPTG induction of $P_{Hyper\text{-}spank}$ was also observed, but at lower total protein-normalized specific activity levels than by xylose-induction (0.31-1.36 U/min). LacZ production under aerobic and anaerobic 1 atm CO_2 conditions represents the first successful use of IPTG-inducible $P_{Hyper\text{-}spank}$ in *B. megaterium.* This development, and verification of strong P_{Xyl} activity under 1 atm CO_2 expands the list of genetic tools available for SR7 engineering, as well as alternative *B. megaterium* strains used for biotechnological applications.

After exhibiting superior total protein-normalized LacZ activity under 1 atm CO_2, the SR7xL strain was investigated for expression under $scCO_2$. Duplicate cultures with and without 0.4% xylose inducer demonstrated robust germination and growth after 21 days under $scCO_2$ conditions, with appearance of vegetative cell morphologies and at least 15-fold increase in cell counts relative

to starting cultures. Duplicate cultures of induced and uninduced reactors showing vegetative cell morphologies, but not robust outgrowth (<10-fold increase in cell counts) were utilized for comparison of LacZ activity in germinated/low-level growth cultures. Both cultures that grew under $scCO_2$ in the presence of xylose showed elevated total protein-normalized LacZ specific activity (0.66-0.90 U/min) relative to uninduced cultures that grew (0.06-0.23 U/min) (Figure 6). Uninduced cultures may display low-level LacZ activity due to minor leakiness of the xylose promoter, as also demonstrated under aerobic conditions (personal communication, Jason Boock). Duplicate induced cultures that did not grow but appeared to have germinated (by microscopy) displayed activity values (0.06-0.17 U/min) on par with the negative control (0.14 U/min), indicating that active growth is required for heterologous enzyme expression under $scCO_2$. Successful LacZ production by SR7 under $scCO_2$ was the first demonstration that exogenous gene expression in a $scCO_2$ headspace bioreactor is possible.

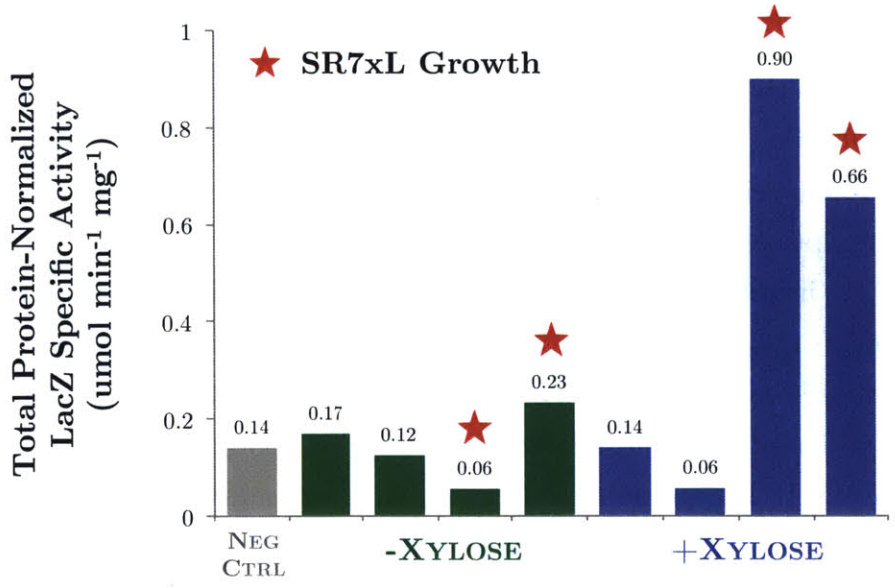


Figure 6. LacZ-specific activity in $scCO_2$ cultures in the absence (green) and presence (blue) of 0.5% xylose, with B-PER negative control. SR7xL cultures demonstrating robust growth are highlighted with red stars. Cultures without stars germinated but did not demonstrate robust outgrowth.

4.3.3 Engineering and expression of a heterologous pathway for biofuel synthesis in $scCO_2$-tolerant strain SR7 under aerobic, 1 atm CO_2 and $scCO_2$ conditions

The final two steps in the production of isobutanol using the valine biosynthesis pathway requires catalytic conversion of α-KIV to isobutyraldehyde by α-ketoisovalerate decarboxylase (KivD), followed by conversion to isobutanol by alcohol dehydrogenase (Figure 1). Since the *kiv*D gene is not present in the SR7 genome (Thesis Chapter 3), it required exogenous introduction in order to be to be expressed. The well-described *Lactococcus lactis* version of keto-isovalerate decarboxylase commonly used for isobutyraldehyde production (de la Plaza *et al.*, 2004) was utilized in this study. Though the SR7 genome indicates that the gene then required for conversion of isobutyraldehyde to isobutanol, alcohol dehydrogenase, is present in the cell, its production is likely lower than if transformed on a plasmid with a strong promoter. As a result, the *E. coli* version, $adh6_{Ec}$, was initially used in SR7 due to laboratory availability.

While upstream optimization may enable efficient conversion of glucose to α-KIV, initial pathway engineering relied on an exogenous supply of α-KIV to constrain heterologous enzyme activity assays to the final two steps of the pathway. Because the isobutanol pathway genes should be functional under both anaerobic and aerobic conditions, induction of the final two steps was first characterized under aerobic and 1 atm CO_2 conditions to validate expression under both conditions. After demonstrating initial activity, subsequent screening for highly active alcohol dehydrogenase enzymes in SR7 under aerobic and 1 atm CO_2 ultimately enabled the use of an optimized construct under $scCO_2$. 1 atm CO_2 passaged cultures of strains SR7xKA6 and SR7x (empty vector control) in LB media grew similarly well 24 hours after gene expression was induced. Based on averaged OD_{600} values, heterologous pathway expression appeared to impose a metabolic burden that results in a 24% decrease in biomass yield relative to the empty vector control (Figure 7). GC-MS analysis verified production of several biofuel products in the 1 atm CO_2 cultures grown in LB after 4 and 24 hours, including expected compounds isobutyraldehyde and isobutanol (Table 2). In a somewhat surprising result, isopentanol and phenethyl alcohol were also produced, indicating that pJBxKA6 genes $kivD_{Ll}$ and $adh6_{Sc}$ appear to redirect

flux of alternative amino acid biosynthesis pathways, including those of leucine (to isopentanol) and phenylalanine (to phenethyl alcohol; Figure 1). No biofuel peaks were detected in either of the P_{Xyl} Empty replicate cultures.

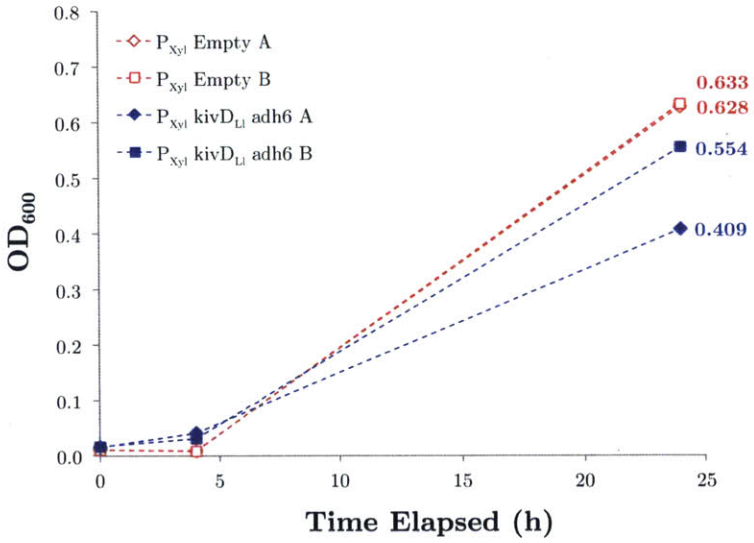


Figure 7. Growth of biofuel and empty vector control strains under 1 atm CO_2

The biofuel strain replicates (A & B) showed accumulation of the intermediate product isobutyraldehyde at the 4-hour time point (A: 1.34, B: 1.66 mM) with trace level accumulation of isobutanol (A & B: 0.01 mM) and isopentanol (A & B: 0.01 mM) and no detectable phenethyl alcohol (Table 2).

Table 2. Summary of bioproducts (mM, mg/l) generated by SR7xKA6 under 1 atm CO_2

[Biofuel] (mM)						
Time	Replicate	Isobutyraldehyde	Isobutanol	Isopentanol	Phenethyl Alcohol	Sum
4	A	1.66	0.01	0.01	0	1.69
4	B	1.34	0.01	0.01	0	1.35
24	A	1.7	4.00	1.95	0.22	7.87
24	B	1.85	4.08	1.98	0.26	8.17
[Biofuel] (mg/L)						
Time	Replicate	Isobutyraldehyde	Isobutanol	Isopentanol	Phenethyl Alcohol	Sum
4	A	23.0	0.1	0.1	0.0	23.3
4	B	18.6	0.1	0.1	0.0	18.8
24	A	23.6	54.0	22.1	1.8	101.5
24	B	25.7	55.0	22.5	2.1	105.3

Culture Conditions
Supplemented: α-IKV: 5 mM
Strain: SR7 P_{Xyl} $kivD_{Li}$ $Adh6_{Sc}$
Media: LB + tet 0.1 ug/mL
Induced: 0.5% xylose

Compounds	Std Curve Conversion y = counts ; x = mM
Isobutyraldehyde	y = 27623x; R = 0.81
Isobutanol	y = 169971x; R = 1.00
Isopentanol	y = 184505x; R = 0.97
Phenethyl Alcohol	y = 434871x; R = 1.00

A marked shift in production was observed in duplicate cultures at the 24-hour time point, with significant accumulation of isobutanol (A: 4.00, B: 4.08 mM), isopentanol (A: 1.95, B: 1.98 mM) and small amounts of phenethyl alcohol (A: 0.22, B: 0.26 mM), while maintaining comparable aldehyde accumulation (A: 1.70, B: 1.85 mM). It therefore appears that while a certain concentration of aldehyde will build up due to the limits of $Adh6_{Sc}$ activity in SR7, by 24 hours the majority of α-KIV substrate has been converted to final biofuel products isobutanol and isopentanol.

1 atm CO_2 cultures of SR7xKA6 generated bioproducts at a slower rate than under aerobic conditions, but final 24-hour titers were similar under both conditions (Figure 8), indicating reduced catalytic efficiency but comparable total substrate conversion. At 4 hours, the sum of 1 atm CO_2 bioproduct concentrations was 28% of the summed concentrations under aerobic conditions, increasing to 88% of the aerobic sum by 24 hours. Specifically with regard to isobutanol, by 24 hours the 1 atm CO_2 incubations generated 93% of the concentration detected in aerobic cultures. Overall, these results suggest that reduced production rates under 1 atm CO_2 relative to aerobic conditions may be due to slower microbial growth/metabolism or diminished enzyme activity.

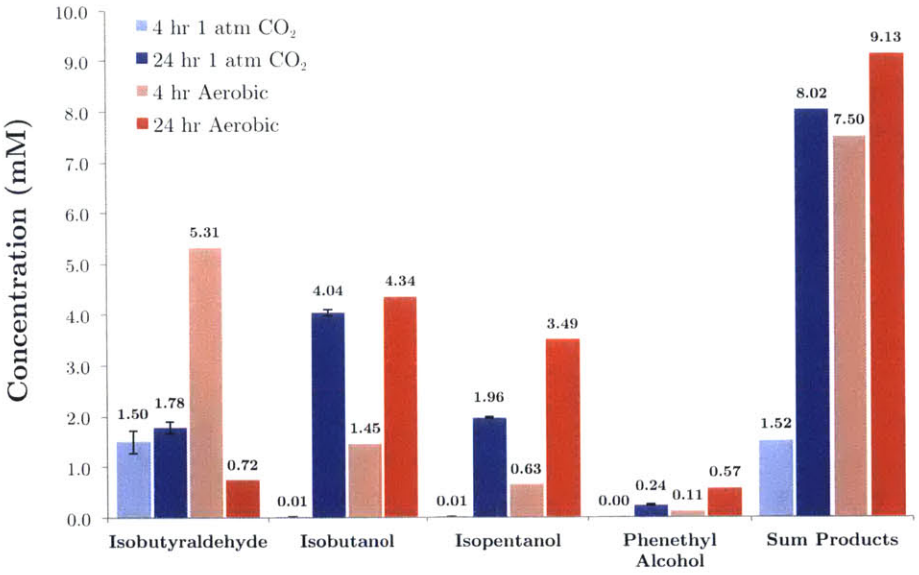


Figure 8. SR7xKA6 bioproduct concentrations under 1 atm CO_2 (light/dark blue) and aerobic (pink/red) conditions 4 and 24 hours after induction. Measurements of biofuel production from aerobically-incubated cultures may be underestimated due to the removal of caps during subsampling of aerobic cultures, which may have resulted in some volatile product losses.

While nearly identical amounts of isobutanol were produced under both aerobic and 1 atm CO_2 conditions, 1 atm CO_2 titers of isopentanol and phenethyl alcohol were about half as concentrated as under aerobic conditions. Therefore, it appears that alternative amino acid pathway enzymes (Figure 1) may be operating at reduced efficiency in siphoning off α-KIV substrate, possibly due to dependence on O_2-dependent co-factor NAD(P)H.

Alcohol dehydrogenase screening

The accumulation of isobutyraldehyde in initial SR7xKA6 cultures under 1 atm CO_2 (Table 2, Figure 8) prompted additional screening of alcohol dehydrogenase gene variants in order to improve the rate and completeness of isobutyraldehyde conversion to isobutanol. This enzymatic reaction is of particular importance in the proposed biofuel production system because isobutyraldehyde is soluble in $scCO_2$ (personal communication, Prof. Mike Timko) and thus premature partitioning of accumulated isobutyraldehyde into the $scCO_2$ headspace would reduce overall yields, titers and purity of the desired isobutanol end product. Vectors constructed with P_{Xyl} *kiv*D_{Ll} alone and with one of five alcohol dehydrogenase variants (*adh*6_{Sc}, *adh*A_{Bm}, *adh*A_{LL}, *adh*P_{Ec}, and *yqh*D_{Ec}) were thus assayed for biofuel production under aerobic and 1 atm CO_2 conditions. Aerobic results for GC-MS detected compounds after 4 and 24 hours are presented in Figure 9.

