Metabolic versatility and interactions of nitrogen cycling microbiomes

By

Christopher Evan Lawson

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Civil and Environmental Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2019

Date of final oral examination: 12/09/2019

This dissertation is approved by the following members of the Final Oral Committee:
Katherine D. McMahon, Professor, Civil and Environmental Engineering and Bacteriology
Daniel R. Noguera, Professor, Civil and Environmental Engineering
Daniel Amador-Noguez, Assistant Professor, Bacteriology
Jennifer L. Reed, Professor, Chemical and Biological Engineering
Christina K. Remucal, Associate Professor, Civil and Environmental Engineering
Sebastian Lücker, Assistant Professor, Microbiology, Radboud University Nijmegen (External)

To my wife, Christine Tam.

Thanks for coming with me on this adventure.

Abstract

Metabolic versatility and interactions of nitrogen cycling microbiomes

By Christopher E. Lawson

The catalytic capabilities of microbial communities ("microbiomes") seem limitless, controlling Earth's biogeochemical cycles and occupying every environmental niche. Engineers have tapped into these capabilities for centuries by harnessing microbiomes to perform important services for society. One important service is the removal of nitrogen from wastewater via engineered biological processes, such as partial nitritation-anammox (PNA). In these systems, a diverse group of nitrogen cycling bacteria remove ammonium from wastewater as dinitrogen gas through their combine metabolic activities that requires precise balancing. However, the metabolic networks and interactions of nitrogen cycling microbiomes remain poorly characterized, limiting the opportunity to improve nitrogen removal biotechnologies using a systems biology approach. This thesis aimed to understanding the metabolic versatility and interactions occurring in nitrogen cycling microbiomes and to develop new computational models that enable their prediction in natural and engineered ecosystems.

In Chapter 1, we provide an overview of the microbial nitrogen cycling network, highlighting key functional guilds, their ecophysiology and interactions, and their application for wastewater treatment. We also outline the systems biology methodology that was used throughout this thesis work. In Chapter 2, we reconstructed the metabolism and interactions of poorly characterized heterotrophic organisms in anammox bioreactors based on metagenomic and metatranscriptomic analysis. This revealed that most heterotrophs were denitrifying bacteria that

exchange carbon and nitrogen substrates with anammox bacteria, possibly allowing for improved nitrogen removal efficiency. In Chapters 3 and 4, we applied cutting-edge metabolomic tools, ^{13}C fluxomics, and metabolic modeling to elucidate the autotrophic and mixotrophic metabolic networks operating in two key nitrogen cycling bacteria: the anammox bacterium *Candidatus* 'Kuenenia stuttgartiensis' and the nitrite-oxidizing bacterium *Nitrospira moscoviensis*. This provided detailed insights on the metabolic networks underlying their versatility and represented the first measurements of metabolic flux in nitrogen cycling microorganism. In Chapter 5, we extended insights from these metabolic networks to build genome-scale models for the anammox bacterium *Brocadia sinica* and the comammox bacterium *Nitrospira nitrosa* to predict their interactions under different environmental conditions. Our simulations predicted novel mechanisms driving the cooperation and competition between comammox and anammox bacteria, which could inform strategies to improve the success of mainstream PNA processes. In Chapter 6, we synthesize common principles and best practices for harnessing microbiomes into a design-build-test-learn (DBTL) cycle that can be used to advance microbiome engineering across disciplines. The cycle outlines top-down and bottom-up design processes, synthetic and self-assembled construction methods, and emerging tools to analyze microbiome function that can be used to improving human and animal health, agriculture and enable the bioeconomy.

The knowledge and tools generated in these studies improves our ability to understand the functions and interactions occurring in nitrogen cycling microbiomes. Moreover, our findings represent the first steps in creating a systems biology approach for the bottom-up prediction of ecosystem function. From our results, we suggest future research needed to extend this systems biology approach for complete prediction and control of nitrogen cycling microbiomes.

Acknowledgements

Obtaining a PhD is not possible without the support of an entire team. For my PhD, it turned out to be two. The list of people that have helped me throughout this journey is probably longer than this thesis itself. I would like to acknowledge some of those people here. For those who I have forgotten to mention by name, thank you for making my time in Madison and the Netherlands such an incredible experience.

My deepest thanks go to my PhD advisors, Trina and Dan. I first approached Trina to pursue a PhD because I saw her as the undisputed "Queen" of biological phosphorus removal and the leading scientist bridging microbial ecology and water engineering. Over the past five years, I've learned that in addition to these well-deserved titles, Trina is also one of the most humble, considerate, and fun people I'll probably ever meet. Through her actions, she has shown me what it takes to be a leader in science and always provided her invisible hand to help me navigate and overcome the different challenges I faced as a graduate student. Thank you for your unwavering encouragement and support in pursuing my personal goals and interests as a scientist. Your belief in me and open-minded nature offered me the opportunity to explore these goals and interests to the fullest, which I am forever grateful for.

During my interview in Madison, I was repeatedly told by different faculty and students that "you have to speak with Dan Noguera!" and was fortunate enough to get an opportunity to do so. From our first conversation and email string (Oct 27th, 2013), Dan has been a key source of inspiration, critical feedback, pragmatism, reflection, and encouragement on a shared vision to integrate systems and synthetic biology approaches into microbial ecology and the engineering of microbial communities ("microbiome engineering"). During my five years in Madison, we iterated over a common vision for how this could look like, which (with the help of others!) ultimately

turned into an important Chapter in this thesis. To this day, you always keep me on my toes with your thoughtful (and detailed) feedback and on-point questions, which has brought the best out of our work together and my personal development as a critical thinker. I look forward to further refining our shared vision for microbiome engineering (together with Trina!) moving forward.

I would also like to thank all the great people in the academic community at UW-Madison. A big thanks goes to Christy, Jennie, and Daniel for serving on my committee and providing me with their advice, teaching, and expertise during my time in Madison. In particular, I would like to extend a special thank you to Daniel for "adopting" me into his lab and providing me the opportunity to learn about metabolomics. The time and training your offered me to integrate metabolomics into my research has made a huge impact on this thesis work and my future plans as a scientist. Thank you to all the folks in the "Dan Lab" who also tolerated me running around the lab during this time and were kind enough to fast-track my learning curve. In particular, I'd like to thank Tyler and Dave for answering my million questions, providing me with hands on training with the Q-Exactive and ^{13}C tracer experiments, and for running samples for me while I was away in the Netherlands (even over Christmas break!).

Special thanks to all the people involved in the CEE Capstone Course, including Tom, Michael, Greg, Charlie, and Mark, for offering me the chance to participate as a mentor for the past 5 years. It has been a wonderful outlet for staying involved in the practical nature of our work, while being inspired by the work of so many bright undergraduates in the program. In particular, I wish to thank Tom for having me help out on all the exciting projects he brings in. You were a knowledgeable, fun, and generous mentor not only to the students, but to me as well.

My deepest appreciation also goes to all my fellow McMahonites and Noguerians, past and present. From the day I stepped foot in Madison, I have been spoiled with the best team and group

of friends I could have ever imagined. Starting with the McMahonites - Ben (the original!), I think you were the first person to really make me feel at home in Madison. Whether it was walking Dukesy, sippin lattes, sharing dinner, or discussing your endless pit of creative ideas, I always enjoyed my time with you. Let me know when we should get those sludge brother tattoos. Travis, thanks for living with me for two years and not killing me. I miss our days at the Main Street place enjoying pool parties, your Jamaican ox tail and beans, and watching PBS news weekend specials. Thanks for always having my back and I'm excited to visit you and Abby in Boston soon. Panchino and Pame! You were always two of the most fun people to be around. Thanks for the positive vibes in the lab, introducing Christine and I to all the wonderful Chilean people in Madison, and for having us out at your beautiful wedding in Chile. Pame, thanks for also letting us borrow Panchino for a while, he did wonders with the cats! Ben (the one and only, aka BenP), you have been quite the successor of the original Ben. We must get along so well because of those luscious locks you have. Thanks for being such a great sounding board in the office and introducing me to your better half Hanna. Joshy boy! Thanks for introducing me to metabolic modeling and also letting me bug you in the office. Looking forward to catching up in Cali! Coty, watching you grow from a smart and inquisitive freshman into a card-carrying scientist has been a highlight of my PhD. Keep crushing at life and I look forward to seeing what you accomplish next! Elizabeth and Amber, the dynamic duo, thanks for always keeping things fun in the lab and enjoy playing with activated sludge and lakes. Patricia, thanks for improving our office with your presence. It's been nice having another Canadian around! Rachel, like Matt I too was surprised you came after we were your tour guides. It's been awesome watching you wrangle the pilot plant and plan your research, best of luck! Robin (aka the Ice Queen) thanks for keep the freezers in tip top shape and always being down for a solid bourbon hour discussion. I wish I could have attended the wedding but it's

happening as I write this sentence! Charlie, although freezers may not be your forte, you seem to have a green thumb for cultivation. It's been awesome having you in the lab and good luck isolating those EETers! Diana and Krysten, you have been amazing additions to the lab. I wish I had more time to get to know you, but I'm glad to see the lab's in good hands. To all past members of the McMahon lab – Sarah, Alex, Christina, Grace, Diego, Andrew, Shaomei – thanks for helping me with computer problems, math problems, and for great discussions over bourbon. You are definitely missed!

On to the Noguerians – Distinguished Professor Scarborough (that's how you wanted to be addressed now, right?) I probably like you so much because you have a Canadian in the family, enjoy hockey, and occasionally offer me fine whisky. Although only starting 6 months before me, you've always been two (maybe three) steps ahead. Thanks for all the support, encouragement, and advice during my time in Madison. Looking forward to working together more and seeing you, Catlin, Claire, and my boy Arlo in Vermont soon! Zach, thanks for trying to teach me how to play frisbee and letting me live in your place. I miss our days together watching The Get Down and talking science. See you in Vermont too! Natalie, thanks for always being so positive, optimistic, and fun. I'm sure you're already schooling them in "real" world out in Colorado and I'm excited to visit soon. Duygu, it was a blast having you in the lab and even better getting to know you out of it. Hope you and Nelis can come visit up in Canada once we make the move back. Miguel, it was great discussing agriculture, waste management, mass spec, or really any subject with you, especially over a beer. I hope one day I get to visit you on your farm in Patagonia. Abel, your positive vibes and chill demeanor have been a great addition to the lab. Looking forward to seeing what you end up turning that manure into! Kevin, thanks for attending my mock teaching session and for inspiring me with your new and improved MCFA reactors. Big thanks also to

Jackie Cooper for keeping things running smoothly at EHall. I'll miss dropping in to chat in your office! And to all the other Noguerians, past and present, including Dan G, Brian, Natalie C, Ming'er, Weiping, Yanjun, Nathan, Rania, and, Kevin M – it was a pleasure working with you. Keep treating wastes and making fuels!

During the first summer of my graduate studies (June 2015), I had the opportunity to attend the 4th International Conference on Nitrification in my hometown (thanks Lisa Stein for allowing me to register late!) where I met then postdoc Sebastian Lücker. Attending that conference and drinking beers with Sebastian et al turned out to be one of the most fortuitous moves of my graduate career. Thank you Sebastian for providing me the opportunity to collaborate with you over the past 4 years and for being so involved my PhD research. Your seemingly endless molecular insights and dedication to mentoring me throughout my PhD has greatly advanced my understanding and appreciation of microbiology. Looking forward to unraveling the mysteries of nitrogen cycling microbes with you for years to come.

While in Nijmegen I also had the privilege of getting to know the professor-in-charge, Mike Jetten. Mike, I greatly enjoyed our many discussions together about anammox physiology and science in general, whether over a nice European beer or in the lab. Thank you for hosting me in Nijmegen for two visits and offering me your mentorship, generosity, and time. If I can build a lab with even half the bioreactors as yours, I'll considerate it a success! Looking forward to more visits and discussions over beers in the future.

My time in Nijmegen also allowed me to interact and learn from some of the world's top microbiology talent. Rob de Graaf, what can I say, ever since I meet you I knew we'd get along. Seeing and learning from the way you work in the lab has taught me to be a more methodical scientist. I've greatly appreciated your friendship and am excited to see what other anammox

mysteries we can uncover together. Guyliane! In addition to teaching me how to run an anammox reactor without them dying you've played a big role in helping carryout our experimental studies. Thanks for the fun times both inside and out of the lab and let's keep it going. Aniela, you have also been an amazing and patient teacher of growing stubborn microbes in the lab. It was a pleasure collaborating with you on our *Nitrospira* work and I hope we will live to see the day where they grow above an optical density of 0.065. Hanna and Laura, it was great having you help out with the metabolomic experiments as well. Enjoy exploring the unknown metabolism of *Nitrospira,* can't wait to see what you uncover. Maartje, thank you for introducing me to your comammox-anammox enrichment culture and for showing me the ropes of ^{15}N stable isotope tracing and GC-MS analysis. We will eventually get that culture to also like ^{13}C tracers and I'm excited to continue working with you on this. Arjan, although we didn't work on any specific project together, I think I bugged you almost every single day for help with the reactors, my experiments, the GC-MS, the gas board, and anything else I couldn't figure out. Thank you for being patient with me and for always offering your help and expertise when I needed it. Theo, Geert, and Jeroen - thank you for all the sequencing and bioinformatic support during my visits. You were a pleasure to work with and I enjoyed our discussions together. Special thanks to Huub, Katinka, Marjan, Rob M., Peter, Sebastian K., Paul, Harry, Suzanne, Laura, Cornielia, Paula, Michiel, Carmen, Tijs, Stijn, Linnea, Wouter L, Judith, Francisca, and Federica. You made me feel at home in the lab and I enjoyed chatting with you in the coffee corner. Last but not least, a huge thank you to the "movie nights crew" and all the other people I enjoyed life with outside of the lab too, including Mo, Dimitra, Guylaine, Arslan, Eric, Wouter, Julia, Christina, Ana, Nunzia, Aniela, Lianna, Anniek, Jeroen, and Stefanie. It was a pleasure having you as friends and please keep in touch!

During my trip to the Netherlands I was fortunate to also collaborate with the brilliant people in the Department of Biotechnology at Delft University of Technology. I would like to sincerely thanks Mark, Robbert, and Martin for hosting me in their labs and offering their time, expertise, and resources to help study the ecophysiology of anammox bacteria. This could not have been done without the day-to-day collaboration with Michele, Aina, and Gerben (i.e. team AinaMox!). It was a pleasure running the parallel bioreactors experiments with you (even at 2 am!) and I look forward to wrapping up our study. A big shout out also goes to David. Since we first met at WEFTEC in 2015 we have been organizing on how to bridge microbial ecology and engineering. Thanks for getting me involved in the several workshops we've run together and collaborating on the perspective article. Looking forward to our future efforts to bring these two disciplines together.

Now for my family. I have many people to thank for instilling in me what it takes to persevere through a PhD. I would like to thank my grandmother Helen for showing me the value of hard work and optimism from a young age. Thanks for always being my number one cheerleader, encouraging me to follow my passion, and enthusiastically letting me show you my research work year after year. To my aunty Chrissy and uncle Tino, thanks for always supporting me over the years and letting me know that I always have a place back home. I've appreciated you always checking in on me and for supporting my dreams even when they meant moving far away. To my uncle Dave and uncle Paddy, you have always played such a huge role in my development and supported my passions now and in all chapters of my life. I hope one day I can return the favor. To my aunty Ade, thanks for being a constant source of encouragement and support, and for staying in touch while I've been away. To my brother Elliott, thanks for always knowing more about the political and social implications of my work and for bringing me up to speed on it.

Staying in touch and getting your songs of the day have been a nice pick me up when I've needed it. To my sister Alicia and family Jeff, Kai, and Teagan, thanks for never forgetting to keep in touch over all these years, it's been amazing watching you grow as a family. Alicia, I've also appreciated you getting me into this genealogy hobby of ours, it's been a fun outlet! To my (twin) brother Keith, thanks for paving the way and inspiring me to tackle the world. You've been an amazing source of guidance and support throughout my studies, and I'm happy we can talk about science until we're old. Looking forward to spending more time with you and Ambica in Toronto.

To my beautiful and intelligent wife, Christine, where do I even begin. They say that behind every man is a greater woman, and for us that couldn't be more true. Since we met more than seven years ago now in Marine Drive you've been an endless source of support for all my passions and dreams. While at UBC, you were the first person to encourage me to pursue a PhD, despite knowing it would mean we may have to be apart. You spent four years funding our back and forth trips between Toronto and Madison so we could see each other before you obtained a visa. And then, you supported my biggest dream of all by saying yes to marrying me. This thesis would not have been possible without you. Thank you for sharing this amazing journey with me and I'm excited to start our next adventures together in California and then Toronto.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.0 Motivation

Population growth and urbanization have created an imbalance in the nitrogen cycle. The discharge of excess nitrogen to the environment has produced hundreds of oxygen deplete "dead zones" and toxic algal blooms in coastal and freshwater bodies. The problem has become so serious that the National Academy of Engineering has listed restoring balance to the nitrogen cycle as a grand challenge for the 21st century. While untreated wastewater is a major source of nitrogen inputs to watersheds, energy and sludge disposal costs associated with nitrogen removal at wastewater treatment plants (WWTPs) greatly increases total operating costs. Microbial communities ("microbiomes") containing anaerobic ammonium-oxidizing (anammox) bacteria represent a promising option for cost-effective and energy-efficient nitrogen removal from wastewater[1]. However, the underlying metabolic networks responsible for nitrogen removal by these microbiomes remain poorly understood, limiting their application for sustainable wastewater treatment. This work aimed to understand the metabolic networks driving the function of nitrogen cycling microbiomes and to create quantitative models that predict their function in natural and engineered ecosystems.

1.1 Harnessing the microbial nitrogen-cycling network

Microorganisms are the gate keepers of the biogeochemical nitrogen cycle, controlling nitrogen

---

Aspects of this introduction have been published and are provided in appendix A as:

Lawson, C. E. & Lucker, S. Complete ammonia oxidation: an important control on nitrification in engineered ecosystems? *Curr. Opin. Biotechnol.* **50**, 158–165 (2018).

processes and fluxes at rates of tetragrams ($1x10^{12}$ grams) of nitrogen per year within terrestrial and marine ecosystems alone[2]. These microorganisms do not exist in isolation, but instead form complex microbiomes that arise from interactions between co-occurring microorganisms to transform nitrogen between its various organic and inorganic forms (Figure 1.1). Engineers have tapped into the unique ability of microorganisms to transform nitrogen compounds for over a century. For example, the microbial conversion of ammonium to nitrate (or "nitrification") has been harnessed for wastewater treatment to prevent fish deaths associated with toxic ammonia releases and the depletion of oxygen in receiving waters[3]. Additionally, the coupling of nitrification with denitrification has been harnessed to prevent algae blooms associated with excess nutrient discharges to aquatic environments[4].

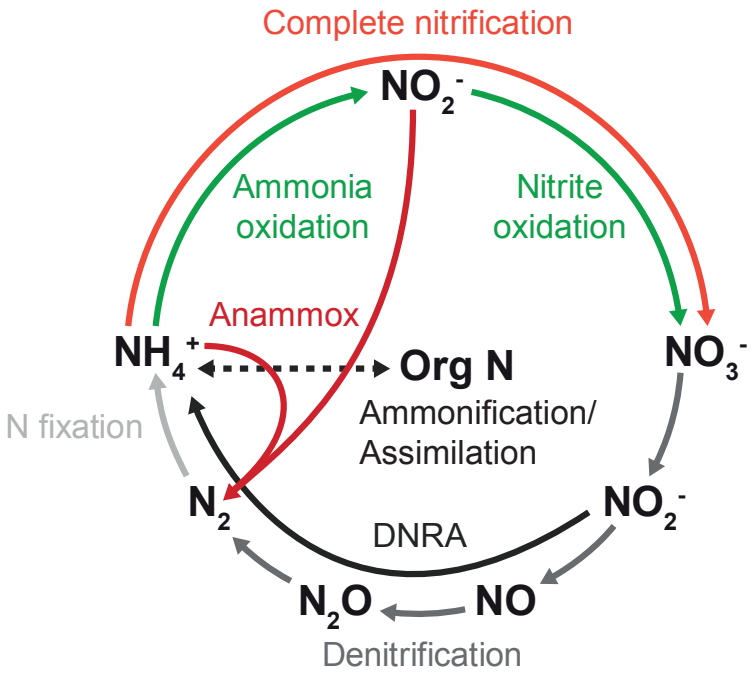

Figure 1.1 The biogeochemical nitrogen cycle mediated by microorganisms. NH_4^+, ammonium; NO_2^-, nitrite; NO_3^-, nitrate; NO, nitric oxide; N_2O, nitrous oxide; N_2, dinitrogen gas; Org N, organic nitrogen.

Although the microbiomes that mediate these processes are highly diverse, a core group of abundant microorganisms can be commonly detected within them[5]. This includes key ammonia oxidizing bacteria (AOB) and nitrite oxidizing bacteria (NOB), such as *Nitrosomonas* spp. and *Nitrospira* spp.[6,7], respectively, as well as a diverse group of denitrifying bacteria, such as members of the Betaproteobacteria[8,9]. Decades of studies on the microbial ecology of nitrogen cycling microbiomes have contributed to our understanding of biological wastewater treatment systems[10, 11,12]. However, the complexity of these microbiomes and the lack of quantitative tools available to investigate and predict their *in situ* metabolic processes has limited the translation of this knowledge to bioprocess engineering and design.

Recently, new discoveries in the nitrogen cycle have coincided with a transition in the wastewater treatment industry towards energy neutrality and resource recovery. In the early 1990's, bacteria that could anaerobically oxidize ammonium to nitrogen gas using nitrite as a terminal electron acceptor (i.e. anammox bacteria) were discovered in a denitrifying reactor[13]. These organisms were previously predicted to exist based on thermodynamic grounds[14], and have now been shown to account for up to 50% of the total nitrogen turnover in the environment[2,15]. Since their discovery, anammox bacteria have been applied for wastewater treatment to achieve more sustainable nitrogen removal, representing one of the most rapid biotechnological advances in wastewater treatment[1,15,16].

Anammox bacteria are currently applied for the sidestream treatment of high strength ammonium wastewaters at mesophilic temperatures (typically exceeding 500 mg NH_3-N/L), such as anaerobic digester centrates that account for ~20-50% of the total nitrogen loading at a WWTP, depending on industrial inputs (e.g. co-digestion of food waste)[17]. Because ammonium is a dominant form of nitrogen in most anaerobic digester centrates, applying anammox typically

requires a nitritation step, where 50% of the influent ammonium is converted to nitrite by aerobic AOB, followed by removal of the remaining ammonium and nitrite to nitrogen gas by anammox bacteria (Figure 1.2)[18,19]. The use of this "partial nitritation-anammox" (PNA) process is more sustainable and cost-effective conventional nitrification-denitrification processes because it reduces oxygen requirements and sludge production by approximately 60% and 90% respectively[15,18,20]. Moreover, PNA processes save additional costs because they require no organic carbon inputs for denitrification, which can instead be used to recover energy as biogas via anaerobic digestion[1].

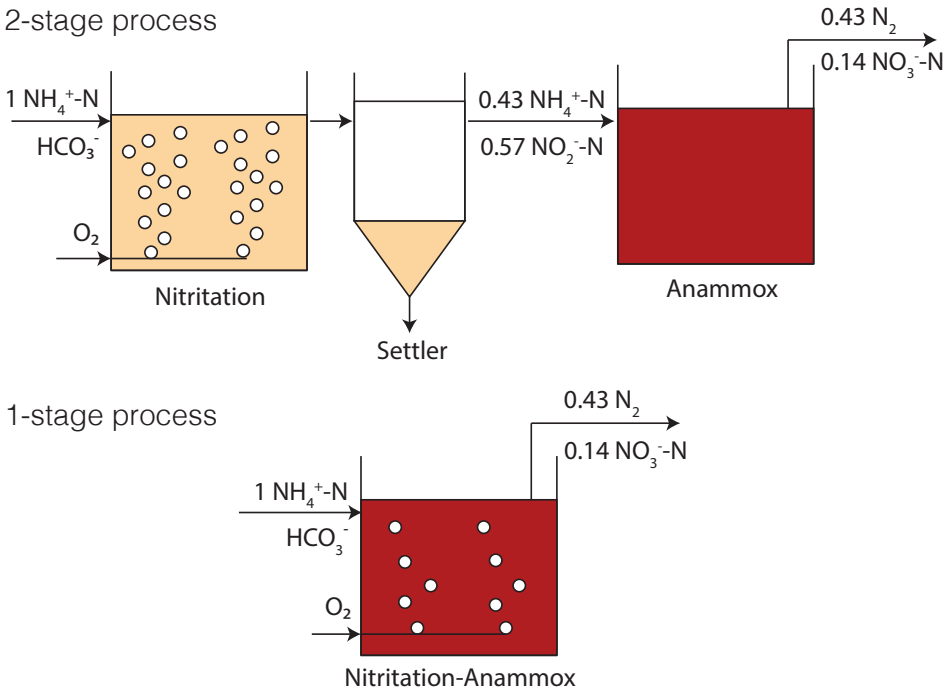

Figure 1.2 Flow diagrams for the partial nitritation-anammox (PNA) process. (Top) Two-stage PNA process with separate reactors for nitritation and anammox. The nitritation reactor receives air and is kept at a low solids retention time (SRT ~1 day) to washout NOB that grow slower at mesophilic temperatures (~35°C). The anammox reactor is kept anaerobic and maintains a long SRT (20+ days) to allow growth of anammox bacteria. (Bottom) Single-stage PNA with nitritation and anammox occurring in the same reactor. Oxygen is maintained at a low flux or pulse aerated to balance nitritation activity by AOB with anammox. Flocs containing low anammox biomass are removed from the reactor by hydrocyclones, sieves, or plate settlers, while slow growing anammox granules are retained.

Application of anammox bacteria to remove 100% of the influent ammonium from wastewater (i.e. mainstream treatment) using PNA has had much less success than sidestream applications. This is partly because the low specific growth and activity rates of anammox bacteria are further reduced at lower mainstream wastewater temperatures (12-25°C), and competition for nitrite with denitrifying bacteria and NOB is harder to control because of higher organic carbon loadings and lower concentration of ammonia (30-50 mg NH_3-N/L)[21]. While some of these challenges can be overcome by process design, for example by including a pretreatment process ("A-stage") for organic carbon recovery prior to nitrogen removal[21], deeper understanding of the metabolic interactions occurring in nitrogen cycling microbiomes under mainstream conditions is likely required to improve process performance.

In addition to anammox bacteria, another major breakthrough in our understanding of the nitrogen cycle has been the recent discovery of complete ammonia-oxidizing ("comammox") bacteria[22,23]. For over a century, nitrification had always been viewed as a two-step process mediated by interactions between AOB and NOB[24] until bacteria affiliated with the genus *Nitrospira* were shown to be capable of complete nitrification and contained pathways for both ammonia and nitrite oxidation[22,23]. Since their discovery, comammox bacteria have been detected in a variety of different habitats, including wastewater treatment reactors[22,25,26]. Because these organisms can use both ammonia and nitrite as substrates for energy generation[22,27], they are expected to be competitors of anammox bacteria. However, van Kessel et al. (2015)[23] have shown that comammox *Nitrospira* and anammox bacteria affiliated with *Brocadia* can co-occur in biomass flocs under low dissolved oxygen concentrations, suggesting they may also cooperate. Whether the control of these interactions could be harnessed for improved nitrogen removal from wastewater has yet to be explored.

1.2 Ecophysiology of nitrogen cycling bacteria

Key to improving the systematic engineering of nitrogen removal biotechnology is understanding the ecophysiology of the major functional guilds controlling nitrogen cycling. This includes improved understanding of their carbon and energy metabolism, metabolic versatility, species diversity, physiological and kinetic properties, environmental controls, and metabolic interactions. A summary of the key ecophysiological properties of functional guilds present in nitrogen removal biotechnologies can be found in Table 1.1 and are described below.

1.2.1 Ecophysiology of ammonia-oxidizing prokaryotes (AOP)

The nitrifying guilds, including the AOB, ammonia-oxidizing archaea (AOA), NOB, and comammox bacteria, are responsible for nitrification in a wide range of natural and engineered ecosystems. In general, organisms within these guilds are chemolithoautotrophs that use carbon dioxide as a carbon source and either ammonia or nitrite as an energy source. They also typically require oxygen as a terminal electron acceptor for energy conservation.

AOB perform the first step in nitrification, oxidizing ammonia to nitrite. A diverse group of AOB species have been identified in wastewater treatment systems, including *Nitrosomonas spp.* and *Nitrosococcus spp.*[28,29]. The catabolism of AOB is driven by sequential enzymatic steps. Ammonia is first oxidized to hydroxylamine via an endergonic reaction catalyzed by ammonia monooxygenase (AMO). Hydroxylamine was then believed to be further oxidized to nitrite via the exergonic reaction catalyzed by hydroxylamine oxidoreductase (HAO), allowing the process to be net energy conserving. However, it has recently been shown that the product of HAO is not nitrite, but nitric oxide, and that an unknown enzyme catalyzes nitric oxide oxidation to nitrite[42].

Table 1.1 Ecophysiological properties of selected chemolithoautotrophic nitrogen cycling microorganisms commonly found in wastewater treatment and freshwater systems

Parameter	*Nitrosomonas europaea*	*Nitrosocosmicus exaquare*	*Nitrospira moscoviensis*	*Nitrospira defluvii*	*Nitrotoga fabula*	*Nitrospira inopinata*	*Kuenenia stuttgartiensis*	*Brocadia sinica*
Functional guild	AOB	AOA	NOB	NOB	NOB	Comammox	Anammox	Anammox
Habitat	Wastewater	Wastewater	Freshwater	Wastewater	Wastewater	Freshwater	Wastewater	Wastewater
Optimal temp (°C)	20-30	33	39	28	24-28	37	30	30
Optimal pH	6.0-9.0	~8.0	7.6-8.0	7.3	7.1 - 7.6	7.6	6.5-9.0	7.0-8.8
Km, NH_3 (µM)	1.9-200	n.d.	n.d.	n.d.	n.d.	0.049	n.d.	28
Ks, NO_2^- (µM)		n.d.	18	48	89	372	0.2-3	34
Ko (µM)	1-40	n.d.	4.06-16.88	n.d.	n.d.	n.d.	n.a.	n.a.
Ammonium tolerance (mM)	>500	15	n.d.	n.d.	n.d.	n.d.	>500	>500
Nitrite tolerance (mM)	5-20	15	15	20-25	30	n.d.	13	<16
Oxygen tolerance (µM)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	0-200	<63
µmax (hrs-1)	0.028-0.068	n.d.	0.022	0.019	0.007**	0.0061	0.0035	0.0041
Y_{bio} (mg protein/mol NH_3)	250	n.d.	213#	122#	n.d.	395	n.d.	849
CO2 fixation pathway	CBB	3HP/4HB	rTCA	rTCA	CBB	rTCA	WLP	WLP
Formate oxidation rate (µmol g-protein^{-1} min^{-1})	+	*	+	*	n.d.	*	5.8	6.5
Acetate oxidation rate (µmol g-protein^{-1} min^{-1})	-	n.d.	-	*	n.d.	*	0.31	0.57
Hydrogen oxidation rate (µmol g-protein^{-1} min^{-1})	+	n.d.	+	*	n.d.	*	n.d.	n.d.
oxygen consumption (mol O2:NH3)	1.5	1.5	0.5	0.5	0.5	2	0	0
references	30,31, 32, 33, 34,35	40	36,37	36,37,38	39	27	41	41

#Y_{bio} based on mol nitrite consumed, + activity present, - activity absent, *genomic potential present, **growth rate estimated from *Vmax* of 28 µmol nitrite mg-protein^{-1} hr^{-1}

In addition to their catabolism, genomic studies have predicted that many AOB use the Calvin-Benson-Bassham (CBB) cycle for CO_2 fixation[43,44,45]. The key enzyme of this cycle, ribulose 1,5 bisphosphate carboxylase (RuBisCO), is suggested to be the most abundant protein on Earth[46], which highlights the widespread use of this pathway for CO_2 fixation. Reducing equivalents (i.e. nicotinamide adenine dinucleotide phosphate, NADPH) for CO_2 fixation are suggested to come from "reverse electron flow" through nicotinamide adenine dinucleotide (NADH) dehydrogenase complex I. This complex consumes energy conserved from ammonia oxidation as a proton motive force to catalyze the unfavorable transfer of electrons from the quinone pool to NAD+[47].

AOA also perform the first step in nitrification. These organisms belong to the phylum Thaumarchaeota and have been shown to play important roles in nitrification in soil and marine ecosystems[48], as well as some wastewater treatment systems[40]. AOA use AMO to oxidize ammonia to hydroxylamine, but do not encode a known HAO enzyme[48]. It has been proposed that hydroxylamine oxidation to nitrite is mediated by reaction of hydroxylamine with nitric oxide catalyzed by a heme-containing cytochrome P460 enzyme and nitrite reductase[49], however this remains to be confirmed. CO_2 fixation by AOA has also been shown to be mediated by unusual biochemistry. AOA have been demonstrated to use a novel variant of the hydroxypropionate/hydroxybutyrate cycle for fixing inorganic carbon[50]. This pathway is more energy efficient than the CBB cycle, which may help AOA thrive in nutrient-limit environments. Additionally, the ammonia oxidation kinetics of AOA suggest that some species may be adapted to oligotrophic nitrification, based on reported half-saturation constants and substrate thresholds. This likely also contributes to their niche differentiation[27,51].

1.2.2. Ecophysiology of nitrite-oxidizing bacteria (NOB)

NOB perform the second step in nitrification, oxidizing nitrite to nitrate. Common NOB detected in wastewater treatment systems have been affiliated with the genus *Nitrospira*[7], although *Nitrotoga* spp. have recently been identified as key NOB as well[39,52]. Nitrite oxidation is mediated by the nitrite oxidoreductase (NXR) enzyme complex. In *Nitrospira* and *Nitrotoga* species, the catalytic NXR subunit is located in the periplasm [39,53], whereas in other NOB such as *Nitrobacter* spp., the catalytic subunit faces the cytoplasm. Differentiation of the NXR may affect competition for nitrite in the environment, as NOB with periplasmic types that do not require substrate transport into the cytoplasm are expected to have higher nitrite affinities[54]. Moreover, the periplasmic NXR is predicted to result in higher energy yields because the two scalar protons produced during nitrite oxidation can contribute to proton motive force generation[54]. This may additionally contribute to niche differentiation among NOB.