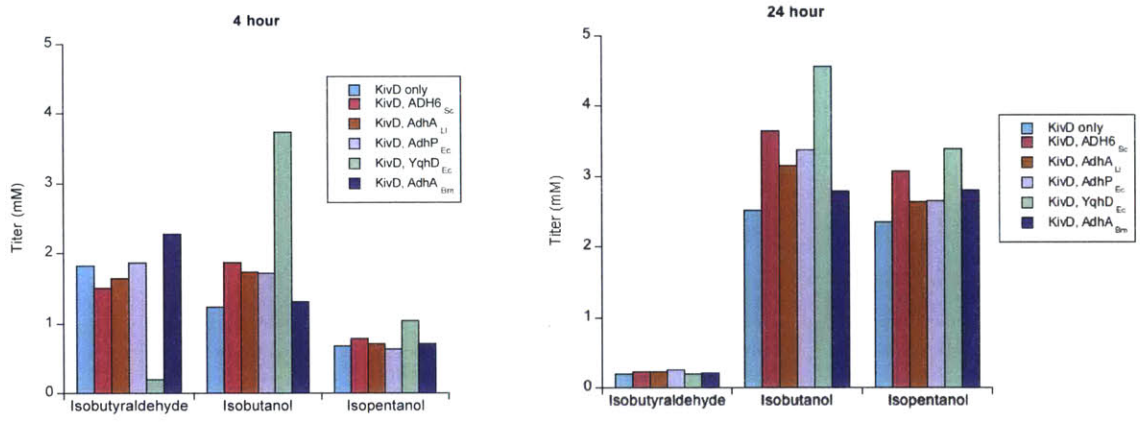


Figure 9. Bioproduct concentrations generated under aerobic conditions by SR7 P_{Xyl} *kiv*D_{Ll} and in tandem with one of five alcohol dehydrogenase variants after **A)** 4 hours (left) and **B)** 24 hours (right)

Results from subsequent alcohol dehydrogenase variant screens (including raw and OD-normalized values) under 1 atm CO_2 are presented in Figure 10.

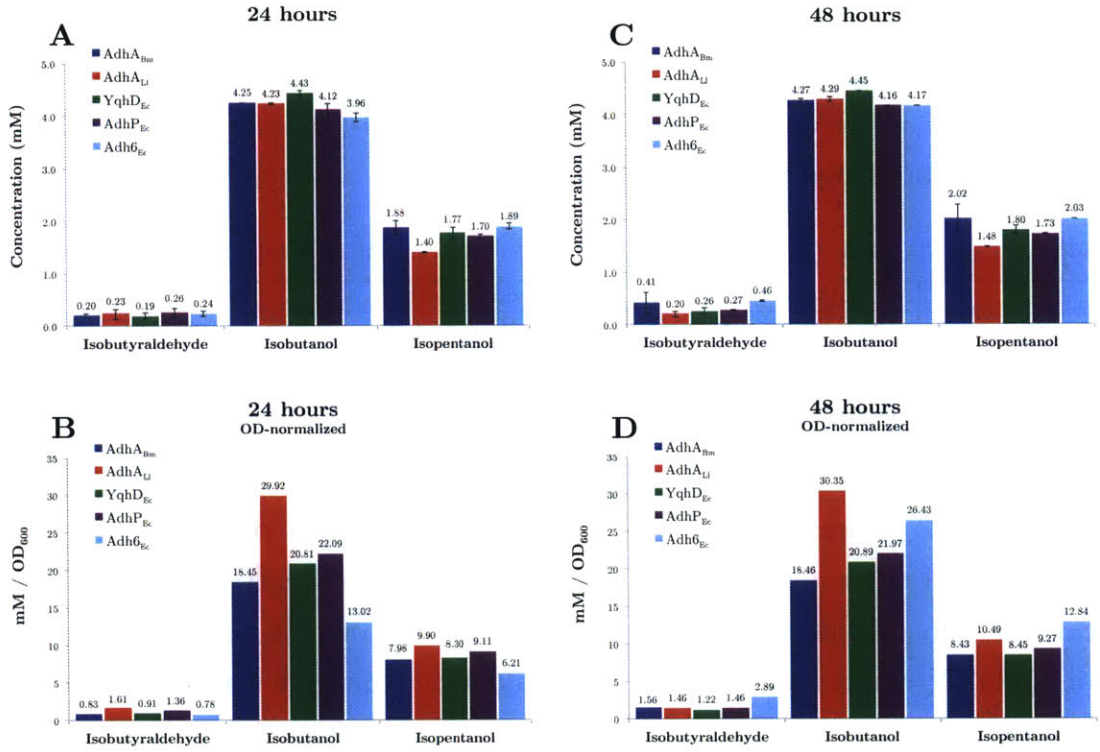


Figure 10. Bioproduct concentrations generated under 1 atm CO_2 by SR7 P_{Xyl} *kiv*D_{Ll} in tandem with five alcohol dehydrogenase variants at 24 hours presented as **A)** raw and **B)** OD-normalized values, and 48 hours as **C)** raw and **D)** OD-normalized values.

Under aerobic conditions concentrations of aldehyde and alcohol products demonstrate that *yqh*D_{Ec} variant cultures (SR7xKY) outperformed all other alcohol dehydrogenases according to several metrics. By 4 hours, while all other variants generated isobutyraldehyde above 1.5 mM, SR7xKY cultures prevented intermediate accumulation by converting nearly all α-KIV substrate to isobutanol and isopentanol (Figure 9A). By 24 hours, while all alcohol dehydrogenase variants had converted isobutyraldehyde to alcohol products, SR7xKY cultures generated the highest titers for both isobutanol (4.6 mM) and isopentanol (3.4 mM). Overall, YqhD$_{Ec}$ results in >90% conversion of α-KIV substrate to biofuel products (Figure 9B).

Under 1 atm CO_2 isobutyraldehyde, isobutanol, and isopentanol concentrations were nearly identical for all strains at both 24 and 48 hours based on raw values (Figure 10A,C), suggesting that effectively all α-KIV substrate had

been converted by 24 hours. The fact that low levels of isobutyraldehyde persist at both 24 and 48 hours also suggests that alcohol dehydrogenase activity may become limited once the aldehyde concentration drops below a threshold level, as all aldehyde concentrations from both time points fell within a narrow range, (0.193-0.457 mM; 0.014-0.033 g/l). The best performing enzyme variants after 24 and 48 hours as determined by maximum alcohol and minimum aldehyde concentrations are listed in Table 3.

Table 3. 1 atm CO_2 alcohol dehydrogenase variant performance summary based on raw concentrations

Time (h)	Enzyme	$[Isobutanol]_{Max}$		Enzyme	$[Isopentanol]_{Max}$		Enzyme	$[Isobutyraldehyde]_{Min}$	
		mM	g/L		mM	g/L		mM	g/L
24	YqhD	4.43	0.342	$Adh6_{Sc}$	1.893	0.167	YqhD	0.193	0.014
48	YqhD	4.448	0.343	$Adh6_{Sc}$	2.025	0.178	$AhdA_{Ll}$	0.203	0.015

OD-normalized product concentrations (Figure 10B,D) suggest that $AdhA_{Ll}$ may be especially efficient at product generation on a per-cell basis, which in addition to displaying the lowest aldehyde concentration at 48 hours indicates it may be one of the better performing variants. In addition to *yqh*D_{Ec} strain SR7xKY demonstrating the fastest aldehyde conversion rates and highest final titers under aerobic conditions (Figure 9), results from 1 atm CO_2 cultures also displayed the highest final titers (Figure 10C), although performance differences under 1 atm CO_2 were marginal relative to aerobic results. With available data especially encouraging for variant $YqhD_{Ec}$, subsequent incubation experiments under $scCO_2$ proceeded with the pJBxKY construct-bearing strain SR7xKY.

4.3.4 Bench scale abiotic isobutanol $scCO_2$ and aqueous phase extractions

In situ extraction via $scCO_2$ relies on the partitioning of a compound from the aqueous phase to the $scCO_2$ phase followed by product recovery. After manifold modifications, batch reactors used for culturing under $scCO_2$ in this study were utilized for bench scale *in situ* extraction, as described in the methods. Standard curves generated for partitioning of isobutanol from aqueous media into the $scCO_2$ phase demonstrated that isobutanol at concentrations from

0.5-5.0 mM (37-371 mg/l) could be quantitatively stripped from the media phase and recovered in the $scCO_2$ phase (Figure 11).

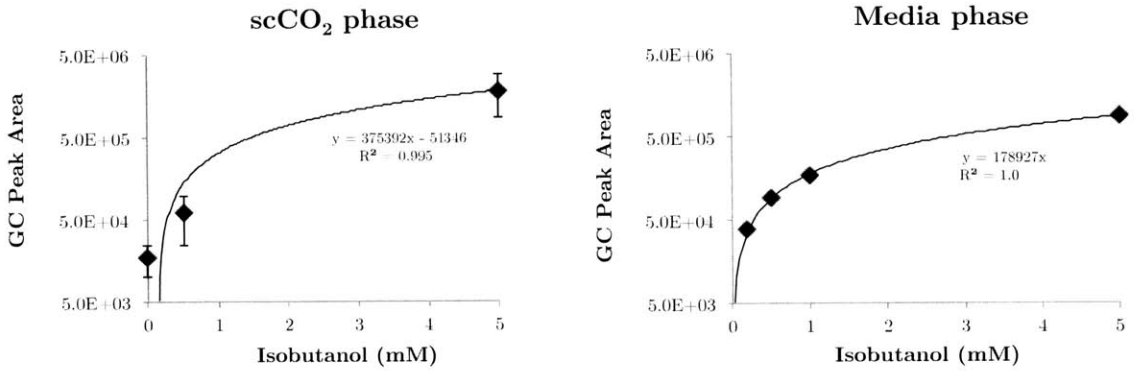


Figure 11. Standard curves based on abiotic **A)** $scCO_2$ phase (left) and **B)** aqueous phase extraction of isobutanol (right)

Process modifications including continuous reactor heating, increased depressurization rates and increased ethyl acetate solvent volume appeared to significantly improve supercritical CO_2 phase recovery efficiencies during a second round of abiotic isobutanol extractions, increasing the percent of total isobutanol recovered from the $scCO_2$ phase by an order of magnitude from 2% to 20% between the initial and second runs. Overall mass balance calculations of the second run demonstrated that between 75-90% of loaded isobutanol concentration was recovered by the sum of aqueous and $scCO_2$ phase products after three-day $scCO_2$ incubations. We note that since the batch bioreactor set up used in this work is not optimized for solvent stripping using $scCO_2$, 2-20% product recovery in $scCO_2$ is satisfactory in the context of this study. However, we expect that further work to optimize the reactor and stripping configuration will enable more efficient *in situ* extraction.

4.3.5 Biosynthesis and *in situ* extraction of natural products and biofuels under $scCO_2$

Having established alcohol dehydrogenase variant $YqhD_{Ec}$ as the best performing enzyme for isobutanol production, cultures loaded with spores of SR7xKY in the presence of xylose inducer were anticipated to generate biofuel products. Conversely, metabolically inactive cultures and LacZ-generating SR7xL

control cultures were not expected to show signatures of alcohol production. Uninduced biofuel strain cultures showing growth were anticipated to generate low-level biofuel concentrations due to the mildly leaky nature of P_{Xyl}. A summary of growth outcomes from $scCO_2$-incubated cultures of genetically modified strains (This Chapter) and wild-type SR7 (Thesis Chapter 3) is presented in Table 4:

Table 4. Summary of growth outcomes for cultures of SR7 wild-type and modified strains in M9A+ media incubated under $scCO_2$.

Inocula	± Xylose	Starting spores/mL	M9A+ Growth Frequency	Max Biomass (cells/mL)
SR7xKY	+	$3x10^5$	33% (5/15)	$5.96x10^7$
	-	$1x10^5$	17% (1/6)	$2.88x10^7$
SR7xL	+	$5x10^5$	13% (2/15)	$1.34x10^7$
	-	$3x10^5$	33% (2/6)	$9.69x10^6$
Wild-type SR7	-	$3x10^4$	64% (16/25)	$1.63x10^7$
Media Control	+	b.d.	b.d.	b.d.

Decreased growth frequencies observed in transformed strains relative to wild-type SR7 may indicate a metabolic burden associated with carrying and expressing the pRBBm34 vector, as observed under 1 atm CO_2 (Figure 7), that reduces germination frequency and/or vegetative outgrowth, though these hypotheses will require additional investigation.

Natural fermentation products were detected by HPLC in the media phase of all reactors demonstrating growth (>10-fold increase in cell counts) over 21-22 day $scCO_2$ incubations, including induced and uninduced cultures of both SR7xL and SR7xKY. Detected compounds were consistent with those generated by wild-type SR7 $scCO_2$ cultures (Thesis Chapter 3), including succinate (up to 73.2 mg/l), lactate (up to 2.8 g/l), and acetate (up to 1.3 g/l) (Figure 12), which reinforces the suggestion of growth via the TCA Cycle and mixed acid fermentation. Total cell-normalized metabolite concentrations demonstrate maximum per cell productivities of 7.6 $x10^{-9}$, 5.3 $x10^{-8}$, and 2.5 $x10^{-8}$ mg product $cell^{-1}$ for succinate, lactate and acetate, respectively, which are also similar to maximum per cell productivities of the wild-type strain under 1 atm CO_2 and $scCO_2$ (Thesis Chapter 3). The quantification of natural metabolites thus has potential for use as an indicator of active growth under $scCO_2$.

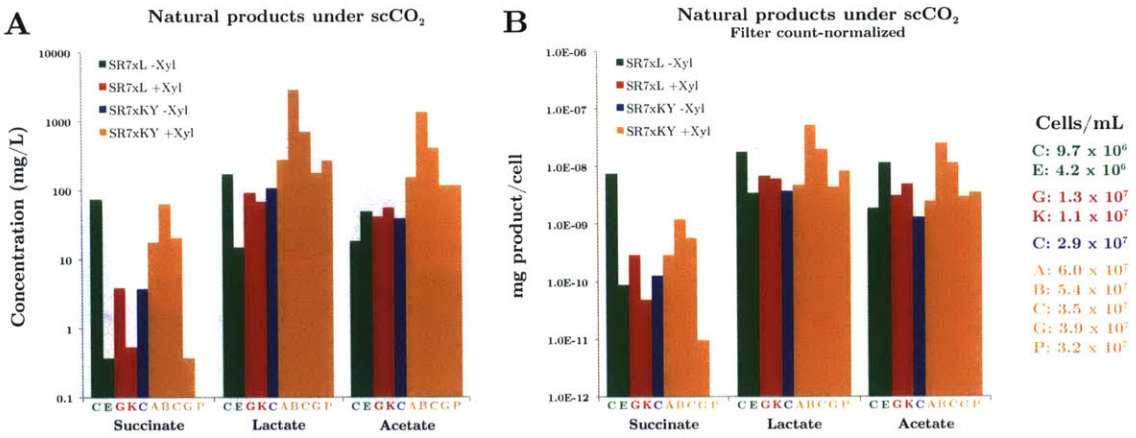


Figure 12. Natural fermentative products generated by SR7xL and SR7xKY cultures under $scCO_2$ showing growth, as detected by HPLC. **A)** total titers and **B)** filter count-normalized per cell metabolite productivity. Final cell concentrations for each sample listed at right. No metabolites were detected in media-only reactors and reactors without cell growth (data not shown).

Biofuels were detected by GC-MS in the media phase of all five reactors loaded with SR7xKY that showed growth and were induced with xylose (Figures 13 and 14; Table 5). Of the two SR7xKY cultures showing growth in the absence of xylose, one showed low level biofuel production (0.3 mg/l isobutanol, 0.1 mg/l isopentanol), putatively as the result of the mildly leaky xylose promoter. No biofuel was generated in any of the reactors that did not show vegetative growth, verifying that metabolic activity (i.e. growth) under $scCO_2$ is required for heterologous compound production (e.g. Figure 6). No biofuels were detected in SR7xL cultures or media only negative controls.

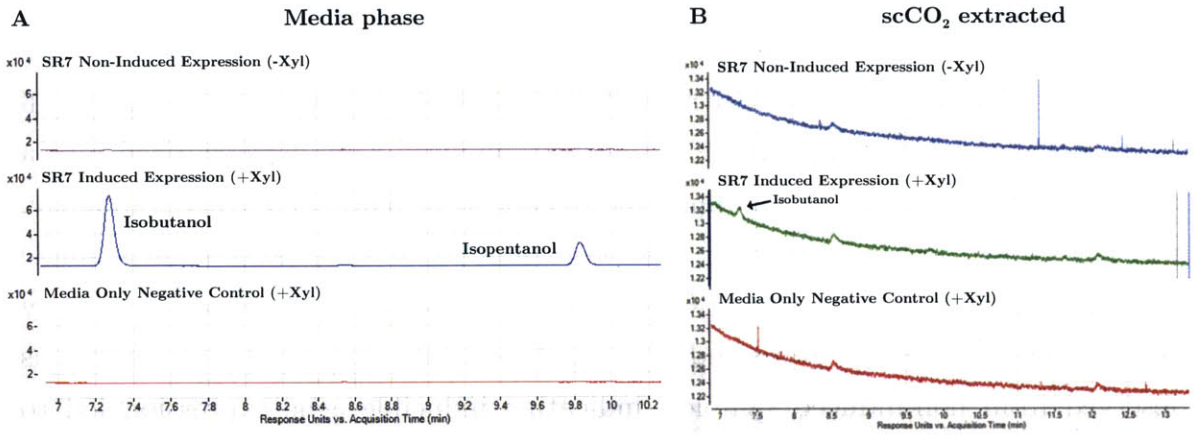


Figure 13. Example GC-FID traces detecting biofuel products isobutanol and isopentanol in the **A)** media phase and **B)** $scCO_2$-extracted phase of $scCO_2$-incubated SR7xKY cultures

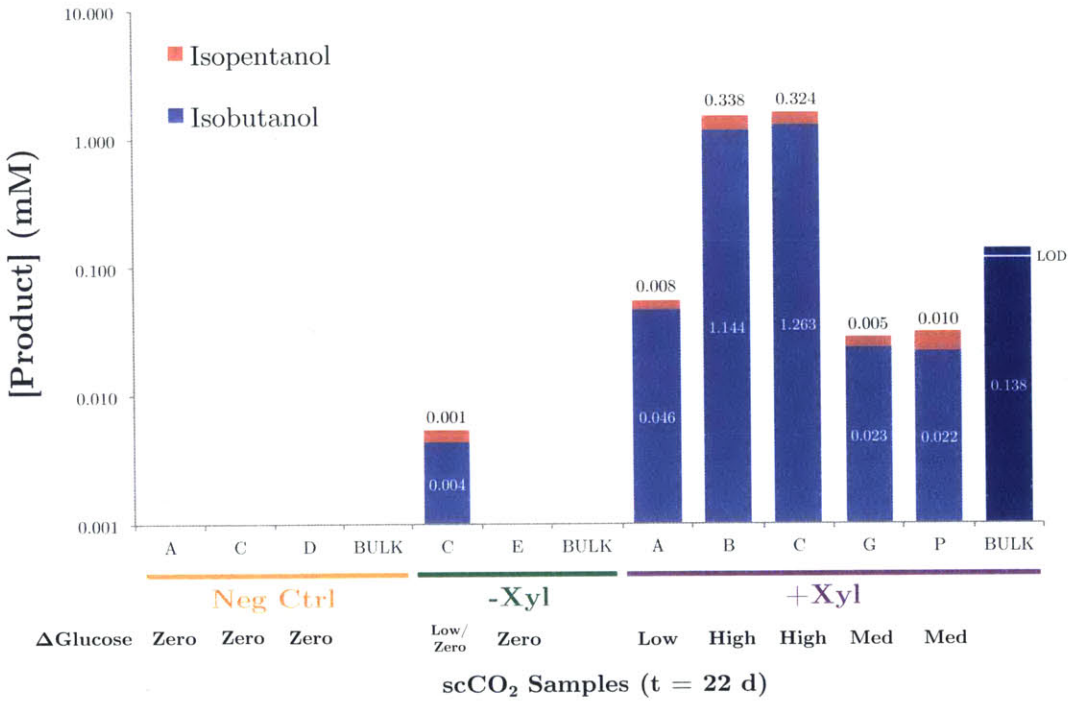


Figure 14. Recovered bioproduct concentrations from cultures showing growth in xylose-induced and uninduced cultures, and media only negative controls. "Bulk" refers to $scCO_2$ phase-extracted headspace volumes from all induced, uninduced, or negative control reactors. Glucose consumption designated as zero, low/zero (0-5% consumed), low (5-20%), medium (20-40%), and high (>40%). "LOD" = limit of detection for $scCO_2$-extracted isobutanol (0.13 mM).

Measured isobutanol concentrations in the aqueous phase of induced cultures ranged from 1.6 to 93.5 mg/l, while isopentanol concentrations varied from 0.5 to 29.7 mg/l. Observed maximum titers of 0.094 g/l isobutanol and 0.030 g/l isopentanol in $scCO_2$ incubations are approximately two orders of magnitude lower than under 1 atm CO_2, possibly due to reduced growth rates and biomass accumulation, or potential redirection of carbon flux under $scCO_2$ conditions.

In order to maximize direct $scCO_2$ phase biofuel compound recovery, all reactors loaded with strain SR7xKY that were induced were depressurized into a single collection solvent to maximize the likelihood of biofuel recovery. Similarly, all reactors loaded with SR7xKY that were uninduced were pooled via a single collection solvent. A pronounced GC-MS peak for isobutanol was observed only for the pooled samples from induced reactors while no isobutanol peak was observed from non-induced samples, indicating inducible gene expression led to biofuel generation under $scCO_2$ (Figures 13B and 14). Based on the abiotic $scCO_2$-extracted isobutanol standard curve (Figure 11), the total $scCO_2$ phase isobutanol concentration was 10.2 mg per liter of media (0.138 mM; Table 5),

which represents 5.2% of the total recovered isobutanol from both the media and $scCO_2$ phases. The detection of biogenic isobutanol in the $scCO_2$ phase represents the first biofuel production under $scCO_2$ conditions, as well as the first harvesting of biofuels from microbial cultures incubated under $scCO_2$.

Table 5. Summary of $scCO_2$-incubation outcomes for SR7xKY-loaded columns that showed increased biomass relative to starting concentrations

Culture Sample	Induced (+Xyl)	Replicate	Filter Count (cells/mL)	Filter Std Dev	Fold Growth [t22]/[t0]	[Glucose] (g/L)	Isobutanol (IBuOH)			Isopentanol (IPnOH)			Sum [Biofuel] mg/L
							GC Area	mM	mg/L	GC Area	mM	mg/L	
Cell-Free M9A+ Control	Yes	A	0.00E+00	0.00E+00	0	3.21	0	0	0	0	0	0	0
		C	0.00E+00	0.00E+00	0	3.14	0	0	0	0	0	0	0
		D	0.00E+00	0.00E+00	0	3.10	0	0	0	0	0	0	0
		Bulk ($scCO_2$)					0	0	0	0	0	0	0
SR7xKY	No	C	2.88E+07	8.13E+06	195.3	3.05	765	0.004	0.316	239	0.001	0.094	0.411
		E	4.36E+05	3.68E+05	3.0	3.14	0	0.000	0.000	0	0.000	0.000	0.000
		Bulk ($scCO_2$)					0	0	0	0	0	0	0
	Yes	A	5.96E+07	1.66E+07	190.0	2.45	8229	0.046	3.403	1698	0.008	0.669	4.072
		B	5.37E+07	1.01E+07	171.3	0.01	204775	1.144	84.690	75485	0.338	29.745	114.435
		C	3.48E+07	6.73E+06	111.0	1.91	225969	1.263	93.455	72353	0.324	28.511	121.966
		G	3.88E+07	9.54E+06	123.5	2.70	4146	0.023	1.715	1181	0.005	0.465	2.180
		P	3.18E+07	1.91E+05	101.5	2.60	3922	0.022	1.622	2126	0.010	0.838	2.460
		Bulk ($scCO_2$)					563	0.138	10.212	0	0	0	10.212

In addition to endogenous metabolites and heterologous biofuels, differentially extracted compounds present in the bulk $scCO_2$ phase of grown cultures that are absent in the aqueous phase may hold promise as SR7 natural products able to be extracted by the $scCO_2$ phase. If these products can be identified, it is possible that optimization of product-generating pathways may enable future production and extraction of these unknown compounds. Peaks differentially present in the $scCO_2$ bulk phase include compounds with retention times of 19.75, 25.34, 25.4, and 26.97 minutes. TIC chromatogram peaks were too small to be reliably compared to the NIST chemical database and thus could not be identified. Repeated experiments and further inspection of GC chromatograms and MS spectral data may aid in elucidating the nature of differentially generated compounds under $scCO_2$.

4.4 Discussion

After demonstrating requisite growth under proposed bioproduction conditions (Thesis Chapter 3), initial attempts were made to confront the challenge of genetically modifying strain SR7 with the goal of generating biofuels

under $scCO_2$ conditions able to extracted by the $scCO_2$ solvent phase. *B. megaterium* has generally been considered a genetically tractable species using a range of transformation approaches (Biedendieck *et al.*, 2011; von Tersch and Robbins, 1990; Biedendieck *et al.*, 2011; Richhardt *et al.*, 2010; Moro *et al.*, 1995). Previous genetic engineering in the species has demonstrated the ability for recombinant overexpression of genes and entire operons through identification and use of strong promoters (Bunk *et al.*, 2010). Furthermore, directed enzyme engineering has enhanced protein stability, and an mRNA inactivation strategy has been developed using antisense RNA to redirect carbon flux from competing pathways (Biedendieck *et al.*, 2010).

Development of a viable protoplast fusion protocol enabled the introduction and maintenance of exogenous plasmid DNA bearing non-native enzymes. Activity assays using two promoters, xylose-inducible P_{Xyl} (Figure 4) and IPTG-inducible $P_{Hyper\text{-}spank}$, established P_{Xyl} as the superior genetic element, generating LacZ specific activity values between 1.26 and 4.41 (U/min) relative to $P_{Hyper\text{-}spank}$ specific activity values of (0.31-1.36 U/min) (Figure 5). While not utilized for additional development in this work, the $P_{Hyper\text{-}spank}$ promoter may be useful in the future as a component of increasingly complex pathways, including those requiring feedback architecture with multiple inducible promoters. After demonstrating the capacity to induce expression of a single enzyme (LacZ) under $scCO_2$ (Figure 6), the logical next step in bioproduction development was expression of a heterologous multi-enzyme pathway generating a compound that can be localized outside the cell for $scCO_2$ extraction. Follow-up efforts thus targeted the production of C4 biofuel isobutanol due to the limited number of required genetic modifications, its ability to partition into $scCO_2$ solvent, and significant global market that is expected to reach $1.18 billion by 2022 (GVR, 2015).

Results demonstrating biofuel production under 1 atm CO_2 (Table 2; Figure 8) were useful as an intermediate proof of concept not only in having generated a non-native product under anaerobic conditions, but also in enabling a higher-throughput mode of experimentation where $scCO_2$ culturing conditions are limiting. These results also established proof of concept that detectable extracellular biofuel compounds may be generated by SR7 through heterologous pathway expression, lending confidence to the hypothesis that production under

$scCO_2$ is an achievable goal. In addition, 1 atm CO_2 titers resembled aerobic concentrations after 24 hours (Figure 8), an encouraging result suggesting that O_2-dependent enzymatic and co-factor function may be overcome through moderately extended incubation. Isobutanol and isopentanol titers in SR7 under aerobic and 1 atm CO_2 conditions are competitive with previous studies attempting to similarly engineer amino acid pathways for biofuel production in *E. coli*. Both aerobic and anaerobic SR7 isobutanol productivities outcompete *E. coli* (Bastian *et al.*, 2011) in one instance, while demonstrating about one-third the productivity of an alternative *E. coli* strain that had been genetically optimized to reduce competing byproduct yields through gene deletions (i.e. *adh*E, *ldh*A, *frd*AB, *fnr*, *pta*; Atsumi *et al.*, 2008) that are currently unavailable in SR7 due to limited genetic tractability. Isopentanol titers were also comparable with a previous amino acid pathway-utilizing study (Connor *et al.*, 2010), which demonstrated nearly identical productivities to those observed in SR7 under aerobic conditions (though 1 atm CO_2 productivities were approximately half the rate of aerobic). Competitive biofuel production and $scCO_2$ exposure resistance therefore uniquely positions SR7 as holding the potential to exploit $scCO_2$ for a novel bioproduction and harvesting platform system if the strain is able to demonstrate heterologous biofuel production under $scCO_2$.

The alcohol dehydrogenase gene initially tested in SR7 ($adh6_{Ec}$) appeared to present a pathway bottleneck as the intermediate aldehyde accumulated after 4 hours in both the aerobic and 1 atm CO_2 incubations (Table 2; Figure 8). As a result, alcohol dehydrogenase variant screens provided a means for boosting pathway kinetics. The results from aerobic testing showed dramatic improvement in substrate conversion rates and final titers for one variant in particular, $YqhD_{Ec}$ (Figure 9). Although screens under 1 atm CO_2 confirmed high final titers by $YqhD_{Ec}$, follow-up experiments with earlier time points and in M9A+ media will provide additional clarity regarding differential alcohol dehydrogenase activity. Previous studies have shown $YqhD_{Ec}$ is a NADP+-dependent dehydrogenase with a preference for alcohols longer than C3 (Sulzenbacher *et al.*, 2004). Because glycolysis exclusively generates NADH under anaerobic conditions, further enzyme co-factor improvement, i.e. conversion from NADP+ to NAD+-dependency, may improve catalytic performance in light of NADH-dependent $AdhA_{Ll}$ (from *Lactococcus lactis*) demonstrating the highest activity

in *E. coli* (Atsumi *et al.*, 2010). As demonstrated by Bastian *et al.* (2011), this could entail overexpression of pyridine nucleotide transhydrogenase (PntAB), which is responsible for transferring hydride from NADH to NADP. Another approach would be to modify enzymes to be NADH-dependent rather than NADPH-dependent. Bastian *et al.* (2010) utilized targeted mutagenesis and recombination to generate two NADH-dependent variants that in conjunction with $AdhA_{Ec}$ produced isobutanol at 100% of the theoretical maximum yield, confirming the utility of co-factor modifications. Utilizing specific methods and targeted mutagenesis developed for *E. coli* it may be possible to modify alcohol dehydrogenase genes for SR7 to further improve kinetics and overall performance.