Pathways for CO_2 fixation by NOB also appear to be differentiated based on genomic evidence and lipid analysis. *Nitrospira* encode all genes for CO_2 fixation via the reductive tricarboxylic acid cycle (rTCA)[53], whereas other NOB such as *Nitrobacter* and *Nitrotoga* encode genes for the CBB cycle[39,44]. As the rTCA cycle is more energy efficient than the CBB cycle[55], this may allow *Nitrospira* spp. to achieve higher growth yields compared to other NOB.

NOB such as *Nitrospria* can exhibit a high degree of functional diversity among closely related species[56]. Outside their chemolithoautotrophic growth, genomic and experimental data have revealed that *Nitrospira* can use alternative substrates to fuel their carbon and energy metabolism[54,57,58], including the oxidation of formate and hydrogen with oxygen or nitrate as terminal electron acceptors[57,58,59]. Genomic analysis also suggests that *Nitrotoga*-affiliated NOB may be capable of hydrogen and sulfite oxidation[39]. This highlights the potential versatility of

NOB that may allow them to survive conditions of nitrite depletion and environmental fluctuations.

1.2.3. Ecophysiology of comammox bacteria

Comammox bacteria can perform both steps of nitrification in a single cell. To date, all identified comammox bacteria have been affiliated with the genus *Nitrospira* and contain a phylogenetically distinct AMO from canonical AOB and AOA that was initially misclassified as a particulate methane monooxygenase (pMMO)[22,23]. For this reason, established AMO polymerase chain reaction (PCR) primer sets had failed to detect their presence in the environment. New PCR-based and metagenomic surveys have shown that comammox *Nitrospira* are broadly distributed across different habitats, including wastewater treatment systems. These organisms contain both pathways for ammonia and nitrite oxidation in their genome, including genes for AMO, HAO, and NXR[22,23]. Like all *Nitrospira*, comammox encode genes for the reductive tricarboxylic acid (rTCA) cycle. Genomic information also indicates that several comammox *Nitrospira* species can use diverse substrates for energy conservation, including formate and hydrogen oxidation[60].

Kinetic theory of optimal pathway length predicted that comammox organisms would have low growth rates, high growth yields, and therefore be adapted to environments with low substrate fluxes[61]. This was recently confirmed based on kinetic analysis of a *Nitrospira inopinata* pure culture, which yielded apparent half-saturation constants for ammonia 4–2500 fold lower than any kinetically characterized AOB and non-marine AOA, and growth yields at least 30% greater than any examined AOP[27]. While these data represent yet only one comammox organism, it may suggest that ammonia availability is a key factor determining niche differentiation of comammox *Nitrospira* and other AOP.

Adaptation to microaerophilic environments is also reflected in comammox *Nitrospira's* electron transport chain, which employs novel cytochrome bd-like terminal oxidases, similar to other *Nitrospira* (described in detail for *Candidatus* "N. defluvii"[53]). These cytochrome c oxidases are speculated to have higher oxygen affinities compared to the low-affinity aa3-type heme-copper oxidases employed by most canonical AOP and may confer comammox *Nitrospira* a selective advantage in environments with low dissolved oxygen. Indeed, this agrees with observations from Camejo et al. (2017)[62], who enriched comammox *Nitrospira* in a low-dissolved oxygen bioreactor inoculated with activated sludge. However, it should be noted that these novel terminal oxidases have not been functionally characterized and both their ability to pump protons and their oxygen affinity remain to be determined experimentally.

1.2.4. Ecophysiology of denitrifying bacteria

Denitrifying bacteria catalyze the reduction of nitrate to dinitrogen gas via the intermediates nitrite, nitric oxide, and nitrous oxide. In wastewater treatment systems, denitrifying bacteria are almost exclusively heterotrophic, using organic substrates as both an electron donor and carbon source. However, autotrophic denitrifying bacteria that use inorganic electron donors (e.g. hydrogen sulfide, ferrous iron) for denitrification and CO_2 as a carbon source are prevalent in other environments[63,64] or wastewater treatment plants that contains suitable electron donors[65]. Denitrification is generally catalyzed by a diverse set of enzymes classified as nitrate reductases (NAR), nitrite reductases (NIR), nitric oxide reductases (NOR), and nitrous oxide reductases (NOS)[66]. For example, several types of NIR can exist, including the copper-dependent enzyme NirK and the cytochrome cd-dependent enzyme NirS that use one electron to reduce nitrite to nitrous oxide[67]. In many environments including wastewater treatment systems, the process

competes with bacteria capable of dissimilatory nitrate reduction to ammonia (DNRA) mediated by NAR and a cytochrome c nitrite reductase (ccNIR)[68].

While many aspects of the biochemistry involved in denitrification have been elucidated[66,68,69,70], the ecophysiology of denitrifying organisms is poorly understood. Many denitrifying organisms do not contain all enzymes involved in denitrification. Instead, denitrification is often segregated across several different species that require their interaction to mediate the entire process[71]. While the mechanisms behind this distributed metabolism are largely unknown, recent studies have shown that cross-feeding of nitrite between two denitrifying bacteria can eliminate intracellular enzyme competition between NAR and NIR for shared resources (e.g. NADH, periplasmic space) and accelerate substrate consumption when nitrite has growth inhibiting effects[72]. In addition to metabolic segregation, the high diversity of organic carbon substrates, including simple compounds (acetate), complex compounds (polysaccharides), and decaying biomass, also contributed to niche differentiation among denitrifiers. This offers considerable functional redundancy to stabilize denitrification processes in engineered ecosystems, but also places additional complexity over their control.

1.2.5. Ecophysiology of anammox bacteria

Anammox bacteria catalyze the anaerobic oxidation of ammonia to nitrogen gas using nitrite as a terminal electron acceptor. Five genera of anammox bacteria have been discovered to date that belong to a deeply branching group of chemolithoautotrophic bacteria within the Planctomycetes, the Brocadiales. These include *Kuenenia*, *Brocadia*, *Anammoxoglobus* and *Jettenia* commonly found in freshwater and wastewater treatment systems, and *Scalindua* commonly found in marine environments. Because anammox bacteria do not exist in pure culture, they all have '*Candidatus'* status and must be grown in laboratory enrichments, such as membrane bioreactors[73].

The catabolism and cell biology of anammox bacteria has received the most intensive study to date. The anammox cell plan consists of a proteinaceous surface layer as the outermost component, as well as a periplasm, a cytoplasm, and an intracellular organelle called the "anammoxosome". The anammoxosome is the location where the anammox reaction occurs and is surrounded by a densely-packed ladderane lipid membrane that is unique to anammox bacteria[74]. The oxidation of ammonium to dinitrogen gas is catalyzed by 3 key enzymes[75,76]. Nitrite is first reduced to the intermediate nitric oxide via nitrite reductase (NIR) in *K. stuttgartiensis*[75], or to hydroxylamine by an unknown enzyme in *Brocadia spp.*[76]. The produced nitric oxide (or hydroxylamine) and ammonia are then combined to form hydrazine via a comproportionation reaction catalyzed by hydrazine synthase (HZS)[75]. Oxidation of the intermediate hydrazine is then catalyzed by hydrazine dehydrogenase (HDH), where the produced electrons are recycled back through the quinone pool and cytochrome *bc1* complex to the first reactions (NIR and HZS), conserving energy via proton motive force generation[75,77,78].

For growth, anammox bacteria are proposed to fix CO_2 via the Wood-Ljungdahl pathway based on measurements of cell carbon isotopic composition, genomic evidence, and gene expression data[77,78,79]. This requires low potential electrons from NADPH and reduced ferredoxin, which are produced by anammox bacteria via the oxidation of nitrite to nitrate catalyzed by NAR. Because the reduction potential of nitrite (+0.43V) sits much higher than NADPH (-0.32V) and ferredoxin (-0.43V), reverse electron flow is required to generate reducing equivalents. This is presumably driven by proton motive force consumption at the cytochrome *bc1* complex, NADH dehydrogenase complex I, and the recently discovered RNF complex in anammox bacteria[80].

In addition to their chemolithoautotrophic lifestyle, anammox bacteria have been observed to be capable of using simple organic substrates for energy conservation, including acetate,

propionate, and formate[78]. It has been proposed that these substrates are fully oxidized to CO_2 and not assimilated directly into biomass[78]. However, details of the metabolic networks driving mixotrophy in anammox bacteria remain poorly understood.

1.2.6. Metabolic interactions between functional guilds

Within the nitrogen cycle, multiple opportunities exist for cooperation and competition between the different functional guilds. In particular, nitrite represents a central node when interactions converge, as nitrite is a substrate for nitrite oxidation, denitrification, nitrite reduction to ammonia, and anammox, and a product of ammonia oxidation and nitrate reduction (Figure 1.1). In engineered ecosystems, balancing these interactions to achieve stable and efficient nitrogen removal is the biggest challenge, especially since most of the organisms involved and the factors modulating their cooperation and competition remain poorly characterized.

An example that illustrates the complexity of these factors are PNA systems (Figure 1.3). Here, anammox bacteria must cooperate with aerobic AOB to obtain nitrite, while also outcompeting NOB and denitrifying bacteria that additionally use nitrite as a substrate. Environmental variables, such as oxygen, organic substrate availability, temperature, and pH, as well as physiological variables such as substrate affinities, enzyme kinetics, and spatial organization properties (e.g. cell cluster size) all have large impacts on interaction outcomes. So does the exchange of carbon metabolites between species, which largely remains a black-box. This underlines the need for a more systematic approach to predict the interactions and functions occurring in nitrogen cycling microbiome and to drive further bioprocess engineering advancements.

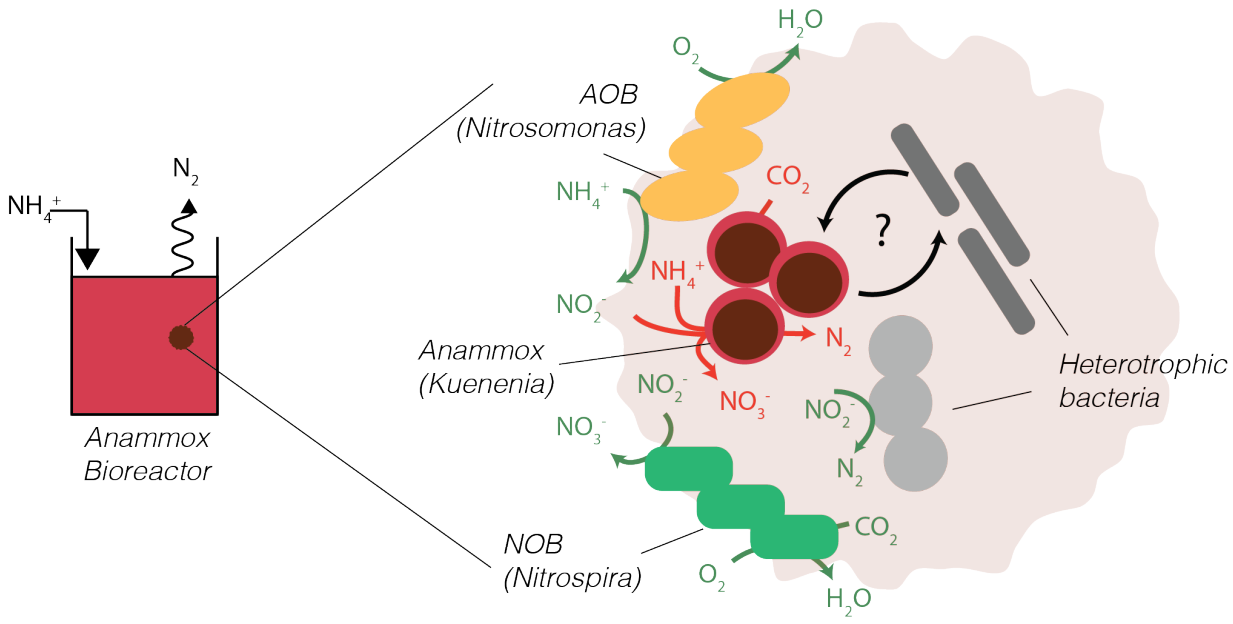

Figure 1.3 Microbial interactions occurring in single-stage PNA systems. Anammox bacteria cooperate with AOB via nitrite exchange but compete with heterotrophic bacteria and NOB for nitrite. "?" denotes unknown carbon exchange between autotrophic and heterotrophic bacteria.

1.3 A systems biology approach for unmasking nitrogen cycle metabolic networks

Translating scientific understanding of nitrogen cycling microbiomes into actionable engineering will require a systems biology approach that links molecular-level processes to ecosystem-level functions. This involves predicting how an individual microorganism's phenotype (or function) emerges from its metabolic network, while also predicting how ecosystem-level functions emerge from microbe-microbe interactions[81]. To create such a framework, quantitative methods from systems biology that enable the reconstruction and analysis of metabolic networks occurring in microbiomes are needed. This includes combining multi-omics tools (metagenomics, metatranscriptomics, metaproteomics, metabolomics) together with isotope tracing techniques and genome-scale metabolic modeling (Figure 1.4).

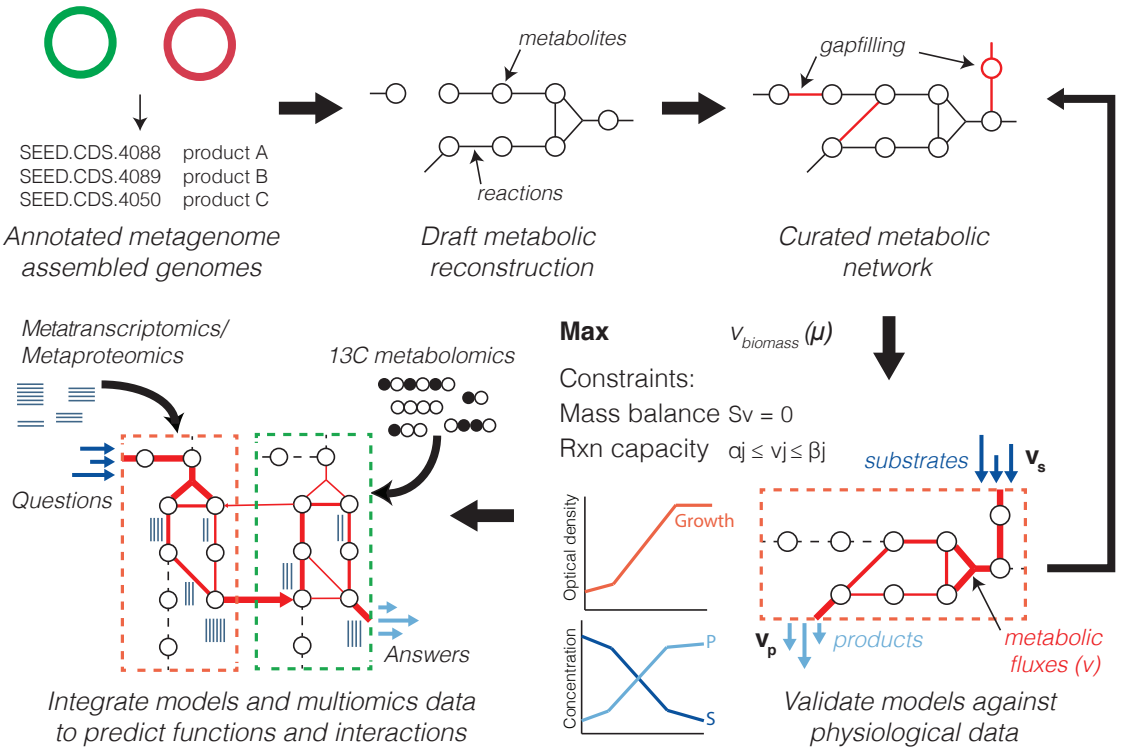

Figure 1.4 Systems biology workflow for metabolic reconstruction and analysis of microbiomes. Isolate genomes are obtained from pure cultures or metagenome-assembled genomes (MAGs) are recovered via binning. Draft metabolic networks are generated via automated reconstruction tools (e.g. KBase[82]) based on genome annotations follow by manual curation against metabolism databases (e.g. MetaCyc[83]) and biochemical literature. Resulting metabolic models are then validated against physiology data and integrated with multi-omic and fluxomic data to predict microbiome functions and interactions.

1.4 Research Needs

To improve the functional predictions of nitrogen cycling microbiomes in PNA systems, a systematic understanding of their metabolic versatility and interactions is needed. While understanding the metabolic interactions between anammox bacteria, AOB, and NOB in PNA systems has produced several strategies to manage these guilds *in situ*[84–86], the function of heterotrophic organisms that can represent up to 50% of the relative community abundance remain largely uncharacterized[87,88]. Clarifying their function and interactions in PNA microbiomes could offer new insights for process control. Additionally, the autotrophic and mixotrophic metabolic

networks driving the anabolism of anammox bacteria and *Nitrospira*-affiliated NOB have not been experimentally resolved beyond genome-based predictions[53,77]. Elucidating this could provide mechanisms underlying their observed metabolic versatility and may also inform PNA reactor organic carbon feeding strategies. Finally, genome-scale metabolic models that predict the function of nitrogen cycling microbiomes are needed to provide a framework for the hypothesis-driven understanding and optimization of PNA systems.

1.5 Research Objectives and Thesis Outline

Based on the knowledge gaps identified above, the chapters presented in this dissertation set out to address the following objectives:

1. Identify the metabolism and interactions of heterotrophic organisms in PNA bioreactors using genome-centric metagenomic and metatranscriptomic analysis.
2. Quantify the autotrophic and mixotrophic metabolic networks of the anammox bacterium *Kuenenia stuttgartiensis* and the NOB *Nitrospira moscoviensis* using ^{13}C isotopic nonstationary metabolic flux analysis.
3. Reconstruct genome-scale metabolic models for anammox, comammox, and NOB bacteria.
4. Predict the environmental variables and mechanisms controlling interactions between comammox and anammox bacteria using dynamic flux balance analysis.
5. Synthesize common principles and best practices for engineering microbiomes into a design-build-test-learn cycle that can be used to harness microbiomes for broad applications.

In **Chapter 2**, we reconstructed the metabolism of a sidestream anammox bioreactor containing anammox bacteria and a diverse group of heterotrophic bacteria and profiled their gene expression using metatranscriptomics. We showed that heterotrophic bacteria in the community contain distributed pathways for denitrification and that the most abundant heterotroph may reduce nitrate produced by anammox bacteria back to nitrite, forming a nitrate-nirtite loop. We also show that heterotrophic bacteria in the community likely use amino acids and peptides derived from anammox bacteria as a carbon and energy source, which may be acquired through the extracellular polymeric matrix or cell decay.

In **Chapter 3**, we experimentally resolved the autotrophic and mixotrophic metabolic networks of the anammox bacterium *"Candidatus"* Kuenenia stuttgartiensis using time-series ^{13}C isotope tracing, metabolomics, and isotopically non-stationary metabolic flux analysis. We show that *K. stuttgartiensis* operates an oxidative TCA cycle, despite the genome not encoding a known citrate synthase. We also elucidate the metabolic networks driving formate and acetate utilization, offering mechanistic insights underlying the observed versatility and mixotrophy of anammox bacteria.

In **Chapter 4**, we build a genome-scale metabolic model of the NOB *Nitrospira moscoviensis* and use flux balance analysis to evaluate the metabolic networks used during chemolithoautotrophic growth and growth on formate. We then perform ^{13}C-tracer experiments with bicarbonate and formate coupled to metabolomics and isotopically nonstationary metabolic flux analysis to experimentally validate model predictions. We experimentally verify that *N. moscoviensis* uses the reductive tricarboxylic acid cycle for CO_2 fixation and also show that *N. moscoviensis* can indirectly use formate as a carbon source by oxidizing it to CO_2 followed by reassimilation via the reductive tricarboxylic acid cycle.

In **Chapter 5**, we use insights gained from Chapters 3 and 4 to reconstruct genome-scale metabolic models for the anammox bacterium *Brocadia sinica* and the comammox bacterium *Nitrospira nitrosa* and simulate their interactions using genome-scale dynamic metabolic modeling. We show that when the ammonia uptake flux is limiting, comammox *Nitrospira* perform complete nitrification and no nitrite is produced to support anammox bacterial growth. However, when the oxygen uptake flux is limiting, comammox bacteria only oxidize ammonia to nitrite, which can be coupled with anammox for the complete removal of ammonia to nitrogen gas. Using metagenomics and metaproteomics, we further show that heterotrophic bacteria in the community are likely denitrifiers that express transporters for amino acid and peptide substrates possibly produced by anammox and comammox bacteria, which provides an independent observation of the interaction mechanism identified in Chapter 2.

In **Chapter 6**, we create a vision for engineering microbiomes using a design-build-test-learn cycle, focusing on generalizable approaches, including top-down and bottom-up design processes, synthetic and self-assembled construction methods, and emerging tools to analyze microbiome function. Key challenges and opportunities of each approach are discussed and synthesized into best practice guidelines for engineering microbiomes.

In **Chapter 7**, we provide a summary of the major conclusions from the dissertation work and outline future research directions.

1.6 References

1. Kartal, B., Kuenen, J. G. & van Loosdrecht, M. C. M. Sewage Treatment with Anammox. *Science (80-.).* **328**, 702–703 (2010).
2. Kuypers, M. M. M., Marchant, H. K. & Kartal, B. The microbial nitrogen-cycling network. *Nat. Rev. Microbiol.* **16**, 263–276 (2018).

3. Ardern, E. & Lockett, W. T. Experiments on the oxidation of sewage without the aid of filters. *J. Soc. Chem. Ind.* **33**, 523–539 (1914).
4. Ludzack, F. J. & Ettinger, M. B. Controlling Operation to Minimize Activated Sludge Effluent Nitrogen. *J. Water Pollut. Control Fed.* **34**, 920–931 (1962).
5. Saunders, A. M., Albertsen, M., Vollertsen, J. & Nielsen, P. H. The activated sludge ecosystem contains a core community of abundant organisms. *ISME J.* **10**, 11–20 (2016).
6. Wagner, M. *et al.* Combining fluorescent in situ hybridization (fish) with cultivation and mathematical modeling to study population structure and function of ammonia-oxidizing bacteria in activated sludge. *Water Sci. Technol.* **37**, 441–449 (1998).
7. Daims, H., Nielsen, J. L., Nielsen, P. H., Schleifer, K. H. & Wagner, M. In Situ Characterization of Nitrospira-Like Nitrite-Oxidizing Bacteria Active in Wastewater Treatment Plants. *Appl. Environ. Microbiol.* **67**, 5273–5284 (2001).
8. Thomsen, T. R., Kong, Y. & Nielsen, P. H. Ecophysiology of abundant denitrifying bacteria in activated sludge. *FEMS Microbiol. Ecol.* **60**, 370–382 (2007).
9. Morgan-Sagastume, F., Larsen, P., Nielsen, J. L. & Nielsen, P. H. Characterization of the loosely attached fraction of activated sludge bacteria. *Water Res.* **42**, 843–54 (2008).
10. Gray, N. F. *Biology of waste water treatment.* (Oxford University Press, 1989).
11. Blackall, L. L. & Burrell, P. C. The microbiology of nitrogen removal in activated sludge systems BT - The Microbiology of Activated Sludge. in (eds. Seviour, R. J. & Blackall, L. L.) 203–226 (Springer Netherlands, 1999). doi:10.1007/978-94-011-3951-9_8
12. Microbial Ecology of Activated Sludge. in (eds. Seviour, R. J. & Nielsen, P. H.) (IWA Publishing, 2010).
13. van de Graaf, A. A. *et al.* Anaerobic oxidation of ammonium is a biologically mediated process. *Appl. Environ. Microbiol.* **61**, 1246 LP – 1251 (1995).
14. Broda, E. Two kinds of lithotrophs missing in nature. *Z. Allg. Mikrobiol.* **17**, 491–493 (1977).
15. Kuenen, J. G. Anammox bacteria : from discovery to application. *Nat. Rev. Microbiol.* **6**, 320–326 (2008).
16. Jetten, M. S. M., Horn, S. J. & van Loosdrecht, M. C. M. Towards a more sustainable municipal wastewater treatment system. *Water Sci. Technol.* **35**, 171–180 (1997).
17. Xie, S., Higgins, M. J., Bustamante, H., Galway, B. & Nghiem, L. D. Current status and

perspectives on anaerobic co-digestion and associated downstream processes. *Environ. Sci. Water Res. Technol.* **4**, 1759–1770 (2018).

18. Sliekers, A. O. *et al.* Completely autotrophic nitrogen removal over nitrite in one single reactor. *Water Res.* **36**, 2475–2482 (2002).
19. Van Dongen, U., Jetten, M. S. M. & Van Loosdrecht, M. C. M. The SHARON®-Anammox® process for treatment of ammonium rich wastewater. *Water Sci. Technol.* **44**, 153–160 (2001).
20. van der Star, W. R. L. *et al.* Startup of reactors for anoxic ammonium oxidation: Experiences from the first full-scale anammox reactor in Rotterdam. *Water Res.* **41**, 4149–4163 (2007).
21. Li, X., Klaus, S., Bott, C. & He, Z. Status, Challenges, and Perspectives of Mainstream Nitritation-Anammox for Wastewater Treatment. *Water Environ. Res.* **90**, 634–649 (2018).
22. Daims, H. *et al.* Complete nitrification by Nitrospira bacteria. *Nature* **528**, 504–509 (2015).
23. van Kessel, M. A. H. J. *et al.* Complete nitrification by a single microorganism. *Nature* **528**, 555–559 (2015).
24. Winogradsky, S. Recherches sur les organismes de la Nitrification. *Comptes rendus l' Académie des Sci.* **110**, 1013–1016 (1890).
25. Beach, N. K. & Noguera, D. R. Design and Assessment of Species-Level qPCR Primers Targeting Comammox . *Frontiers in Microbiology* **10**, 36 (2019).
26. Cotto, I. *et al.* Long solids retention times and attached growth phase favor prevalence of comammox bacteria in nitrogen removal systems. *Water Res.* **169**, 115268 (2020).
27. Kits, K. D. *et al.* Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature* **549**, 269–272 (2017).
28. Juretschko, S. *et al.* Combined Molecular and Conventional Analyses of Nitrifying Bacterium Diversity in Activated Sludge : Nitrosococcus mobilis and Nitrospira -Like Bacteria as Dominant Populations Combined Molecular and Conventional Analyses of Nitrifying Bacterium Diversity. *Appl. Environ. Microbiol.* **64**, 3042–3051 (1998).
29. Okabe, S., Satoh, H. & Watanabe, Y. In Situ Analysis of Nitrifying Biofilms as Determined by In Situ Hybridization and the Use of Microelectrodes. *Appl. Environ.*

Microbiol. **65**, 3182 LP – 3191 (1999).

30. Caranto, J. D. & Lancaster, K. M. Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase. *Proc. Natl. Acad. Sci.* **114**, 8217 LP – 8222 (2017).
31. Krümmel, A. & Harms, H. Effect of organic matter on growth and cell yield of ammonia-oxidizing bacteria. *Arch. Microbiol.* **133**, 50–54 (1982).
32. Stüven, R., Vollmer, M. & Bock, E. The impact of organic matter on nitric oxide formation by Nitrosomonas europaea. *Arch. Microbiol.* **158**, 439–443 (1992).
33. Schmidt, I., Spanning, R. J. M. Van & Jetten, M. S. M. Denitrification and ammonia oxidation by Nitrosomonas europaea wild-type , and NirK- and NorB-deficient mutants Printed in Great Britain. *Microbiology* **150**, 4107–4114 (2004).
34. Stein, L. Y. & Arp, D. J. Loss of ammonia monooxygenase activity in nitrosomonas europaea upon exposure to nitrite. *Appl. Environ. Microbiol.* **64**, 4098–4102 (1998).
35. Laanbroek, H. J. & Gerards, S. Competition for limiting amounts of oxygen between Nitrosomonas europaea and Nitrobacter winogradskyi grown in mixed continuous cultures. *Arch. Microbiol.* **159**, 453–459 (1993).
36. Park, H. D. & Noguera, D. R. Characterization of two ammonia-oxidizing bacteria isolated from reactors operated with low dissolved oxygen concentrations. *J. Appl. Microbiol.* **102**, 1401–1417 (2007).
37. Nowka, B. & Daims, H. Comparison of Oxidation Kinetics of Nitrite-Oxidizing Bacteria : Nitrite Availability as a Key Factor in Niche Differentiation. *Appl. Environ. Microbiol.* **81**, 745–753 (2015).
38. Off, S., Alawi, M. & Spieck, E. Enrichment and Physiological Characterization of a Novel Nitrospira-Like Bacterium Obtained from a Marine Sponge. *Appl. Environ. Microbiol.* **76**, 4640 LP – 4646 (2010).
39. Park, M. R., Park, H. & Chandran, K. Molecular and Kinetic Characterization of Planktonic Nitrospira spp. Selectively Enriched from Activated Sludge. *Environ. Sci. Technol.* **51**, 2720–2728 (2017).
40. Kitzinger, K. *et al.* Characterization of the First "Candidatus Nitrotoga" Isolate Reveals Metabolic Versatility and Separate Evolution of Widespread Nitrite-Oxidizing Bacteria. *MBio* **9**, e01186-18 (2018).

41. Sauder, L. A. *et al.* Cultivation and characterization of Candidatus Nitrosocosmicus exaquare, an ammonia-oxidizing archaeon from a municipal wastewater treatment system. *ISME J.* **11**, 1142–1157 (2017).
42. Oshiki, M., Satoh, H. & Okabe, S. Ecology and physiology of anaerobic ammonium oxidizing bacteria. *Environ. Microbiol.* **18**, 2784–2796 (2016).
43. Chain, P. *et al.* Complete Genome Sequence of the Ammonia-Oxidizing Bacterium and Obligate Chemolithoautotroph Nitrosomonas europaea. *J. Bacteriol.* **185**, 2759–2773 (2003).
44. Starkenburg, S. R. *et al.* Genome Sequence of the Chemolithoautotrophic Nitrite-Oxidizing Bacterium Nitrobacter winogradskyi Nb-255. *Appl. Environ. Microbiol.* **72**, 2050 LP – 2063 (2006).
45. Arp, D. J., Chain, P. S. G. & Klotz, M. G. The Impact of Genome Analyses on Our Understanding of Ammonia-Oxidizing Bacteria. *Annu. Rev. Microbiol.* **61**, 503–528 (2007).
46. Ellis, R. J. The most abundant protein in the world. *Trends Biochem. Sci.* **4**, 241–244 (1979).
47. Ward, B. B., Arp, D. J. & Klotz, M. G. *Nitrification*. (ASM Press, 2011). doi:10.1128/9781555817145
48. Stahl, D. a & de la Torre, J. R. Physiology and diversity of ammonia-oxidizing archaea. *Annu. Rev. Microbiol.* **66**, 83–101 (2012).
49. Kozlowski, J. A., Stieglmeier, M., Schleper, C., Klotz, M. G. & Stein, L. Y. Pathways and key intermediates required for obligate aerobic ammonia-dependent chemolithotrophy in bacteria and Thaumarchaeota. *ISME J.* **10**, 1836–1845 (2016).
50. Könneke, M. *et al.* Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *Proc. Natl. Acad. Sci.* **111**, 8239 LP – 8244 (2014).
51. Martens-Habbena, W., Berube, P. M., Urakawa, H., de la Torre, J. R. & Stahl, D. A. Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* **461**, 976–979 (2009).
52. Lücker, S. *et al.* Nitrotoga-like bacteria are previously unrecognized key nitrite oxidizers in full-scale wastewater treatment plants. *ISME J.* **9**, 708–720 (2015).
53. Lücker, S. *et al.* A Nitrospira metagenome illuminates the physiology and evolution of

globally important nitrite-oxidizing bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13479–84 (2010).

54. Daims, H., Lücker, S. & Wagner, M. A New Perspective on Microbes Formerly Known as Nitrite-Oxidizing Bacteria. *Trends Microbiol.* **24**, 699–712 (2016).
55. Berg, I. A. Ecological aspects of the distribution of different autotrophic CO2 fixation pathways. *Appl. Environ. Microbiol.* **77**, 1925–1936 (2011).
56. Gruber-Dorninger, C. *et al.* Functionally relevant diversity of closely related Nitrospira in activated sludge. *ISME J.* 1–13 (2014). doi:10.1038/ismej.2014.156
57. Koch, H. *et al.* Growth of nitrite-oxidizing bacteria by aerobic hydrogen oxidation. *Science* **345**, 1052–4 (2014).
58. Koch, H. *et al.* Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus Nitrospira. *Proc. Natl. Acad. Sci.* **112**, 201506533 (2015).
59. Ehrich, S., Behrens, D., Lebedeva, E. V, Ludwig, W. & Bock, E. A New Obligately Chemolithoautotrophic, Nitrite-Oxidizing Bacterium, Nitrospira-Moscoviensis Sp-Nov and Its Phylogenetic Relationship. *Arch. Microbiol.* **164**, 16–23 (1995).
60. Lawson, C. E. & Lücker, S. Complete ammonia oxidation: an important control on nitrification in engineered ecosystems? *Curr. Opin. Biotechnol.* **50**, 158–165 (2018).
61. Costa, E., Pérez, J. & Kreft, J. U. Why is metabolic labour divided in nitrification? *Trends Microbiol.* **14**, 213–219 (2006).
62. Camejo, P. Y., Domingo, J. S., Noguera, D. R. & McMahon, K. D. Genome-Enabled Insights into the Ecophysiology of the Comammox Bacterium "Candidatus Nitrospira nitrosa". *mSystems* **2**, e00059-17 (2017).
63. Galán, A., Faúndez, J., Thamdrup, B., Santibáñez, J. F. & Farías, L. Temporal dynamics of nitrogen loss in the coastal upwelling ecosystem off central Chile: Evidence of autotrophic denitrification through sulfide oxidation. *Limnol. Oceanogr.* **59**, 1865–1878 (2014).
64. Hawley, A. K., Brewer, H. M., Norbeck, A. D., Pasa-Toli, L. & Hallam, S. J. Metaproteomics reveals differential modes of metabolic coupling among ubiquitous oxygen minimum zone microbes. *Proc. Natl. Acad. Sci.* **111**, 11395–11400 (2014).
65. Vaiopoulou, E., Melidis, P. & Aivasidis, A. Sulfide removal in wastewater from petrochemical industries by autotrophic denitrification. *Water Res.* **39**, 4101–4109 (2005).