In support of our focus on genetically engineering wild-type environmental isolate *B. megaterium* SR7 for biofuels production under $scCO_2$, we converted our bench scale bioreactor to an *in situ* stripping device and demonstrated that through very minimal system/process optimization, 5.5% of isobutanol in the system could be recovered from the $scCO_2$ phase. We expect that further engineering of the stripping method with specialized reactors that include mixing by impeller and continuous $scCO_2$ sparging will allow enhanced rates of mass transfer. Demonstration of direct isobutanol extraction by $scCO_2$ is representative of a breakthrough aspect of this work: beyond conferring self-sterilizing capacity (Spilimbergo and Bertucco, 2003; Peet *et al.*, 2015) and reducing end product toxicity effects, $scCO_2$ solvent extraction generates a "dry" product upon $scCO_2$ depressurization that requires limited additional dehydration processing.

Unlike most current bioreactor harvesting technologies (Oudshoorn, 2012) because water shows very limited solubility in $scCO_2$ (~1 mol%; Sabirzyanov *et al.*, 2002), compounds extracted in the $scCO_2$ phase are inherently dry. Water and isobutanol form an azeotrope, which is a mixture of two liquids that cannot be separated by conventional distillation because boiling generates a vapor phase with the same proportions of the liquid mixture constituents, i.e. isobutanol has a boiling point of 108°C, water of 100°C, and the mixture of 90°C. As a result, $scCO_2$ extraction of (iso)butanol from an aqueous phase provides the opportunity to "break" the azeotrope in a single step rather than as a multi-step process where the isobutanol-water azeotrope is first removed from the growth medium before subsequent steps to purify and dehydrate the isobutanol phase (Vane, 2008). The infrastructure and energy costs of the current fermented isobutanol harvesting

methods have prevented this approach from being brought to scale in any industrial setup to date. A novel process of isobutanol production-to-harvest could thus occur in our proposed reactor system while utilizing an inexpensive, low-energy, environmentally benign solvent.

Current bioreactor schemes are plagued by contamination events and end product accumulation toxicity that require the cessation of product generation and subsequent re-sterilization of the entire system. Bioreactor incubations that include a $scCO_2$ phase not only enable unique extractive chemistries, but also protects against colonization by an outside organism. Due to contamination concerns, bioreactors are almost always set up as batch or fed-batch reactors (Cinar *et al.*, 2003). However, inherent to batch schemes is that as metabolic byproducts build up in the culturing medium, microbes are unable to withstand the toxic effects of these end products, causing them to die. Therefore, continuous removal of metabolites that will preferentially partition into the $scCO_2$ phase will relieve end-product accumulation, extending the temporal capacity to continue product generation beyond conventional limits. Furthermore, as the consistency of *in vitro* biocatalysis may be limited by variable stability of enzymes under $scCO_2$, the use of a microorganism as host provides enzymes additional layers of protection from direct $scCO_2$ exposure that *in vitro* incubations do not allow. This protective feature is especially effective within Gram-positive bacteria (e.g. *Bacillus spp.*) whose thick peptidoglycan cell wall is thought to curtail the ability of $scCO_2$ to penetrate the cellular membrane (Mitchell *et al.*, 2008). As a result, the capacity to culture *B. megaterium* SR7 under $scCO_2$ appears to solve several challenges associated with $scCO_2$ biocatalysis.

Since a vast range of recombinant proteins have been produced in *B. megaterium* (Bunk *et al.*, 2010, Martens *et al.*, 2002) for pharmaceutical and industrial applications, including via CO_2-consuming carboxylation reactions, compounds with vastly different chemistries and uses may be considered for future bioproduction (Matsuda *et al.*, 2001, 2008). As a result, in order to narrow the focus of target compounds, several aspects of prospective products must be taken into account for potential biphasic media-$scCO_2$ bioreactor production using strain SR7, including 1) does the compound have a favorable (i.e. non-polar, hydrophobic) partition coefficient into the $scCO_2$ phase from water-based

media, 2) does the product require a prohibitively high number of genetic elements, and 3) has the pathway been demonstrated in closely-related species? An additional consideration is whether a feasible economic landscape exists for the bioproduction and purified extraction of specific compounds using the SR7 system, including from a range of alternative feedstocks. Considering that Bacilli-generated enzymes comprise around 60% of the ~$2 billion global market for homologous and heterologous enzyme production (Ferrari and Miller, 1999), it is likely that entry to a commercialized market by proper chemical vetting would enable industrial development of the complete production and harvesting system using SR7.

4.5 Conclusions

By integrating process and genetic engineering developments established under 1 atm CO_2 with proven culturing methods under $scCO_2$ (Thesis Chapter 3), heterologous *lacZ* expression and biofuel production by actively growing cultures under $scCO_2$ conditions establishes an exciting new field of biotechnological development with significant implications for the future of sustainable biofuel and biochemical production. While low-frequency growth (Peet *et al.*, 2015) and enzymatic catalysis (Matsuda *et al.*, 2001) have previously been shown under $scCO_2$, this work represents the first combined demonstration of heterologous enzyme production and robust vegetative growth under $scCO_2$ conditions. Growth compatibility and protein production in SR7 now provides a conduit to exploiting the $scCO_2$ solvent phase, which was previously thought inaccessible for microbial-mediated product generation and solvent extraction (Knutson *et al.*, 1999). Now that this specialized growth system has proven feasible, attention turns to further engineered pathway complexity, reactor and extraction design, and potential economically viable applications of this novel technology.

4.6 Acknowledgements

I thank Dr. Jason Boock for his significant contributions to the work presented in this chapter. Specifically, Dr. Boock optimized the protoplast fusion protocol, constructed vectors, transformed modified strains (except SR7JR1), and conducted aerobic screens of alcohol dehydrogenase gene variants.

5

Conclusions and Future Work

The reason that I wrote this thesis - and did this research - was because the world is currently on an irrevocable path to global climate change by continued reliance on fossil fuel-based energy sources. In particular, the emission of carbon dioxide throughout the course of industrialized human history is the single largest contributing driver of anthropogenic-induced climate change (IPCC, 2007). Therefore, our society must consider how we are going to *use* carbon dioxide for productive purposes or at the very least find methods for storing CO_2 in a safe and effective manner. This thesis thus aimed to survey different methods by which microbial life may play a role in a sustainable CO_2 future. Specifically, I explored the role of the deep biosphere in a GCS analog system and how living microbes may harness the solvent power of $scCO_2$ as a component of energy efficient microbial compound biosynthesis and extraction.

Microbes hold significant potential for helping to mitigate the migration and leakage of injected $scCO_2$ for long-term GCS. The taxonomic identification and laboratory isolation of bacterial species associated with deep subsurface fluids from McElmo Dome may inform new research into the development of $scCO_2$-tolerant strains with deep subsurface utility. Biofilm barriers have previously been developed to reduce aquifer permeability by supplying nutrients to native or introduced microbial populations in order to enable robust growth and EPS formation (Komlos et al., 2004; Seo et al., 2009; Cunningham et al., 2009; Mitchell et al., 2008; Mitchell et al., 2009). Furthermore, *B. megaterium* cells in particular have been shown to induce the carbonate mineralization of CO_2 (Lee et al., 2014), which if applied *in situ* in GCS formations could increase the efficiency of subsurface trapping. As a result it is possible that upon co-injection with $scCO_2$, certain microbial strains including SR7, other spore-forming Bacilli, or species native to natural $scCO_2$ formations may enhance the $scCO_2$ trapping capacity of deep subsurface formations.

Additional research and development of microbial bioproduction in $scCO_2$-exposed bioreactors holds significant potential for increasingly sustainable compound generation and purified extraction. Although the work in this thesis focused on batch reactor production, continuous flow reactors are generally preferable due to the capacity for higher volume production, increased potential for process automation, reduced stress on instrumentation due to lack of frequent sterilization, and a more highly tunable system in which minor process modifications may be continuously evaluated to enable product titer improvements (Mascia et al., 2013; Poechlauer et al., 2012, Roberge et al., 2005). Therefore, the work done in this thesis was done with the expectation that upon proof of concept, the production system would be ported into a highly optimized continuous flow reactor. Unlike most continuous flow reactors that are plagued by major contamination events, the presence of $scCO_2$ will severely limit the potential that any non-target organisms will be able to colonize a scaled up bioreactor.

A $scCO_2$-stripping chemostat that we specifically designed for use on this project is being constructed by the Timko Lab at Worcester Polytechnic Institute (Figure 1). After having shown the ability to produce isobutanol and isopentanol under $scCO_2$ in batch reactors, SR7 may now be utilized for continuous flow production and compound harvesting. After growth occurs, feed dilution rates will be set to balance kinetics of batch growth and the relationship between

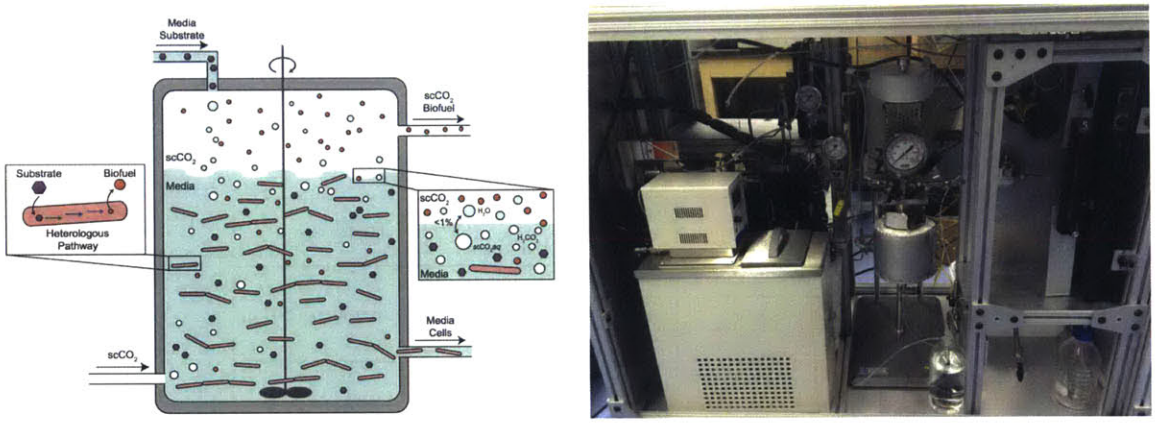


Figure 1. A) Overall heterologous bioproduction and $scCO_2$ extraction scheme (left) **B)** Customized continuous flow reactor (right) built at WPI (Timko Group) for use with $scCO_2$-resistant environmental isolate strain *B. megaterium* SR7.

dilution rate and steady-state cell density will be determined. The aqueous phase flow rate will also be adjusted in order to maximize the ability to sweep away secondary products that might adversely affect cell growth or target compound production.

Additional development of SR7 for use in the optimized reactor involves engineering increasingly complex and varied pathways for higher value products, including 1-butanol and 4-methyl-pentanol. The possibility of incorporating the bicarbonate-fixing phosphoenolpyruvate carboxylase enzyme endogenous to SR7 for $scCO_2$-consuming reactions holds exciting potential for $scCO_2$ to serve a dual purpose as both solvent and dissolved substrate. There is also the potential to test enzymes recovered as DNA sequences from the McElmo Dome metagenome (or other high pCO_2 systems) that may be especially well adapted to function under high pCO_2 conditions. Developing protocols for genomic integration of heterologous pathway enzymes may confer additional stability in the context of scaled up industrial production. Furthermore, transcriptomic analysis of SR7 growth under a variety of headspace and media conditions would provide crucial insight into competing pathways that may be knocked out/down in order to improve product yields and titers. Transcriptomics may also enable a genome-wide promoter map with broad pathway engineering applications.

The genetic and process development of SR7 within a highly-customized continuous flow biphasic aqueous-$scCO_2$ chemostat is also exciting due to its potential energy savings relative to conventional extraction methods, as demonstrated by preliminary analysis by the Timko Group at WPI and Oudshoorn (2012) with respect to butanol (Table 1).

Table 1. Energy balance for butanol extraction methods

Extraction Method	**Energy Requirement (MJ/kg butanol)**	**Reference**
Gas Stripping	22	Oudshoorn (2012)
Liquid-Liquid Extraction	9	
Distillation	24	
Supercritical CO_2 Extraction (1 bar, 25 °C)	3.9	Timko Group

Even when fully depressurized to 1 bar, $scCO_2$ extraction still represents a 57% energy savings over the next most efficient method (liquid-liquid extraction). SR7

bioproduction in an optimized reactor context may also provide new conduits for the production and harvesting of broad classes of compounds that are currently reliant on environmentally harmful organic solvents for extraction. Based on published partition coefficients and current market values (Table 2; personal communication, Mike Timko), short-to-medium chain aldehydes are attractive as end products for use in the $scCO_2$ bioproduction system, specifically isobutanal (isobutyraldehyde), n-pentanal, n-butanal, and n-propanal. Encouragingly, as SR7 has already demonstrated the capacity for isobutyraldehyde production by introduction of the *kiv*D_{LI} gene, the current state of strain development is already

Table 2 Economics of fermentative products, including water/$scCO_2$ ($K_{w/c}$) coefficients indicating modeled and experimental partitioning ratios from water into the $scCO_2$ phase (Timko *et al.*, 2004).

Chemical	Price US$/kg	Ref.	Global Capacity (tons/year)	Ref.	$K_{w/c}$ calculated, (expt)	Uses
n-butanol	1.42-1.54	ICIS 2006	3,924,200	SRI data 2011	5 (2.2)	Fuel, extraction of oils, antibiotics
n-butanal	1.19-1.28	ICIS 2006	1,100,000	Delft report	61	Intermediates
			>500,000 (US)	EPA 2002		
isobutanol	1.32-1.42	ICIS 2006	552,400	Grand view research 2015	5 (2)	Fuel, Chemical feedstock, solvent,
isobutanal	1.60-4.80	Alibaba.com	223,500 (US)	SRI 1991	93 (est)	intermediates
n-pentanol	3.00-8.00	Alibaba.com	25,000 (US)	US EPA 2002	7	Solvent, biofuel, lubricants
n-pentanal	1.00-2.00	Lookchem.com	25,000-50,000 (US)	US EPA 2002	90	flavorings, resin chemistry, and rubber accelerators
n-propanol	1.39-1.48	ICIS 2006	110,000	Kirk-Othmer Encyclopedia of Chemical Technology, 1993	4	Solvent, resins
n-propanal	2.00-4.00	Alibaba.com	25,000-50,000 (US)	US EPA 2002	44	Chemical precursor in resins, organic synthesis
4-methyl-2-pentanol	2.20-2.30	Alibaba.com			9	Solvent, brake fluid, plastercizers
isopropanol	1.56-2.16	ICIS 2006	>500,000 (US)	US EPA 2002	3	Solvent, chemical intermediate, rubbing alcohol,
n-hexanol	1.94	ICIS 2006	5,000-25,000	US EPA 2002	10	Perfumes, adhesives, lubricants, solvents, fuels
vanillin	17.86-18.74	ICIS 2006	13,200-17,600	ICIS 2006	-1.5	Flavoring for food and beverage, pharmaceuticals

equipped to generate an additional compound beyond isobutanol and isopentanol that is able to take advantage of $scCO_2$ solvent properties.