66. Payne, W. J. *Denitrification.* (John Wiley & Sons Inc., 1981).
67. Heylen, K. *et al.* The incidence of nirS and nirK and their genetic heterogeneity in cultivated denitrifiers. *Environ. Microbiol.* **8**, 2012–2021 (2006).
68. Einsle, O. *et al.* Structure of cytochrome c nitrite reductase. *Nature* **400**, 476–480 (1999).
69. Brown, K. *et al.* A novel type of catalytic copper cluster in nitrous oxide reductase. *Nat. Struct. Biol.* **7**, 191–195 (2000).
70. Godden, J. W. *et al.* The 2.3 angstrom X-ray structure of nitrite reductase from Achromobacter cycloclastes. *Science (80-.).* **253**, 438 LP – 442 (1991).
71. Van de Pas-Schoonen, K. T. *et al.* Complete conversion of nitrate into dinitrogen gas in co-cultures of denitrifying bacteria. *Biochem. Soc. Trans.* **33**, 205–209 (2005).
72. Lilja, E. E. & Johnson, D. R. Segregating metabolic processes into different microbial cells accelerates the consumption of inhibitory substrates. *ISME J.* **10**, 1568–78 (2016).
73. van der Star, W. R. L. *et al.* The membrane bioreactor: a novel tool to grow anammox bacteria as free cells. *Biotechnol. Bioeng.* **101**, 286–94 (2008).
74. Damste, J. S. S. *et al.* Linearly concatenated cyclobutane lipids form a dense bacterial membrane. *Nature* **419**, 708–712 (2002).
75. Kartal, B. *et al.* Molecular mechanism of anaerobic ammonium oxidation. *Nature* **479**, 127–30 (2011).
76. Oshiki, M., Ali, M., Shinyako-Hata, K., Satoh, H. & Okabe, S. Hydroxylamine-dependent Anaerobic Ammonium Oxidation (Anammox) by " Candidatus Brocadia sinica". *Environ. Microbiol.* (2016). doi:10.1111/1462-2920.13355
77. Strous, M. *et al.* Deciphering the evolution and metabolism of an anammox bacterium from a community genome. *Nature* **440**, 790–4 (2006).
78. Kartal, B. *et al.* How to make a living from anaerobic ammonium oxidation. *FEMS Microbiol. Rev.* **37**, 428–61 (2013).
79. Schouten, S. *et al.* Stable Carbon Isotopic Fractionations Associated with Inorganic Carbon Fixation by Anaerobic Ammonium-Oxidizing Bacteria. *Appl. Environ. Microbiol.* **70**, 3785–3788 (2004).
80. Almeida, N. M. De *et al.* Membrane-bound electron transport systems of an anammox bacterium : A complexome analysis. *BBA - Bioenerg.* **1857**, 1694–1704 (2016).
81. Zengler, K. & Palsson, B. O. A road map for the development of community systems

(CoSy) biology. *Nat. Rev. Microbiol.* **10**, 366–372 (2012).

82. Palumbo, A. *et al.* KBase : An Integrated Knowledgebase for Predictive Biology and Environmental Research. 7
83. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res.* **46**, D633–D639 (2017).
84. Winkler, M. K. H., Kleerebezem, R., Kuenen, J. G., Yang, J. & Loosdrecht, M. C. M. Van. Segregation of Biomass in Cyclic Anaerobic / Aerobic Granular Sludge Allows the Enrichment of Anaerobic Ammonium Oxidizing Bacteria at Low Temperatures. *Environ. Sci. Technol.* **45**, 7330–7337 (2011).
85. Ma, B. *et al.* Suppressing Nitrite-oxidizing Bacteria Growth to Achieve Nitrogen Removal from Domestic Wastewater via Anammox Using Intermittent Aeration with Low Dissolved Oxygen. *Sci. Rep.* **5**, 13048 (2015).
86. Regmi, P. *et al.* Ammonia-based intermittent aeration control optimized for efficient nitrogen removal. *Biotechnol. Bioeng.* **112**, 2060–2067 (2015).
87. Gonzalez-Gil, G., Sougrat, R., Behzad, A. R., Lens, P. N. L. & Saikaly, P. E. Microbial Community Composition and Ultrastructure of Granules from a Full-Scale Anammox Reactor. *Microb. Ecol.* **70**, 118–131 (2014).
88. Gonzalez-Martinez, A. *et al.* Comparison of bacterial diversity in full scale anammox bioreactors operated under different conditions. *Biotechnol. Prog.* **31**, 1464–1472 (2015).

2. Metabolic network analysis reveals microbial community interactions in anammox granules

A version of this chapter has been published as:

Lawson, C. E., Wu, S., Bhattacharjee, A.S., Hamilton, J.J., McMahon, K.D., Goel, R., Noguera, D.R. Metabolic network analysis reveals microbial community interactions in anammox granules. *Nat. Commun.* **8**, 15416 (2017).

All supplementary data files are available online with the published article.

Author contributions

C.E.L., K.D.M., R.G., and D.R.N. designed the study. S.W. and A.S.B. operated the bioreactor and performed the sampling and sequencing. C.E.L. analyzed the data and wrote the manuscript. J.J.H. and A.S.B. contributed to the data analysis. J.J.H., A.S.B., K.D.M., R.G., and D.R.N. reviewed and provided valuable edits to the manuscript.

2.0 Abstract

Microbial communities mediating anaerobic ammonium oxidation (anammox) represent one of the most energy-efficient environmental biotechnologies for nitrogen removal from wastewater. However, little is known about the functional role heterotrophic bacteria play in anammox granules. Here, we use genome-centric metagenomics to recover 17 draft genomes of anammox and heterotrophic bacteria from a laboratory-scale anammox bioreactor. We combine metabolic network reconstruction with metatranscriptomics to examine the gene expression of anammox and heterotrophic bacteria and to identify their potential interactions. We find that Chlorobi-affiliated bacteria may be highly active protein degraders, catabolizing extracellular peptides while recycling nitrate to nitrite. Other heterotrophs may also contribute to scavenging of detritus and peptides produced by anammox bacteria, and potentially use alternative electron donors, such as H_2, acetate, and formate. Our findings improve the understanding of metabolic activities and interactions between anammox and heterotrophic bacteria and offer the first transcriptional insights on ecosystem function in anammox granules.

2.1 Introduction

Microbial communities mediating anaerobic ammonium oxidation (anammox) represent one of the most energy-efficient environmental biotechnologies for nitrogen removal from wastewater. The process is typically used to treat high strength ammonium wastewaters and offers significant cost savings compared to conventional nitrogen removal processes that require energy-intensive aeration for nitrification and also consume large quantities of organic carbon during denitrification[1]. In practice, anammox-based wastewater treatment systems are combined with a nitritation step in either a single-stage[2] or two-stage bioreactor system[3]. In these engineered

ecosystems, a fraction of the ammonium is first oxidized to nitrite by aerobic ammonia oxidizing bacteria (AOB). Subsequently, anammox bacteria anaerobically oxidize the remaining ammonium directly to nitrogen gas using the produced nitrite as a terminal electron acceptor[4,5].

Five genera of anammox bacteria have been discovered to date, including *Kuenenia*, *Brocadia*, *Anammoxoglobus* and *Jettenia* commonly found in activated sludge, and *Scalindua* commonly found in marine environments[6]. These lineages all have *'Candidatus'* status as they do not exist in pure culture and must be grown in laboratory enrichments. Because anammox bacteria have a slow growth rate[7], either biofilm reactors that use carrier media or granular sludge reactors are used to retain sufficient biomass in the system[8]. Such reactor configurations support the formation of dense microbial communities that can be readily separated from the liquid wastewater and enriched in the bioreactor. Anammox granules consist of a mixture of cell aggregates and abiotic particles embedded within a matrix of organic extracellular polymeric substances (EPS)[9,10,11]. EPS present in anammox granules has been found to contain high amounts of protein and polysaccharides, where increased hydrophobic amino acid content has been observed to be a main factor determining granule aggregation ability[11]. The high aggregate density also serves to limit oxygen diffusion into the granule interior, allowing for the proliferation of anammox and other anaerobic bacteria.

While most studies have focused on understanding the ecophysiology of anammox bacteria[4,6,12] and their interactions with autotrophic nitrifying bacteria[13], little is known about the activity of heterotrophic bacteria in anammox bioreactors. Previous studies based on 16S ribosomal RNA (rRNA) gene clone libraries and amplicon sequencing have shown that heterotrophic bacteria affiliated with the phyla Chlorobi, Bacteroidetes, Chloroflexi, and Proteobacteria comprise a large fraction of the microbial community in anammox

bioreactors[14,10,15]. Despite differences in bioreactor influent composition, these heterotrophic bacteria appear to share high phylogenetic similarity across different anammox wastewater treatment systems[16], suggesting that a more universal interaction exists between them and anammox bacteria.

The exact role of heterotrophic bacteria in anammox systems has not yet been determined, though a few clues have been uncovered. A recent metagenomic study revealed that most of the heterotrophic organisms in anammox granules encode the ability to respire nitrate via partial denitrification, possibly completing a nitrite loop with anammox and nitrite oxidizing bacteria (NOB) by reducing nitrate back to nitrite[16] . This activity could contribute to the removal of excess nitrate produced from the system during anammox growth or nitrite oxidation by NOB. Hydrolysis of EPS into soluble compounds and/or the secretion of soluble microbial products by anammox bacteria during cell growth is believed to support denitrification and heterotrophic growth, particularly for anammox bioreactors that receive no external organic carbon substrates[17,18]. However, the specific metabolite exchange reactions promoting interactions between heterotrophic and anammox bacteria remain poorly understood.

Here, we combine metagenomic and metatranscriptomic data to examine the gene expression of anammox and heterotrophic bacteria in a laboratory-scale anammox bioreactor and to identify their potential interactions. Metagenomic binning was used to recover near-complete population genomes from members of the microbial community inhabiting anammox granules. Resulting population genomes were used to reconstruct each organism's metabolic network and served as reference platforms for the profiling of gene expression at the community scale. We find that Chlorobi-affiliated bacteria may be highly active protein degraders, catabolizing extracellular peptides bound in the EPS matrix, while respiring nitrate produced during anammox bacterial

growth to nitrite. Other heterotrophic bacteria may also contribute to the scavenging of detritus and peptides produced by anammox bacteria, and potentially use alternative electron donors, such as H_2, acetate, and formate to fuel their energy metabolism. These findings improve the understanding of major metabolic activities and interactions occurring between anammox and heterotrophic bacteria and offer the first transcriptional insights on ecosystem function in anammox granules.

2.2 Results

2.2.1 Metagenomic sequencing and binning

Sequencing of whole community DNA extracted from two separate biomass samples yielded a total of 41,063,036 reads after quality filtering (Table 2.1, Table 2.2). Co-assembly of the resulting reads using CLC Genomics Workbench generated a total of 83,770 contigs with an N50 of 2,376 bp and an N20 of 20,848 bp, accounting for 96% of the quality filtered DNA reads obtained from the second sampling event (Table 2.3). Approximately 74% of the quality filtered mRNA reads could be mapped to this assembly, indicating that it also captured a large proportion of the metatranscriptome. Contigs were subsequently binned into population genomes based on tetranucleotide frequency and differential coverage using MetaBAT. This resulted in the recovery of 17 draft metagenome-assembled genomes (MAGs) affiliated with the phyla Planctomycetes, Chlorobi, Bacteroidetes, Chloroflexi, Proteobacteria, and the candidate phylum Microgenomates (OP11) (Table 2.1). Together, these genomes accounted for approximately 82% and 59% of the total quality filtered DNA and mRNA reads obtained from the biomass samples, respectively, and therefore represented a major fraction of the microbial community present in the granular sludge reactor (Table 2.3).

Figure 2.1 shows a phylogenetic tree of the recovered population genomes based on the protein sequences of 37 conserved bacterial marker genes. Many of the recovered genomes, particularly CHB1 and CFX5, had high similarity to genomes recently recovered from a single-stage reactor that were shown to share high 16S rRNA gene similarity to organisms detected in other anammox systems[16]. Surprisingly, the Chlorobi CHB1 genome was virtually identical to the Chlorobi OLB4 genome recovered by Speth et al.[16], sharing an average nucleotide identity (ANI) of 99.8% despite the significant differences in reactor operation (two-stage versus single-stage), wastewater influent composition (potato-processing wastewater versus anaerobic digester filtrate), and geographical location (USA versus Netherlands) (Supplementary Data 1).

Table 2.1 - Metagenomic and metatranscriptomic sequencing statistics.

Sample	DNA_09_04_13_S1	DNA_09_09_15	RNA_09_09_15
total reads	4,123,584	45,715,259	22,056,078
filtered reads	4,116,701	36,946,335	18,437,366
merged reads	1,816,852	34,274,103	11,068,638
merged length	300-590	125-240	125-240
mRNA	-	-	8,780,658
rRNA	-	-	2,287,980
Description	MiSeq metagenome	HiSeq metagenome	HiSeq metatranscriptome
BioSample No.	SAMN05785373	SAMN05785375	SAMN05785376

2.2.2 Microbial community abundance and gene expression

We used the reads per kilobase per million mapped reads (RPKM) values for metagenomic reads and transcripts that mapped to each MAG as proxies for relative abundance and gene expression, respectively (Figure 2.2). A summary of the metatranscriptomic sequencing statistics and read mapping can be found in Table 2.3. Overall, gene expression corresponded with abundance in the anammox granules. Genomes affiliated with Brocadia (AMX1) and Chlorobi (CHB1) dominated

the abundance and gene expression of the microbial community in the anammox granules (Figure 2.2). AMX1 had a relative abundance and gene expression of approximately 62% and 54%, respectively whereas CHB1 had a relative abundance and gene expression of approximately 21% and 26%, respectively. Other organisms that displayed moderate abundance and gene expression in the anammox ecosystem were affiliated with the phyla Chlorobi, Bacteroidetes, Chloroflexi, and Proteobacteria (Figure 2.2). Interestingly, several low abundance genomes were also observed to be particularly active based on gene expression (i.e. PRO1, PRO2, PLA1), as has been previously observed in other activated sludge ecosystems[19].

2.2.3 Metabolic reconstruction of microbial community

To examine the functional potential and gene expression of the anammox community, ORFs were predicted and annotated across each MAG. Subsequently, mRNA transcripts were mapped against all ORFs to identify functions that were highly expressed by the community during steady-state bioreactor operation (Supplementary Data 2). Metabolic pathways were then reconstructed for each of eight near-complete MAGs that displayed high abundance and/or gene expression in the community (>0.5%) using MetaPathways 2.5[20]. A complete list of the inferred pathways and associated enzymes of each MAG can be found in Supplementary Data 3.

2.2.4 Anammox metatranscriptomic insights

As expected, genes involved in anammox metabolism from AMX1 were among the highest genes expressed in the community (Supplementary Data 2). Most notably, UTAMX1_1243, UTAMX1_1246, and UTAMX1_1249 annotated as hydrazine dehydrogenase (*hdh*), hydroxylamine oxidoreductase (*hao*), and hydrazine synthase subunit A (*hzsA*) were highly expressed, maintaining 10-20 fold coverage above median gene expression levels in the AMX1

genome (Figure 2.3). Six other *hao*-like genes were also identified in the AMX1 genome and displayed above median gene expression levels. Reciprocal best-BLAST searches confirmed that these genes were orthologous to *hao*-like genes from publically available anammox genomes (Figure 2.3). Among the Hao-like proteins, UTAMX1_1246 had the highest gene expression, suggesting it played an important role in anammox metabolism. This gene is orthologous to previously characterized Hao proteins from '*Ca.* Kueneia stuttgartiensis' (kustc1061)[5,21], '*Ca.* Brocadia anammoxidans'[22], and '*Ca.* Jettenia caeni'[23], which have been shown to oxidize hydroxylamine to nitric oxide and were also observed to be among the most abundant proteins expressed in these organisms. Phylogenetic analysis showed that AMX1 was closely related to both '*Ca.* Brocadia sinica'[24] and '*Ca.* Brocadia fulgida'[25] (Figure 2.1). Similar to '*Ca.* B. sinica', neither *nirK* nor *nirS* were found in the AMX1 genome. This is consistent with the recently proposed hydroxylamine-dependent anammox mechanism in '*Ca.* B. sinica' that first reduces nitrite to hydroxylamine (instead of nitric oxide), and subsequently converts hydroxylamine and ammonium to hydrazine[26] (Figure 2.3). It has been proposed that Hao-like proteins lacking a crosslinking tyrosine in the c-terminus may be involved in the initial nitrite reduction reaction to either nitric oxide[5] or hydroxylamine[26], for which an enzyme has not yet been identified. The hao-like genes UTAMX1_1996, UTAMX1_1759, and UTAMX1_1192 expressed in the AMX1 genome that lack a crosslinking tyrosine might fulfill this role (Figure 2.3).

Table 2.2 - Genome statistics of metagenome-assembled genomes (MAGs) recovered from the anammox community. GenBank accession numbers for each MAG can be found in Table 2.4. * indicates incomplete MAGs not submitted to GenBank.

Bin ID	Locus_tag	Taxonomy	Completeness (%)	Contamination (%)	Genome size (bp)	# scaffolds	N50 (scaffolds)	GC	predicted genes
AMX1	UTAMX1	Bacteria; Planctomycetes; Planctomycetia; Planctomycetales; Planctomycetaceae; Brocadia	96	1	3,142,066	36	137,788	42.3	2,724
AMX2	UTAMX2	Bacteria; Planctomycetes; Planctomycetia; Planctomycetales; Planctomycetaceae; Brocadia	96	2	3,437,337	148	32,519	45.1	3,064
PLA1	UTPLA1	Bacteria; Planctomycetes	82	2	4,435,946	642	8,518	60.8	4,062
BCD1	UTBCD1	Bacteria; Bacteroidetes; Sphingobacteriia; Sphingobacteriales	96	1	3,757,364	58	91,406	41.0	3,180
CFX1	UTCFX1	Bacteria; Chloroflexi; Anaerolineae	91	2	2,777,515	34	125,597	52.6	2,576
CFX2	UTCFX2	Bacteria; Chloroflexi; Anaerolineae	81	5	2,924,761	111	48,865	57.0	2,622
CFX3	UTCFX3	Bacteria; Chloroflexi; Anaerolineae	82	1	2,655,948	281	12,640	61.3	2,676
CFX4	UTCFX4	Bacteria; Chloroflexi	81	5	4,654,268	672	8,164	54.8	4,526
CFX5	UTCFX5	Bacteria; Chloroflexi; Anaerolineae	84	2	3,776,973	364	12,671	62.8	3,432
CHB1	UTCHB1	Bacteria; Chlorobi; Ignavibacteria; Ignavibacteriales; Melioribacteraceae	95	0	2,424,598	19	211,700	37.6	2,083
CHB2	UTCHB2	Bacteria; Chlorobi; Ignavibacteria; Ignavibacteriales; Ignavibacteriaceae	95	0	4,062,427	225	36,474	33.0	3,556
CHB3	UTCHB3	Bacteria; Chlorobi; Ignavibacteria; Ignavibacteriales	83	2	2,943,821	349	10,790	42.1	2,526
CHB4*	UTCHB4	Bacteria; Chlorobi; Ignavibacteria; Ignavibacteriales	43	0	1,189,945	296	4,153	34.7	1,311
CHB5*	UTCHB5	Bacteria; Chlorobi; Ignavibacteria; Ignavibacteriales	6	2	1,684,758	362	4,608	34.4	1,621
PRO1	UTPRO1	Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales; Polyangiaceae	74	2	3,818,019	45	114,904	69.8	3,321
PRO2	UTPRO2	Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales; Rhodocyclaceae; Sulfuritalea	68	1	2,355,954	380	7,446	66.9	2,730
CPR1	UTCPR1	Bacteria; Microgenomates	68	0	892,243	91	13,172	37.7	959

Table 2.3 - Metagenomic and metatranscriptomic read mapping statistics.

MAG	genome size	read count	coverage (RPKM)	relative abundance	total reads (%)	mapped reads (%)	mRNA count	expression (RPKM)	relative expression	total mRNA (%)	mapped mRNA (%)
AMX1	3,142,066	19,308,825	185.41	62.0%	56.3%	58.6%	3,057,308	153.72	54.8%	34.8%	47.2%
CHB1	2,424,598	4,872,772	62.29	20.8%	14.2%	14.8%	1,099,805	71.16	25.4%	12.5%	17.0%
CHB4	1,189,945	677,700	17.25	5.8%	2.0%	2.1%	49,708	6.44	2.3%	0.6%	0.8%
CHB5	1,684,758	490,899	9.02	3.0%	1.4%	1.5%	37,025	3.48	1.2%	0.4%	0.6%
CHB2	4,062,427	782,510	5.40	1.8%	2.3%	2.4%	131,573	5.41	1.9%	1.5%	2.0%
BCD1	3,757,364	509,896	4.14	1.4%	1.5%	1.5%	130,039	5.33	1.9%	1.5%	2.0%
CFX1	2,777,515	274,783	3.14	1.1%	0.8%	0.8%	148,199	9.00	3.2%	1.7%	2.3%
CFX2	2,924,761	260,104	2.71	0.9%	0.8%	0.8%	47,987	2.49	0.9%	0.5%	0.7%
PRO2	2,355,954	252,380	1.99	0.7%	0.7%	0.8%	129,168	5.21	1.9%	1.5%	2.0%
PRO1	3,818,019	123,619	1.59	0.5%	0.4%	0.4%	67,184	4.32	1.5%	0.8%	1.0%
AMX2	3,437,337	190,916	1.68	0.6%	0.6%	0.6%	84,967	3.98	1.4%	1.0%	1.3%
CFX3	2,655,948	100,255	1.14	0.4%	0.3%	0.3%	31,808	1.73	0.6%	0.4%	0.5%
CFX5	3,776,973	150,431	1.21	0.4%	0.4%	0.5%	23,103	0.98	0.3%	0.3%	0.4%
PLA1	4,435,946	152,921	1.04	0.3%	0.4%	0.5%	105,111	3.72	1.3%	1.2%	1.6%
CPR1	892,243	8,398	0.27	0.1%	0.0%	0.0%	9,895	1.64	0.6%	0.1%	0.2%
CFX4	4,654,268	35,973	0.23	0.1%	0.1%	0.1%	36,599	1.19	0.4%	0.4%	0.6%
CHB3	2,943,821	11,249	0.11	0.0%	0.0%	0.0%	3,767	0.19	0.1%	0.0%	0.1%

Total DNA reads mapping to metagenomic assembly: 32,968,582 (96%)
Total mRNA reads mapping to metagenomic assembly: 6,473,379 (74%)

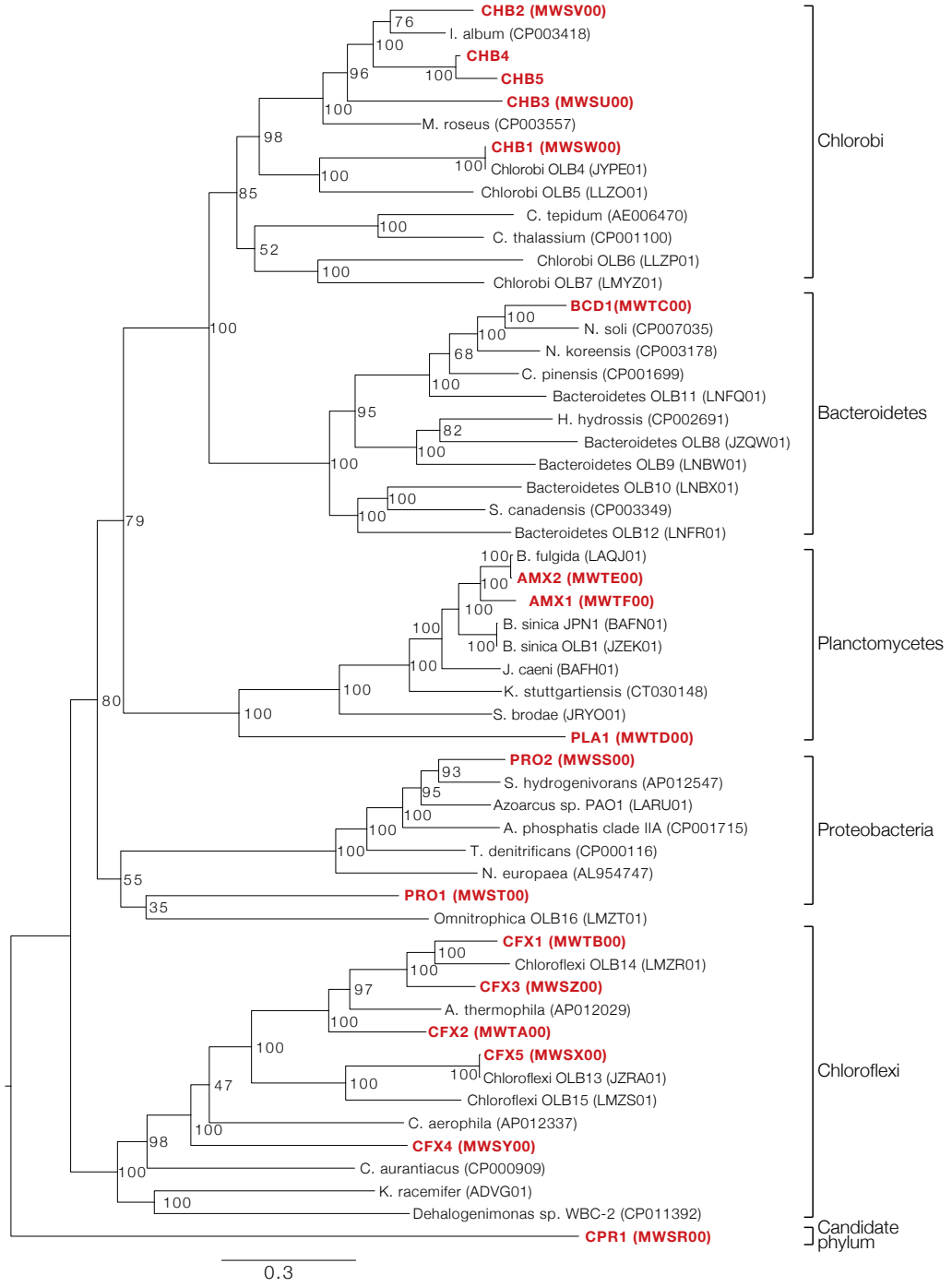

Figure 2.1 - Phylogenetic tree of all recovered draft genomes from the anammox bioreactor. Tree includes metagenome-assembled genomes recovered from this study (red) and closely related genomes downloaded from the NCBI genome repository. GenBank accession numbers for each genome are provided in parentheses. Branch node numbers represent bootstrap support values. The tree was constructed using RAxML based on a set of 37 concatenated universal single-copy marker genes.

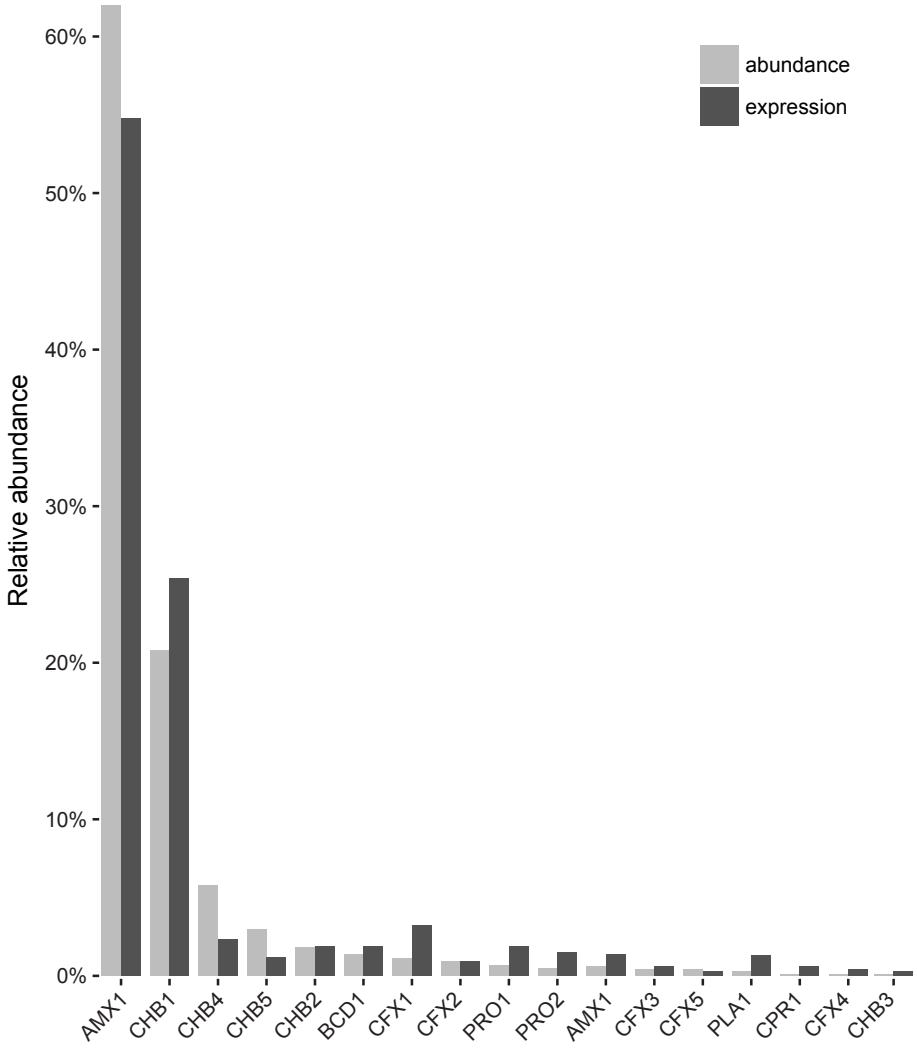

Figure 2.2 - Abundance and gene expression of organisms represented by the MAGs recovered from the anammox bioreactor. Abundance and gene expression estimates were based on RPKM values of metagenomic reads and transcripts that mapped to each MAG, respectively. See Table 2.3 for mapping details.

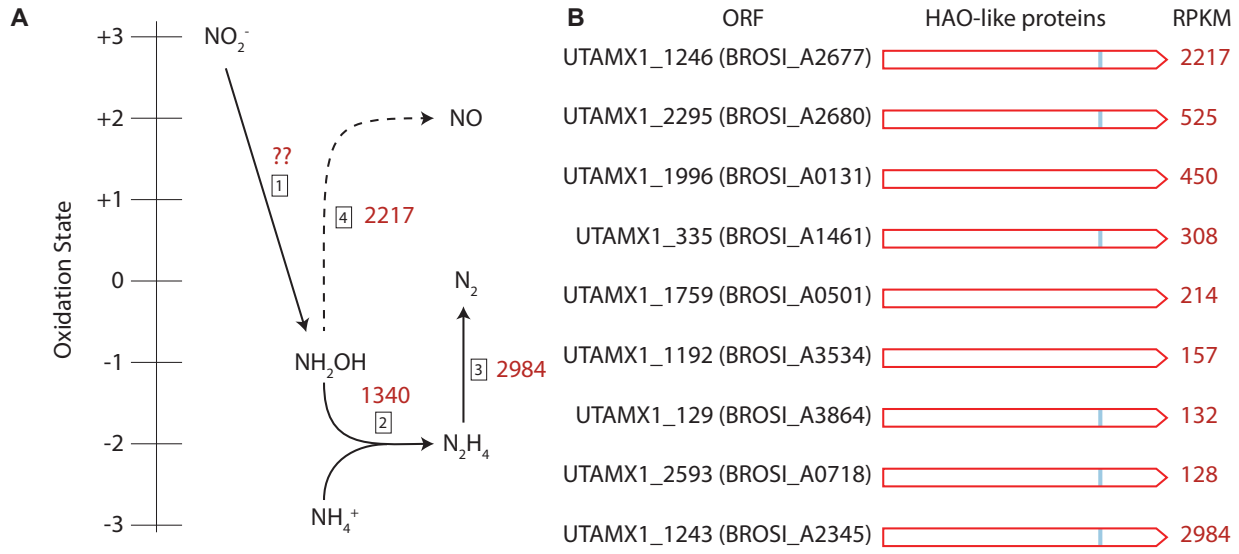

Figure 2.3 - A. Expression of genes involved in the anammox metabolic pathway recently proposed by Oshiki et al. (2016). Genes: [1] unidentified HAO-like protein involved in nitrite reduction, [2] hydrazine synthase subunit A, [3] hydrazine dehydrogenase, [4] hydroxylamine oxidoreducase (kust1061 ortholog). Red text indicates gene expression values (RPKM). B. HAO-like proteins identified in the UTAMX1 genome. '*Ca*. B. sinica' ortholog shown in brackets. Red text indicates gene expression values (RPKM). Blue line indicates cross-linking tyrosine residue.

2.2.5 Denitrification gene expression in anammox granules

All heterotrophic organisms in the anammox granules encoded capabilities for partial or full denitrification (Figure 2.4). This was similar to a recent report from a single-stage reactor by Speth et al.[16] Genes involved in denitrification were highly expressed in the genomes, consistent with oxidative phosphorylation coupled to nitrate respiration being a dominant form of heterotroph energy metabolism in the anammox granules (Supplementary Data 2; Supplementary Data 4). All heterotrophs expressed respiratory nitrate reductase genes (*narGHIJ*) that reduce nitrate to nitrite and require the transport of nitrate into the cytoplasm (Figure 2.4; Supplementary Data 4). However, it should be noted that the heme-containing membrane anchor subunit (*narI*) gene had no detectable expression in CHB2 or CFX2, suggesting that this enzyme had low activity in these

organisms. A periplasmic nitrate reductase (*napABCGH*) that does not contain a coupling site for proton motive force (PMF) generation was also expressed by PRO2. Pathways for dissimilatory nitrate reduction to ammonia (DNRA) via pentaheme nitrite reductase genes (*nrfHA*) were expressed by the Chlorobi genome CHB2 and the Chloroflexi genome CFX2 (Figure 2.4; Supplementary Data 4). Chloroflexi (CFX1, CFX2) and Proteobacteria (PRO1, PRO2) genomes also expressed genes for the reduction of nitrite to nitric oxide, either through a copper containing nitrite reductase (*nirK*) or cytochrome *cd1* nitrite reductase (*nirS*) (Figure 2.4; Supplementary Data 4). The reduction of nitric oxide to nitrous oxide, either through the cytochrome *c*-dependent nitric oxide reductase (*norBC*) or the quinol-dependent nitric oxide reductase (*norZ*), was expressed by the BCD1, CFX2, and PRO1 genomes. Finally, Chlorobi (CHB1, CHB2), Proteobacteria (PRO1, PRO2), and the Bacteroidetes (BCD1) genomes expressed genes capable of reducing nitrous oxide to nitrogen gas via nitrous oxide reductase (*nosZ*). Together, these functions could facilitate a nitrite loop with anammox bacteria or support complete denitrification, thus enhancing overall nitrogen removal performance in the bioreactor.