Overall, this thesis has shown that the microbial biosphere may access supercritical CO_2 in new and unforeseen ways, opening the door to additional research and development of geoengineering and bioengineering applications. I believe strongly that innovative approaches must be taken in order to find novel uses for the greenhouse gas carbon dioxide in order to diminish the effects of global climate change. This thesis was my attempt to contribute to that effort.

When applying to doctoral programs, I wrote in my personal statement that my proposed graduate research aim was "the engineering of methods that isolate and convert byproduct wastes, utilizing their vast potential as productive, valuable substrates. Combining these methods with renewable energy resources represents my vision of the next generation of energy production, water purification, and clean industrial processes." I believe that my research has delivered on my initial goals, and will continue to provide new avenues by which our society's future may come to treat "waste" as a outdated concept that fails to acknowledge the inherent utility of all compounds as substrates, especially that of carbon dioxide.

References

Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH. "Genome Sequences of Rare, Uncultured Bacteria Obtained by Differential Coverage Binning of Multiple Metagenomes." *Nature Biotechnology* 31 (2013): 533-38.

Allis R, Chidsey T, Gwynn W, Morgan C, White S, Adams M, Moore J. "Natural CO2 Reservoirs on the Colorado Plateau and Southern Rocky Mountains: Candidates for CO2 Sequestration." Paper presented at the Proceedings of First National Conference on Carbon Sequestration, Washington, D.C., 2001.

Aloisi G, Gloter A, Kruger M, Wallmann K, Guyot F, Zuddas P. "Nucleation of Calcium Carbonate on Bacterial Nanoglobules." *Geology* 34 (2006): 1017-20.

Anbu P, Kang CH, Shin YJ, So JS. "Formations of Calcium Carbonate Minerals by Bacteria and Its Multiple Applications." *SpringerPlus* 5 (2016): 1-26.

Arioli S, Monnet C, Guglielmetti S, Mora D "Carbamoylphosphate Synthetase Activity Is Essential for the Optimal Growth of Streptococcus Thermophilus in Milk." *Journal of Applied Microbiology* 107 (2009): 348-54.

Atsumi S, Cann AF, Connor MR, Shen CR, Smith KM, Brynildsen MP, Chou KJY, Hanai T, Liao JC. "Metabolic Engineering of Escherichia Coli for 1-Butanol Production." *Metabolic Engineering* 10 (2008): 305-11.

Atsumi S, Hanai T, Liao JC. "Non-Fermentative Pathways for Synthesis of Branched-Chain Higher Alcohols as Biofuels." *Nature Letters* 451 (2008): 86-90.

Atsumi S, Wu TY, Ecki EM, Hawkins SD, Buelter T, Liao JC. "Engineering the Isobutanol Biosynthetic Pathway in Escherichia Coli by Comparison of Three Aldehyde Reductase/Alcohol Dehydrogenase Genes." *Applied Microbiology and Biotechnology* 85 (2010): 651-57.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. "The Rast Server: Rapid Annotations Using Subsystems Technology." *BMC Genomics* 9 (2008): 75.

Baier D, Reineke K, Doehner I, Mathys A, Knorr D. "Fluorescence-Based Methods for the Detection of Pressure-Induced Spore Germination and Inactivation." *High Pressure Research* 31 (2010): 110.

Baines SJ, Worden RH, ed. *The Long-Term Fate of CO2 in the Subsurface: Natural Analogues for CO2 Storage.* Edited by Worden RH Baines SJ Vol. 233, Geological Storage of Carbon Dioxide. London, United Kingdom: Geological Society Special Publications, 2004.

Ballivet-Tkatchenko D, Chambrey S, Keiski R, Ligabue R, Plasseraud L, Richard P, Turunen H. "Direct Synthesis of Dimethyl Carbonate with Supercritical Carbon Dioxide: Characterization of a Key Organotin Oxide Intermediate." *Catalysis Today* 115 (2006): 80-87.

Barker WW, Welch SA, Chu S, Banfield JF. "Experimental Observations of the Effects of Bacteria on Aluminosilicate Weathering." *American Mineralogist* 83 (1998): 1551-63.

Bastian S, Liu X, Meyerowitz JT, Snow CD, Chen MM, Arnold FH. "Engineered Ketol-Acid Reductoisomerase and Alcohol Dehydrogenase Enable Anaerobic 2-Methylpropan-1-Ol Production at Theoretical Yield in Escherichia Coli." *Metabolic Engineering* 13 (2011): 345–52.

Biddle JF, Fitz-Gibbon S, Schuster SC, Brenchley JE, House CH. "Metagenomic Signatures of the Peru Margin Subseafloor Biosphere Show a Genetically Distinct Environment." *Proceedings of the National Academy of Sciences* 105 (2008): 10583-88.

Biedendieck R, Malten M, Barg H, Bunk B, Martens JH, Deery E, Leech H, Warren MJ, Jahn D. "Metabolic Engineering of Cobalamin (Vitamin B12) Production in Bacillus Megaterium." *Microbial Biotechnology* 3 (2010): 24-37.

Boone DR, Johnson RL, Liu Y "Diffusion of the Interspecies Electron Carriers H2 and Formate in Methanogenic Ecosystems and Its Implications in the Measurement of Km for H2 or Formate Uptake." *Applied and Environmental Microbiology* 55 (1989): 1735-41.

Boonyaratanakornkit BB, Miao LY, Clark DS. "Transcriptional Responses of the Deep-Sea Hyperthermophile Methanocaldococcus Jannaschii under Shifting Extremes of Temperature and Pressure." *Extremophiles* 11, no. 3 (2007): 495-503.

Braissant O, Cailleau G, Dupraz C, Verrecchia AP. "Bacterially Induced Mineralization of Calcium Carbonate in Terrestrial Environments: The Role of Exopolysaccharides and Amino Acids." *Journal of Sedimentary Research* 73 (2003): 485-90.

Brock TD, Freeze H. "Thermus Aquaticus Gen. N. And Sp. N., a Nonsporulating Extreme Thermophile." *Journal of Bacteriology* 98 (1969): 28-297.

Brockman IM, Prather KLJ. "Dynamic Metabolic Engineering: New Strategies for Developing Responsive Cell Factories." *Biotechnology Journal* 10, no. 9 (2015): 1360-69.

Buchfink B, Xie C, Huson DH. "Fast and Sensitive Protein Alignment Using Diamond." *Nature Methods* 12 (2015): 59-60.

Buday Z, Linden JC, Karim MN. "Improved Acetone-Butanol Fermentation Analysis Using Subambient HPLC Column Temperature." *Enzyme and Microbial Technology* 12 (1990): 24-27.

Budisa N, Schulze-Makuch D. "Supercritical Carbon Dioxide and Its Potential as a Life-Sustaining Solvent in a Planetary Environment." *Life* 4 (2014): 331-40.

Buerger S, Spoering A, Gavrish E, Leslin C, Ling L, Epstein SS. "Microbial Scout Hypothesis, Stochastic Exit from Dormancy, and the Nature of Slow Growers." *Applied and Environmental Microbiology* 78 (2012): 3221-28.

Bunk B, Biedendieck R, Jahn D, Vary PS. *Bacillus Megaterium and Other Bacilli: Industrial Applications.* Encyclopedia of Industrial Biotechnology. John Wiley & Sons, Inc., 2010.

Camarena L, Bruno V, Euskirchen G, Poggio S, Snyder M. "Molecular Mechanisms of Ethanol-Induced Pathogenesis Revealed by RNA-Sequencing." *Plos Pathogens* 6 (2010): e1000834.

Cappa J, Rice D. "Carbon Dioxide in Mississippian Rocks of the Paradox Basin and Adjacent Areas, Colorado, Utah, New Mexico, and Arizona." edited by 2000-H U.S. Geological Survey Bulletin, 1995.

Carlson CA, Ingraham JL. "Comparison of Denitrification by Pseudomonas Stutzeri, Pseudomonas Aeruginosa, and Paracoccus Denitrificans." *Applied and Environmental Microbiology* 45 (1983): 1247-53.

Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. "Act: The Artemis Comparison Tool." *Bioinformatics* 21 (2005): 3422-23.

Chang J, Lee S, Kim K. "Arsenic in an as-Contaminated Abandoned Mine Was Mobilized from Fern-Rhizobium to Frond-Bacteria Via the Ars Gene." *Biotechnology and Bioprocess Engineering* 15, no. 5 (2010): 862-73.

Chien A, Edgar DB, Trela JM. "Deoxyribonucleic Acid Polymerase from the Extreme Thermophile Thermus Aquaticus." *Journal of Bacteriology* 127 (1976): 1550-57.

Chivian D, Brodie EL, Alm EJ, Culley DE, Dehal PS, DeSantis TZ, Gihring TM, Lapidus A, Lin LH, Lowry SR, Moser DP, Richardson PM, Southam G, Wanger G, Pratt LM, Andersen GL, Hazen TC, Brockman FJ, Arkin AP, Onstott TC. "Environmental Genomics Reveals a Single-Species Ecosystem Deep within Earth." *Science* 322 (2008): 275-78.

Cinar A, Parulekar SJ, Undey C, Birol G. *Batch Fermentation: Modeling: Monitoring, and Control.* New York, New York: Marcel Dekker, Inc., 2003.

Colwell F, Onstott T, Delwiche M, Chandler D, Fredrickson J, Yao Q, McKinley J, Boone D, Griffiths R, Phelps T, Ringelberg D, White D, LaFreniere L, Balkwill D, Lehman R, Konisky J, Long P. "Microorganisms from Deep, High Temperature Sandstones: Constraints on Microbial Colonization." *FEMS Microbial Ecology* 20 (1997): 425-35.

Connor MR, Cann A, Liao JC. "3-Methyl-1-Butanol Production in Escherichia Coli: Random Mutagenesis and Two-Phase Fermentation." *Applied Microbiology and Biotechnology* 86 (2010): 1155-64.

Connor MR, Liao JC. "Microbial Production of Advanced Transportation Fuels in Non-Natural Hosts." *Current Opinion in Biotechnology* 20 (2009): 307–15.

Coppi MV, Leang C, Sandler SJ, Lovley DR "Development of a Genetic System for Geobacter Sulfurreducens." *Applied and Environmental Microbiology* 67 (2001): 3180-87.

Cronin UP, Martin GK. "The Use of Flow Cytometry to Study the Germination of Bacillus Cereus Endospores." *International Society for Analytical Cytology* 71A (2007): 143-53.

Crump BC, Kling GW, Bahr M, Hobbie JE "Bacterioplankton Community Shifts in an Arctic Lake Correlate with Seasonal Changes in Organic Matter Source." *Applied and Environmental Microbiology* 69 (2003): 2253–68.

Cunliffe M. "Correlating Carbon Monoxide Oxidation with Cox Genes in the Abundant Marine Roseobacter Clade." *ISME J* 5, no. 4 (2011): 685-91.

Cunningham AB, Gerlach R, Spangler L, Mitchell AC. "Microbially Enhanced Geologic Containment of Sequesterd Supercritical CO2." *Energy Procedia* 1 (2009): 3245-52.

David F, Westphal R, Bunk B, Jahn D, Franco-Lara E. "Optimization of Antibody Fragment Production in Bacillus Megaterium: The Role of Metal Ions on Protein Secretion." *Journal of Biotechnology* 150, no. 115-124 (2010).

de Beer D, Haeckel M, Neumann J, Wegener G, Inagaki F, Boetius A. "Saturated CO2 Inhibits Microbial Processes in CO2-Vented Deep-Sea Sediments." *Biogeosciences* 10 (2013): 5639–49.

De Bok FAM, Plugge CM, Stams AJM "Interspecies Electron Transfer in Methanogenic Propionate Degrading Consortia." *Water Research* 38 (2004): 1368-75.

de la Plaza M, Fernandez de Palencia P, Pelaez C, Requena T. "Biochemical and Molecular Characterization of Alpha-Ketoisovalerate Decarboxylase, an Enzyme Involved in the Formation of Aldehydes from Amino Acids by Lactococcus Lactis." *FEMS Microbiology Letters* 238 (2004): 367-74.

Desimone JM, Tumas W, ed. *Green Chemistry Using Liquid and Supercritical Carbon Dioxide.* Oxford, UK: Oxford University Press, 2003.

Dinneen B. "Today's U.S. Ethanol Industry." news release, 2007, https://www.eia.gov/conference/2008/conf_pdfs/Tuesday/Dinneen.pdf.

Dodsworth JA, Hungate BA, Hedlund BP. "Ammonia Oxidation, Denitrification and Dissimilatory Nitrate Reduction to Ammonium in Two Us Great Basin Hot Springs with Abundant Ammonia-Oxidizing Archaea." *Environmental Microbiology* 13, no. 8 (2011): 2371-86.

Edgar RC. "UPARSE: Highly Accurate OTU Sequences from Microbial Amplicon Reads." *Nature Methods* 10 (2013): 996-98.

Ehrenberg SN, Jacobsen KG. "Plagioclase Dissolution Related to Biodegradation of Oil in Brent Group Sandstones (Middle Jurassic) of Gullfaks Field, Northern North Sea." *Sedimentology* 48 (2001): 703-22.

Eichler B, Schink B. "Oxidation of Primary Aliphatic Alcohols by Acetobacterium Carbinolicum Sp. Nov., a Homoacetogenic Anaerobe." *Archives of Microbiology* 140 (1984): 147-52.

Eisenmann E, Beuerle J, Sulger K, Kroneck P, Schumacher W. "Lithotrophic Growth of Sulfurospirillum Deleyianum with Sulfide as Electron Donor Coupled to Respiratory Reduction of Nitrate to Ammonia." *Archives of Microbiology* 164 (1995): 180-85.

EJ, Beckman. "A Challenge for Green Chemistry: Designing Molecules That Readily Dissolve in Carbon Dioxide." *Chemical Communications* 17 (2004): 1885-88.

El-Tarabily KA, Soaud AA, Saleh ME, Matsumoto S. "Isolation and Characterisation of Sulfur-Oxidising Bacteria, Including Strains of Rhizobium, from Calcareous Sandy Soils and Their Effects on Nutrient Uptake and Growth of Maize (Zea Mays L.)." *Australian Journal of Agricultural Research* 57 (2006): 101-11.

Emerson JB, Thomas BC, Alvarez W, Banfield JF. "Metagenomic Analysis of a High Carbon Dioxide Subsurface Microbial Community Populated by Chemolithoautotrophs and Bacteria and Archaea from Candidate Phyla." *Environmental Microbiology* Epub ahead of print (2015).

Engelhardt T, Sahlberg M, Cypionka H, Engelen B. "Biogeography of Rhizobium Radiobacter and Distribution of Associated Temperate Phages in Deep Subseafloor Sediments." *ISME J* 7 (2013): 199-209.