2.2.6 Carbon and energy metabolism of heterotrophic bacteria

Figure 4 shows the gene expression profiles of major carbon and energy metabolic pathways across the anammox community. Only two low abundance heterotrophic organisms, BCD1 and PRO2, encoded known pathways for CO_2 fixation that would permit mixotrophic growth (Figure 2.5; Supplementary Data 3). BCD1 displayed above median gene expression of PEP carboxylase (UTBCD1_1894), which allows CO_2 fixation into oxaloacetate (Supplementary Data 2); PRO2 displayed above median gene expression of Ribulose-1,5-bisphosphate carboxylase (RuBisCO, UTPRO2_2197), which would permit CO_2 fixation via the Calvin-Benson-Bassham cycle (Figure 2.5; Supplementary Data 3). While all heterotrophs, except CHB1, expressed genes encoding

pyruvate:flavodoxin/ferredoxin oxidoreductase (PFOR) and 2-oxoglutarate ferredoxin oxidoreductase (KFOR) that are commonly implicated in the reductive TCA cycle, other key genes of this pathway, such as ATP-citrate lyase and fumarate reductase, were missing from the genomes. As PFOR and KFOR are known to participate in other metabolic pathways (for example, KFOR may participate in the oxidative TCA cycle[27]), it is likely that these genes are not used for CO_2 fixation by heterotrophs in the community.

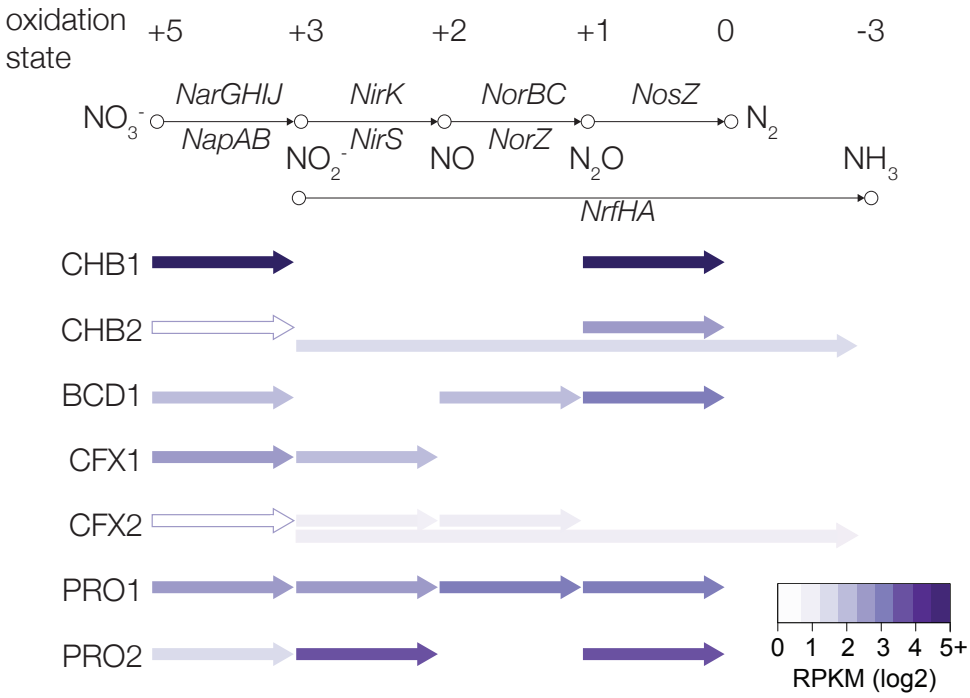

Figure 2.4 - Presence and expression of denitrification genes across the recovered heterotrophic genomes. Purple arrows indicate gene presence. Color intensity represents gene expression (log2 RPKM), based on mapping of metatranscriptomic reads to the metagenomic assembly. An outlined white arrow indicates one or more enzyme subunits had no detectable gene expression. A summary of genes involved in denitrification across the recovered genomes can be found in Supplementary Data 4.

All organisms expressed genes involved in central carbon metabolic pathways, including glycolysis/gluconeogenesis, glycogen synthesis/degradation, the TCA cycle, and the pentose

phosphate pathway (PPP) (Figure 2.5; Supplementary Data 3). Surprisingly, the oxidative branch of the PPP was completely missing in both Chlorobi bacterium genomes (CHB1 and CHB2), suggesting that other enzymes are important for NADPH generation in these organisms. Genes for the oxidative PPP were also confirmed to be missing from the closely related Chlorobi bacterium OLB4 and OLB6 genomes[16], although they were present in the genomes of *Ignavibacterium album*[28] and *Melioribacter roseus*[29]. An alternative route for NADPH generation in CHB1 and CHB2 is likely through $NAD(P)^+$ transhydrogenase[30], which transfers electrons from NADH to $NADP^+$ and had above median gene expression in both genomes (Supplementary Data 2).

Genes encoding pathways for pyruvate fermentation to acetate and CO_2 via PFOR, phosphate acetyltransferase (*pta*) and acetate kinase (*ack*) were expressed in the CHB2 and PRO2 genomes (Figure 2.5; Supplementary Data 3). While these genes could facilitate ATP synthesis via substrate level phosphorylation, it is also possible that the *pta-ack* reactions encoded by CHB2 and PRO2 are used for acetate consumption rather than production. These reactions form a lower affinity pathway for acetate assimilation compared to the high-affinity acetyl-CoA synthetase (*acs*) pathway[31]. Under this scenario however, only PRO2 could use acetate as a sole carbon source because it is the only genome that encodes the glyoxylate cycle (i.e., malate synthase and isocitrate lyase). The AMP-forming *acs* gene was expressed in all genomes, whereas the ADP-forming *acs* gene was expressed in CHB1, CHB2, CFX1, CFX2, and PRO2 (Supplementary Data 2; Supplementary Data 3). Since these reactions are also reversible, with acetate formation resulting in ATP conservation, it is unclear whether the *acs* and/or *pta-ack* acetate pathways expressed by the microbial community are involved in acetate formation, acetate assimilation, or both.

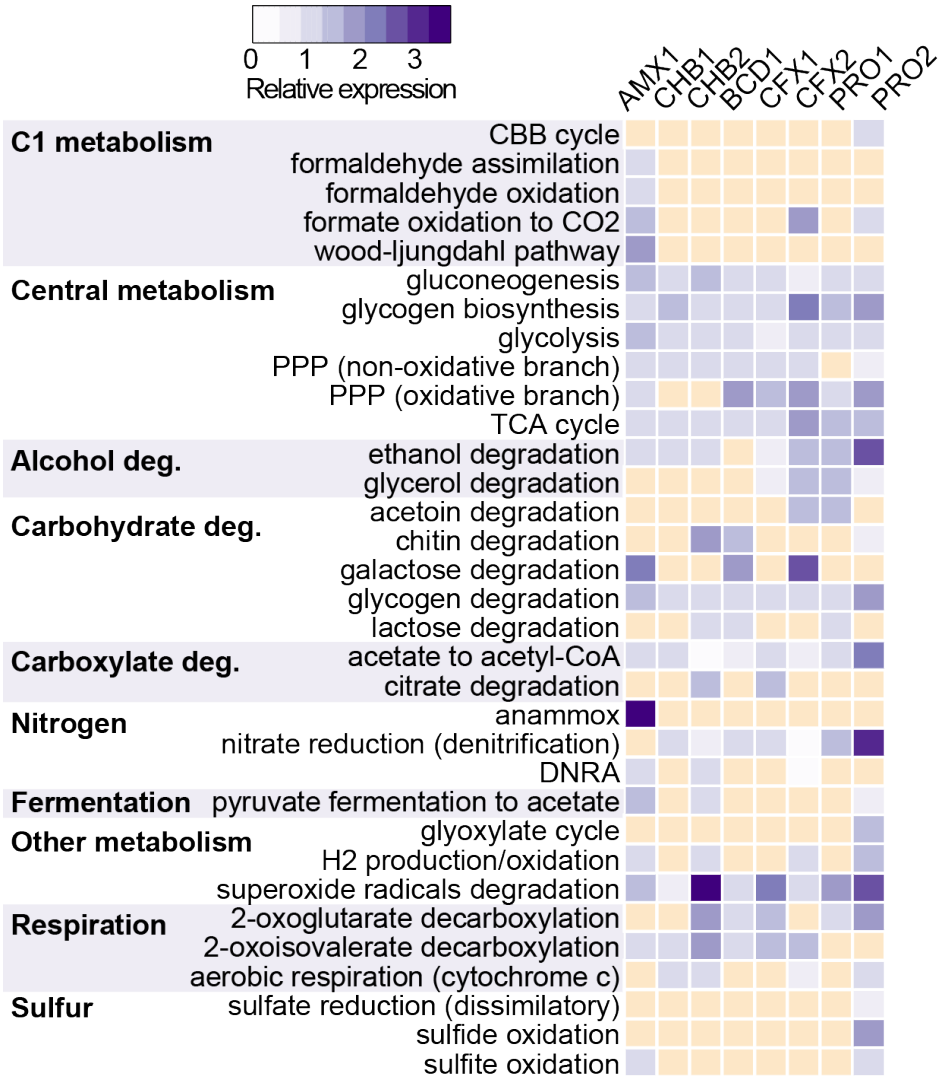

Figure 2.5 - Relative gene expression of major carbon and energy metabolic pathways encoded by each MAG. Color intensity represents gene expression, based on mapping of metatranscriptomic reads to the metagenomic assembly. Gene expression was relativized by median RPKM values calculated across each open reading frame in a given MAG (see Methods). A value of 1 equals median expression in a given genome. Yellow box indicates pathway absence. Anammox metabolic pathway in the AMX1 genome had a relative expression value of 15. See Supplementary Data 3 for a detailed summary of all reconstructed metabolic pathways.

The CHB2 genome also expressed an eight-gene cluster (UTCHB2_2599 to UTCHB2_2606) ferredoxin:NAD+ oxidoreductase (*rnfCDGEAB)* complex and two putative electron-bifurcating hydrogenase gene clusters (*hydABC)* (UTCHB2_709-711 and UTCHB2_1049-1051). The Rnf

complex could couple the formation of a proton gradient with ATP synthesis by oxidizing reduced ferredoxin produced from the electron-bifurcating hydrogenase and/or PFOR[32]. This energy conservation mode was also identified in the closely related *I. album*[28] and *M. roseus*[29] genomes, however no orthologs were detected in the CHB1 genome.

The ability to use H_2, sulfite, formate, ethanol, and other simple organic compounds as electron donors could also support the energy metabolism of the community (Figure 2.5; Supplementary Data 3). CHB2, CFX2, and PRO2 expressed genes encoding a NAD-dependent hydrogenase that could be involved in H_2 oxidation. CFX2 also expressed genes encoding a formate hydrogenlyase (FML) complex (UTCFX2_2421-2429) that directly links formate oxidation to H_2 production[33]. Organisms encoding the ability to use formate for energy conservation included CFX2 and PRO2 (Figure 2.5; Supplementary Data 3). These organisms expressed genes encoding NAD-dependent formate dehydrogenase that catalyzes the oxidation of formate to CO_2, donating electrons to NAD+ that could be used to generate a PMF. PRO2 also expressed genes encoding formate dehydrogenase O (FDH-O) (UTPRO2_1096 – UTPRO2_1099) that would donate electrons to the quinone pool and could also be used for PMF generation[34]. Because no organism encoded the ability to produce formate via pyruvate formate lyase (PFL), it is possible that formate is made available to the community via CO_2 reduction to formate by AMX1, which is the first step in CO_2 fixation via the reductive acetyl-CoA pathway.

Unexpectedly, genes encoding aerobic respiration machinery were expressed by some community members, despite the lack of oxygen in the bioreactor. CHB1, CFX2, and PRO2 expressed genes encoding a low-affinity aa3-type terminal cytochrome c oxidases; CHB2 expressed a high-affinity cytochrome d ubiquinol oxidase gene (Figure 2.5; Supplementary Data 3). While this could allow these organisms to respire oxygen under aerobic conditions, it is also

possible that these oxidases have evolved for protection against oxygen[35], perform some unknown function, or are post-transcriptionally repressed. All organisms were also found to express genes encoding superoxide dismutase (SOD) (Figure 2.5; Supplementary Data 3), which may additionally function to protect against oxidative stress.

2.2.7 Amino acid and carbohydrate catabolism

The breakdown of EPS produced by anammox bacteria has been proposed to be a major organic carbon source supporting heterotrophic growth in anammox granules[11,17]. In agreement with this, genes encoding a wide range of peptidases were expressed in the heterotrophic genomes (Figure 2.6; Supplementary Data 5). In particular, CHB1 and CHB2 encoded many extracellular subtilisin-like serine peptidases and metallopeptidases (Figure 2.6; Supplementary Data 5). These genes were among the highest expressed ORFs in the genomes and may be involved with growth on proteinaceous substrates[36,37]. Consistent with this function, the CHB1 and CHB2 genomes expressed genes for the transport and catabolism of peptides and amino acids to central carbon intermediates (Figure 2.6; Figure 2.7; Supplementary Data 2; Supplementary Data 3). This would allow CHB1 and CHB2 to use amino acids as a carbon and energy source, in addition to direct assimilation into protein biosynthesis. Interestingly, ranking of the amino acids according to their biosynthetic cost[38] revealed that all organisms in the anammox granules lacked catabolic pathways for many amino acids with a high biosynthetic cost (Figure 2.7; Supplementary Data 3). This may suggest that heterotrophic community members have been selected to preferentially degrade amino acids that are less costly to synthesize, while directly incorporating more costly amino acids into protein synthesis. Other heterotrophic genomes (BCD1, CFX1, CFX2, PRO2, and PRO1) also expressed genes involved in the transport and degradation of amino acids (Figure 2.6; Figure 2.7;

Supplementary Data 2). These organisms could potentially take advantage of substrates made available by CHB1 and CHB2, or contribute to extracellular protein degradation as well.

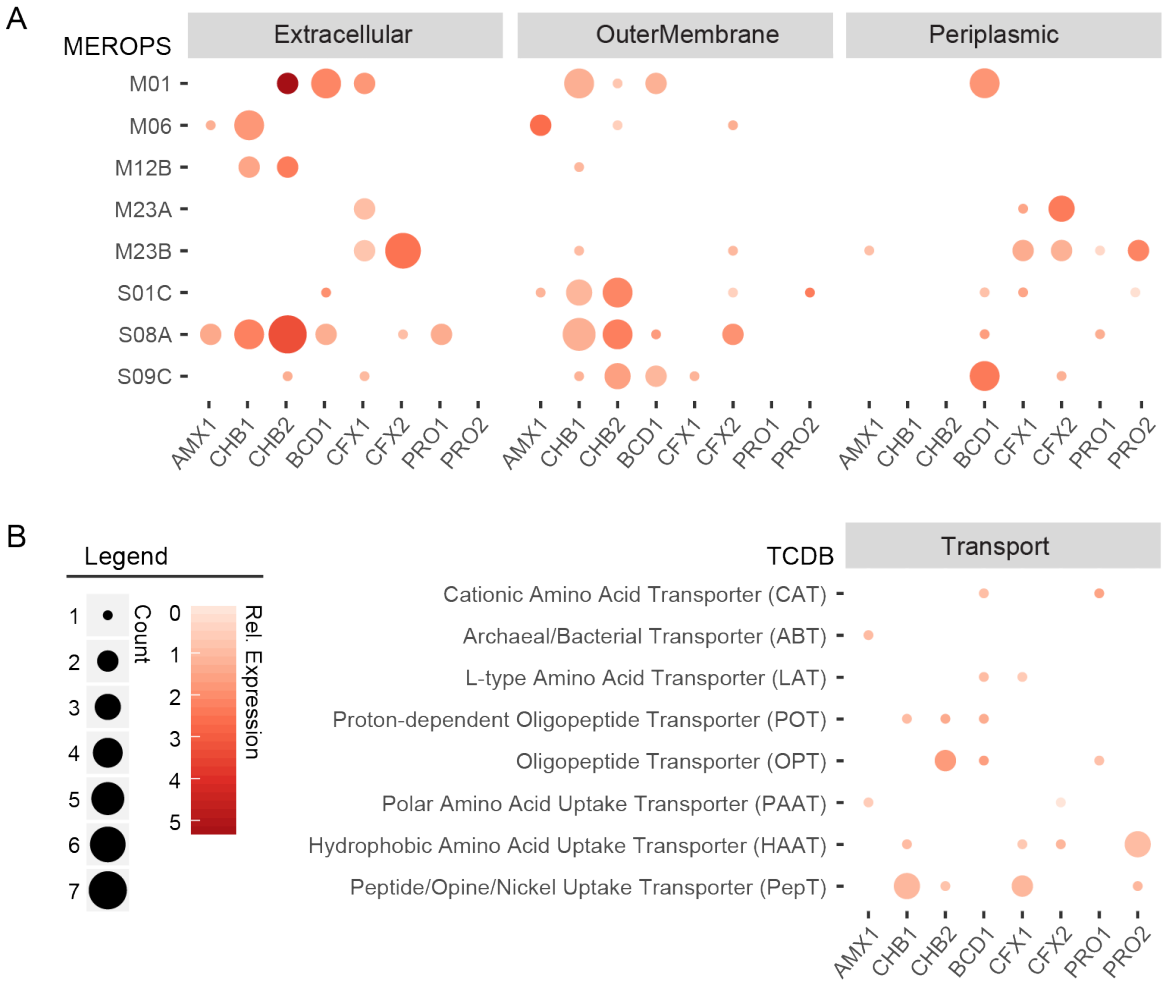

Figure 2.6 - Predicted peptidases and amino acid transporters recovered from the MAGs. (A) Number (bubble diameter) and relative gene expression (bubble color intensity) of selected peptidases possibly involved in EPS matrix protein degradation. Peptidases were annotated against the MEROPS database[66]. The subcellular location (extracellular, outer membrane, or peroplasm) of each peptidase was predicted using the subcellular localization predictor (CELLO)[69]. A summary of all predicted peptidases can be found in Supplementary Data 5. (B) Number (bubble diameter) and relative gene expression (bubble color intensity) of amino acid and peptide transporters predicted across the recovered genomes. Transporters were annotated against the transporter classification database (TCDB)[70] and can be found in Supplementary Data 2.

Microbial community members also expressed diverse genes involved in the hydrolysis of carbohydrate bonds (Supplementary Data 6). In particular, CHB2 and BCD1 expressed many ORFs encoding glycoside hydrolases, some which were predicted to be extracellular. The high number of these genes in the CHB2 genome appeared to be a major function differentiating it from the CHB1 genome, which had the lowest number of predicted glycoside hydrolyases in the community (Supplementary Data 6). CHB2 expressed several genes that may act on glycosidic bonds found in polysaccharides, whereas BCD1 expressed a range of glycoside hydrolase genes that may use oligosaccharides, mucins, and glycolipids as substrates and play an important role in carbohydrate breakdown for the community.

2.2.8 Vitamin and amino acid auxotrophy

Metabolite exchange of amino acids and vitamins is known to shape microbial community assembly[39]. We observed that several of the abundant heterotrophic bacteria, namely CHB1, CHB2, CFX1, and CFX2, were missing pathways for the synthesis of many hydrophobic amino acids (Figure 2.7; Supplementary Data 3). This was in contrast to the AMX1 genomes and other heterotrophic genomes (BCD1, PRO1, PRO2) that expressed pathways for the synthesis of most amino acids.

Several of the heterotrophic bacteria were also missing key genes involved in B-vitamin biosynthesis. The CHB1, CHB2, and CFX2 genomes were missing key genes involved in thiamin (vitamin B1) biosynthesis (thiamine-phosphate synthase and thiamine-monophosphate kinase), biotin (vitamin B7) biosynthesis (Adenosylmethionine-8-amino-7-oxononanoate aminotransferase and bioin synthase), and adenosylcobalamin (vitamin B12) biosynthesis (cobalamin synthase and adenosylcobinamide-phosphate synthase) (Supplementary Data 3). CFX1, BCD1, PRO2, and PRO1 also lacked key genes involved in adenosylcobalamin *de novo*

synthesis. These genes were expressed in the AMX1 genome, suggesting that anammox bacteria may support B-vitamin requirements for the community. Genes encoding thiamin ABC-type transporters were expressed in the PRO1 and PRO2 genomes, whereas the CFX1, CFX2, PRO1, and PRO2 genomes expressed ABC-type transporters for vitamin B12 (*btuFCD*) (Supplementary Data 2). The AMX1 (UTAMX1_2357), CHB1 (UTCHB1_731), CHB2 (UTCHB2_2713), and BCD1 (UTBCD1_3076) genomes expressed genes encoding the TonB-dependent vitamin B12 transporter (*btuB*), which facilitates vitamin B12 translocation across the outer membrane and may also be involved in uptake.

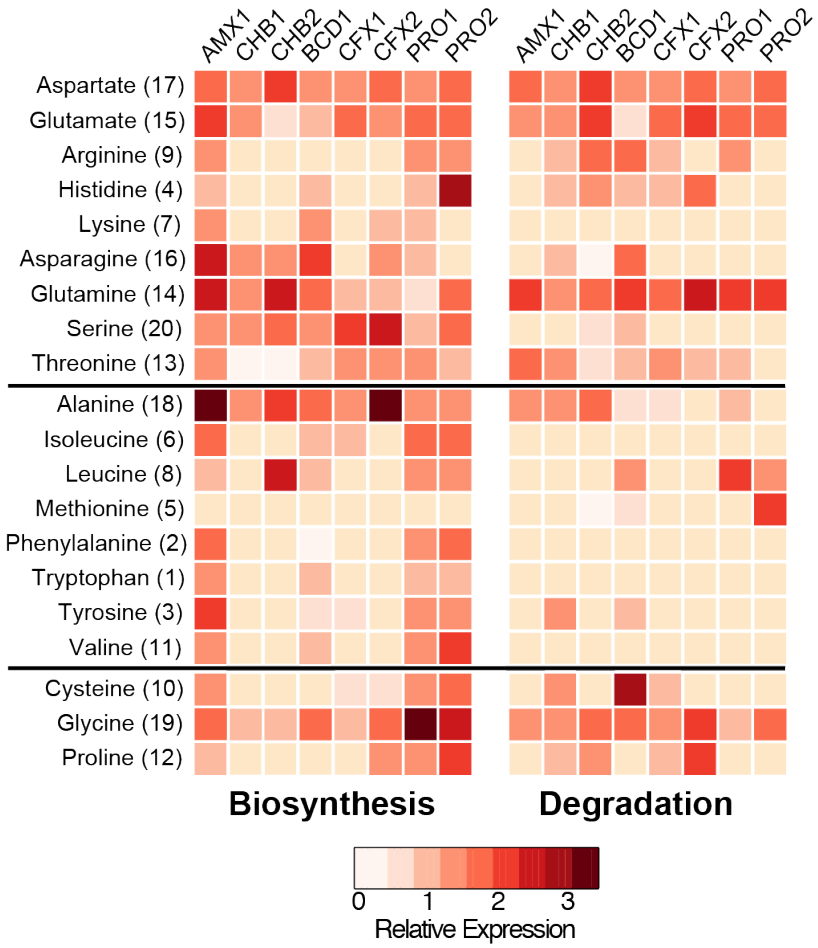

Figure 2.7 - Relative gene expression of amino acid biosynthetic and degradation pathways encoded by each MAG. Red color intensity represents gene expression, which was relativized

by median RPKM values calculated across each open reading frame in a given MAG. A value of 1 equals median expression. Top, middle, and bottom panels separate amino acids by hydrophilic, hydrophobic, and special structured side chains, respectively. Bracketed numbers rank the metabolic cost of amino acid biosynthesis based on values reported by Akashi and Gojobori[38], with 1 being the most costly.

2.3 Discussion

The integration of metagenomic and metatranscriptomic sequencing allowed us to examine gene expression and microbial interactions in a lab-scale anammox community at the ecosystem-scale. Our results revealed that anammox bacteria affiliated with Brocadia (AMX1) and heterotrophic bacteria affiliated with Chlorobi (CHB1) dominated the abundance and gene expression of the anammox granule community. Other less abundant heterotrophic bacteria affiliated with Chloroflexi, Bacteroidetes, and Proteobacteria also displayed significant gene expression that may contribute to ecosystem function. The high similarity between these genomes and the genomes recently recovered from a single-stage anammox reactor[16] suggests that a core microbiome exists in anammox-based wastewater treatment systems, despite differences in reactor operation and influent wastewater composition.

Remarkably, the CHB1 genome was nearly identical to the Chlorobi bacterium OLB4 genome recovered by Speth et al[16] (99.8% ANI) that achieved the second highest abundance in a full-scale partial-nitritation anammox reactor from the Netherlands[16,40], suggesting that this organism interacts consistently with *Brocadia sp.* in anammox-based wastewater treatment systems. Figure 2.8 summarizes the major metabolic interactions proposed between AMX1 and CHB1 in anammox granules, based on the metabolic reconstruction and gene expression results obtained in this study. We posit that the ecological role of CHB1 in anammox granules is to degrade and catabolize extracellular peptides bound in the EPS matrix, while respiring nitrate

produced during anammox bacterial growth to nitrite. This is supported by the high gene expression of respiratory nitrate reductase and extracellular peptidases in the CHB1 genome, as well as transporters and genes involved in amino acid catabolism. Interestingly, the CHB1 genome lacked catabolic pathways for amino acids that carry a high biosynthetic cost, suggesting that less costly amino acids (e.g., glutamine) may be preferentially used as carbon and energy sources by heterotrophic bacteria in anammox granules. Such preferential degradation of amino acids has also been observed in anoxic water columns with sinking particulate organic carbon[41].

Because CHB1 lacks many pathways for amino acid biosynthesis, it is likely that peptide and amino acid substrates used as carbon and energy sources originate from *Brocadia sp.*, which derives carbon via CO_2 fixation and expressed pathways for the synthesis of nearly all amino acids (Figure 2.7). This is consistent with ^{14}C-bicarbonate tracer experiments that showed heterotrophic bacteria in anammox bioreactors fed with no external organic carbon compounds degraded and utilized cellular components produced by anammox bacteria[42]. The high transcript abundance of CHB1 in the metatranscriptome (25% relative community expression) also suggests that proteolysis by this organism is a significant component of carbon flux in anammox granules, possibly making amino acids available to other microbial community members. The number and high gene expression of amino acid transporters across the other heterotrophic genomes supports this interaction, suggesting that protein degradation by CHB1 may be a significant contributor to community assembly in anammox granules, similar to particle degrading taxa in marine ecosystems[43]. It is also possible that *Brocadia sp.* supply B-vitamins to members of the microbial community, based on the expression of key genes involved in their biosynthesis and auxotrophies present across the heterotrophic genomes. The exchange of B-vitamins has been postulated to shape the structure and function of some microbial communities, such as those harboring the

human gut[44]. The specific role vitamin exchange plays in shaping anammox community interactions currently remains unknown, and should therefore be investigated further in future studies. The oxidization of amino acids coupled to the reduction of nitrate to nitrite by CHB1 would serve to enhance overall nitrogen removal in the ecosystem through the completion of a nitrite loop with *Brocadia sp*., as proposed by Speth et al.[16] Amino acid catabolism by CHB1 would also result in the release of ammonia, allowing it to be removed from the system by *Brocadia sp*., thus further enhancing overall nitrogen removal. It is also possible that nitrate reduction to nitrite by CHB1 and other heterotrophic organisms helps stimulate anammox metabolism. Recent microsensor studies have shown that nitrite concentrations can become limiting in the interior of anammox granules[45]. Therefore, the ability of CHB1 to rapidly recycle nitrate back to nitrite may be critical for supporting anammox growth, which is required for their own proliferation.

In addition to a nitrite loop, heterotrophic bacteria in anammox ecosystems may facilitate complete denitrification, either alone or based on metabolite exchange of denitrification intermediates with other organisms. Indeed, the segregation of denitrification intermediates across different organisms has been shown to reduce the accumulation of growth-inhibiting metabolites (e.g. nitrite) and potentially accelerate denitrification rates[46]. The close spatial association of organisms in anammox granules may encourage such a distributed metabolic network[47,48]. While complete denitrification would also serve to remove nitrogen from the ecosystem, it may potentially be detrimental to anammox bacterial growth because it would consume nitrite, rather than make it available for *Brocadia sp.* Specifically, the heterotrophic organisms capable of using nitrite as an electron acceptor (CHB2, CFX1, CFX2, PRO1, PRO2) via nitrite reductase (NirK/S

or NrfHA) may compete with *Brocadia sp.*, similar to nitrite oxidizing bacteria (NOB) in single-stage partial nitritation-anammox bioreactors.

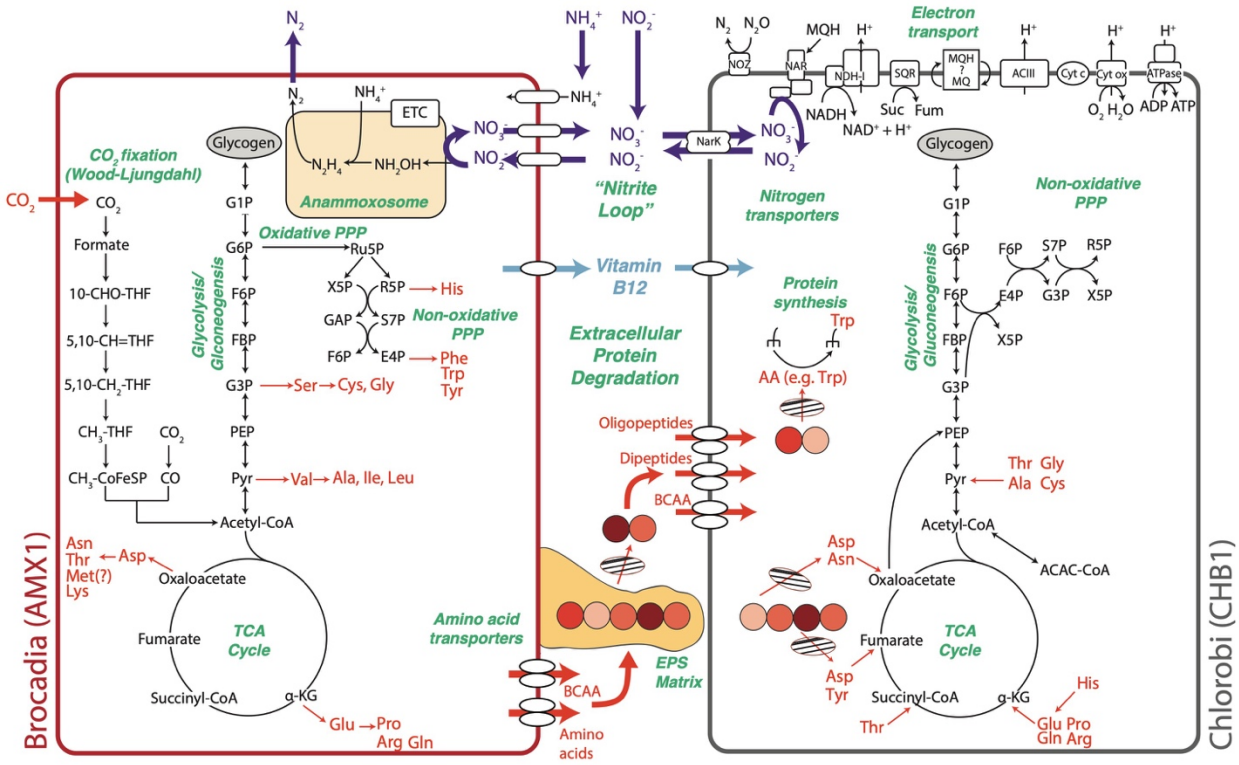

Figure 2.8 - Proposed metabolic interactions between *Brocadia* (AMX1) and *Chlorobi* (CHB1) in anammox granules. AMX1 fixes CO_2 and synthesizes amino acids and extracellular polymeric substances (EPS). CHB1 degrades proteins bound in EPS using extracellular peptidases and subsequently transports and catabolizes short peptides (circles) to central carbon intermediates. Nitrite oxidation and reduction by AMX1 and CHB1, respectively, results in a distributed nitrite loop. Purple arrows indicate nitrogen cycling; orange arrows indicate carbon cycling; light blue arrows indicate vitamin B metabolite exchange. Hatched ovals indicate peptidases. BCAA, branched-chain amino acids. The presence of a periplasm has been ignored for clarity of the schematic.

Aside from proteinaceous substrates, many organisms in the anammox community were capable of using other electron donors for energy conservation, including acetate, H_2, formate, or simple organic compounds. It is possible that these substrates are produced during fermentation

reactions and/or the breakdown of polysaccharide compounds present in the EPS. Chlorobi organisms related to CHB2 may be the biggest contributor to this process, based on expression of PFOR, hydrogenases, and ADP-forming ACS involved in H_2 and acetate production. The availability of such substrates may contribute to niche differentiation by heterotrophs in the community, while also providing additional reducing equivalents to fuel anammox metabolism[49]. Nevertheless, oxidative phosphorylation coupled to nitrate respiration, rather than fermentation, was found to be the main energy metabolism expressed across the heterotrophic genomes, consistent with the high availability of nitrate and nitrite in the ecosystem. As such, the significance of H_2 and acetate metabolism in anammox granules still remains to be determined.