Eppinger M, Bunk B, Johns MA, Edirisinghe JN, Kutumbaka KK, Koenig SS, Creasy HH, Rosovitz MJ, Riley DR, Daugherty S, Martin M, Elbourne LD, Paulsen I, Biedendieck R, Braun C, Grayburn S, Dhingra S, Lukyanchuk V, Ball B, Ul-Qamar R, Seibel J, Bremer E, Jahn D, Ravel J, Vary PS. "Genome Sequences of the Biotechnologically Important Bacillus Megaterium Strains Qm B1551 and Dsm319." *Journal of Bacteriology* 193 (2011): 4199-213.

Ezeji TC, Milne C, Price ND, Blaschek HP. "Achievements and Perspectives to Overcome the Poor Solvent Resistance in Acetone and Butanol-Producing Microorganisms." *Applied Microbiology and Biotechnology* 85 (2010): 1697–712.

Ezeji TC, Qureshi N, Blaschek HP. "Acetone-Butanol-Ethanol Production from Concentrated Substrate: Reduction in Substrate Inhibition by Fed-Batch Technique and Product Inhibition by Gas Stripping." *Applied Microbiology and Biotechnology* 63 (2004): 653-58.

Ezeji TC, Qureshi N, Blaschek HP. "Bioproduction of Butanol from Biomass: From Genes to Bioreactors." *Current Opinion in Biotechnology* 18, no. 3 (2007): 220-27.

Fernandes SO, Bonin PC, Michotey VD, LokaBharathi PA. "Denitrification: An Important Pathway for Nitrous Oxide Production in Tropical Mangrove Sediments (Goa, India)." *Journal of Environmental Quality* 39 (2010): 1507-16.

Ferrari E, Miller B. "Bacillus Expression: A Gram Positive Model." In *Gene Expression Systems*, edited by Hoeffler JP Fernandez JM, 66-94. San Diego, California: Academic Press Carlsbad, 1999.

Ferris JP, Hill Jr AR, Liu R, Orgel LE. "Synthesis of Long Prebiotic Oligomers on Mineral Surfaces." *Nature* 381 (1996): 59-61.

Finster K, Liesack W, Tindall BJ. "Sulfurospirillum Arcachonense Sp. Nov., a New Microaerophilic Sulfur-Reducing Bacterium." *International Journal of Systematic and Evolutionary Microbiology* 47, no. 4 (1997): 1212-17.

Fischer CR, Klein-Marcuschamer D, Stephanopoulos G. "Selection and Optimization of Microbial Hosts for Biofuels Production." *Metabolic Engineering* 10 (2008): 295–304.

Foster JW. "When Protons Attack: Microbial Strategies of Acid Adaptation." *Current Opinion in Microbiology* 2 (1999): 170-74.

Gagen EJ, Denman SE, Padmanabha J, Zadbuke S, Jassim RA, Morrison M, McSweeney CS. "Functional Gene Analysis Suggests Different Acetogen Populations in the Bovine Rumen and Tammar Wallaby Forestomach." *Applied and Environmental Microbiology* 76, no. 23 (2010): 7785-95.

Gaidenko TA, Price CW. "General Stress Transcription Factor Simgab and Sporulation Transcription Factor Sigmah Each Contribute to Survival of Bacillus Subtilis under Extreme Growth Conditions." *Journal of Bacteriology* 180 (1998): 3730-33.

Gao P, Tian H, Wang Y, Li Y, Li Y, Xie J, Zeng B, Zhou J, Li G, Ma T. "Spatial Isolation and Environmental Factors Drive Distinct Bacterial and Archaeal Communities in Different Types of Petroleum Reservoirs in China." *Scientific Reports* 6 (2016): 20174.

Geesey G, Jang L. "Extracellular Polymers for Metal Binding." In *Microbial Mineral Recovery*, edited by Brierley CL Ehrlich HL, 223-47. New York, New York: McGraw Hill, 1990.

Ghosh S, Setlow P. "Isolation and Characterization of Superdormant Spores of Bacillus Species." *Journal of Bacteriology* 191 (2009): 1787-97.

Gilfillan SMV, Ballentine CJ, Holland G, Blagburn D, Lollar BS, Stevens S, Schoelld M, Cassidye M. "The Noble Gas Geochemistry of Natural CO2 Gas Reservoirs from the Colorado Plateau and Rocky Mountain Provinces, USA." *Geochimica et Cosmochimica Acta* 72, no. 4 (2008): 1174-98.

Gilfillan SMV, Lollar BS, Holland G, Blagburn D, Stevens S, Schoell M, Cassidy M, Ding Z, Zhou Z, Lacrampe-Couloume G, Ballentine CJ. "Solubility Trapping in Formation Water as Dominant CO2 Sink in Natural Gas Fields." *Nature Letters* 458 (2009): 614-18.

Goris T, Schubert T, Gadkari J, Wubet T, Tarkka M, Buscot F, Adrian L, Diekert G. "Insights into Organohalide Respiration and the Versatile Catabolism of Sulfurospirillum Multivorans Gained from Comparative Genomics and Physiological Studies." *Environmental Microbiology* 16, no. 11 (2014): 3562-80.

Grigoriev A. "Analyzing Genomes with Cumulative Skew Diagrams." *Nucleic Acids Research* 26, no. 10 (1998): 2286-90.

Guérout-Fleury AM, Frandsen N, Stragier P. "Plasmids for Ectopic Integration in Bacillus Subtilis." *Gene* 180 (1996): 57-61.

GVR. "Isobutanol Market Analysis by Product (Synthetic, Bio-Based), Application (Oil & Gas, Solvents & Coatings, Chemical Intermediates) and Segment Forecasts to 2022." edited by Grand View Research, 2015.

Hamady M, Lozupone C, Knight R. "Fast Unifrac: Facilitating High-Throughput Phylogenetic Analyses of Microbial Communities Including Analysis of Pyrosequencing and Phylochip Data." *ISME J* 4 (2009): 17-27.

Hammond DA, Karel M, Klibanov AM, Krukonis VJ. "Enzymatic-Reactions in Supercritical Gases." *Applied Biochemistry and Biotechnology* 11 (1985): 393-400.

Hartman FC, Harpel MR. "Structure, Function, Regulation and Assembly of D-Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase." *Annual Review of Biochemistry* 63 (1994): 197-234.

Haszeldine RS, Quinn O, England G, Wilkinson M, Shipton ZK, Evans JP, Heath J, Crossey L, Ballentine CJ, Graham C. "Natural Geochemical Analogues for Carbon Dioxide Storage in Deep Geological Porous Reservoirs, a United Kingdom Perspective." *Oil and Gas Science and Technology* 60, no. 1 (2005): 33-49.

Hattori S, Luo H, Shoun H, Kamagata Y. "Involvement of Formate as an Interspecies Electron Carrier in a Syntrophic Acetate-Oxidizing Anaerobic Microorganism in Coculture with Methanogens." *Journal of Bioscience and Bioengineering* 91, no. 3 (2001): 294-98.

Holloway S., Pearce JM, Ohsumi T, Hards VL. "A Review of Natural CO2 Occurrences and Their Relevance to CO2 Storage." edited by International Energy Agency, 124, 2005.

Hu B, Lidstrom ME. "Metabolic Engineering of Methylobacterium Extorquens Am1 for 1-Butanol Production." *Biotechnology for Biofuels* 7, no. 156 (2014): 1-10.

Hubert C, Voordouw G. "Oil Field Souring Control by Nitrate-Reducing Sulfurospirillum Spp. That Outcompete Sulfate-Reducing Bacteria for Organic Electron Donors." *Applied and Environmental Microbiology* 73 (2007): 2644-52.

Hyatt MT, Levinson HS. "Conditions Affecting Bacillus Megaterium Spore Germination in Glucose or Various Nitrogenous Compounds." *Journal of Bacteriology* 83 (1962): 1231-37.

Inui M, Suda M, Kimura S, Yasuda K, Suzuki H, Toda H, Yamamoto S, Okino S, Suzuki N, Yukawa H "Expression of Clostridium Acetobutylicum Butanol Synthetic Genes in Escherichia Coli." *Applied Microbiology and Biotechnology* 77 (2008): 1305-16.

IPCC, ed. *Climate Change 2007: The Physical Science Basis.* Edited by Qin D Solomon S, Manning M, Chen Z, Marquis M, Averyt KB, et al., Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel of Climate Change. Cambridge, United Kingdom: Cambridge University Press, 2007.

Irwin H, Curtis CD, Coleman ML. "Isotopic Evidence for Source of Diagenetic Carbonate Formed During the Burial of Organic Rich Sediments." *Nature* 269 (1977): 209-13.

Isenschmid A, Marison WI, Stockar VU. "The Influence of Pressure and Temperature of Compressed CO2 on the Survival of Yeast Cells." *Journal of Biotechnology* 39 (1995): 229-37.

Itävaara M, Nyyssönen M, Kapanen A, Nousiainen A, Ahonen L, Kukkonen I. "Characterization of Bacterial Diversity to a Depth of 1500 M in the Outokumpu Deep Borehole, Fennoscandian Shield." *FEMS Microbial Ecology* 77, no. 2 (2011): 295-309.

John M, Rubick R, Schmitz RP, Rakoczy J, Schubert T, Diekert G. "Retentive Memory of Bacteria: Long-Term Regulation of Dehalorespiration in Sulfurospirillum Multivorans." *Journal of Bacteriology* 191 (2009): 1650-55.

Johnson DB. "Biodiversity and Ecology of Acidophilic Microorganisms." *FEMS Microbiology Ecology* 27 (1998): 307-17.

Kaden J, Galushko AS, Schink B. "Cysteine-Mediated Electron Transfer in Syntrophic Acetate Oxidation by Cocultures of Geobacter Sulfurreducens and Wolinella Succinogenes." *Archives of Microbiology* 178, no. 1 (2002): 53-58.

Katano Y, Fujinami S, Kawakoshi A, Nakazawa H, Oji S, Iino T, Oguchi A, Ankai A, Fukui S, Terui Y, Kamata S, Harada T, Tanikawa S, Suzuki K, Fujita N. "Complete Genome Sequence of Oscillibacter Valericigenes Sjm18-20t (=Nbrc 101213t) ". *Standards in Genomic Sciences* 6 (2012): 406-14.

Kataoka N, Tajima T, Kato J, Rachadech W, Vangnai AS. "Development of Butanol-Tolerant Bacillus Subtilis Strain Grsw2-B1 as a Potential Bioproduction Host." *AMB Express* 1 (2011): 1-11.

Kawanami H, Ikushima Y. "Chemical Fixation of Carbon Dioxide to Styrene Carbonate under Supercritical Conditions with Dmf in the Absence of Any Additional Catalysts." *Chemical Communications* (2000): 2089-90.

Kelly DP, Wood AP, ed. *The Chemolithotrophic Prokaryotes.* Edited by Falkow S Dworkin M, Rosenberg E, Schleifer K, Stackebrandt E, The Prokaryotes: Vol. 2: Ecophysiology and Biochemistry. Singapore: Springer, 2006.

Kharaka YK, Cole DR, Kovorka SD, Gunter WD, Knauss KG, Freifeld BM "Gas-Water-Rock Interactions in Frio Formation Following CO2 Injection: Implications for the Storage of Greenhouse Gases in Sedimentary Basins." *Geology* 34 (2006): 577-80.

Khosravi-Darani K, Vasheghani-Farahani E. "Application of Supercritical Fluid Extraction in Biotechnology." *Critical Reviews in Biotechnology* 25 (2005): 231-42.

Kieft TL, Ringelberg DB, White DC. "Changes in Ester-Linked Phospholipid Fatty Acid Profiles of Subsurface Bacteria During Starvation and Desiccation in a Porous Medium." *Applied and Environmental Microbiology* 60 (1994): 3292-99.

Kim H, Goepfert J. "A Sporulation Medium for Bacillus Anthracis." *Journal of Applied Microbiology* 37 (1974): 265-67.

Kiran E, Debenedetti PG, Peters CJ. *Supercritical Fluids Fundamentals and Applications.* Dordrecht: Kluwer Academic Publishers, 2000.

Kirk MF. "Variation in Energy Available to Populations of Subsurface Anaerobes in Response to Geologic Carbon Storage." *Environmental Science & Technology* 45 (2011): 6676-82.

Klein W, Weber MH, Marahiel MH. "Cold Shock Response of Bacillus Subtilis: Isoleucine-Dependent Switch in the Fatty Acid Branching Pattern for Membrane Adaptation to Low Temperatures." *Journal of Bacteriology* 181 (1999): 5341-49.

Knoshaug EP, Zhang M. "Butanol Tolerance in a Selection of Microorganisms." *Applied Biochemistry and Biotechnology* 153 (2009): 13–20.

Knutson BL, Strobel HJ, Nokes SE, Dawson KA, Berberich JA, Jones CR. "Effect of Pressurized Solvents on Ethanol Production by the Thermophilic Bacterium Clostridium Thermocellum." *Journal of Supercritical Fluids* 16 (1999): 149-56.

Komlos J, Cunningham AB, Camper AK, Sharp RR. "Biofilm Barriers to Contain and Degrade Dissolved Trichloroethylene." *Environmental Progress* 23 (2004): 69-77.

Kopke M, Held C, Hujer S, Liesegang H, Wiezer A, Wollherr A, Ehrenreich A, Liebl W, Gottschalk G, Durre P. "Clostridium Ljungdahlii Represents a Microbial Production Platform Based on Syngas." *Proceedings of the National Academy of Sciences* 107 (2010): 13087–92.

Korneli C, David F, Biedendieck R, Jahn D, Wittmann C. "Getting the Big Beast to Work--Systems Biotechnology of Bacillus Megaterium for Novel High-Value Proteins." *Journal of Biotechnology* 163 (2013): 87-96.

LaBelle EV, Marshall CW, Gilbert JA, May HD. "Influence of Acidic pH on Hydrogen and Acetate Production by an Electrosynthetic Microbiome." *Plos One* 9, no. 10 (2014): e109935.

Laitinen A, Kaunisto J "Supercritical Fluid Extraction of 1-Butanol from Aqueous Solutions." *Journal of Supercritical Fluids* 15 (1999): 245-52.

Lal R. "Carbon Sequestration." *Philosophical Transactions of the Royal Society B: Biological Sciences* 363, no. 1492 (2008): 815–30.

Lam FH, Ghaderi A, Fink GR, Stephanopoulos G. "Engineering Alcohol Tolerance in Yeast." *Science* 346 (2014): 71-75.

Lee JY, Kim CG, Mahanty B. "Mineralization of Gaseous CO2 by Bacillus Megaterium in Close Environment System." *Water, Air, & Soil Pollution* 225 (2014): 1787.

Lee WH, Seo SO, Bae YH, Nan H, Seo JH. "Isobutanol Production in Engineered Saccharomyces Cerevisiae by Overexpression of 2-Ketoisovalerate Decarboxylase and Valine Biosynthetic Enzymes." *Bioprocess and Biosystems Engineering* 35, no. 9 (2012): 1467-75.