In conclusion, our combined approach of genome-centric metagenomics with metatranscriptomics allowed us to obtain a much deeper understanding of ecosystem function in anammox granules that has broad implications for anammox-based wastewater treatment. This study provides a robust analysis of the metabolic networks underlying several poorly characterized taxa frequently detected in anammox granules (Planctomycetes, Chlorobi, Bacteroidetes, Chloroflexi, and Proteobacteria) and also sheds light on their potential ecological roles and interactions. Our results implicate members of the Chlorobi affiliated with CHB1 that are broadly found in anammox granules as highly active protein degraders, possibly liberating amino acids bound in the EPS matrix for themselves and other organisms in the community. These amino acids likely originate from anammox bacteria, which also provide nitrate and potentially essential B-vitamins to the community, supporting anaerobic respiration and growth. Other heterotrophic community members also appear to contribute to the scavenging of detritus and peptides produced by anammox bacteria, and may potentially use alternative electron donors, such as H_2, acetate, and

formate. These substrates may be made available through the hydrolysis and fermentation of EPS and detritus.

While further studies are still required to determine the specific niche space occupied by members of the community, the metabolic characterization provided here advances our insight on the ecological roles played by both anammox and heterotrophic bacteria in anammox granules. The high expression of enigmatic functions by the community, including hypothetic proteins and HAO-like proteins, together with our incomplete understanding of the biochemical pathways underlying EPS biosynthesis, highlights the need for continued work on these novel and industrially important ecosystems.

2.4 Methods

2.4.1 Bioreactor operation

A 5L continuously fed sequencing batch reactor (anammox bioreactor) has been operating in the Goel Laboratory at the University of Utah for the past 6 years, achieving stable anammox performance[50]. The anammox bioreactor was originally inoculated with anammox biomass from the City College of New York (Civil Engineering Department) that was enriched from activated sludge. The bioreactor was fed with anaerobic digester filtrate obtained from the belt filter press of a local wastewater treatment plant (CVWRF, Salt Lake City, UT). Prior to entering the anammox reactor, the filtrate was first passed through a nitritation reactor, where approximately half of the influent ammonia was oxidized to nitrite. The nitritation reactor had a working volume of 2L with a hydraulic retention time of 1 day[50]. Dissolved oxygen in the reactor was maintained at 0.5±0.5 mg/L and bicarbonate present in the influent filtrate kept the pH buffered between 6.5-8.5. The anammox bioreactor was operated at room temperature with a hydraulic retention time (HRT) of 2 days and no biomass wasting, which encouraged the formation of dense granules.

The pH of the anammox bioreactor was maintained at between 6.8-8.0 and the reactor contents were completely mixed using a stir bar and gas sparging. Anaerobic conditions in the anammox bioreactor were maintained by continuously purging a mixture of 95% nitrogen gas and 5% carbon dioxide[50]. Table 2 summarizes the process performance data of the nitritation and anammox bioreactors during steady-state operation.

2.4.2 DNA and RNA sequencing

Two independent biomass samples were collected from the anammox bioreactor at different time points during periods of high total nitrogen removal efficiency (> 80%). DNA was extracted separately from each biomass sample to improve the recovery of metagenome-assembled genomes (MAGs) and total RNA was extracted from the second biomass sample to examine steady-state community gene expression. 500 mg of biomass was collected from the reactor, centrifuged at 4500 rpm for 5 minutes at 4^{o}C to remove the supernatant, and instantly flash frozen in liquid nitrogen. Total genomic DNA and RNA was extracted from biomass pellets using the PowerMax™ Soil DNA Isolation Kit (MoBio Laboratories, USA) and the PureLink RNA mini kit (Life Technology, NY), respectively, according to the manufacturers' protocols. Genomic DNA was quality-checked using agarose gel electrophoresis and a Nanodrop ND-2000c (Thermo Fisher Scientific, US). Following RNA extraction, residual genomic DNA was removed from total RNA using an on-column PureLink DNase set (Life Technologies, NY). Total RNA quality and quantity was subsequently checked using the Bioanalyzer RNA 6000 Nano Assay (Agilent, Santa Clara, CA, USA) to ensure only high-quality nucleic acids were used for downstream analysis. 100ng of total RNA was used to construct strand specific RNA-Seq libraries with the Encore® Complete Prokaryotic RNA-Seq DR Multiplex System (NuGEN, San Carlos, CA). Non-rRNA in RNA-Seq libraries were enriched by selective priming during the first strand cDNA synthesis reaction, as

well as in the final library construction steps using manufacturer's protocols. DNA from the first sample was sequenced on the Illumina MiSeq platform (Illumina, California, USA) to generate 300 bp paired-end reads (550 bp mean insert size). DNA and RNA from the second sample were sequenced on the Illumina HiSeq 2000 platform to generate 125 bp paired-end reads (180 bp mean insert size) at greater sequencing depth. All DNA and RNA sequencing was performed at the Huntsman Cancer Institute (HCI), University of Utah. Raw DNA and RNA sequences can be found on the National Center for Biotechnology Information (NCBI) website under BioProject PRJNA343219.

2.4.3 Metagenomic assembly and binning

Raw paired-end reads from the MiSeq and HiSeq platforms were initially filtered using Sickle v1.33[51] based on a minimum quality score of 20, a minimum sequence length of 100 bp, and allowing for no ambiguous bases. Paired-end reads were then merged using FLASH v1.2.11[52] and co-assembled using the *de novo* assembler of CLC Genomics Workbench v7.0.3 (CLCbio, Arhus, Denmark) based on default parameters (word size = 20, bubble size = 50). The per-base coverage depth across all contigs was calculated by mapping raw reads from each sample against the co-assembled contigs using BBMap v35.92 (https://sourceforge.net/projects/bbmap/) with the parameters *'minid=0.95'* and *'ambig=random'*. Resulting mapping files were subsequently used by MetaBAT v0.26.3 to bin metagenomic contigs into draft genomes, based on the *'sensitive'* parameters[53]. CheckM v1.0.3[54] was used to estimate the contamination and completeness of each draft genome based on 111 essential single-copy marker genes[55]. Draft genomes have been deposited to GenBank under the accession numbers provided in Table 2.4.

Table 2.4 - Metagenome-assembled genome DDBJ/ENA/GenBank accession numbers.

MAG	BioSample	GenBank Accession	Organism
AMX1	SAMN06342761	MWTF00000000	Candidatus Brocadia sp. UTAMX1
AMX2	SAMN06342762	MWTE00000000	Candidatus Brocadia sp. UTAMX2
BCD1	SAMN06342764	MWTC00000000	Sphingobacteriales bacterium UTBCD1
CFX1	SAMN06342765	MWTB00000000	Anaerolineae bacterium UTCFX1
CFX2	SAMN06342766	MWTA00000000	Anaerolineae bacterium UTCFX2
CFX3	SAMN06342767	MWSZ00000000	Anaerolineae bacterium UTCFX3
CFX4	SAMN06342768	MWSY00000000	Chloroflexi bacterium UTCFX4
CFX5	SAMN06342769	MWSX00000000	Anaerolineae bacterium UTCFX5
CHB1	SAMN06342770	MWSW00000000	Ignavibacteriales bacterium UTCHB1
CHB2	SAMN06342771	MWSV00000000	Ignavibacteriales bacterium UTCHB2
CHB3	SAMN06342772	MWSU00000000	Ignavibacteriales bacterium UTCHB3
CPR1	SAMN06342775	MWSR00000000	Microgenomates bacterium UTCPR1
PLA1	SAMN06342763	MWTD00000000	Planctomycetes bacterium UTPLA1
PRO1	SAMN06342773	MWST00000000	Polyangiaceae bacterium UTPRO1
PRO2	SAMN06342774	MWSS00000000	Rhodocyclaceae bacterium UTPRO2

2.4.4 Phylogenetic analysis of recovered draft genomes

Phylogenetic analysis of the recovered draft genomes was accomplished using Phylosift v1.0.1, based on a set of 37 universal single-copy marker genes[56]. The taxonomic affiliation of each draft genome was determined using the Phylosift 'all' command. Marker genes were also identified in 33 publicly available genomes closely related to the recovered draft genomes and used to build a phylogenetic tree. Marker genes were concatenated and aligned using Phylosift and a maximum likelihood tree was generated using RAxML v8.2.4 with the automatic protein model assignment algorithm (PROTGAMMAAUTO) and 100 bootstraps.

2.4.5 Genome annotation and metabolic reconstruction

Metabolic reconstruction of the recovered draft genomes was performed using MetaPathways v2.5[20]. Briefly, open reading frames (ORFs) were predicted using Prodigal v2.0[57], based on a

minimum nucleotide length of 60, and queried against the SEED subsystems (accessed March 2013), Clusters of Orthologous Groups (COG, accessed December 2013), RefSeq (accessed September 2015), and MetaCyc (accessed October 2011) protein databases using the optimized LAST algorithm (E value, E-6) for functional annotation[58]. Contig nucleotide sequences were also queried against the SILVA small subunit (SSU) 123 database to identify the taxonomy of recovered 16S ribosomal rRNA (rRNA) genes. Annotated genomes were then used to reconstruct the metabolic network of each organism using Pathway Tools and the MetaCyc database[59]. Pathway/Genome Databases (PGDB) were created for each genome, where pathway inference was based on a set of rules used by the Pathway Tools prediction algorithm Pathologic, including the presence of all key reactions and the completeness of the reconstructed pathway[60]. All inferred pathways were then manually curated to verify predictions made by Pathologic.

2.4.6 Metatranscriptomic analysis

Metatranscriptomic reads were quality filtered and merged as described above. Subsequently, rRNA sequences were filtered from the metatranscriptomic dataset using SortMeRNA v2.0, based on multiple rRNA databases for bacterial, archaeal, and eukaryotic sequences[61]. Resulting non-rRNA reads were mapped back to all assembled metagenomic contigs using BBMap v35.92 with the parameters *'minid=0.95'*, which specifies a minimum alignment identify of 95% corresponding to the well-established criteria for identifying microbial species using average nucleotide identify[62] and *'ambig=random'*, which ensures reads with multiple top-scoring mapping locations are assigned randomly to a single location. Read counts were calculated for each predicted open reading frame (ORF) using *htseq-count* v0.6.1 with the *'intersection strict'* parameter[63] and normalized by sequencing depth and ORF length and expressed as reads per kilobase per million mapped reads (RPKM) values[64]. This allowed for the profiling of gene

expression across the recovered draft genomes. Relative gene expression within each genome was calculated by relativizing the expression (RPKM value) of each ORF by the median RPKM value calculated across the genome. Pathway expression levels were calculated based on averaging the RPKM values for each reaction in a given pathway. For reactions with multiple genes, the highest expressed gene between multi-copy genes was selected, whereas the lowest expressed gene between multi-gene enzyme complexes (e.g. hydrazine synthase, *hzsABC*) was selected.

The relative abundance and gene expression of the recovered genomes was calculated from the total number of DNA or mRNA reads that mapped to the genome, divided by the genome length (read count / genome size).

2.4.7 Carbohydrate hydrolase and peptidase identification

Carbohydrate hydrolases and peptidases potentially involved in the breakdown of EPS were identified in each genome based on BLASTP searches against the CAZy (accessed September 2014)[65] and MEROPS release 10.0[66] databases, respectively. Manual searches were also performed across all genomes based on gene annotations, and protein domains with hydrolytic activity were subsequently confirmed based on queries against the NCBI's Conserved Domain Database (CDD)[67] and Pfam database[68] using MOTIF (http://www.genome.jp/tools/motif/). The subcellular location of identified proteins was predicted using CELLO v2.5, which classifies proteins using support vector machines trained by multiple feature vectors based on *n*-peptide composition[69].

2.4.8 Data availability statement

All data associated with this project can be found at the NCBI under BioProject PRJNA343219. Illumina HiSeq metagenomic data can be found under BioSample SAMN05785375, Illumina

MiSeq metagenomic data can be found under BioSample SAMN05785373, and Illumina HiSeq metatranscriptomic data can be found under BioSample SAMN05785376. Annotated GenBank files for the whole genome sequences (MAGs) described in this study can be found under accession numbers listed in the Supplementary Information.

2.4.9 Supplementary Data

Supplementary Data 1. Genome comparison of the CHB1 and OLB4 genomes recovered from the lab-scale reactor employed in this study and the Olburgen PNA reactor (Speth et al., 2016), respectively.

Supplementary Data 2. Predicted open reading frame annotations and gene expression for each MAG recovered from the anammox community.

Supplementary Data 3. MetaCyc pathway/genome database for each MAG recovered from the anammox community.

Supplementary Data 4. Summary of genes and corresponding gene expression values (RPKM) involved in denitrification across MAGs.

Supplementary Data 5. Summary of predicted peptidases across MAGs based on annotations against the MEROPs database. * indicates peptidases identified using the CDD database or MOTIF.

Supplementary Data 6. Summary of genes predicted to be involved in carbohydrate metabolism based on annotations against the CAZy database.

Supplementary Information accompanies this chapter at http://www.nature.com/ncomms/

2.5 Acknowledgements

This work was supported by funding from the Central Valley Water Reclamation Facility (CVWRF) and the North Davis Sewer District awarded to RG and SW. Partial support was also provided by funding from the National Science Foundation (CBET-1435661 and MCB-1518130) awarded to DRN and KDM and a Postgraduate Scholarship-Doctoral (PGS-D) awarded by the National Sciences and Engineering Research Council of Canada (NSERC) to CEL. The authors acknowledge Masaru K. Nobu for helpful discussion with the manuscript.

2.6 References

1. Kartal, B., Kuenen, J. G. & van Loosdrecht, M. C. M. Sewage Treatment with Anammox. *Science (80-.).* **328,** 702–703 (2010).
2. Sliekers, A. O. *et al.* Completely autotrophic nitrogen removal over nitrite in one single reactor. *Water Res.* **36,** 2475–2482 (2002).
3. Van Dongen, U., Jetten, M. S. M. & Van Loosdrecht, M. C. M. The SHARON®-Anammox® process for treatment of ammonium rich wastewater. *Water Sci. Technol.* **44,** 153–160 (2001).
4. Strous, M. *et al.* Deciphering the evolution and metabolism of an anammox bacterium from a community genome. *Nature* **440,** 790–4 (2006).
5. Kartal, B. *et al.* Molecular mechanism of anaerobic ammonium oxidation. *Nature* **479,** 127–30 (2011).
6. Kartal, B. *et al.* How to make a living from anaerobic ammonium oxidation. *FEMS Microbiol. Rev.* **37,** 428–61 (2013).
7. Strous, M. *et al.* Missing lithotroph identified as new planctomycete. *Nature* **400,** 446–449 (1999).
8. Abma, W. R. *et al.* Full-scale granular sludge Anammox process. *Water Sci. Technol.* **55,** 27–33 (2007).
9. Ni, B. J. *et al.* Microbial and physicochemical characteristics of compact anaerobic ammonium-oxidizing granules in an upflow anaerobic sludge blanket reactor. *Appl. Environ. Microbiol.* **76,** 2652–2656 (2010).

10. Gonzalez-Gil, G., Sougrat, R., Behzad, A. R., Lens, P. N. L. & Saikaly, P. E. Microbial Community Composition and Ultrastructure of Granules from a Full-Scale Anammox Reactor. *Microb. Ecol.* **70,** 118–131 (2014).
11. Hou, X., Liu, S. & Zhang, Z. Role of extracellular polymeric substance in determining the high aggregation ability ofanammox sludge. *Water Res.* **75,** 51–62 (2015).
12. Ali, M. *et al.* Physiological characterization of anaerobic ammonium oxidizing bacterium 'Candidatus Jettenia caeni'. *Environ. Microbiol.* **17,** n/a-n/a (2015).
13. Ma, Y., Sundar, S., Park, H. & Chandran, K. The effect of inorganic carbon on microbial interactions in a biofilm nitritation-anammox process. *Water Res.* **70,** 246–254 (2015).
14. Li, X. R. *et al.* The bacterial diversity in an anaerobic ammonium-oxidizing (anammox) reactor community. *Syst. Appl. Microbiol.* **32,** 278–289 (2009).
15. Gonzalez-Martinez, A. *et al.* Microbial community analysis of a full-scale DEMON bioreactor. *Bioprocess Biosyst. Eng.* **38,** 499–508 (2015).
16. Speth, D. R., in 't Zandt, M. H., Guerrero-Cruz, S., Dutilh, B. E. & Jetten, M. S. M. Genome-based microbial ecology of anammox granules in a full-scale wastewater treatment system. *Nat. Commun.* **7,** 11172 (2016).
17. Ni, B. J., Ruscalleda, M. & Smets, B. F. Evaluation on the microbial interactions of anaerobic ammonium oxidizers and heterotrophs in Anammox biofilm. *Water Res.* **46,** 4645–4652 (2012).
18. Liu, Y. *et al.* Assessment of Heterotrophic Growth Supported by Soluble Microbial Products in Anammox Biofilm using Multidimensional Modeling. *Sci. Rep.* **6,** 27576 (2016).
19. Lawson, C. E. *et al.* Rare taxa have potential to make metabolic contributions in enhanced biological phosphorus removal ecosystems. *Environ. Microbiol.* **17,** 4979–4993 (2015).
20. Konwar, K. M. *et al.* MetaPathways v2.5: Quantitative functional, taxonomic and usability improvements. *Bioinformatics* **31,** 3345–3347 (2015).
21. Maalcke, W. J. *et al.* Structural basis of biological no generation by octaheme oxidoreductases. *J. Biol. Chem.* **289,** 1228–1242 (2014).
22. Schalk, J., De Vries, S., Kuenen, J. G. & Jetten, M. S. M. Involvement of a novel hydroxylamine oxidoreductase in anaerobic ammonium oxidation. *Biochemistry* **39,** 5405–5412 (2000).

23. Shimamura, M. *et al.* Another multiheme protein, hydroxylamine oxidoreductase, abundantly produced in an anammox bacterium besides the hydrazine-oxidizing enzyme. *J. Biosci. Bioeng.* **105,** 243–248 (2008).
24. Oshiki, M., Shinyako-Hata, K., Satoh, H. & Okabe, S. Draft Genome Sequence of an Anaerobic Ammonium-Oxidizing Bacterium,'Candidatus Brocadia sinica'. *Genome Announc.* **3,** 3–4 (2015).
25. Ferousi, C. *et al.* Identification of the type II cytochrome c maturation pathway in anammox bacteria by comparative genomics. *BMC Microbiol.* **13,** 265 (2013).
26. Oshiki, M., Ali, M., Shinyako-Hata, K., Satoh, H. & Okabe, S. Hydroxylamine-dependent Anaerobic Ammonium Oxidation (Anammox) by ' Candidatus Brocadia sinica'. *Environ. Microbiol.* (2016). doi:10.1111/1462-2920.13355
27. Zhang, S. & Bryant, D. The tricarboxylic acid cycle in cyanobacteria. *Science (80-.).* **7002,** 1551–1554 (2011).
28. Liu, Z. *et al.* Complete genome of Ignavibacterium album, a metabolically versatile, flagellated, facultative anaerobe from the phylum Chlorobi. *Front. Microbiol.* **3,** 1–15 (2012).
29. Kadnikov, V. V. *et al.* Genomic Analysis of Melioribacter roseus, Facultatively Anaerobic Organotrophic Bacterium Representing a Novel Deep Lineage within Bacteriodetes/Chlorobi Group. *PLoS One* **8,** (2013).
30. Spaans, S. K., Weusthuis, R. A., van der Oost, J. & Kengen, S. W. M. NADPH-generating systems in bacteria and archaea. *Front. Microbiol.* **6,** 1–27 (2015).
31. Wolfe, A. J. The Acetate Switch The Acetate Switch. *Microbiol. Mol. Biol. Rev.* **69,** 12–50 (2005).
32. Buckel, W. & Thauer, R. K. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na+ translocating ferredoxin oxidation. *Biochim. Biophys. Acta - Bioenerg.* **1827,** 94–113 (2013).
33. McDowall, J. S. *et al.* Bacterial formate hydrogenlyase complex. *Proc. Natl. Acad. Sci. U. S. A.* **111,** E3948-56 (2014).
34. Sawers, G. The hydrogenases and formate dehydrogenases of Escherichia coli. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **66,** 57–88 (1994).
35. Richardson, D. J. Bacterial respiration: A flexibe process for a changing environment.

Microbiology **146,** 551–571 (2000).

36. Häse, C. C. & Finkelstein, R. A. Bacterial extracellular zinc-containing metalloproteases. *Microbiol. Rev.* **57,** 823–37 (1993).
37. Siezen, R. J. & Leunissen, J. A. Subtilases: the superfamily of subtilisin-like serine proteases. *Protein Sci.* **6,** 501–23 (1997).
38. Akashi, H. & Gojobori, T. Metabolic efficiency and amino acid composition in the proteomes of Escherichia coli and Bacillus subtilis. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 3695–3700 (2002).
39. Mee, M. T., Collins, J. J., Church, G. M. & Wang, H. H. Syntrophic exchange in synthetic microbial communities. *Proc. Natl. Acad. Sci. U. S. A.* **111,** E2149–E2156 (2014).
40. Abma, W. R., Driessen, W., Haarhuis, R. & Van Loosdrecht, M. C. M. Upgrading of sewage treatment plant by sustainable and cost-effective separate treatment of industrial wastewater. *Water Sci. Technol.* **61,** 1715–1722 (2010).
41. Van Mooy, B. A. S., Keil, R. G. & Devol, A. H. Impact of suboxia on sinking particulate organic carbon: Enhanced carbon flux and preferential degradation of amino acids via denitrification. *Geochim. Cosmochim. Acta* **66,** 457–465 (2002).
42. Kindaichi, T., Yuri, S., Ozaki, N. & Ohashi, A. Ecophysiological role and function of uncultured Chloroflexi in an anammox reactor. *Water Sci. Technol.* **66,** 2556–2561 (2012).
43. Datta, M. S., Sliwerska, E., Gore, J., Polz, M. & Cordero, O. X. Microbial interactions lead to rapid micro-scale successions on model marine particles. *Nat. Commun.* **7,** 11965 (2016).
44. Degnan, P. H., Taga, M. E. & Goodman, A. L. Vitamin B12 as a modulator of gut microbial ecology. *Cell Metab.* **20,** 769–778 (2014).
45. Ali, M. *et al.* Source identification of nitrous oxide emission pathways from a single-stage nitritation-anammox granular reactor. *Water Res.* **102,** 147–157 (2016).
46. Lilja, E. E. & Johnson, D. R. Segregating metabolic processes into different microbial cells accelerates the consumption of inhibitory substrates. *ISME J.* **10,** 1568–78 (2016).
47. Kindaichi, T. *et al.* In situ activity and spatial organization of anaerobic ammonium-oxidizing (anammox) bacteria in biofilms. *Appl. Environ. Microbiol.* **73,** 4931–4939 (2007).

48. Vlaeminck, S. E. *et al.* Aggregate size and architecture determine microbial activity balance for one-stage partial nitritation and anammox. *Appl. Environ. Microbiol.* **76,** 900–909 (2010).
49. Kartal, B. *et al.* Candidatus 'Brocadia fulgida': An autofluorescent anaerobic ammonium oxidizing bacterium. *FEMS Microbiol. Ecol.* **63,** 46–55 (2008).
50. Kotay, S. M., Mansell, B. L., Hogsett, M., Pei, H. & Goel, R. Anaerobic ammonia oxidation (ANAMMOX) for side-stream treatment of anaerobic digester filtrate process performance and microbiology. *Biotechnol. Bioeng.* **110,** 1180–1192 (2013).
51. Joshi, N. & Fass, J. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). *[Software]* **Available,** (2011).
52. Salzberg, S. L. FLASH : Fast Length Adjustment of Short Reads to Improve Genome Assemblies Tanja Magoˇ. *Bioinformatics* **27,** 2957–2963 (2011).
53. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3,** e1165 (2015).
54. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25,** 1043–55 (2015).
55. Dupont, C. L. *et al.* Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J.* **6,** 1186–99 (2012).
56. Darling, A. E. *et al.* PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ* **2,** e243 (2014).
57. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11,** 119 (2010).
58. Konwar, K., Hanson, N., Page, A. & Hallam, S. J. MetaPathways: A modular pipeline for constructing Pathway/Genome Databases from environmental sequence information. *BMC Bioinformatics* **14,** (2013).
59. Karp, P. D. *et al.* Pathway Tools version 19.0 update: software for pathway/genome informatics and systems biology Peter. *Brief. Bioinform.* (2015). doi:10.1093/bib/bbp043
60. Karp, P. D., Latendresse, M. & Caspi, R. The pathway tools pathway prediction algorithm. *Stand. Genomic Sci.* **5,** 424–9 (2011).

61. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28,** 3211–3217 (2012).
62. Konstantinidis, K. T. & Tiedje, J. M. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 2567–2572 (2005).
63. Anders, S., Pyl, P. T. & Huber, W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31,** 166–169 (2015).
64. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5,** 621–628 (2008).
65. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42,** 490–495 (2014).
66. Rawlings, N. D., Barrett, A. J. & Finn, R. Twenty years of the *MEROPS* database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* **44,** D343-50 (2015).
67. Marchler-Bauer, A. *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43,** D222–D226 (2015).
68. Finn, R. D. *et al.* The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **44,** D279–D285 (2016).
69. Yu, C.-S., Lin, C.-J. & Hwang, J.-K. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. *Protein Sci.* **13,** 1402–6 (2004).
70. Saier, M. H. *et al.* The Transporter Classification Database (TCDB): Recent advances. *Nucleic Acids Res.* **44,** D372–D379 (2016).

3. Autotrophic and mixotrophic metabolic network fluxes suggest versatile lifestyle for the anammox bacterium *Candidatus* 'Kuenenia stuttgartiensis'

A version of this chapter has been submitted for publication and is available online at bioRxiv:

Lawson, C.E., de Graaf, R.M., Nuijten, G.H.L., Jacobson, T.B., Pabst, M., Stevenson, D.M., Jetten, M.S.M., Noguera, D.R., McMahon, K.D., Amador-Noguez, D., Lücker, S. Autotrophic and mixotrophic metabolic network fluxes suggest versatile lifestyle for the anammox bacterium *Candidatus* 'Kuenenia stuttgartiensis'. *bioRxiv* 835298 (2019). doi:10.1101/835298

All supplementary data files are available online at:
https://www.biorxiv.org/content/10.1101/835298v1.supplementary-material

Author Contributions

C.E.L., S.L., D.A-N., K.D.M., and D.R.N. designed the study. C.E.L., G.H.L.N., R.M.G. performed the ^{13}C tracer experiments. C.E.L., T.B.J., and D.M.S. performed the metabolomic analysis. M.P. and C.E.L. performed the metaproteomic analysis. C.E.L. wrote the manuscript. All authors provided valuable feedback and edits on the manuscript.

3.0 Abstract

Anaerobic ammonium-oxidizing (anammox) bacteria mediate a key step in the biogeochemical nitrogen cycle and have been applied worldwide for the energy-efficient removal of nitrogen from wastewater. However, outside their core energy metabolism, little is known about the metabolic networks driving anammox bacterial anabolism and mixotrophy beyond genome-based predictions. Here, we experimentally resolved the central carbon metabolism of the anammox bacterium *Candidatus* Kuenenia stuttgartiensis using time-series ^{13}C isotope tracing, metabolomics, and isotopically nonstationary metabolic flux analysis (INST-MFA). Our findings confirm predicted metabolic pathways used for CO_2 fixation, central metabolism, and amino acid biosynthesis in *K. stuttgartiensis*, and reveal several instances where genomic predictions are not supported by *in vivo* metabolic fluxes. This includes the use of an incomplete oxidative tricarboxylic acid cycle, despite the genome not encoding a known citrate synthase. We also demonstrate that *K. stuttgartiensis* is able to directly assimilate formate via the Wood-Ljungdahl pathway instead of oxidizing it completetly to CO_2 followed by reassimilation. In contrast, our data suggests that acetate is fully oxidized to CO_2 via reversal of the Wood-Ljungdahl pathway and partial TCA cycle activity, followed by reassimilation of the produced CO_2. Together, these findings highlight the versatility of central carbon metabolism in anammox bacteria and will enable the construction of accurate metabolic models that predict their function in natural and engineered ecosystems.

3.1 Introduction

Anaerobic ammonium oxidation (anammox) is a key step in the biogeochemical nitrogen cycle and represents a novel treatment process for the sustainable removal of nitrogen from wastewater.

The process is mediated by a deeply branching group of chemolithoautotrophic bacteria within the Planctomycetes, the Brocadiales, that couple the anaerobic oxidiation of ammonium to nitrite reduction and dinitrogen gas formation[1,2,3,4]. The discovery of this unique metabolism and subsequent translation to full-scale applications represents one of the most rapid biotechnological advances in wastewater treatment[5,6,7]. However, the metabolic networks controlling anammox metabolism remain poorly understood, which limits the prediction of their function in natural and engineered ecosystems.

Metagenomic sequencing together with experimental studies have begun to unravel the metabolic potential of anammox bacteria[2,3,8]. A combination of molecular approaches have been used to reveal the key enzymes and reactions involved in anammox catabolism, which include hydrazine (N_2H_4) and nitric oxide (NO) as volatile intermediates in the anammox bacterium *Candidatus* Kuenenia stuttgartiensis (hereafter, *K. stuttgartiensis*)[3,9,10]. These reactions are localized within a specialized intracellular organelle, the anammoxosome, which is believed to be dedicated to energy conservation[11,12] and also contains membrane-bound respiratory complexes of *K. stuttgartiensis*' electron transport chain, including complex I, ATP synthase, and an NAD+:ferredoxin oxidoreductase (RNF)[13]. Experimental studies together with genomic evidence have also suggested that anammox bacteria are much more versatile than initially assumed, and can use alternative electron donors to ammonium, such as formate, acetate, and propionate for energy conservation with nitrite or nitrate as electron acceptors[2,8,14,15,16,17]. Intriguingly, it has been proposed that these organic substrates are fully oxidized to CO_2 and not directly assimilated into cell biomass, suggesting that anammox bacteria adhere to their autotrophic lifestyle[4].

Based on measurements of cell carbon isotopic composition, genomic evidence, and gene expression data, it has been proposed that anammox bacteria fix CO_2 to acetyl-CoA via the Wood-

Ljungdahl pathway[2,4,18,19]. Four additional carboxylation reactions are also predicted to incorporate CO_2 into central carbon metabolism based on *K. stuttgartiensis'* genome annotations, catalyzed by pyruvate:ferredoxin oxidoreductase (PFOR), 2-oxoglutarate:ferredoxin oxidoreductase (OFOR), pyruvate carboxylase, and phosphoenolpyruvate carboxylase[2,4]. Products from these reactions are proposed to flow through the tricarboxylic acid (TCA) cycle, gluconeogenesis, and the pentose phosphate pathway to synthesize all biomass precursor metabolites[2,4]. Since *K. stuttgartiensis* apparently lacks a citrate synthase gene, it has been hypothesized that the TCA cycle operates in the reductive direction via OFOR to synthesize essential precursor metabolites, such as alpha-ketoglutarate[4]. However, these genome-based predictions of *K. stuttgartiensis'* metabolic network remain to be tested.

Here, we experimentally resolved the central carbon metabolism of a planktonic *K. stuttgartiensis* cell culture (more than 95% enriched) using time-series ^{13}C isotope tracing, metabolomics, and isotopically nonstationary metabolic flux analysis (INST-MFA). Our results show that several of the metabolic predictions summarized above, which were primarily based on genomic evidence, are not supported by the flux of metabolites experimentally observed. For instance, *K. stuttgartiensis* operates an incomplete oxidative TCA cycle despite having no predicted citrate synthase gene. We also demonstrate that *K. stuttgartiensis* is able to directly assimilate formate via the Wood-Ljungdahl pathway instead of oxidizing it to CO_2 before assimilation. On the contrary, we show that acetate is fully oxidized to CO_2 via a reversed Wood-Ljungdahl pathway and a partial TCA cycle, followed by reassimilation of the produced CO_2. These findings highlight the versatility of carbon metabolism in *K. stuttgartiensis* and provide fundamental insights on the metabolic networks controlling anammox bacterial anabolism.

3.2 Results

3.2.1 Mapping anammox autotrophic metabolism.

To elucidate the central carbon metabolic network of *K. stuttgartiensis* under chemolithoautotrophic growth conditions, we first performed time-series isotopic tracer experiments with ^{13}C-bicarbonate coupled to metabolomic analysis. Planktonic *K. stuttgartiensis* cells were cultivated under steady-state conditions in a continuous-flow membrane bioreactor using minimal media supplemented with ammonium and nitrite. Subsequently, ^{13}C-labelled bicarbonate was rapidly introduced into the bioreactor to a concentration of 30 mM (approximately 65% ^{13}C-dissolved inorganic carbon, DIC), which incorporated into *K. stuttgartiensis*' metabolome over time. Samples were collected over a 3-hour period to trace metabolic network structure based on rates of metabolite ^{13}C-enrichment.

Based on proposed carbon assimilation pathways for anammox bacteria[2,4], we expected that CO_2 fixation would largely occur through the Wood-Ljungdahl pathway and PFOR, resulting in fast labelling of acetyl-CoA and pyruvate, followed by phosphoenolpyruvate and other downstream metabolites. However, despite the almost immediate labeling of phosphoenolpyruvate, ^{13}C-enrichment of acetyl-CoA and pyruvate was slow during the 3-hour experiment (Figure 3.1a; Figure 3.1b). One hypothesis for this observation is substrate channeling, where the product of one enzymatic reaction is directly passed to the next enzyme without opportunity to equilibrate within the cytoplasm[20]. Substrate channeling has been previously reported as a mechanism to protect highly labile intermediates of the Wood-Ljungdahl pathway[21,22] or regulate acetyl-CoA biosynthesis[23], and in *K. stuttgartiensis* it could be combined with PFOR to form a channel from CO_2 to phosphoenolpyruvate. An alternative explanation could be that different pools of acetyl-CoA and pyruvate exist through compartmentation. For example, the

Wood-Ljungdahl pathway and PFOR activities could occur in one compartment or specific cytoplasmic location[24], where other pools of acetyl-CoA and pyruvate that do not get labelled exist in another, diluting the overall ^{13}C metabolite measurements. Consistent with the latter, amino acids synthesized from pyruvate (i.e., valine, and alanine) showed faster labelling and higher ^{13}C-enrichment (Figure 3.1a; Figure 3.1b).

Acetyl-CoA and pyruvate are expected to enter the TCA cycle and gluconeogenesis to produce biomass precursors. Since *K. stuttgartiensis'* genome does not encode a citrate synthase required to operate the oxidative TCA cycle, it is hypothesized that synthesis of key precursor metabolites, including succinyl-CoA and alpha-ketoglutarate, occurs via the reductive direction[4]. If this hypothesis is correct, we would expect to observe high ^{13}C-labelling of oxaloacetate, succinate, and alpha-ketoglutarate. While fast labeling of aspartic acid, which was used as a surrogate for oxaloacetate labelling, implied high activity of phosphoenolpyruvate (or pyruvate) carboxylase, ^{13}C-labelling of succinate was much less and slower than the labelling of alpha-ketoglutarate (Figure 3.1b; Figure 3.1c). This suggested that OFOR and other reductive TCA cycle enzymes were not operating in *K. stuttgartiensis* to synthesize alpha-ketoglutarate.

Other biomass precursors are additionally predicted to be synthesized from gluconeogenesis and the pentose phosphate pathway in *K. stuttgartiensis*[2]. Consistent with this, we observed fast ^{13}C-labeling of the gluconeogenic intermediates 3-phosphoglycerate, fructose 6-phosphate, and glucose 6-phosphate, as well as pentose phosphate pathway intermediates, such as sedheptulose 7-phosphate and ribose 5-phosphate (Figure 3.1b; Figure 3.1c; Supplementary Dataset 1).

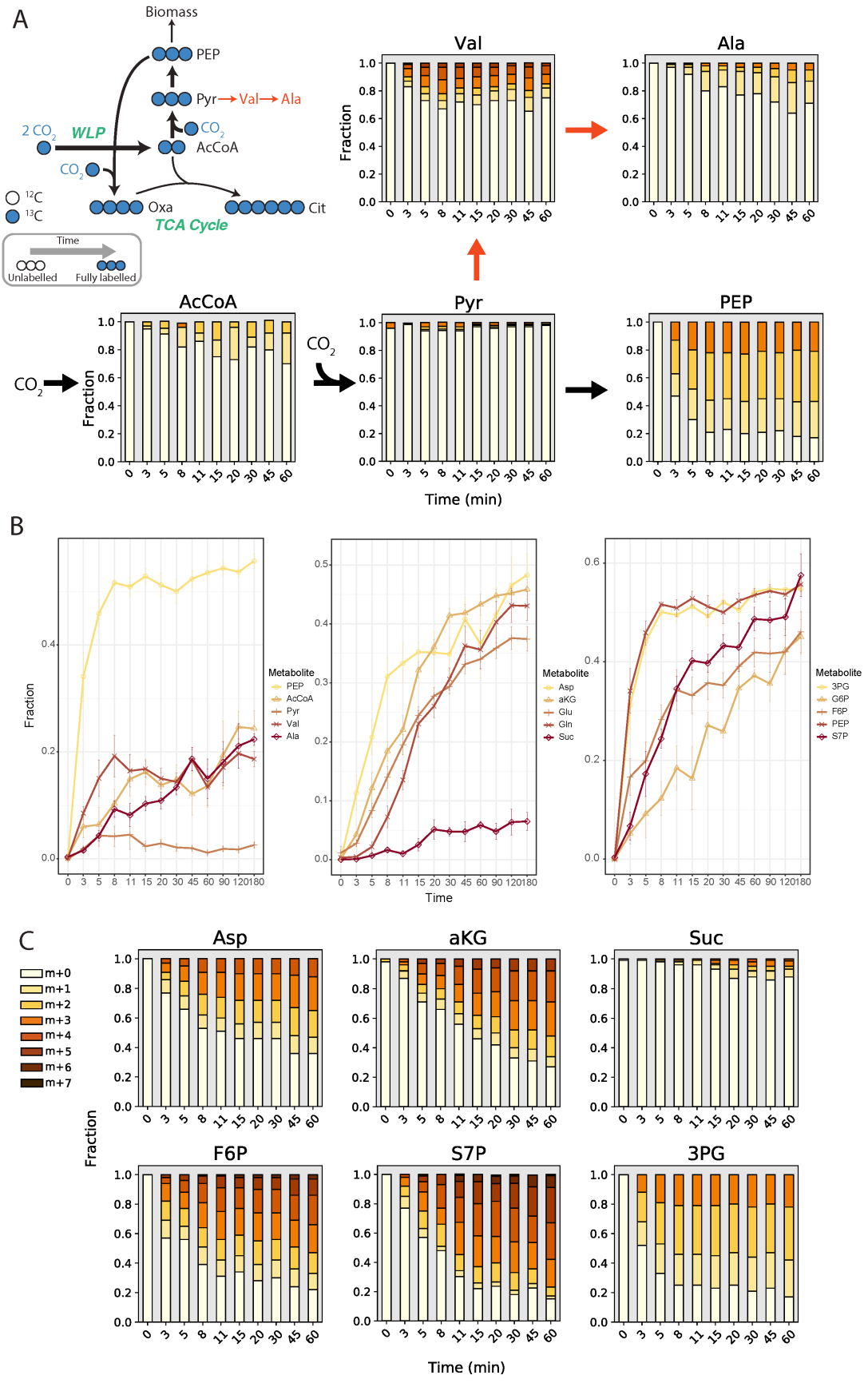
A
Biomass
PEP
Pyr → Val → Ala
2 CO2
WLP
CO2
AcCoA
CO2
12C
13C
Oxa
Cit
TCA Cycle
Time
Unlabelled
Fully labelled
Val
Ala
AcCoA
Pyr
PEP
CO2
CO2
Fraction
Time (min)
B
Fraction
Metabolite
PEP
AcCoA
Pyr
Val
Ala
Asp
aKG
Glu
Gln
Suc
3PG
G6P
F6P
PEP
S7P
Time
C
m+0
m+1
m+2
m+3
m+4
m+5
m+6
m+7
Asp
aKG
Suc
F6P
S7P
3PG
Fraction
Time (min)

Figure 3.1 - ^{13}C-enrichment of selected metabolites during ^{13}C-bicarbonate dynamic tracing experiments. (A) Mass isotopomer distributions (MID) for selected metabolites illustrating potential substrate channeling through the Wood-Ljungdahl Pathway and PFOR. (B) ^{13}C enrichment of metabolites associated with (left) initial CO_2 fixation reactions (Wood-Ljungdahl Pathway, pyruvate:ferredoxin oxidoreductase) and metabolites downstream of pyruvate; (center) TCA cycle metabolites; (right) gluconeogenesis and pentose phosphate pathway metabolites. (C) Selected mass isotopomer distributions for metabolites of the TCA cycle, gluconeogenesis, and the pentose phosphate pathway. All measured metabolite MIDs and standard errors can be found in Supplementary Dataset 1.

3.2.2 ^{13}C-formate tracing confirms formate assimilation pathways and oxidative TCA cycle in K. stuttgartiensis.

We further probed central carbon metabolism with ^{13}C-formate. While it has been proposed that anammox bacteria fully oxidize organic substrates, such as formate, to CO_2[4], we hypothesized that formate could be assimilated by *K. stuttgartiensis* via the methyl branch of the Wood-Ljundahl pathway. This would result in a positionally labelled acetyl-CoA pool that would provide additional information on metabolic network activity (Figure 3.2a).

We tested this hypothesis by rapidly introducing ^{13}C-formate into fresh continuous cultures of *K. stuttgartiensis* to a concentration of 50 mM followed by metabolome sampling over 180 minutes (14 timepoints total). Within 1.5 minutes of ^{13}C-formate introduction, we observed steady-state labelling of several central metabolites, including phosphoenolpyruvate (Figure 3.2b) and 3-phosphoglycerate (Figure 3.3b), consistent with direct assimilation of formate. In agreement with formate assimilation via the Wood-Ljungdahl pathway, only the M+1 mass isotopomer of acetyl-CoA became enriched during the experiment (Figure 3.2c). M+1 mass isotopomers of phosphoenolpyruvate and aspartic acid (oxaloacetate surrogate) were also dominant (Figure 3.2c), consistent with their synthesis from acetyl-CoA via the sequential reactions of PFOR, phosphoenolpyruvate synthase, and phosphoenolpyruvate (or pyruvate) carboxylase, respectively.

Since only a very minor fraction of the M+2 mass isotopomer were detected in these metabolites (<3% over initial 45 minutes), it can be concluded that intracellular ^{13}C-CO_2 concentrations remained low during the experiment. Consistent with this, measured ^{13}C-DIC content in the liquid media produced from ^{13}C-formate oxidation was low, increasing to only 5% over 45 minutes (Figure 3.2b). This supports the inference that ^{13}C-inorganic carbon incorporation into metabolites was insignificant compared to the rate of incorporation via ^{13}C-formate. Similar to ^{13}C-bicarbonate experiments, slower labelling of acetyl-CoA and pyruvate was observed during the ^{13}C-formate tracer experiments (Figure 3.2b; Figure 3.2c). This further supports the hypothesis that separate intracellular pools of these metabolites may exist in *K. stuttgartiensis*.

^{13}C-formate labelling experiments also allowed us to analyze operation of the TCA cycle. If the reductive TCA cycle was operating in *K. stuttgartiensis* only a single carbon in alpha-ketoglutarate would be labelled (from oxaloacetate), while two carbons would be labelled if alpha-ketoglutarate was produced oxidatively (from oxaloacetate and acetyl-CoA). Consistent with the latter route, mass isotopomer distributions for citrate and alpha-ketoglutarate consisted largely of M+2 mass isotopomers (Figure 3.2c). This clearly demonstrates that alpha-ketoglutarate was produced via an oxidative TCA cycle in *K. stuttgartiensis*, and not via the reductive TCA cycle. On the contrary, malate, fumarate, and succinate pools were largely comprised of M+1 mass isotopomers (Figure 3.2c), which suggests that a bifurcated TCA cycle was operating. The labelling patterns of TCA cycle metabolites suggest that *K. stuttgartiensis* uses a novel or highly divergent enzyme for citrate synthesis. While no citrate synthase is annotated in the *K. stuttgartiensis* genome, several acyltransferase candidates exist, including genes annotated as (R)-citramalate synthase (KSMBR1_RS19040) believed to be involved in isoleucine biosynthesis[25] and redundant copies of 2-isopropylmalate synthase (KSMBR1_RS18315 and

KSMBR1_RS10820). In particular, one of the 2-isopropylmalate synthase genes (KSMBR1_RS10820) is phylogenetically related (55.1% identity) to *Re*-citrate synthase identified in *Clostridium kluyveri*[26] and is located next to an ADP-forming succinate-CoA ligase of the oxidative TCA cycle. Therefore, we posit that this gene encodes a dedicated *Re*-citrate synthase that allows the oxidative TCA cycle to be operational in *K. stuttgartiensis*.

3.2.3 Multiple pathways for sugar phosphate biosynthesis?

Results from the ^{13}C-formate tracer experiments also allowed for closer examination of pentose and hexose sugar phosphate biosynthesis in *K. stuttgartiensis.* Because the labelled pools of 3-phosphoglycerate and dihydroxyacetone phosphate were largely comprised of M+1 mass isotopomers, we expected fructose 6-phosphate to largely consist of M+2 mass isotopomers based on gluconeogenesis reactions (Figure 3.3a). However, a considerable fraction of fructose 6-phosphate (Figure 3.3b) and glucose 6-phosphate (Figure 3.4) were consistently present as M+1 mass isotopomers during the ^{13}C-formate tracer experiment (~25-45%). While it is possible that this labeling pattern was produced during the period when the M+1 isotopomers of dihydroxyacetone phosphate and glyceraldehyde 3-phosphate were ~50% (<1.5 minutes), it more likely suggests that alternative pathways exist for sugar phosphate biosynthesis in *K. stuttgartiensis*.

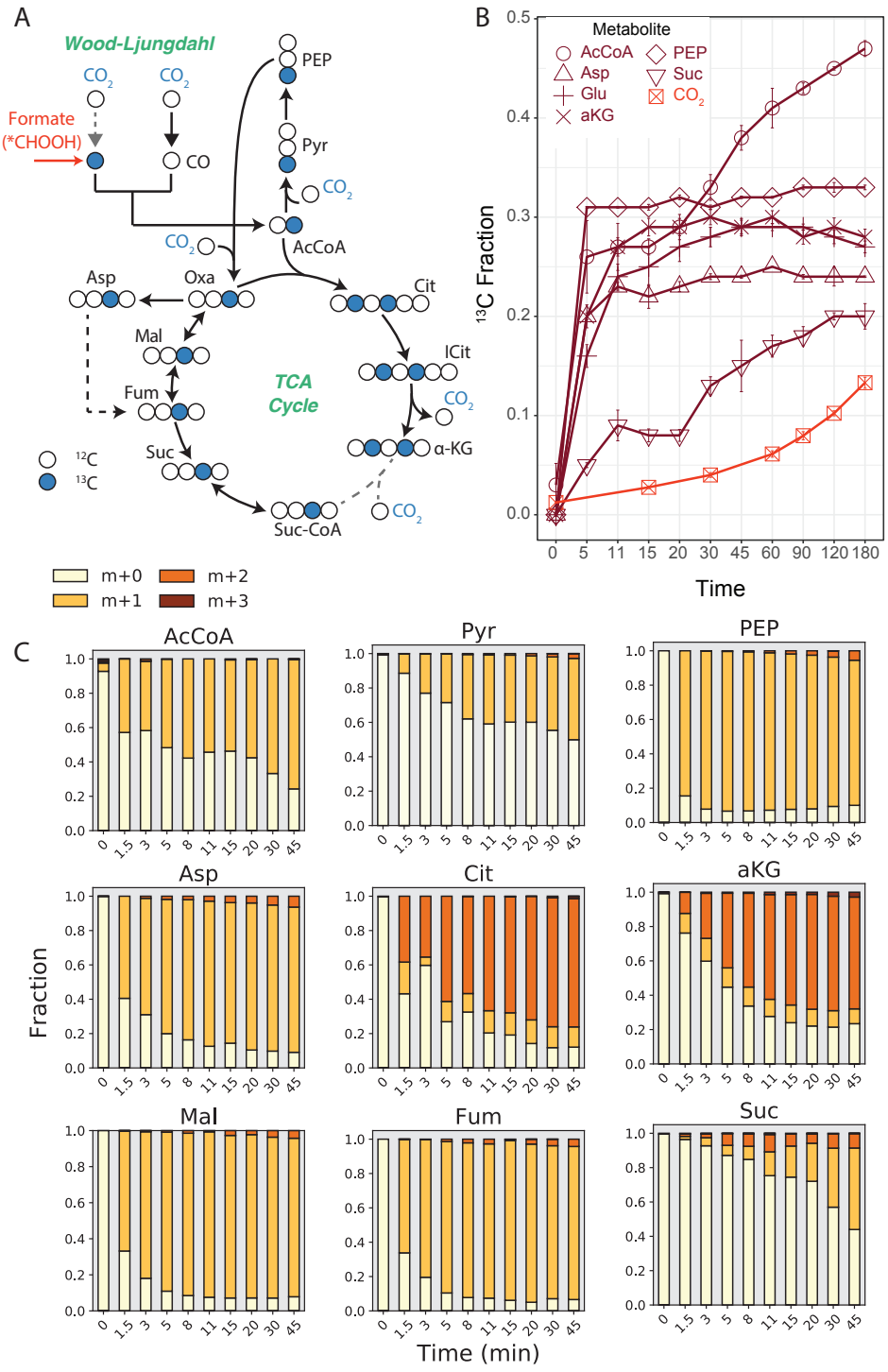

Figure 3.2 - Elucidating TCA cycle of *K. stuttgartiensis* with ^{13}C-formate. (A) Proposed labelling of TCA cycle metabolites with ^{13}C-formate. (B) ^{13}C-enrichment of selected metabolites during isotope tracer experiments with ^{13}C-formate. (C) Time-series mass isotopomer distributions of selected TCA cycle metabolites during isotope tracer experiments with ^{13}C-formate. All measured metabolite MIDs and standard errors can be found in Supplementary Dataset 1.

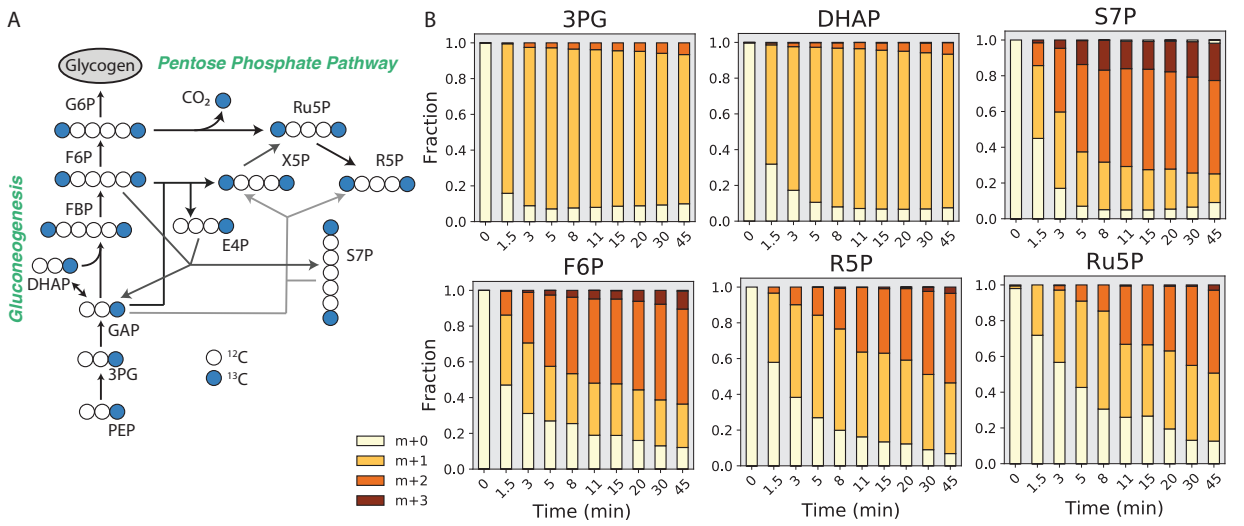

Figure 3.3 - Operation of gluconeogenesis and pentose phosphate pathway in *K. stuttgartiensis* revealed by ^{13}C-formate dynamic labelling experiments. (A) proposed atom mapping of gluconeogenesis and pentose phosphate pathway from ^{13}C-formate labelled phosphoenolpyruvate at steady-state. (B) Time-series mass isotopomer distributions of selected gluconeogenesis and pentose phosphate pathway metabolites during dynamic isotope tracer experiments with ^{13}C-formate. All measured metabolite MIDs and standard errors can be found in Supplementary Dataset 1.

3.2.4 Amino acid biosynthetic pathways

^{13}C-formate tracer results were also used to confirm major amino acid biosynthetic pathways in *K. stuttgartiensis*. Our data supports the synthesis of aspartate, asparagine, and threonine via canonical routes from oxaloacetate; the synthesis of glutamate, glutamine, proline, and arginine from alpha-ketogluturate; and the synthesis of serine from 3-phosphoglycerate (Figure 3.4). Labelling patterns for valine, alanine, and leucine support their production via canonical branched chain amino acid biosynthesis pathways from pyruvate (Figure 3.4). Interestingly, isoleucine biosynthesis was not supported by canonical routes from threonine, but rather via a recently described citramalate-dependent pathway from acetyl-CoA and pyruvate[27,28] (Figure 3.4). Finally, labeling patterns supported the synthesis of the aromatic amino acids, phenylalanine and tyrosine from erythrose 4-phosphate and phosphoenolpyruvate via the shikimate pathway (Figure 3.4).

These amino acid biosynthetic pathways were consistent with pathways predicted from the *K. stuttgartiensis* genome annotation.

Despite the *K. stuttgartiensis* genome lacking an annotated pathway for methionine biosynthesis, methionine was labelled during both ^{13}C-formate and ^{13}C-bicarbonate experiments. Canonical precursors for methionine biosynthesis include aspartic acid and methyl-THF (from formate via methyl-branch of Wood Ljungdahl pathway), thus if this pathway were operating, we would expect to see mainly M+2 methionine. However, a considerable pool of M+1 methionine was consistently observed in our experiments (Figure 3.4), suggesting that a potentially novel pathway is operating to synthesize methionine in *K. stuttgartiensis* that remains to be elucidated.

3.2.5 Acetate oxidation pathway of anammox bacteria.

In addition to formate, we also examined the impact of acetate on *K. stuttgartiensis*' metabolic network. While it has been proposed that anammox bacteria can oxidize acetate to CO_2[8,29], the pathways used for acetate oxidation and whether or not acetate is assimilated into biomass have yet to be resolved. If acetate were oxidized to CO_2, we expected that it would initially be incorporated into acetyl-CoA based on previous enzymatic studies with AMP-forming acetyl-CoA synthetase (KSMBR1_RS14485)[30], followed by oxidation to CO_2 via either the oxidative TCA cycle or reversal of the Wood-Ljungdahl pathway, as previously suggested for other anaerobic chemolithoautotrophic bacteria[31,32] (Figure 3.5a).

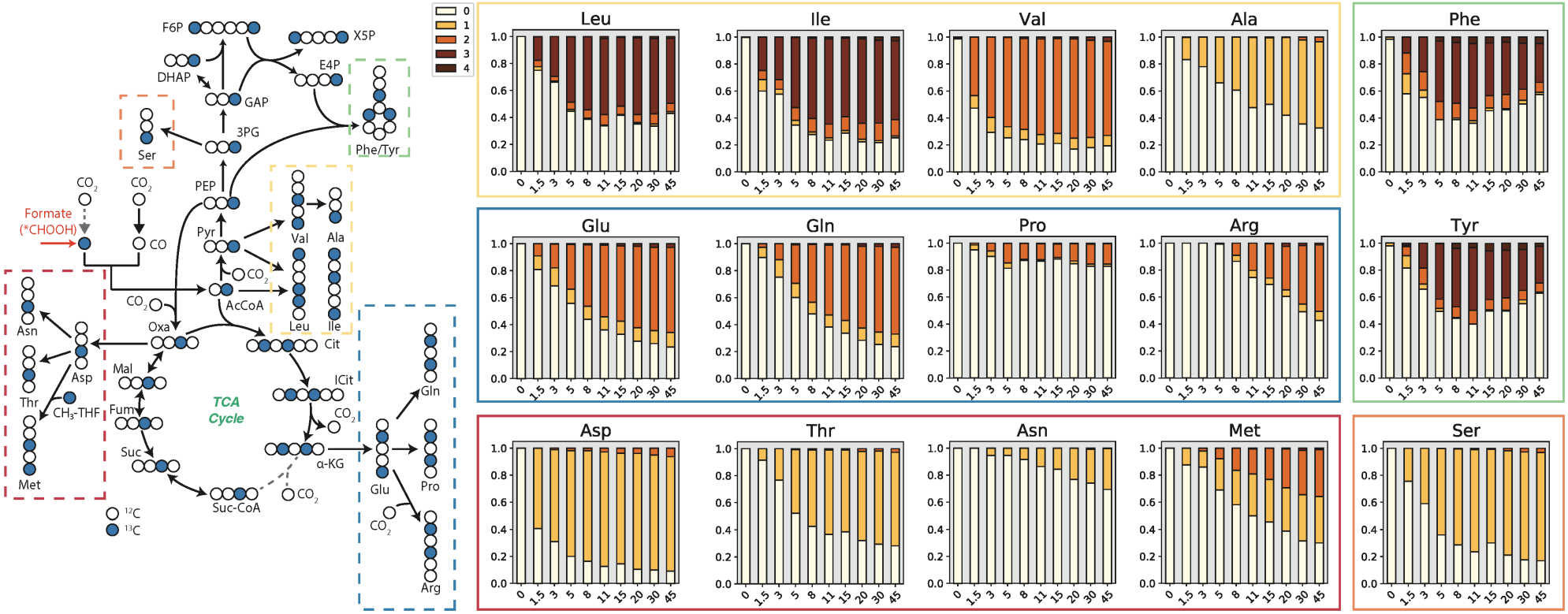

Figure 3.4 - Confirmation of ^{13}C-labelled substrate incorporation into the proteome of *K. stuttgartiensis.* Distribution of relative isotope abundances for identified peptides assigned to the *K. stuttgartiensis* proteome during [2-^{13}C]acetate, ^{13}C-formate, and ^{13}C-bicarbonate tracer experiments after 0 and 72 hours.

To elucidate metabolic pathways involved in acetate metabolism, we rapidly introduced [2-^{13}C]acetate into active continuous cultures of *K. stuttgartiensis* to a final concentration of 10 mM and sampled the metabolome over 180 minutes (12 timepoints). Within 1.5 minutes after [2-^{13}C]acetate addition, we observed steady-state enrichment of the M+1 mass isotopomer of acetyl-CoA, indicating its synthesis via CoA acetylation (Figure 3.5c). Considerable ^{13}C-labelling of citrate and glutamate was also observed, suggesting partial oxidative TCA cycle activity (Figure 3.5c). However, the mass isotopomer distributions for citrate and glutamate contained heavier mass isotopomers (up to M+4) and appeared more evenly distributed, an observation more consistent with the labeling patterns observed during ^{13}C-bicarbonate tracing (Figure 3.1c) versus ^{13}C-formate tracing (Figure 3.2c). This pattern can be explained by additional acetate oxidation via the reverse Wood-Ljungdahl pathway to ^{13}C-CO_2, followed by reincorporation of ^{13}C-CO_2 into central metabolites. In agreement with this, mass isotopomer distributions for all other measured metabolites, including phosphoenolpyruvate, 3-phosphoglycerate, dihydroxyacetone phosphate, and fructose 6-phosphate, had relatively even distributions of ^{13}C-labelled mass isotopomers, suggesting they were also synthesized from re-incorporated ^{13}C-CO_2 (Figure 3.5c). Since ^{13}C-enrichment for most of these metabolites were approximately 2-10 times higher than that of ^{13}C-CO_2 in the media (Figure 3.5b), we conclude that ^{13}C-CO_2 re-incorporation was faster than the efflux of ^{13}C-CO_2.

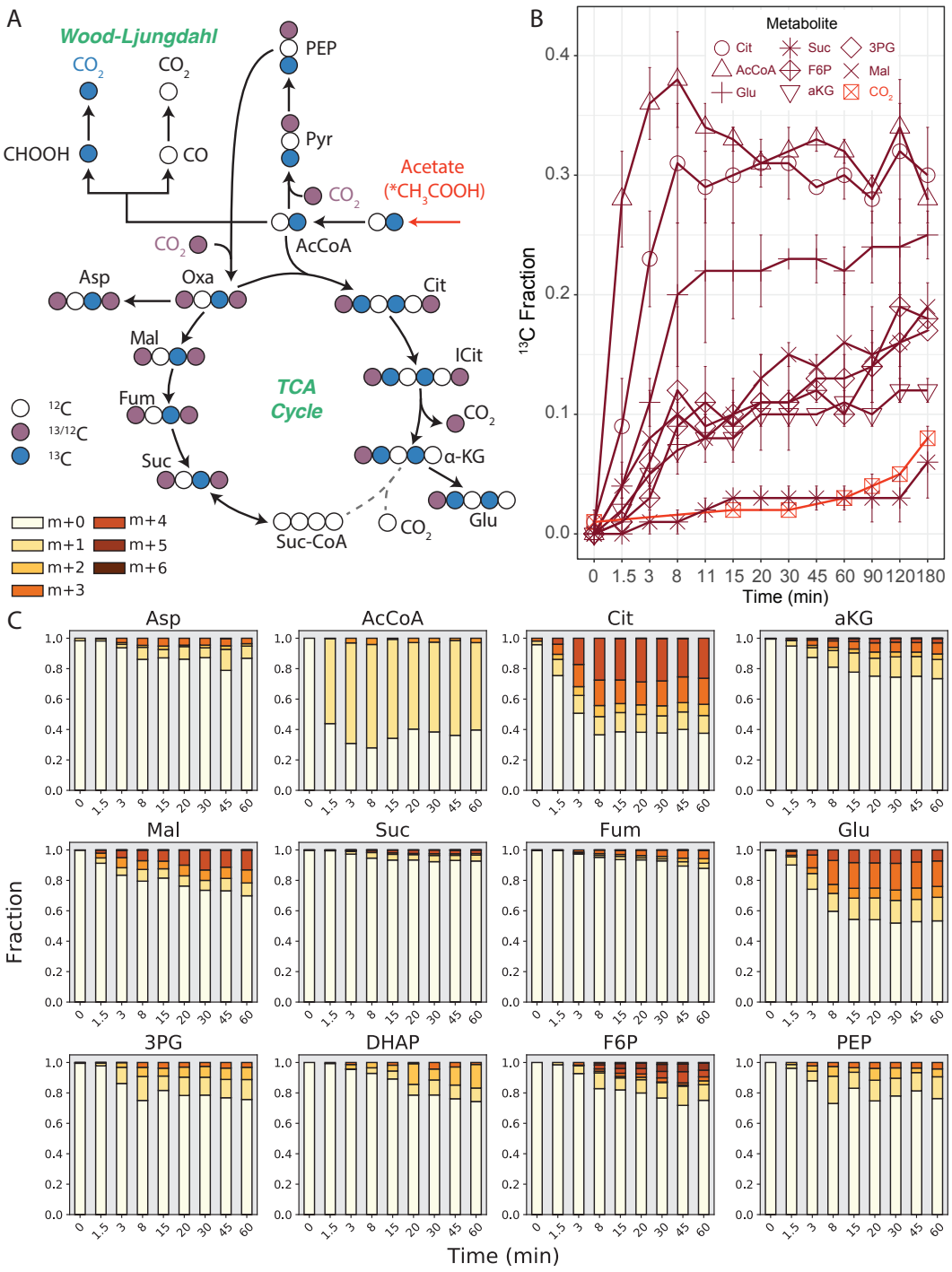

Figure 3.5 - Reverse Wood-Ljungdahl pathway oxidizes acetate in *K. stuttgartiensis*. (A) Proposed labelling of TCA cycle metabolites with [2-^{13}C]acetate. (B) ^{13}C-enrichment of selected metabolites during isotope tracer experiments with [2-^{13}C]acetate (red). (C) Time-series mass isotopomer distributions of selected TCA cycle metabolites during isotope tracer experiments with ^{13}C-acetate. All measured metabolite MIDs and standard errors can be found in Supplementary Dataset 1.

Surprisingly, metabolites that labelled rapidly during ^{13}C-bicarbonate and ^{13}C-formate experiments, including phosphoenolpyruvate and 3-phosphoglycerate, labelled slowly with ^{13}C-acetate (Figure 3.5c). This could be additional evidence that spatial separation of carbon metabolism occurs in *K. stuttgartiensis*, contributing to different pools of acetyl-CoA and pyruvate. Separation of the AMP-forming acetyl-CoA synthetase and potentially other central metabolism enzymes to the outer membrane and periplasm has been reported in the chemolithoautotrophic archeon *Ignicoccus hospitalis*[33]. Interestingly, this enzymes shares homology with the AMP-forming acetyl-CoA synthetase found in *K. stuttgartiensis*[30] and both organisms have ATPases localized to their outer membranes[34,35]. While this raises the possibility that similar spatial separation of metabolism may be present in *K. stuttgartiensis,* further experimental validation is required.

3.2.6 ^{13}C protein stable isotope probing confirms substrate uptake by K. stuttgartiensis.

To confirm uptake of labelled substrates into the biomass of *K. stuttgartiensis* cells, we performed shotgun proteomics on peptides extracted from bioreactor cell pellets collected during ^{13}C-labelling experiments. Metaproteomic analysis of samples collected after 0 and 72 hours confirmed that ^{13}C-bicarbonate was incorporated into the *K. stuttgartiensis* proteome, increasing at a median relative isotope abundance of ~50%, consistent with the ^{13}C-DIC content of the liquid media (Figure 3.6). Incorporation of ^{13}C-formate and [2-^{13}C]acetate into the *K. stuttgartiensis* proteome was also detected after 72 hours at median relative isotope abundances of ~30% and ~10%, respectively (Figure 3.6). These values are consistent with the use of the Wood-Ljungdahl pathway for formate assimilation, and with re-assimilation of ^{13}C-CO_2 produced from acetate oxidation, as the ^{13}C-DIC in the liquid media held at ~11% between 5 and 72 hours (Figure 3.5b).

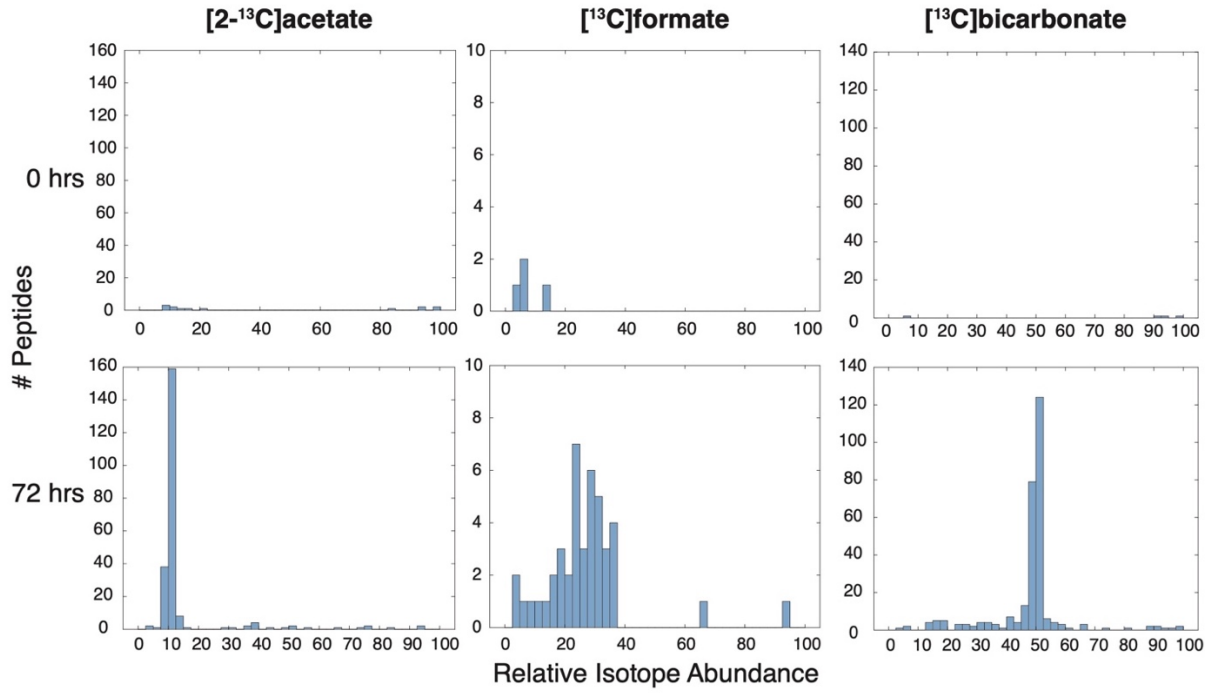

Figure 3.6 - Confirmation of ^{13}C-labelled substrate incorporation into the proteome of *K. stuttgartiensis.* Distribution of relative isotope abundances for identified peptides assigned to the *K. stuttgartiensis* proteome during [2-^{13}C]acetate, ^{13}C-formate, and ^{13}C-bicarbonate tracer experiments after 0 and 72 hours.