Lee YJ, Romanek CS, Wiegel J. "Desulfosporosinus Youngiae Sp. Nov., a Spore-Forming, Sulfate-Reducing Bacterium Isolated from a Constructed Wetland Treating Acid Mine Drainage." *International Journal of Systematic and Evolutionary Microbiology* 59 (2009): 2743-46.

Leitner W. "Supercritical Carbon Dioxide as a Green Reaction Medium for Catalysis." *Accounts of Chemical Research* 35, no. 9 (2002): 746-56.

Letain TE, Kane SR, Legler TC, Salazar EP, Agron PG, Beller HR. "Development of a Genetic System for the Chemolithoautotrophic Bacterium Thiobacillus Denitrificans." *Applied and Environmental Microbiology* 73 (2007): 3265-71.

Levinson HS, Hyatt MT. "Activation Energy for Glucose-Induced Germination of Bacillus Megaterium Spores." *Journal of Bacteriology* 103 (1970): 269-70.

Li S, Wen J, Jia X. "Engineering Bacillus Subtilis for Isobutanol Production by Heterologous Ehrlich Pathway Construction and the Biosynthetic 2-Ketoisovalerate Precursor Pathway Overexpression." *Applied Microbiology and Biotechnology* 91, no. 3 (2011): 577-89.

Liao H, Zhang F, Hu X, Liao X. "Effects of High-Pressure Carbon Dioxide on Proteins and DNA in Escherichia Coli." *Microbiology* 157 (2011): 709-20.

Liu A, Garcia-Dominguez E, Rhine ED, Young LY. "A Novel Arsenate Respiring Isolate That Can Utilize Aromatic Substrates." *FEMS Microbial Ecology* 48 (2004): 323-32.

Liu L, Li Y, Zhang J, Zou W, Zhou Z, Liu J, Li X, Wang L, Chen J. "Complete Genome Sequence of the Industrial Strain Bacillus Megaterium Wsh-002." *Journal of Bacteriology* 193 (2011): 6389-90.

Liu S, Qureshi N. "How Microbes Tolerate Ethanol and Butanol." *New Biotechnology* 26 (2009): 117-21.

Lloyd JR, Oremland RS. "Microbial Transformations of Arsenic in the Environment: From Soda Lakes to Aquifers." *Elements* 2 (2006): 85-90.

Lovley DR, Chapelle FH. "Deep Subsurface Microbial Processes." *Review of Geophysics* 33 (1995): 365-81.

Magge A, Setlow B, Cowan AE, Setlow P. "Analysis of Dye Binding by and Membrane Potential in Spores of Bacillus Species." *Journal of Applied Microbiology* 106 (2009): 814-24.

Magnuson JK, Stern RV, Gossett JM, Zinder SH, Burris DR. "Reductive Dechlorination of Tetrachloroethene to Ethene by a Two-Component Enzyme Pathway." *Applied and Environmental Microbiology* 64 (1998): 1270-75.

Malten M, Hollmann R, Deckwer WD, Jahn D. "Production and Secretion of Recombinant leuconostoc Mesenteroides dextransucrase Dsrs In bacillus Megaterium." *Biotechnology and Bioengineering* 89 (2005): 206-18.

Mangelsdorf K, Finsel E, Liebner S, Wagner D. "Temperature-Dependent Molecular Cell Membrane Adaptation of Microbial Populations from a Permafrost Region." In *Goldschmidt 2007.* Cologne, Germany, 2007.

Marshall CW, Ross DE, Fichot EB, Norman RS, May HD. "Long-Term Operation of Microbial Electrosynthesis Systems Improves Acetate Production by Autotrophic Microbiomes." *Environmental Science & Technology* 47 (2013): 6023-29.

Martens JH, Barg H, Warren MJ, Jahn D. "Microbial Production of Vitamin B12." *Applied and Environmental Microbiology* 58, no. 3 (2002): 275-85.

Martin-Galiano AJ, Ferrandiz MJ, de la Campa AG. "The Promoter of the Operon Encoding the F0f1 ATPase of Streptococcus Pneumoniae Is Inducible by Ph." *Molecular Microbiology* 41 (2001): 1327-38.

Marty A, Chulalaksananukul W, Willemot RM, Condoret JS. "Kinetics of Lipase-Catalyzed Esterification in Supercritical CO2." *Biotechnology and Bioengineering* 39, no. 3 (1992): 273-80.

Mascia S, Heider PL, Zhang H, Lakerveld R, Benyahia B, Barton PI, Braatz RD, Cooney CL, Evans JMB, Jamison TF, Jensen KF, Myerson AS, Trout BL. "End-to-End Continuous Manufacturing of Pharmaceuticals: Integrated Synthesis, Purification, and Final Dosage Formation." *Angewandte Chemie International Edition* 52 (2013): 12359-63.

Matsuda T, Harada T, Nakamura K. "Alcohol Dehydrogenase Is Active in Supercritical Carbon Dioxide." *Chemical Communications* 15 (2000): 1367-68.

Matsuda T, Harada T, Nakamura K. "Biocatalysis in Supercritical CO2." *Current Organic Chemistry* 9 (2005): 299-315.

Matsuda T, Marukado R, Koguchi S, Nagasawa T, Mukouyama M, Harada T, Nakamura K. "Novel Continuous Carboxylation Using Pressurized Carbon Dioxide by Immobilized Decarboxylase." *Tetrahedron Letters* 49 (2008): 6019-20.

Matsuda T, Ohashi Y, Harada T, Yanagihara R, Nagasawa T, Nakamura K. "Conversion of Pyrrole to Pyrrole-2-Carboxylate by Cells of Bacillus Megaterium in Supercritical CO2." *Chemical Communications*, no. 21 (2001): 2194-95.

Matsuda T, Watanabe K, Harada T, Nakamura K. "Enzymatic Reactions in Supercritical CO2: Carboxylation, Asymmetric Reduction and Esterification." *Catalysis Today* 96 (2004): 103-11.

Mitchell AC, Phillips AJ, Hamilton MA, Gerlach R, Hollis WK, Kaszuba JP, Cunningham AB. "Resilience of Planktonic and Biofilm Cultures to Supercritical CO2." *The Journal of Supercritical Fluids* 47 (2008): 318-25.

Mitchell AC, Phillips AJ, Hiebert R, Gerlach R, Spangler LH, Cunningham AB. "Biofilm Enhanced Geologic Sequestration of Supercritical CO2." *International Journal of Greenhouse Gas Control* 3 (2009): 90-99.

Mizrahi-Man O, Davenport ER, Gilad Y. "Taxonomic Classification of Bacterial 16s rRNA Genes Using Short Sequencing Reads: Evaluation of Effective Study Designs." *Plos One* 8 (2013): e53608.

Moro A, Sánchez JC, Serguera C. "Transformation of Bacillus Megaterium by Electroporation." *Biotechnology Techniques* 9, no. 8 (1995): 589-90.

Morozova D, Wandrey M, Alawi M, Zimmer M, Vieth-Hillebrand A, Zettlitzer M, Würdemann H. "Monitoring of the Microbial Community Composition of the Saline Aquifers During CO2 Storage by Fluorescence in Situ Hybridisation." *International Journal of Greenhouse Gas Control* 4, no. 6 (2010): 981-89.

Morozova D, Zettlitzer M, Let D, Wurdemann H, CO2SINK Group. "Monitoring of the Microbial Community Composition in Deep Subsurface Saline Aquifers During CO2 Storage in Ketzin, Germany." *Energy Procedia* 4 (2011): 4362-70.

Morris BE, Henneberger R, Huber H, Moissl-Eichinger C. "Microbial Syntrophy: Interaction for the Common Good." *FEMS Microbial Ecology* 37, no. 3 (2013): 384-406.

Mu A, Boreham C, Leong HX, Haese RR, Moreau JW. "Changes in the Deep Subsurface Microbial Biosphere Resulting from a Field-Scale CO2 Geosequestration Experiment." *Frontiers in Microbiology* 5, no. 209 (2014): 1-11.

Mukhopadhyay A, He Z, Alm EJ, Arkin AP, Baidoo EE, Borglin SC, Chen W, Hazen TC, He Q, Holman HY, Huang K, Huang R, Joyner DC, Katz N, Keller M, Oeller P, Redding A, Sun J, Wall J, Wei J, Yang Z, Yen HC, Zhou J, Keasling JD. "Salt Stress in Desulfovibrio Vulgaris Hildenborough: An Integrated Genomics Approach." *Journal of Bacteriology* 11 (2006): 4068-78.

Nakamura K, Chi YM, Yoshinobu Y, Yano T. "Lipase Activity and Stability in Supercritical Carbon Dioxide." *Chemical Engineering Communications* 45 (1986): 207-12.

Nakayama T, Kamikawa R, Tanifuji G, Kashiyama Y, Ohkouchi N, Archibald JM, Inagaki, Y. "Complete Genome of a Nonphotosynthetic Cyanobacterium in a Diatom Reveals Recent Adaptations to an Intracellular Lifestyle." *Proceedings of the National Academy of Sciences of the United States of America* 111, no. 31 (2014): 11407-12.

Nedwell DB, Banat IM. "Hydrogen as an Electron Donor for Sulfate-Reducing Bacteria in Slurries of Salt Marsh Sediment." *Microbial Ecology* 7, no. 4 (1981): 305-13.

NETL. "Carbon Dioxide Enhanced Oil Recovery: Untapped Domestic Energy Supply and Long Term Carbon Storage Solution." edited by US Department of Energy National Energy Technology Laboratory, 2010.

Nielsen DR, Leonard E, Yoon SH, Tseng HC, Yuan C, Prather KLJ. "Engineering Alternative Butanol Production Platforms in Heterologous Bacteria." *Metabolic Engineering* 11 (2009): 262–73.

Nigam PS, Singh A. "Production of Liquid Biofuels from Renewable Resources." *Progress in Energy and Combustion Science* 37 (2011): 52-68.

Ogasawara H, Shinohara S, Yamamoto K, Ishihama A. "Novel Regulation Targets of the Metal-Response Bass-Basr Two-Component Sysytem of Escherichia Coli." *Microbiology* 158 (2012): 1482-92.

Onstott TC. "Impact of CO2 Injections on Deep Subsurface Microbial Ecosystems and Potential Ramifications for the Surface Biosphere." In *The CO2 Capture and Storage Project (CCP) for Carbon Dioxide Storage in Deep Geologic Formations for Climate Change Mitigation, Geologic Storage of Carbon Dioxide with Monitoring and Verification*, edited by Benson SM, 1217–50. Oxford, UK: Elsevier Publishing, 2005.

Oppermann BI, Michaelis W, Blumenberg M, Frerichsc J, Schulzc HM, Schippers A, Beaubiend SE, Krüger M. "Soil Microbial Community Changes as a Result of Long-Term Exposure to a Natural CO2 Vent." *Geochimica et Cosmochimica Acta* 74, no. 9 (2010): 2697-716.

Orr FM. "Onshore Geologic Storage of CO2." *Science* 325 (2009): 1656-58.

Ortuño C, Martínez-Pastor MT, Mulet A, Benedito, J. "Supercritical Carbon Dioxide Inactivation of Escherichia Coli and Saccharomyces Cerevisiae in Different Growth Stages." *The Journal of Supercritical Fluids* 63 (2012): 8-15.

Oudshoorn A. "Recovery of Bio-Based Butanol." Technische Universiteit Delft, 2012.

Oulé MK, Dickman M, Joseph A. "Microbicidal Effect of Pressurized CO2 and the Influence of Sensitizing Additives." *European Journal of Scientific Research* 41, no. 4 (2010): 569-81.

Oulé MK, Kablan T, Bernier AM, Arul, J. "Escherichia Coli Inactivation Mechanism by Pressurized CO2." *Canadian Journal of Microbiology* 52 (2006): 1208-17.

Peet KC, Freedman AJE, Hernandez HH, Britto V, Boreham C, Ajo-Franklin JB, Thompson JR. "Microbial Growth under Supercritical CO2." *Applied and Environmental Microbiology* 81, no. 8 (2015): 2881-92.

Pham VD, Hnatow LL, Zhang S, Fallon RD, Jackson SC, Tomb JF, DeLong EF, Keeler SJ. "Characterizing Microbial Diversity in Production Water from an Alaskan Mesothermic Petroleum Reservoir with Two Independent Molecular Methods." *Environmental Microbiology* 11 (2009): 176-87.

Phelps TJ, Murphy EM, Pfiffner SM, White DC. "Comparison between Geochemical and Biological Estimates of Subsurface Microbial Activities." *Microbial Ecology* 28 (1994): 335-49.

Poechlauer P, Manley J, Broxterman R, Gregertsen B, Ridemark M. "Continuous Processing in the Manufacture of Active Pharmaceutical Ingredients and Finished Dosage Forms: An Industry Perspective." *Organic Process Research & Development* 16 (2012): 1586-90.

Preheim SP, Perrotta AR, Martin-Platero AM, Gupta A, Alm EJ "Distribution-Based Clustering: Using Ecology to Refine the Operational Taxonomic Unit." *Applied and Environmental Microbiology* 79 (2013): 6593-603.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. "The Silva Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools." *Nucleic Acids Research* 41, no. D590–D596 (2013).

Rabinowitz D, Janowiak M. "Reasonable, Foreseeable Development: Oil, Natural Gas, and Carbon Dioxide in Canyons of the Ancients National Monument." edited by San Juan Public Lands Center Bureau of Land Management, 66, 2005.

Rastogi G, Osman S, Kukkadapu R, Engelhard M, Vaishampayan PA, Andersen GL, Sani RK. "Microbial and Mineralogical Characterizations of Soils Collected from the Deep Biosphere of the Former Homestake Gold Mine, South Dakota." *Microbial Ecology* 60 (2010): 539-50.

Richard H, Foster JW. "Escherichia Coli Glutamate- and Arginine-Dependent Acid Resistance Systems Increase Internal pH and Reverse Transmembrane Potential." *Journal of Bacteriology* 186, no. 6032-6041 (2004).

Richhardt J, Larsen M, Meinhardt F. "An Improved Transconjugation Protocol for Bacillus Megaterium Facilitating a Direct Genetic Knockout." *Applied Microbiology and Biotechnology* 86 (2010): 1959-65.

Roberge DM, Ducry L, Bieler N, Cretton P, Zimmermann B. "Microreactor Technology: A Revolution for the Fine Chemical and Pharmaceutical Industries?". *Chemical Engineering & Technology* 28 (2005): 318-23.

Rodriguez-Contreras A, Koller M, Miranda-de Sousa Dias MMS, Calafell-Monfort M, Braunegg G, Marques-Calvo MS. "High Production of Poly(3-Hydroxybutyrate) from a Wild Bacillus Megaterium Bolivian Strain." *Journal of Applied Microbiology* 114 (2013): 1378-87.

Romão CV, Pereira IA, Xavier AV, LeGall J, Teixeira M. "Characterization of the [Nife] Hydrogenase from the Sulfate Reducer Desulfovibrio Vulgaris Hildenborough." *Biochemical and Biophysical Research Communications* 240 (1997): 75-79.

Roseboom W, De Lacey AL, Fernandez VM, Hatchikian EC, Albracht SPJ. "The Active Site of the [Fefe]-Hydrogenase from Desulfovibrio Desulfuricans. Ii. Redox Properties, Light Sensitivity and Co-Ligand Exchange as Observed by Infrared Spectroscopy." *Journal of Biological Inorganic Chemistry* 11 (2006): 102-18.

Roth M. "Helium Head Pressure Carbon Dioxide in Supercritical Fluid Extraction and Chromatography: Thermodynamic Analysis of the Effects of Helium." *Analytical Chemistry* 70 (1998): 2104-09.

Roth NG, Lively DH. "Germination of Spores of Certain Aerobic Bacilli under Anaerobic Conditions." *Journal of Bacteriology* 71 (1956): 162-66.

Sabirzyanov AN, Il'in AP, Akhunov AR, Gumerov FM. "Solubility of Water in Supercritical Carbon Dioxide." *High Temperature* 40 (2002): 203-06.

Salgin U, Salgin S, Takac S. "The Enantioselective Hydrolysis of Racemic Naproxen Methyl Ester in Supercritical CO2 Using Candida Rugosa Lipase." *Journal of Supercritical Fluids* 43 (2007): 310-16.

Salter S, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman N, Walker AW. "Reagent Contamination Can Critically Impact Sequence-Based Microbiome Analyses." *BMC Biology* 12 (2014): 1-12.

Seo Y, Lee WH, Sorial G, Bishop PL. "The Application of a Mulch Biofilm Barrier for Surfactant Enhanced Polycyclic Aromatic Hydrocarbon Bioremediation." *Environmental Pollution* 157 (2009): 95-101.

Sulzenbacher G, Alvarez K, Van Den Heuvel RH, Versluis C, Spinelli S, Campanacci V, Valencia C, Cambillau C, Eklund H, Tegoni M. "Crystal Structure of E. coli Alcohol Dehydrogenase YqhD: Evidence of a Covalently Modified NAPD Coenzyme." *Journal of Molecular Biology* 342 (2004): 489-502.

Santillan EU, Shanahan TM, Omelon CR, Major JR, Bennett PC. "Isolation and Characterization of a CO2-Tolerant Lactobacillus Strain from Crystal Geyser, Utah, U.S.A.". *Frontiers in Earth Science* 3 (2015): 1-11.

Santos S., Neto IFF, Machado MD, Soares HMVM. "Siderophore Production by Bacillus Megaterium: Effect of Growth Phase and Cultural Conditions." *Applied Biochemistry and Biotechnology* 172 (2014): 549-60.

Setlow B, Swaroopa A, Kitchel R, Koziol-Dube K, Setlow P. "Role of Dipicolinic Acid in Resistance and Stability of Spores of Bacillus Subtilis with or without DNA-Protective Alpha/Beta-Type Small Acid-Soluble Proteins." *Journal of Bacteriology* 188 (2006): 3740-47.

Setlow P. "Spore Germination." *Current Opinion in Microbiology* 6 (2003): 550–56.

Setlow P. "Spores of Bacillus Subtilis: Their Resistance to and Killing by Radiation, Heat and Chemicals." *Journal of Applied Microbiology* 101 (2006): 514-25.

Sheppard MJ, Kunjapur AM, Wenck SJ, Prather KLJ. "Retro-Biosynthetic Screening of a Modular Pathway Design Achieves Selective Route for Microbial Synthesis of 4-Methyl-Pentanol." *Nature Communications* 5 (2014): 1-10.

Shipton Z, Evans JP, Kirchner D, Kolesar PT, Williams AP, Heath J. "Analysis of CO2 Leakage Along Faults from Natural Reservoirs in the Colorado Plateau, Us." In *Geologic Storage of Carbon Dioxide*, edited by Worden RH Baines SJ, 43-58. London, United Kingdom: Geological Society Special Publications, 2004.

Sinclair L, Osman OA, Bertilsson S, Eiler A. "Microbial Community Composition and Diversity Via 16s rRNA Gene Amplicons: Evaluating the Illumina Platform." *Plos One* 10 (2015): e0116955.

Smith JT, Ehrenberg SN. "Correlation of Carbon Dioxide Abundance with Temperature in Clastic Hydrocarbon Reservoirs: Relationship to Inorganic Chemical Equilibrium." *Marine and Petroleum Geology* 6 (1989): 129-35.

Smith KS, Ferry JG. "Prokaryotic Carbonic Anhydrases." *FEMS Microbial Ecology* 24, no. 4 (2000): 335-66.

Solomon KV, Sanders TM, Prather KLJ. "A Dynamic Metabolite Valve for the Control of Central Carbon Metabolism." *Metabolic Engineering* 14, no. 6 (2012): 661-71.

Sowden, R. J.; Sellin, M. F.; Blasio, N. D.; Cole-Hamilton, D. J. "Carbonylation of Methanol in Supercritical CO2 Catalysed by a Supported Rhodium Complex." *Chemical Communications* (1999): 2511-12.

Spilimbergo S, Bertucco A. "Non-Thermal Bacteria Inactivation with Dense CO2." *Biotechnology and Bioengineering* 84 (2003): 627-38.

Spilimbergo S, Mantoan D, Quaranta A, Della Mea G. "Real-Time Monitoring of Cell Membrane Modification During Supercritical CO2 Pasteurization." *Journal of Supercritical Fluids* 48 (2009): 93-97.

SRES. "Special Report on Emissions Scenarios: A Special Report of Working Group Iii of the Intergovernmental Panel on Climate Change." edited by Swart R Nakicenovic N. Cambridge, United Kingdom, 2000.

Steen EJ, Chan R, Prasad N, Myers S, Petzold CJ, Redding A, Ouellet M,, and Keasling JD. "Metabolic Engineering of Saccharomyces Cerevisiae for the Production of N-Butanol." *Microbial Cell Factories* 7, no. 36 (2008).

Stevens SH, Pearce JM, Rigg AA. "Natural Analogs for Geologic Storage of CO2: An Integrated Global Research Program." In *First National Conference on Carbon Sequestration.* Washington, D.C.: U.S. Department of Energy, National Energy Technology Laboratory, 2001.

Stolz JF, Ellis DJ, Blum JS, Ahmann D, Lovley DR, Oremland RS. "Sulfurospirillum Barnesii Sp. Nov. And Sulfurospirillum Arsenophilum Sp. Nov., New Members of the Sulfurospirillum Clade of the Epsilon Proteobacteria." *International Journal of Systematic and Evolutionary Microbiology* 49 (1999): 1177-80.

Suess J, Sass H, Cypionka H, Engelen B. "Widespread Distribution of Rhizobium Radiobacter in Mediterranean Sediments." *Progr. Abstr., Jnt. Int. Symp. Subsurf. Microbiol. (ISSM 2005) and Environ. Biogeochem. (ISEB XVII)* (2005): 63.

Sugimura K, Kuwabata S, Yoneyama, H. "Electrochemical Fixation of Carbon Dioxide in Oxoglutaric Acid Using an Enzyme as an Electrocatalyst." *Journal of the American Chemical Society* 111 (1989): 2361-62.

Sugimura K, Kuwabata S, Yoneyama H. "Electrochemical Fixation of Carbon Dioxide in Pyruvic Acid to Yield Malic Acid Using Malic Enzyme as an Electrocatalyst." *Journal of Electroanalytical Chemistry* 299 (1990): 241.

Szulczewski ML, MacMinn CW, Herzog HJ, Juanes R. "Lifetime of Carbon Capture and Storage as a Climate-Change Mitigation Technology." *Proceedings of the National Academy of Sciences* 109, no. 14 (2012): 5185-89.

Takai K, Nakamura K, Toki T, Tsunogai U, Miyazaki M, Miyazaki J, Hirayama H, Nakagawa S, Nunoura T, Horikoshi K. "Cell Proliferation at 122 Degrees C and Isotopically Heavy Ch4 Production by a Hyperthermophilic Methanogen under High-Pressure Cultivation." *Proceedings of the National Academy of Sciences of the United States of America* 105 (2008): 10949–54.

Tan BF, Foght J. "Draft Genome Sequences of Campylobacterales (Epsilonproteobacteria) Obtained from Methanogenic Oil Sands Tailings Pond Metagenomes." *Genome Announcements* 2 (2014): e01034-14.

Timko MT, Nicholson BF, Steinfeld JI, Smith KA, Tester JW. "Partition Coefficients of Organic Solutes between Supercritical Carbon Dioxide and Water: Experimental Measurements and Empirical Correlations." *Journal of Chemical and Engineering Data* 49 (2004): 768-78.

Timóteo CG, Guilherme M, Penas D, Folgosa F, Tavares P, Pereira AS "Desulfovibrio Vulgaris Bacterioferritin Uses H2O2 as a Co-Substrate for Iron Oxidation and Reveals DPS-Like DNA Protection and Binding Activities." *Biochemical Journal* 446 (2012): 125-33.

Turner S, Pryer KM, Miao VPW, Palmer JD. "Investigating Deep Phylogenetic Relationships among Cyanobacteria and Plastids by Small Subunit rRNA Sequence Analysis." *Journal of Eukaryotic Microbiology* 46 (1999): 327-38.

Ulmer HM, Burger D, Ganzle MG, Engelhardt H, Vogel RF. "Effect of Compressed Gases on the High Pressure Inactivation of Lactobacillus Plantarum Tmw 1.460." *Progress in Biotechnology* 19 (2002): 317-24.

Van Ooij C, Losick R. "Subcellular Localization of a Small Sporulation Protein in Bacillus Subtilis." *Journal of Bacteriology* 185 (2003): 1391-98.

Vane L. "Separation Technologies for the Recovery and Dehydration of Alcohols from Fermentation Broths." *Biofuels, Bioproducts and Biorefining* 2 (2008): 553-88.

Vary JC. "Germination of Bacillus Megaterium Spores after Various Extraction Procedures. Journal of Bacteriology." *Journal of Bacteriology* 95 (1973): 1327-34.

Vary PS, Biedendieck R, Fuerch T, Meinhardt F, Rohde M, Deckwer WD, Jahn D. "Prime Time for Bacillus Megaterium." *Microbiology* 140 (1994): 1001-13.

Vary PS, Biedendieck R, Fuerch T, Meinhardt F, Rohde M, Deckwer WD, Jahn D.. "Bacillus Megaterium--from Simple Soil Bacterium to Industrial Protein Production Host." *Applied Microbiology and Biotechnology* 76 (2007): 957-67.

Wang Q, Garrity GM, Tiedje JM, Cole JR. "Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy." *Applied and Environmental Microbiology* 73 (2007): 5261-67.

Wei J, Shah IM, Ghosh S, Dworkin J, Hoover DG, Setlow P. "Superdormant Spores of Bacillus Species Germinate Normally with High Pressure, Peptidoglycan Fragments, and Bryostatin." *Journal of Bacteriology* 192 (2010): 1455-58.

West JM, McKinley IG, Palumbo-Roe B, Rochelle CA. "Potential Impact of CO2 Storage on Subsurface Microbial Ecosystems and Implications for Groundwater Quality." *Energy Procedia* 4 (2011): 3163-70.

White A, Burns D, Christensen TW. "Effective Terminal Sterilization Using Supercritical Carbon Dioxide." *Journal of Biotechnology* 123 (2006): 504-15.

Wieser M, Fujii N, Yoshida T, Nagasawa T. "Carbon Dioxide Fixation by Reversible Pyrrole-2-Carboxylate Decarboxylase from Bacillus Megaterium Pyr2910." *Journal of Molecular Catalysis B: Enzymatic* 257 (1998): 495-99.

Wimmer Z, Zarevúcka M. "A Review on the Effects of Supercritical Carbon Dioxide on Enzyme Activity." *International Journal of Systematic and Molecular Sciences* 11, no. 1 (2010): 233-53.

Wycherley H, Fleet A, Shaw H, Wilkinson J. "Origins of Large Volumes of Carbon Dioxide Accumulations in Sedimentary Basins." In *Geofluids II '97 Extended Abstracts*, edited by Carey P Hendry J, Parnell J, Ruffell A, Worden RH, 264-68. Belfast, Northern Ireland, 1997.

Xiong ZQ, Wang JF, Hao YY, Wang Y. "Recent Advances in the Discovery and Development of Marine Microbial Natural Products." *Marine Drugs* 11 (2013): 700-17.

Yanagawa K, Morono Y, De Beer D, Haeckel M, Sunamura M, Futigami T, Hoshino T, Terada H, Nakamura KI, Urabe T, Rehder G, Boetius A, Inagaki F. "Metabolically Active Microbial Communities in Marine Sediment under High-CO2 and Low-pH Extremes." *ISME J* 7 (2012): 555-67.

Yang X, Wang S, Zhou L. "Effect of Carbon Source, C/N Ratio, Nitrate and Dissolved Oxygen Concentration on Nitrite and Ammonium Production from Denitrification Process by Pseudomonas Stutzeri D6." *Bioresource Technology* 104 (2012): 65-72.

Yardley BWD. *An Introduction to Metamorphic Petrology.* Longman Earth Science Series. Prentice Hall, 1989.

Yoshida M, Hara N, Okuyama S. "Catalytic Production of Urethanes from Amines and Alkyl Halides in Supercritical Carbon Dioxide." *Chemical Communications* (2000): 151-52.

Yu ZC, Zhao DL, Ran LY, Mi ZH, Wu ZY, Pang X, Zhang XY, Su HN, Shi M, Song XY, Xie BB, Qin QL, Zhou BC, Chen XL, Zhang YZ. "Development of a Genetic System for the Deep-Sea Psychrophilic Bacterium pseudoalteromonas Sp. Sm9913." *Microbial Cell Factories* 13 (2014): 1-9.

Zamarreno DV, Inkpen R, May E. "Carbonate Crystals Precipitated by Freshwater Bacteria and Their Use as a Limestone Consolidant." *Applied and Environmental Microbiology* 75 (2009): 5981-90.

Zhang F, She YH, Chai LJ, Banat IM, Zhang XT, Shu FC, Wang ZL, Yu LJ, Hou DJ. "Microbial Diversity in Long-Term Water-Flooded Oil Reservoirs with Different in Situ Temperatures in China." *Scientific Reports* 2 (2012): 1-10.

Zhang J, Davis TA, Matthews MA, Drews MJ, LaBerge M, An YH. "Sterilization Using High-Pressure Carbon Dioxide." *The Journal of Supercritical Fluids* 38 (2006): 354-72.

Zhuang WQ, Yi S, Bill M, Brisson VL, Feng X, Men Y, Conrad ME, Tang YJ, Alvarez-Cohen L. "Incomplete Wood-Ljungdahl Pathway Facilitates One-Carbon Metabolism in Organohalide-Respiring Dehalococcoides Mccartyi." *Proceedings of the National Academy of Sciences of the United States of America* 111, no. 17 (2014): 6419-24.