3.2.7 Isotopically non-stationary metabolic flux analysis of autotrophic growth.

To quantitatively examine *K. stuttgartiensis*' central carbon metabolism and obtain intracellular flux measurements, we performed INST-MFA by fitting measured, time-resolved metabolite mass isotopomer distributions from ^{13}C-formate tracer experiments to an isotopomer network model[36]. This provided a quantitative systems-level flux map of *K. stuttgartiensis'* inferred central carbon metabolism (Figure 3.7). Flux values were normalized to a net CO_2 uptake rate, which was estimated from the growth rate and cell carbon content: 0.0042 hrs^{-1} x 45 mmol-C/gDW = 0.186 mmol-C/gDW/hr. The resulting flux map reproduces the high intracellular flux anticipated through the Wood-Ljungdahl pathway, PFOR, and phosphoenolpyruvate (or pyruvate) carboxylase, which are the main CO_2 fixation reactions that we observed in *K. stuttgartiensis* (Figure 3.7). INST-MFA

also supported alpha-ketoglutarate production via the oxidative TCA cycle. Moreover, instead of running a bifurcated TCA cycle, the INST-MFA analysis predicts that the M+1 isotopomers of fumarate, succinate, and malate were indirectly derived from aspartic acid as a result of histidine and arginine biosynthesis (Figure 3.7). This suggests that the TCA cycle in *K. stuttgartiensis* operates incompletely, essentially functioning to produce alpha-ketoglutarate (amino acid precursor) and recycle intermediates of amino acid biosynthesis. Considerable oxidative pentose phosphate pathway flux was also measured (Figure 3.7). As no transhydrogenase could be identified in the genome, it is likely that this pathway is key for NAPDH generation in *K. stuttgartiensis*.

INST-MFA also allowed us to query alternative reactions for the unexpected labelling patterns of sugar phosphates during ^{13}C-formate tracer experiments. The genome annotation of *K. stuttgartiensis* has genes coding for hexulose 6-phosphate synthase and 6-phospho-3-hexuloisomerase (KSMBR1_RS05220 and KSMBR1_RS18790, respectively). These are key enzymes of the ribulose monophosphate (RuMP) pathway, a formaldehyde assimilation pathway in many methylotrophic bacteria[37]. Together, these reactions fix formaldehyde to fructose 6-phosphate via a hexulose 6-phosphate intermediate[37]. We hypothesize that these reactions, as well as an unidentified formaldehyde dehydrogenase, could explain the considerable M+1 pentose and hexose phosphate isotopomers observed during ^{13}C-formate labelling (Figure 3.8). Including these reactions in our INST-MFA improved the models fit by approximately 19% (SSR of 802.1 versus 988.7, 95% confidence), accounting for approximately 23% of the flux synthesizing fructose 6-phosphate (Figure 3.7).

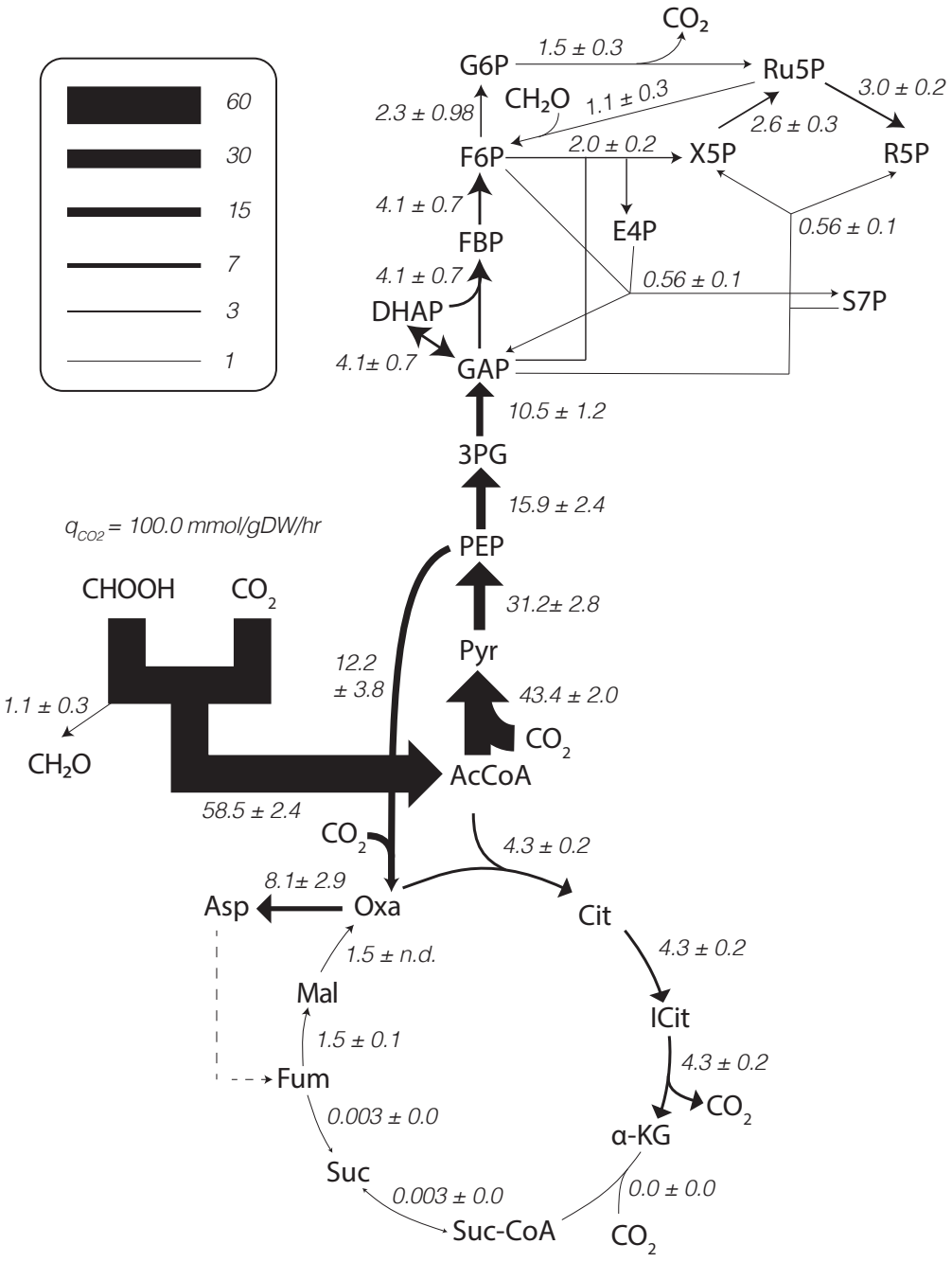

Figure 3.7 - *K. stuttgartiensis* flux map generated by ^{13}C INST-MFA. *K. stuttgartiensis* flux map under anaerobic, continuous flow, ammonium and nitrite medium conditions determined by fitting metabolites labelled with ^{13}C-formate tracers to a single, statistically acceptable isotopomer network model. Flux values represent the net flux through a given reaction +/- standard error defined at 95% confidence. All fluxes are normalized to a net CO_2 uptake rate of q=100 mmol-C/gDW/hr (actual CO_2 uptake rate was 0.186 mmol-C/gDW/hr). All isotopomer network model reactions are provided in Supplementary Dataset 2. INST-MFA solutions are provided in Supplementary Dataset 3.

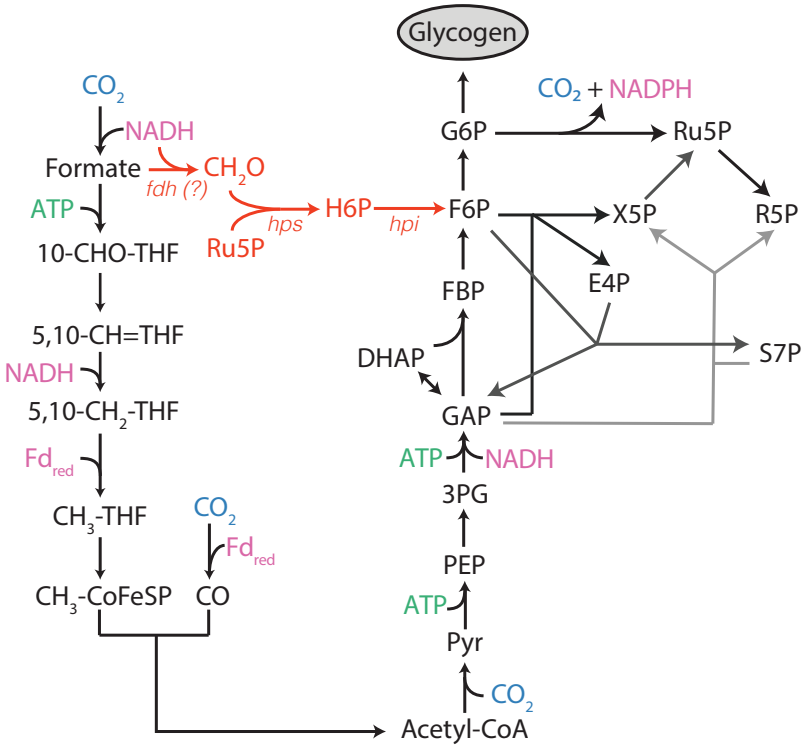

Figure 3.8 - Proposed synthesis of sugar phosphates from gluconeogenesis and the RuMP cycle in *K. stuttgartiensis.* RuMP cycle reactions that synthesize fructose 6-phosphate (F6P) from formaldehyde (CH_2O) and ribulose 5-phosphate (Ru5P) are shown in orange. Production of CH_2O from formate via an unknown formaldehyde dehydrogenase is also show in orange. Reducing equivalents shown in pink; ATP shown in green; CO_2 shown in blue. *fdh*: formaldehyde dehydrogenase; *hps*: hexulose 6-phosphate synthase; *hpi*: hexulose 6-phosphate isomerase; *?* indicates no gene annotation present.

3.3 Discussion

Elucidating the *in vivo* metabolic network of *K. stuttgartiensis* represents a major advance towards predicting the function of anammox bacteria in natural and engineered ecosystems. Our study offers the first measurements of metabolic flux via INST-MFA in a chemolithoautotrophic organism, providing a systems-level flux map for *K. stuttgartiensis* that can be used to understand anammox bacterial central carbon metabolism. The discovery of an incomplete oxidative TCA cycle operating in *K. stuttgartiensis*, likely mediated by a novel *Re*-citrate synthase, avoids the

energetically costly use of reduced ferredoxin for alpha-ketoglutarate biosynthesis via the reductive TCA cycle. Furthermore, the considerable flux measured through the oxidative pentose phosphate pathway highlights an important link between carbon and energy metabolism for generating reducing equivalents (i.e. NADPH) in anammox bacteria. Our analysis validated the use of the Wood-Ljungdahl pathway, PFOR, and phosphoenolpyruvate/pyruvate carboxylase for CO_2 fixation in *K. stuttgartiensis* and provided first evidence of possible compartmentalization and/or metabolic channeling in these pathways. This may enable faster pathway kinetics, avoid degradation of unstable tetrahydrofolate-based intermediates, or limit competition between competing reactions, as has been shown with other pathways[38,39].

We also elucidated the role of the Wood-Ljungdahl pathway for formate assimilation by *K. stuttgartiensis* and show that reversal of this pathway can additionally be used for acetate oxidation. This may further support the metabolism of *K. stuttgartiensis* in their anaerobic habitats, where these compounds likely exist as fermentation products[40,41]. Together, these findings provide insight into the mechanisms underlying the observed versatility of anammox bacteria[42] and may inform strategies to control anammox-based bioprocesses via organic substrate addition[43].

Our elucidation of *K. stuttgartiensis' in vivo* metabolic network fluxes will spur further quantitative studies on anammox metabolism and enable the construction of accurate genome-scale metabolic models to predict anammox bacterial physiology. We believe that when integrated with metabolic models of other nitrogen cycling bacteria[44], drastic improvements in the prediction and control of anammox-based biotechnologies and biogeochemical processes will be possible.

3.4 Materials and Methods

3.4.1 Cultivation of K. stuttgartiensis cells

A high enrichment of planktonic *K. stuttgartiensis* cells were cultivated in a continuous flow membrane bioreactor (MBR) on mineral salts medium[45] supplemented with 45 mM of both ammonium and nitrite. Cultures were maintained under steady-state conditions at an OD_{600} of 1.0-1.1 via continuous biomass removal and the bioreactor was continuously sparged with Ar/CO_2 (95%/5% v/v) at a rate of 10 ml/min to maintain anaerobic conditions. The reactor hydraulic and solids retention times were approximately 46 hours and 10.5 days, respectively. The temperature and pH of the reactor were controlled at 30°C and 7.3 using a heat exchanger and 1 M $KHCO_3$ buffer, respectively. The reactor was continuously stirred at 600 rpm. Nitrite concentrations were checked daily to ensure nitrite-limited conditions (Nitrite test strips MQuant®, Merck, Darmstadt, Germany).

3.4.2 ^{13}C isotope tracer experiments

Isotope tracing experiments with ^{13}C-labelled substrates ([^{13}C]sodium bicarbonate, [2-^{13}C]sodium acetate, and [^{13}C]sodium formate; Cambridge Isotopes Laboratories, MA, USA) were performed separately on continuous cultures of *K. stuttgartiensis* cells harvested from the MBR system. ^{13}C-labelled substrates were rapidly introduced (within 1 minute) into the bioreactor containing *K. stuttgartiensis* cells growing under steady-state conditions. Initial reactor concentrations of bicarbonate, acetate and formate were approximately 30 mM, 5 mM and 50 mM respectively. Following ^{13}C-label introduction, 5 ml samples were rapidly withdrawn from the reactor at timepoints 0, 1.5, 3, 5, 8, 11, 15, 20, 30, 45, 60, 90, 120, and 180 minutes. Samples were immediately filtered (Millipore 0.45 µm hydrophilic nylon filter HNWPO4700) using a vacuum pump to remove the medium, and filters were placed face down in 1.5 ml of -80°C extraction

solvent (40:40:20 acetonitrile:methanol:water) for cell quenching and metabolite extraction. Samples were then centrifuged (10,000 rpm, 4°C, 5 mins) and 1 ml of cell-free supernatant was collected and stored at -80°C for metabolomic analysis. The time 0 min sample corresponded to the period directly before ^{13}C-label addition. The ratio of $^{13}C/^{12}C$ DIC remained constant during the course of the experiment as determined by gas chromatography coupled with mass spectrometry (GC-MS) analysis (See method below).

3.4.3 Metabolomic analysis

Samples were analysed using a high-performance HPLC–MS system consisting of a Vanquish™ UHPLC system (Thermo Scientific) coupled by electrospray ionization (ESI; negative polarity) to a hybrid quadrupole high-resolution mass spectrometer (Q Exactive Orbitrap, Thermo Scientific) operated in full scan mode for detection of targeted compounds based on their accurate masses. Properties of Full MS–SIM included a resolution of 140,000, AGC target of 1E6, maximum IT of 40 ms and scan range from 70 to 1,000 m/z. LC separation was achieved using an ACQUITY UPLC BEH C18 column (2.1 × 100 mm column, 1.7 μm particle size; part no. 186002352; serial no. 02623521115711, Waters). Solvent A was 97:3 water:methanol with 10 mM tributylamine (TBA) adjusted to pH 8.1–8.2 with 9 mM acetic acid. Solvent B was 100% methanol. Total run time was 25 min with the following gradient: 0 min, 5% B; 2.5 min, 5% B; 5 min, 20% B; 7.5 min, 20% B; 13 min, 55% B; 15.5 min, 95% B; 18.5 min, 95% B; 19 min, 5% B; 25 min, 5% B. Flow rate was 200 μl min^{-1}. The autosampler and column temperatures were 4°C and 25°C, respectively. Mass isotopomer distributions were corrected for natural abundance using the method of Su et al (2017)[46] and ^{13}C enrichment values were calculated using the formula $(1/N)\ \sum_{i=1}^{N} Mi\ \times\ i$, where N is the number of carbon atoms in the metabolite and Mi is the fractional abundance of the i^{th} mass isotopomer.

To improve separation and measurement sensitivity of specific central carbon metabolites and intracellular amino acids, samples were first derivatized with either aniline[47,48] or benzyl chloroformate[49], respectively. For aniline derivatization, samples were resuspended in 50 µl HPLC-grade water, 5 µl aniline (6M, pH 4.5), and 5 µl N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride, EDC, (200 mg/ml). After 2 hours of incubation at room temperature, 1 ul of triethylamine was added to stop the reaction. For benzl chloroformate derivatization, samples were resuspended in 10 µl HPLC-grade water, 40 µl methanol, 5 µl of triethylamine, and 1 µl benzyl chloroformate and incubated at room temperature for 30 minutes.

3.4.4 GC-MS analysis of dissolved inorganic carbon isotopic fractions

Isotopic fractions of DIC in the liquid media were measured based on a modified headspace method[50]. 3 ml of liquid culture were collected from the bioreactor with a syringe and directly filtered through a sterile 0.45 µm filter (Whatmann, celluloacetate) and 26G needle into a 120 ml bottle containing 1 ml 6M HCl and crimp sealed with a rubber stopper. Prior to adding the liquid sample, bottles and HCl were flushed with either 100% N_2 or Ar gas to void the headspace of background CO_2. Samples were equilibrated with the acid in the bottles for at least 1 hour at room temperature to drive all DIC into the gas phase. 50 µl of the bottles headspace was then injected with a gas tight syringe (Hamilton) into a gas chromatograph (Agilent 6890 equipped with 6 ft Porapak Q columns) at 80°C with helium as a carrier gas at a flow rate of 24 ml/min, coupled to a mass spectrometer (Agilent 5975C MSD; Agilent, Santa Clara, CA) to determine the isotopic fractions of $^{12}CO_2$ and $^{13}CO_2$.

3.4.5 Isotopic non-stationary metabolic flux analysis

Intracellular metabolic fluxes were estimated from the measured metabolite isotope labelling dynamics via INST-MFA using the elementary metabolite unit method[36] implemented in the INCA software package v1.8[51]. Metabolic fluxes and pool sizes were estimated by minimizing the lack-of-fit between measured and computationally simulated metabolite mass isotopomer distributions using least-squares regression. All metabolite mass isotopomer distribution measurements and model reactions used for flux determination are provided in Supplementary Datasets 1 and 2, respectively. The biomass equation was based on experimental measurements of the amino acid composition obtained from *K. stuttgartiensis* biomass pellets (Table 3.1). Pseudofluxes were added to the model for specific metabolites to account for inactive metabolite pools that did not participate in metabolism, but contributed to measured metabolite labelling patterns, similar to Ma et al. (2014)[52]. Chi-squared statistical tests were performed on resulting flux distributions to assess goodness-of-fit, and accurate 95% confidence intervals were computed for all estimated parameters by evaluating the sensitivity of the sum-of-squared residuals to parameter variations[53].

3.4.6 Amino acid composition analysis

Cultures were centrifuged (10,000 rpm, 15 mins, 4°C) to obtain cell pellets, which were subsequently freeze-dried prior to analysis. Total protein concentration was determined using the Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific) and amino acid composition was determined according to Carnicer et al. (2009)[54] using a Varian 920-LC high performance liquid chromatography amino acid analyzer.

Table 3.1 - K. stuttgartiensis biomass amino acid composition.

Amino Acid	Mass (umol/mgDW)	Std Dev (umol/mgDW)
Ala	375	39
Arg	279	57
Asp	130	10
Glu	108	9
Gly	779	117
His	42	8
Ile	268	43
Leu	393	55
Lys	155	44
Met	219	58
Phe	205	43
Pro	460	96
Ser	437	94
Thr	381	87
Val	438	76

3.4.7 ^{13}C protein stable isotope probing

Proteins were extracted from bioreactor cell pellets using glass bead beating (acid, washed, 0.1 mm diameter) in a suspension containing B-PER reagent (Thermo Scientific, Germany) and TEAB buffer (50 mM TEAB, 1% (w/w) NaDOC at pH 8). Following DTT reduction and alkylation using iodo acetamide (IAA) protein extracts were subject to proteolytic digestion using trypsin. Resulting peptides were solid phase extraction purified using an Oasis HLB 96 well plate (Waters, UK), according to the manufacturer protocols. The purified peptide fraction was analysed via a one-dimensional reverse phase separation (Acclaim PepMap RSLC RP C18, 50 µm x 150 mm, 2µm, 100A) coupled to a Q-Exactive plus Orbitrap mass spectrometer (Thermo Scientific, Germany) operating in data dependent acquisition mode (DDA, shot-gun proteomics). The flow rate was maintained at 300 nL/min over a linear gradient from 5% to 30% over 90 minutes and finally to 75% B over 25 minutes. Solvent A was H_2O containing 0.1% formic acid, and solvent

B consisted of 80% acetonitrile in H_2O and 0.1% formic acid. The Orbitrap was operated in DDA mode acquiring peptide signals form 350-1400 m/z, where the top 10 signals (with a charge between 2-7) were isolated at a window of 2.0 m/z and fragmented using a NCE of 30. The AGC target was set to 1e5, at a max IT of 54 ms and 17.5 K resolution. Protein identification and relative isotope abundances were determined from Tandem-MS data using PEAKS Studio X (BSI, Canada) and MetaProSIP (OpenMS, Univ Tuebingen/Berlin, Germany)[55] integrated into the KNIME 4.0.1 analysis platform (Zurich, Switzerland), respectively. All peptide spectra were matched against a protein database generated from predicted open reading frames from the *K. stuttgartiensis* genome.

3.4.8 Supplementary Materials

Supplementary Dataset 1. Average metabolite mass isotopomer distributions and associated standard errors during ^{13}C-bicarbonate, ^{13}C-formate, and [2-^{13}C]acetate tracer experiments. (Sheet 1A) Average mass isotopomer distributions for selected metabolites during 13C-bicarbonate tracing, (Sheet 1B) Mass isotopomer distributions standard error values for selected metabolites during 13C-bicarbonate tracing, (Sheet 2A) Average mass isotopomer distributions for selected metabolites during 13C-formate tracing, (Sheet 2B) Mass isotopomer distributions standard error values for selected metabolites during 13C-formate tracing, (Sheet 3A) Average mass isotopomer distributions for selected metabolites during [2-13C]acetate tracing, (Sheet 3B) Mass isotopomer distributions standard error values for selected metabolites during [2-13C]acetate tracing.

Supplementary Dataset 2. *K. stuttgartiensis* isotopomer network model. Letters within brackets indicate carbon atom transitions of each metabolite for a given reaction.

Supplementary Dataset 3. INST-MFA model results. Metabolite MIDs used for model fitting were Pro, Asn, Ala, Thr, aKG, Ser, Suc, Asp, Glu, R5P, PEP, 3PG, Cit, Mal, Ru5P, Fum, F6P, Pyr, G6P, Val, CO2, and Gln at timepoints 0, 1.5, 3, 5, 8, 11, 15, 20, 30, and 45 minutes. All metabolite MIDs can be found in Supplementary Dataset 1.

3.5 Acknowledgements

The authors would like to acknowledge Patricia van Dam and Carol de Ram for help with metaproteomic sample preparation, Katinka van de Pas-Schoonen for help with bioreactor maintenance, Paul van der Ven and Sebastian Krosse for help with amino acid analysis, and Arjan Pol and Huub Op den Camp for helpful discussions. Funding was provided by the National Science Foundation (CBET-1435661, CBET-1803055 and MCB-1518130), the Netherlands Organization for Scientific Research (Grants 016.Vidi.189.050 and SIAM Gravitation Grant 024.002.002), the European Research Council (ERC Advanced Grant ECO-MOM 339880), a Wisconsin Distinguished Graduate Fellowship, a Postgraduate Scholarship-Doctoral (PGS-D) by the National Sciences and Engineering Research Council of Canada (NSERC), and the UW-Madison Office of the Vice Chancellor for Research and Graduate Education through the Microbiome Initiative.

3.6 Reference

1. Strous, M. *et al.* Missing lithotroph identified as new planctomycete. *Nature* **400**, 446–449 (1999).
2. Strous, M. *et al.* Deciphering the evolution and metabolism of an anammox bacterium from a community genome. *Nature* **440**, 790–4 (2006).
3. Kartal, B. *et al.* Molecular mechanism of anaerobic ammonium oxidation. *Nature* **479**, 127–30 (2011).
4. Kartal, B. *et al.* How to make a living from anaerobic ammonium oxidation. *FEMS*

Microbiol. Rev. **37**, 428–461 (2013).

5. Jetten, M. S. M., Horn, S. J. & van Loosdrecht, M. C. M. Towards a more sustainable municipal wastewater treatment system. *Water Sci. Technol.* **35**, 171–180 (1997).
6. Kartal, B., Kuenen, J. G. & van Loosdrecht, M. C. M. Sewage Treatment with Anammox. *Science (80-.).* **328**, 702–703 (2010).
7. Lackner, S. *et al.* Full-scale partial nitritation/anammox experiences – An application survey. *Water Res.* **55**, 292–303 (2014).
8. Kartal, B. *et al.* Anammox bacteria disguised as denitrifiers: nitrate reduction to dinitrogen gas via nitrite and ammonium. *Environ. Microbiol.* **9**, 635–42 (2007).
9. Oshiki, M., Ali, M., Shinyako-Hata, K., Satoh, H. & Okabe, S. Hydroxylamine-dependent Anaerobic Ammonium Oxidation (Anammox) by " Candidatus Brocadia sinica". *Environ. Microbiol.* (2016). doi:10.1111/1462-2920.13355
10. Hu, Z., Wessels, H. J. C. T., van Alen, T., Jetten, M. S. M. & Kartal, B. Nitric oxide-dependent anaerobic ammonium oxidation. *Nat. Commun.* **10**, 1244 (2019).
11. van Niftrik, L. & Jetten, M. S. M. Anaerobic Ammonium-Oxidizing Bacteria: Unique Microorganisms with Exceptional Properties. *Microbiol. Mol. Biol. Rev.* **76**, 585 LP – 596 (2012).
12. Neumann, S. *et al.* Isolation and characterization of a prokaryotic cell organelle from the anammox bacterium Kuenenia stuttgartiensis. *Mol. Microbiol.* **94**, 794–802 (2014).
13. Almeida, N. M. De *et al.* Membrane-bound electron transport systems of an anammox bacterium : A complexome analysis. *BBA - Bioenerg.* **1857**, 1694–1704 (2016).
14. Güven, D. *et al.* Propionate Oxidation by and Methanol Inhibition of Anaerobic Ammonium-Oxidizing Bacteria. *Appl. Environ. Microbiol.* **71**, 1066 LP – 1071 (2005).
15. Kartal, B. *et al.* Candidatus 'Anammoxoglobus propionicus' a new propionate oxidizing species of anaerobic ammonium oxidizing bacteria. *Syst. Appl. Microbiol.* **30**, 39–49 (2007).
16. Kartal, B. *et al.* Candidatus 'Brocadia fulgida': An autofluorescent anaerobic ammonium oxidizing bacterium. *FEMS Microbiol. Ecol.* **63**, 46–55 (2008).
17. Narita, Y. *et al.* Enrichment and physiological characterization of an anaerobic ammonium-oxidizing bacterium 'Candidatus Brocadia sapporoensis'. *Syst. Appl. Microbiol.* **40**, 448–457 (2017).

18. Schouten, S. *et al.* Stable Carbon Isotopic Fractionations Associated with Inorganic Carbon Fixation by Anaerobic Ammonium-Oxidizing Bacteria. *Appl. Environ. Microbiol.* **70**, 3785–3788 (2004).
19. Ali, M. *et al.* Physiological characterization of anaerobic ammonium oxidizing bacterium 'Candidatus Jettenia caeni'. *Environ. Microbiol.* **17**, 2172–2189 (2015).
20. Sweetlove, L. J. & Fernie, A. R. substrate channelling in metabolic regulation. *Nat. Commun.* **9**, 2136 (2018).
21. Ragsdale, S. W. & Pierce, E. Acetogenesis and the Wood – Ljungdahl pathway of CO2 fixation. *Biochim. Biophys. Acta* **1784**, 1873–1898 (2008).
22. Fuchs, G. Alternative Pathways of Carbon Dioxide Fixation : Insights into the Early Evolution of Life ? *Annu. Rev. Microbiol.* **65**, 631–658 (2011).
23. Lieber, D. J. *et al.* A multienzyme complex channels substrates and electrons through acetyl-CoA and methane biosynthesis pathways in Methanosarcina. *PLoS One* **9**, e107563–e107563 (2014).
24. Meyer, P., Cecchi, G. & Stolovitzky, G. Spatial localization of the first and last enzymes effectively connects active metabolic pathways in bacteria. *BMC Syst. Biol.* **8**, 131 (2014).
25. Howell, D. M., Xu, H. & White, R. H. (R)-Citramalate Synthase in Methanogenic Archaea. *J. Bacteriol.* **181**, 331 LP – 333 (1999).
26. Li, F., Hagemeier, C. H., Seedorf, H., Gottschalk, G. & Thauer, R. K. Re -Citrate Synthase from Clostridium kluyveri Is Phylogenetically Related to Homocitrate Synthase and Isopropylmalate Synthase Rather Than to Si -Citrate Synthase. *J. Bacteriol.* **189**, 4299–4304 (2007).
27. Risso, C., Dien, S. J. Van, Orloff, A., Lovley, D. R. & Coppi, M. V. Elucidation of an Alternate Isoleucine Biosynthesis Pathway in Geobacter sulfurreducens. *J. Bacteriol.* **190**, 2266–2274 (2008).
28. Tang, Y. J. *et al.* Investigation of Carbon Metabolism in Dehalococcoides ethenogenes Strain 195 by Use of Isotopomer and Transcriptomic Analyses. *J. Bacteriol.* **191**, 5224 LP – 5231 (2009).
29. Oshiki, M., Shimokawa, M., Fujii, N., Satoh, H. & Okabe, S. Physiological characteristics of the anaerobic ammonium-oxidizing bacterium ' Candidatus Brocadia sinica '. *Microbiology* **157**, 1706–1713 (2011).

30. Russ, L. *et al.* Genome analysis and heterologous expression of acetate-activating enzymes in the anammox bacterium Kuenenia stuttgartiensis. *Arch. Microbiol.* **194**, 943–948 (2012).
31. Spormann, A. M. & Thauer, R. K. Anaerobic acetate oxidation to CO2 by Desulfotomaculum acetoxidans. *Arch. Microbiol.* **150**, 374–380 (1988).
32. Schauder, R., Preu, A., Jetten, M. & Fuchs, G. Oxidative and reductive acetyl CoA/carbon monoxide dehydrogenase pathway in Desulfobacterium autotrophicum. *Arch. Microbiol.* **151**, 84–89 (1989).
33. Mayer, F. *et al.* AMP-Forming Acetyl Coenzyme A Synthetase in the Outermost Membrane of the Hyperthermophilic Crenarchaeon Ignicoccus hospitalis. *J. Bacteriol.* **194**, 1572 LP – 1581 (2012).
34. Van Niftrik, L. *et al.* Intracellular localization of membrane-bound ATPases in the compartmentalized anammox bacterium 'Candidatus Kuenenia stuttgartiensis'. *Mol. Microbiol.* **77**, 701–715 (2010).
35. Küper, U., Meyer, C., Müller, V., Rachel, R. & Huber, H. Energized outer membrane and spatial separation of metabolic processes in the hyperthermophilic Archaeon Ignicoccus hospitalis *Proc. Natl. Acad. Sci.* **107**, 3152 LP – 3156 (2010).
36. Young, J. D., Walther, J. L., Antoniewicz, M. R., Yoo, H. & Stephanopoulos, G. An elementary metabolite unit (EMU) based method of isotopically nonstationary flux analysis. *Biotechnol. Bioeng.* **99**, 686–99 (2008).
37. Kato, N., Yurimoto, H. & Thauer, R. K. The Physiological Role of the Ribulose Monophosphate Pathway in Bacteria and Archaea. *Biosci. Biotechnol. Biochem.* **70**, 10–21 (2006).
38. W, F. P., Williams, T. C. R., Sweetlove, L. J. & Ratcliffe, R. G. Capturing Metabolite Channeling in Metabolic. *Plant Physiol.* **157**, 981–984 (2011).
39. Bulutoglu, B., Garcia, K. E., Wu, F., Minteer, S. D. & Banta, S. Direct Evidence for Metabolon Formation and Substrate Channeling in Recombinant TCA Cycle Enzymes. *ACS Chem. Biol.* (2016). doi:10.1021/acschembio.6b00523
40. Speth, D. R., in 't Zandt, M. H., Guerrero-Cruz, S., Dutilh, B. E. & Jetten, M. S. M. Genome-based microbial ecology of anammox granules in a full-scale wastewater

treatment system. *Nat. Commun.* **7**, 11172 (2016).

41. Lawson, C. E. *et al.* Metabolic network analysis reveals microbial community interactions in anammox granules. *Nat. Commun.* **8**, 15416 (2017).
42. Winkler, M. K. H. *et al.* Nitrate reduction by organotrophic Anammox bacteria in a nitritation / anammox granular sludge and a moving bed biofilm reactor. *Bioresour. Technol.* **114**, 217–223 (2012).
43. Le, T. *et al.* Nitrate residual as a key parameter to efficiently control partial denitrification coupling with anammox. *Water Environ. Res.* **91**, 1455–1465 (2019).
44. Mellbye, B. L. *et al.* Genome-Scale, Constraint-Based Modeling of Nitrogen Oxide Fluxes during Coculture of Nitrosomonas europaea and Nitrobacter winogradskyi. *mSystems* **3**, e00170-17 (2018).
45. Graaf, A. A. Van De, Bruijn, P. De, Robertson, L. A., Jetten, M. M. & Kuenen, J. G. Autotrophic growth of anaerobic ammonium-oxidzing micro-organisms in a fluidized bed reactor. *Microbiology* **142**, 2187–96 (1996).
46. Su, X., Lu, W. & Rabinowitz, J. D. Metabolite Spectral Accuracy on Orbitraps. *Anal. Chem.* **89**, 5940–5948 (2017).
47. Yang, W.-C. *et al.* Simultaneous Quantification of Metabolites Involved in Central Carbon and Energy Metabolism Using Reversed-Phase Liquid Chromatography−Mass Spectrometry and in Vitro 13C Labeling. *Anal. Chem.* **80**, 9508–9516 (2008).
48. Jannasch, A., Sedlak, M. & Adamec, J. Quantification of Pentose Phosphate Pathway (PPP) Metabolites by Liquid Chromatography-Mass Spectrometry (LC-MS) BT - Metabolic Profiling: Methods and Protocols. in (ed. Metz, T. O.) 159–171 (Humana Press, 2011). doi:10.1007/978-1-61737-985-7_9
49. Kamphorst, J. J. *et al.* Human Pancreatic Cancer Tumors Are Nutrient Poor and Tumor Cells Actively Scavenge Extracellular Protein. **75**, 544–554 (2015).
50. Åberg, J. & Wallin, B. Evaluating a fast headspace method for measuring DIC and subsequent calculation of pCO2 in freshwater systems. *Inl. Waters* **4**, 157–166 (2014).
51. Young, J. D. INCA : a computational platform for isotopically non-stationary metabolic flux analysis. *Bioinformatics* **30**, 1333–1335 (2014).
52. Ma, F., Jazmin, L. J., Young, J. D. & Allen, D. K. Isotopically nonstationary 13C flux analysis of changes in Arabidopsis thaliana leaf metabolism due to high light acclimation.

Proc. Natl. Acad. Sci. **111**, 16967 LP – 16972 (2014).

53. Antoniewicz, M. R., Kelleher, J. K. & Stephanopoulos, G. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab. Eng.* **8**, 324–337 (2006).
54. Carnicer, M. *et al.* Macromolecular and elemental composition analysis and extracellular metabolite balances of Pichia pastoris growing at different oxygen levels. *Microb. Cell Fact.* **8**, 65 (2009).
55. Sachsenberg, T. *et al.* MetaProSIP: Automated Inference of Stable Isotope Incorporation Rates in Proteins for Functional Metaproteomics. *J. Proteome Res.* **14**, 619–627 (2015).

4. Investigating the chemolithoautotrophic and formate metabolism of *Nitrospira moscoviensis* by constraint-based reconstruction and analysis

This chapter has been formatted as a manuscript for submission with the following authors:

Christopher E. Lawson, Aniela B. Mundinger, Hanna Koch, Laura Hesp, Mike S. M. Jetten, Daniel Amador-Noguez, Daniel R. Noguera, Katherine D. McMahon, Sebastian Lücker

All supplementary materials are available online at:
https://github.com/celawson87/phdthesis/tree/master/chapter4

Author Contributions

C.E.L., S.L., K.D.M., and D.R.N. designed the study. C.E.L. built the models and performed the metabolomic analysis. C.E.L., A.B.M., H.K., and L.H. perform the ^{13}C isotopic tracer experiments. H.K. performed the ^{13}C-formate batch experiments. C.E.L. wrote the manuscript. S.L., K.D.M., D.R.N., M.S.M.J., A.B.M., H.K., and L.H. provided valuable feedback and edits on the manuscript.

4.0 Abstract

Nitrite-oxidizing bacteria belonging to the genus *Nitrospira* mediate a key step in nitrification and play important roles in the biogeochemical nitrogen cycle and wastewater treatment. While these organisms have recently been shown to exhibit metabolic flexibility beyond their chemolithoautotrophic lifestyle, including the use of simple organic compounds to fuel their energy metabolism, the metabolic networks controlling their autotrophic and mixotrophic growth remain poorly understood. Here, we reconstructed a genome-scale metabolic model for *Nitrospira moscoviensis* (*iNmo691*) and used constraint-based analysis to evaluate the metabolic networks controlling their lithoautotrophic growth and growth on formate. Subsequently, ^{13}C-tracer experiments with bicarbonate and formate coupled to metabolomics analysis were performed to experimentally validate model predictions. Our findings confirm that *N. moscoviensis* uses the reductive tricarboxylic acid cycle for CO_2 fixation. We also show that *N. moscoviensis* can indirectly use formate as a carbon source by oxidizing it to CO_2 followed by reassimilation via the reductive tricarboxylic acid cycle. Our study offers the first measurements of *Nitrospira's in vivo* central carbon metabolism and provides a quantitative tool that can be used for understanding and predicting their metabolic processes.

4.1 Introduction

The oxidation of nitrite to nitrate is a key step in nitrification and the global nitrogen cycle. The process is a critical control point counteracting nitrogen loss to the atmosphere and is mediated by a phylogenetically diverse functional guild known as the nitrite-oxidizing bacteria (NOB)[1]. The genus *Nitrospira* constitutes the most diverse and abundant NOB based on marker genes (16S ribosomal RNA and nitrite oxidoreductase (NXR) and metagenomic surveys. *Nitrospira* consist

of six lineages that mediate nitrite oxidation across various habitats, including soil, freshwater, marine, terrestrial and engineered ecosystems[1,2]. *Nitrospira* must be flexible enough to survive in the wide range of fluctuating environmental conditions characteristic of these habitats, suggesting their ecophysiology and ecological niches extend beyond those initially defined by their chemolithoautotrophic lifestyle.

Despite their recalcitrance to cultivation, recent insights driven by metagenomics have shed light on the unique features of *Nitrospira's* carbon and energy metabolism[3]. *Nitrospira* encode novel respiratory chain enzymes for energy conservation, including an evolutionarily distinct membrane-bound periplasmic nitrite oxidoreductase (NXR) and a putative cytochrome *bd*-like oxidase that may allow them to adapt to low dissolved oxygen environments[3]. *Nitrospira* also encode all genes for CO_2 fixation via the reductive tricarboxylic acid cycle (rTCA) and lack the two key genes (ribulose bisphosphate carboxylase and phosphoribulokinase) needed to operate the Calvin-Benson-Bassham cycle (CBB) common in other NOB[3]. Outside their chemolithoautotrophic growth, genomic and experimental data have revealed that *Nitrospira* can use alternative substrates to fuel their carbon and energy metabolism[1,4,5]. In addition to nitrite, some *Nitrospira* species have been experimentally shown to use formate, hydrogen, and ammonia as electron donors with oxygen or nitrate as terminal electron acceptors[4,5,6,7,8]. For example, *N. moscoviensis* contains a soluble formate dehydrogenase and NiFe hydrogenase that allows growth with formate or H_2 respectively[4,5], whereas *N. nitrosa* contains pathways for both ammonia and nitrite oxidation that enables growth via complete nitrification[7]. In addition to their energy metabolism, fluorescent *in situ* hybridization combined with microautoradiography (FISH-MAR) experiments have also suggested that *Nitrospira* populations present within activated sludge microbial communities can assimilate pyruvate[2] and formate[9] although the carbon assimilation

pathways for these substrates have yet to be determined. While this expanded metabolic versatility suggests that *Nitrospira* are adapted to dynamic environmental conditions, our ability to predict their function in natural and engineered ecosystems is constrained by the limited understanding of their metabolic network and the lack of quantitative tools to study their metabolism.

Genome-scale metabolic modeling is a powerful method for analyzing and predicting the biochemical pathways driving microbial metabolism. Such modeling approaches calculate the flow of metabolites through a reconstructed metabolic network based on relevant constraints (e.g. network stoichiometry, thermodynamics, measured fluxes) using a technique called flux balance analysis (FBA)[10]. This provides a quantitative framework for analyzing metabolism and predicting phenotypes when combined with physiological data. Moreover, genome-scale models can be used to generate testable hypotheses on the functional capabilities of organisms under defined conditions that can subsequently be tested; for example, using ^{13}C isotopic tracing combined with metabolomics[11].

Here, we provide the first constraint-based metabolic reconstruction and analysis of *N. moscoviensis* representing the ubiquitous *Nitrospira* lineage II. We examine the metabolism of *N. moscoviensis* growing under canonical chemolithoautotrophic conditions and also during growth with formate as a substrate. Taking advantage of recent advances to cultivate *Nitrospira* in continuous flow membrane bioreactors[12], we subsequently validate *N. moscoviensis'* predicted metabolic network using isotopic tracers combined with quantitative metabolomic analysis. Our ^{13}C metabolomics results support the use of the rTCA for carbon fixation by *N. moscoviensis* during chemolithoautotrophic growth. We further show that *N. moscoviensis* does not assimilate formate directly, but instead re-assimilates CO_2 produced via formate oxidation using the rTCA

cycle. The resulting genome-scale model (GEM), *iNmo691*, will serve as a knowledge base for understanding and predicting the function of *Nitrospira* in both natural and engineered ecosystems.

4.2 Results

4.2.1 Genome-scale metabolic reconstruction of Nitrospira moscoviensis

The genome-scale metabolic network of *N. moscoviensis* (*iNmo691)* was reconstructed from the most recent *N. moscoviensis* genome annotation (NCBI accession number NZ_CP011801), aided by reaction annotations in the MetaCyc[13] and ModelSEED databases available through the Department of Energy Systems Biology Knowledgebase[14]. The final reconstruction contained a total of 673 reactions, 636 metabolites, and 691 genes (Supplementary Dataset 1). Reactions for *Nitrospira's* respiratory chain were reconstructed based on existing models of electron flow[1,3] and assuming that the two protons produced during nitrite oxidation in the periplasm contribute to proton motive force generation. Cytoplasmic reactions for formate dehydrogenase and hydrogenase that catalyze formate and hydrogen oxidation to CO_2, respectively, were also included based on genomic and experimental evidence[4,5], as well as formate transport and assimilation reactions mediated by a reductive glycine pathway[15] (Supplementary Dataset 1) to evaluate the possibility of directly using formate as a carbon source, as suggested by previous FISH-MAR observations[9].

Since the stoichiometry for proton translocation in *Nitrospira's* respiratory chain complexes is unknown, values were assumed from model organisms (Supplementary Dataset 1; Figure 4.1). The mechanism for reducing low-potential electron carriers, such as ferredoxin, required for carbon fixation in *Nitrospira* is also unknown, but it has been hypothesized that the 2M-type NADH dehydrogenase complex in *Nitrospira* performs this function[16]. Thus, the model

included this mechanism for ferredoxin reduction. All annotated biosynthetic and biodegradation reactions for amino acids, nucleic acids, carbohydrates, lipids, and cofactors were included in the reconstruction (Supplementary Data 1). This encompassed all predicted central carbon metabolic reactions annotated in the genome, including reactions of the rTCA cycle, gluconeogenesis, the pentose phosphate pathway, anaplerotic reactions, one-carbon metabolism, and fatty acid metabolism. Gaps in the network were identified and filled manually to ensure that *iNmo691* could grow on minimal NOB media (see Materials and Methods).

To obtain qualitative and quantitative outputs from a genome-scale model via FBA, an objective function is required. This is typically accomplished based on a biomass objective function, which assumes that maximization of biomass growth is the cellular objective[17]. Therefore, to generate a representative biomass objective function for *N. moscoviensis* we experimentally determined its biomass composition (i.e. macromolecular components) during steady-state growth (see Materials and Methods).

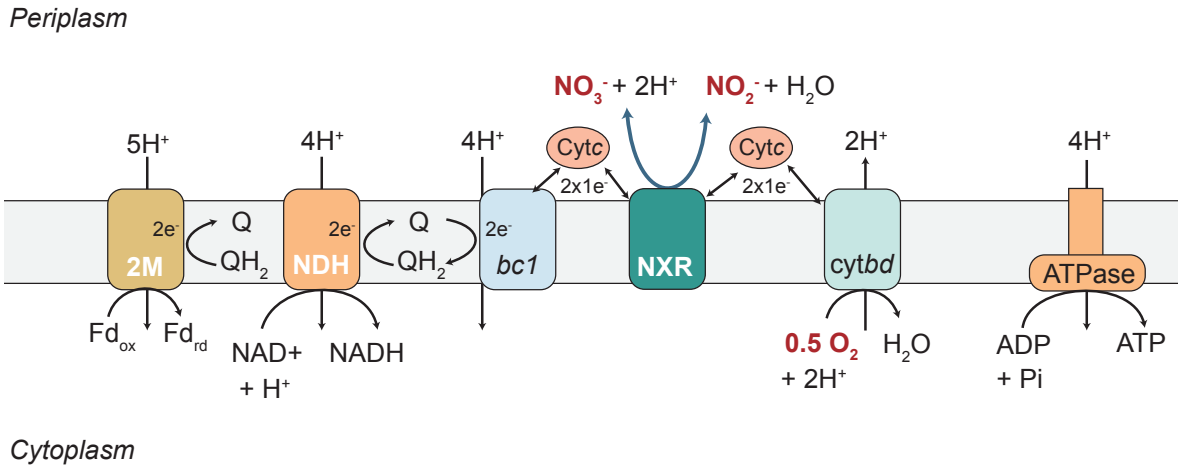

Figure 4.1 - Theoretical model for the energy metabolism of *Nitrospira moscoviensis* hypothesized in this study. Q: quinone; QH2: quinonol; Fd: ferredoxin; 2M: 2M-type complex I, NuoL; NDH: NADH dehydrogenase (complex I); bc1: cytochrome bc1 complex (complex III); NXR: nitrite oxidoreductase; cytbd: cytochrome c oxidase (complex IV).

A summary of the biomass composition for *N. moscoviensis* is presented in Tables 4.1 and 4.2. These measurements, together with the genome sequence data and published fatty acid composition data[18], were used to formulate *N. moscoviensis'* biomass objective function (Supplementary Dataset 1).

Table 4.1 - *N. moscoviensis* biomass composition

Molecule	Average (% DW)	Std Dev (% DW)
RNA	2.4	0.04
DNA[1]	2.4	n.d.
Proteins	46.7	4.7
Carbohydrates	26.9	1.8
Lipids	16.3	3.2
Inorganics[1]	5.0	n.d.

[1]Values for DNA and inorganics were assumed based on values from Neidhardt et al. (1991)[19]. n.d. not determined.

Growth and non-growth associated maintenance energy requirements were estimated to be 725 mmol ATP gDW^{-1} and 0.90 mmol ATP gDW^{-1} hr^{-1}, respectively, by plotting the experimentally determined nitrite uptake rate against the growth rate and using a net ATP yield of 1.0 mmol ATP mmol NO_2-N^{-1} determined from the model (Figure 4.2).

Table 4.2 - *N. moscoviensis* amino acid composition

Amino Acid	Mass (µmol/mgDW)	Std Dev (µmol/mgDW)
Ala	301	46
Arg	312	64
Asp	82	12
Glu	85	16
Gly	825	116
His	38	3

Ile	199	62
Leu	309	67
Lys	106	22
Met	194	23
Phe	224	108
Pro	641	227
Ser	298	63
Thr	277	70
Tyr	42	n.d.
Val	374	59

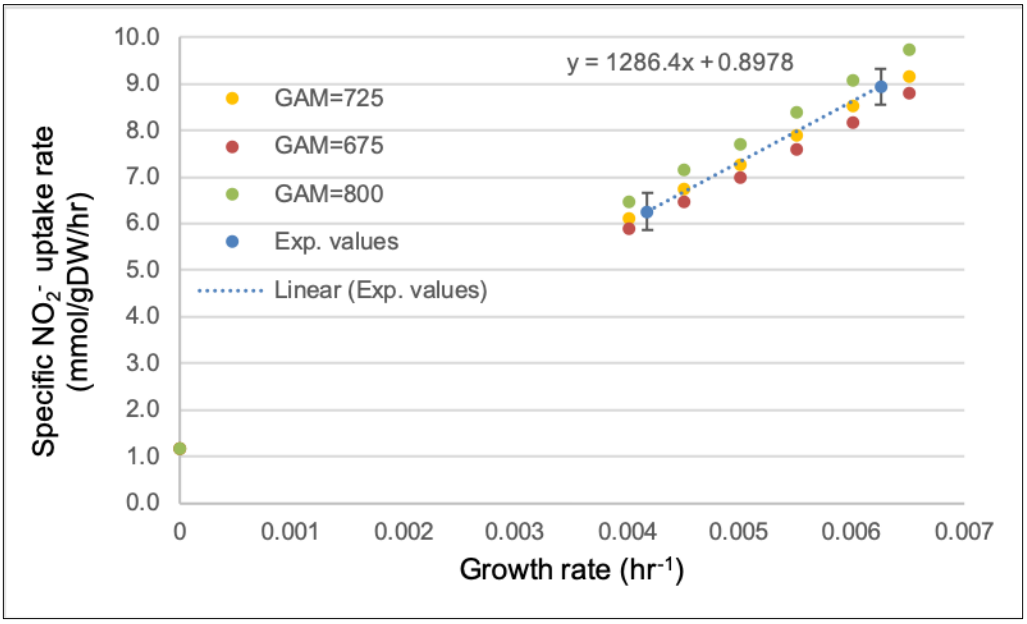

Figure 4.2 - Maintenance energy determination based on plotting the experimentally determined nitrite uptake rate versus the growth rate. Experimental values were taken from Mundinger et al., (2019)[12]. Non-growth associated maintenance (NGAM) was determined by multiplying the y-intercept value (0.90 mmol NO_2^-/gDW-hr) by the model determined ATP yield (1 mmol ATP per 1 mmol NO_2^-). Growth associated maintenance (GAM) was determined by matching the slope of the curve to that determined by the model, based on adjusting the value for GAM and setting the biomass growth rate to different values while minimizing the nitrite uptake rate.

4.2.2 Analysis of chemolithoautotrophic growth on nitrite

The *iNmo691* genome-scale model was first used to quantify the carbon and electron flux distribution in *N. moscoviensis* during aerobic growth on minimal media with nitrite as an electron donor and CO_2 as a carbon source. The model was constrained by the nitrite uptake rate measured during chemostat cultivation, and the objective function was maximizing growth. Consistent with experimental data, when the nitrite uptake rate was set to 8.5 mmol NO_2^- gDW^{-1} hr^{-1} the specific biomass growth rate equaled 0.006 hrs^{-1} (~116 hr doubling time; Supplementary Figure 1). Under these conditions, approximately 63% of the oxidized NO_2^- is used for growth associated maintenance (GAM, polymerization of amino acids into proteins, RNA error checking, etc.) and 15% is oxidized for non-growth associated maintenance (NGAM, maintenance of chemical gradients, turgor pressure, etc.) in the form of ATP (Figure 4.3). The remaining 22% is used to generate ATP (4%) and reducing equivalents (18%) required for CO_2 fixation and synthesis of macromolecular building blocks (e.g. central metabolites, amino acids, nucleotides, fatty acids).

The predicted carbon flux distribution in *N. moscoviensis* during chemolithoautotrophic growth showed that 2-oxoglutarate:ferredoxin oxidoreductase (OFOR), succinyl-CoA synthetase (SCS) and pyruvate:ferredoxin oxidoreductase (PFOR) carried the highest carbon flux, consistent with their primary role in CO_2 fixation (Figure 4.4). These flux values agree with the high gene expression levels of rTCA cycle enzymes recently reported by Mundinger et al., (2019)[12]. Pyruvate carboxylase also carried high carbon flux, which is required to replenish TCA cycle intermediates used as precursors for biosynthesis (e.g. oxaloacetate). Generation of reducing equivalents for carbon fixation in *N. moscoviensis* is predicted to result from reverse electron flow from the cytochrome *c* pool[3]. Overall, the model predicts that approximately 9% of the oxidized nitrite is used to generate quinol (largely for nitrite assimilation via octaheme nitrite reductase), 8.5% is

used to generate NADH (via complex I), and 4.5% is used to generate reduced ferredoxin (via the 2M-type complex I) (Figure 4.3).

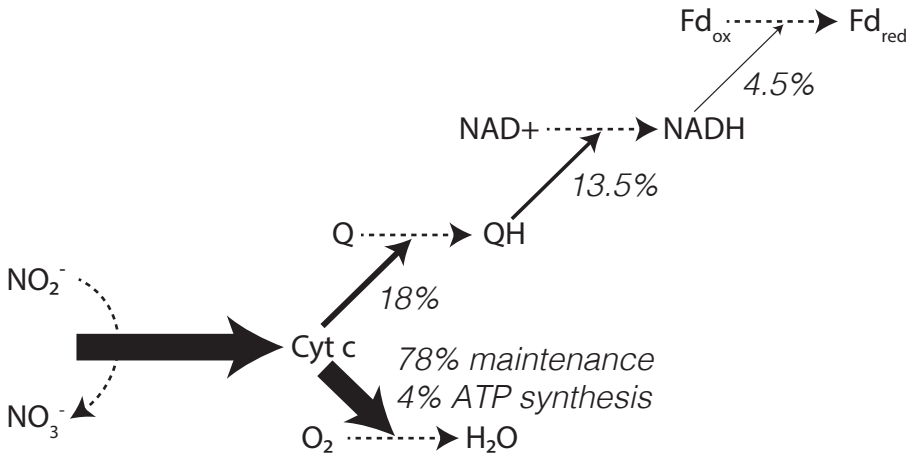

Figure 4.3 - Electron flux distribution in *N. moscoviensis.* Solid lines indicate electron flow, dashed lines indicate redox half-reactions. FBA solutions can be found in Supplementary Dataset 2.

Reduction of NADP+ to NADPH, the main electron carrier for biosynthetic reactions, was predicted to occur through NAD(P)(+) transhydrogenase. Three NAD(P) transhydrogenases are encoded in *N. moscoviensis*, however, only one had above median gene expression in the published transcriptomic dataset (NITMOv2_RS05045, 2.6 fold above median, Mundinger et al., 2019[12]). Therefore, we posit that NITMOv2_RS05045 is the main transhydrogenase active during chemolithoautotrophic growth.

4.2.3 Confirmation of autotrophic metabolism with ^{13}C-bicarbonate tracer experiments

To confirm the biosynthetic pathways predicted by genome-scale modeling, isotopic tracers combined with high-resolution metabolomics was used to follow ^{13}C-labelled bicarbonate

incorporation into the metabolome of *N. moscoviensis.* Cells were grown in a membrane bioreactor under steady-state conditions and ^{13}C-bicarbonate was rapidly introduced into the bioreactor, which equilibrated with the dissolved inorganic carbon (DIC) pool in the liquid media to approximately 65% ^{13}C enrichment. Following ^{13}C-bicarbonate addition, multiple metabolite samples were collected from the bioreactor over a 2-hour time period by rapid quenching and extraction of *N. moscoviensis* cells. A time zero sample corresponding to the period immediately before ^{13}C-bicarbonate addition was also collected as a control. Consistent with the operation of the rTCA cycle, high ^{13}C label incorporation was measured in phosphoenolpyruvate (PEP), acetyl-CoA, succinate, and aspartic acid (used as oxaloacetate surrogate) (Figure 4.5). Other TCA cycle metabolites showed considerable ^{13}C enrichment, including malate, fumarate, and alpha-ketoglutarate, although to a lower extent than PEP, acetyl-CoA, succinate, and aspartic acid (Figure 4.5). This suggested that the ^{13}C enrichment of malate, fumarate, and alpha-ketoglutarate may have been diluted by inactive pools of these metabolites that were not labeled. Further evidence for this was observed from the faster labelling of glutamate and glutamine than alpha-ketoglutarate, which was unexpected because these amino acids are derived from alpha-ketoglutarate (Figure 4.5; Supplementary Dataset 3). We hypothesize that these patterns may reflect substrate channelling in the rTCA cycle of *N. moscoviensis*, which has previously been observed in the TCA cycle of other organisms[20, 21, 22].

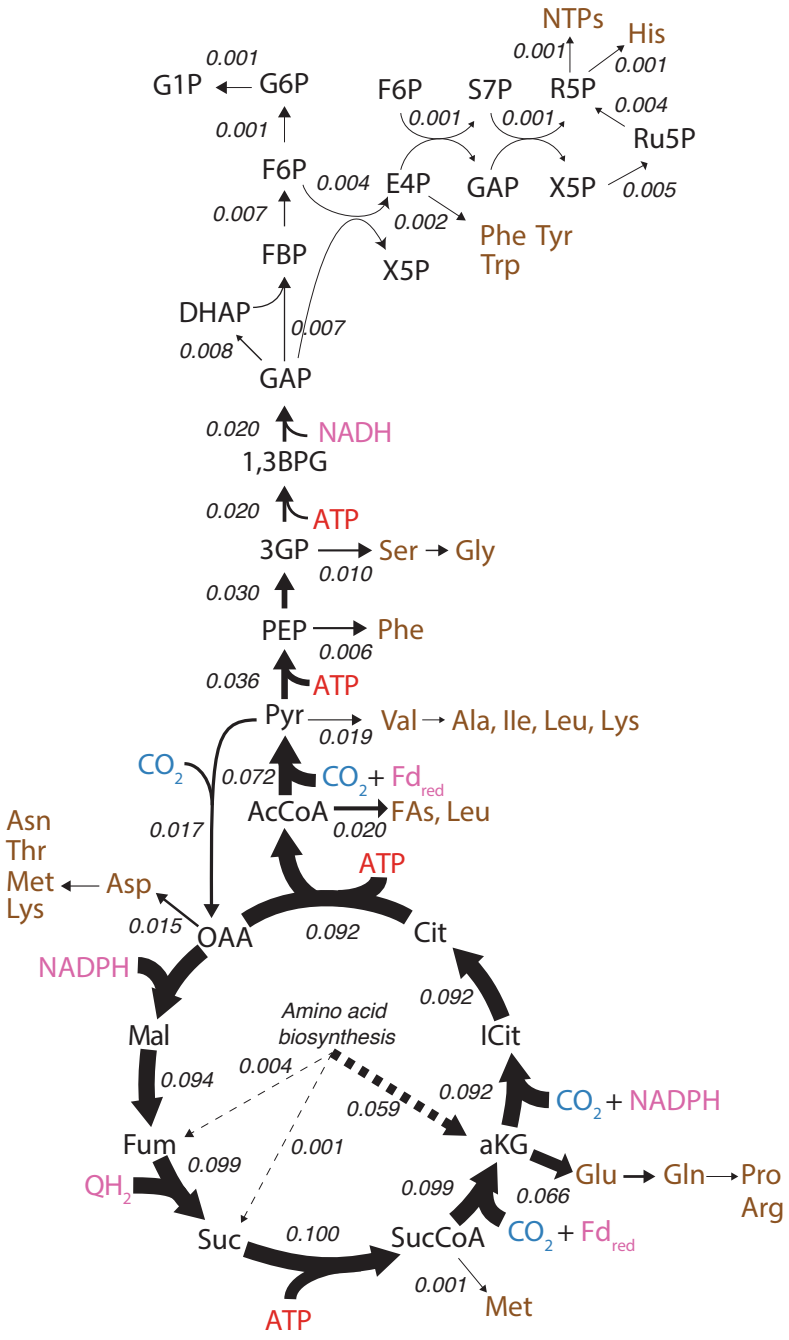

Figure 4.4 - Carbon flux distribution predicted via flux balance analysis during chemolithoautotrophic growth on nitrite. The model was constrained to a nitrite uptake rate of 8.5 mmol gDW^{-1} hr^{-1} and the biomass growth rate was 0.006 hrs^{-1}. CO_2 is shown in blue, amino acids and other biomass precursors are shown in brown, ATP is shown in red, and reducing equivalents are shown in pink. Numerical values are calculated fluxes in units of mmol gDW^{-1} hr^{-1}. Model reactions and compounds can be found in Supplementary Dataset 1. FBA solutions can be found in Supplementary Dataset 2.

In addition to the TCA cycle, fast labelling of 3-phosphoglycerate, fructose 6-phosphate, glucose 6-phosphate, and sedheptulose 7-phosphate was observed (Figure 4.5; Supplementary Dataset 3). This is consistent with the use of gluconeogenesis and the pentose phosphate pathway for synthesis of precursor metabolites in *N. moscoviensis.*

4.2.4 Analysis of growth on formate

It has been demonstrated that *N. moscoviensis* can use formate as an energy source and that the genome contains a formate transporter (focA) and a soluble NAD^+-reducing formate dehydrogenase (fds), which catalyses the oxidation of formate to CO_2 with concomitant reduction of NAD^+ to NADH[5,24]. It has also been shown that formate can also be assimilated as a carbon source[9], although the pathway for assimilation currently remains unclear. Metabolic reconstruction suggested that formate could potentially be assimilated either indirectly through oxidation to CO_2 and reassimilation via the rTCA cycle or directly based on genes encoding a reductive glycine pathway (Supplementary Dataset 1). Therefore, we used *iNmo691* to explore the growth and metabolism of *N. moscoviensis* on formate compared to chemolithoautotrophic conditions.

Growth comparisons were done on minimal NOB media with the following substrates: 1) nitrite and O_2, 2) formate and O_2, and 3) nitrite + formate and O_2. For Case 2, we further examined the growth benefits of using either the reductive glycine pathway or the rTCA cycle for formate assimilation by turning on/off key reactions in the model. The formate uptake rate in the model was set equal to the nitrite uptake rate, based on previously published data[5]. The model predicted that formate or formate + nitrite utilization would increase the growth rate of *N. moscoviensis* by approximately 2.6 or 3.8 fold, respectively (Figure 4.6). Here, electrons derived from formate oxidation were predicted to drive energy conservation via the electron transport chain, while also providing reducing equivalents for biosynthesis.

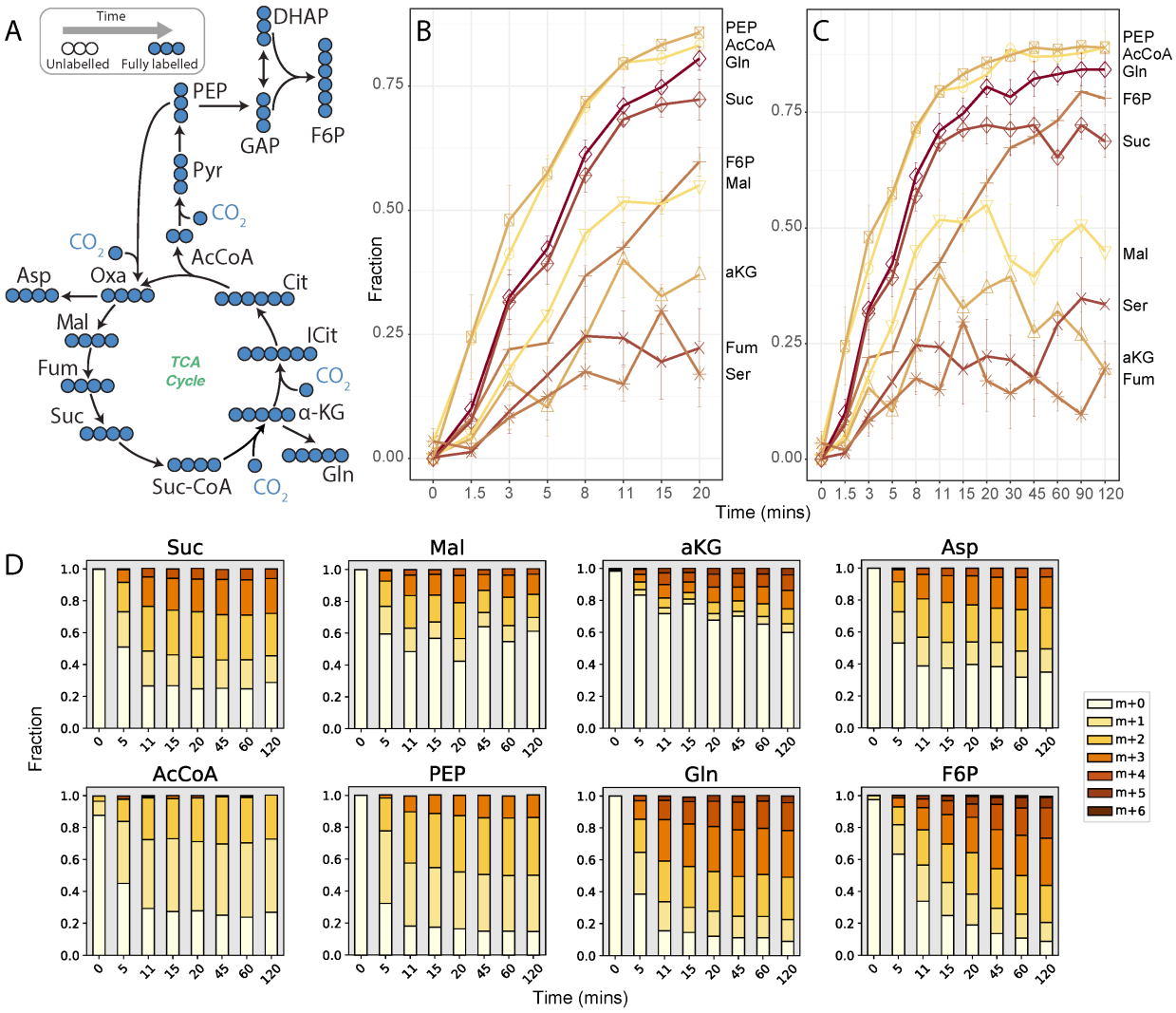

Figure 4.5 - Dynamic ^{13}C-labelling of selected central carbon metabolites during ^{13}C-bicarbonate tracer experiments. (A) Expected ^{13}C labeling of central metabolism over time. ^{13}C enrichment of selected metabolites over (B) 20 minutes and (C) 120 mins. ^{13}C-enrichment values were normalized to a tracer ^{13}C fraction of 1. (D) Mass isotopomer distributions for selected metabolites. AcCoA: acetyl-CoA, aKG: alpha-ketoglutarate, Asp: aspartic acid, F6P: fructose 6-phosphate, Fum: fumarate, Gln: glutamine, Mal: malate, PEP: phosphoenolpyruvate, Ser: serine, Suc: succinate. All measured metabolite mass isotopomer distributions can be found in Supplementary Dataset 3.

FBA also predicted that *N. moscoviensis* growth rates would improve approximately 4% by directly assimilating formate via the reductive glycine pathway versus indirectly via the rTCA cycle (Figure 4.6). This is because the reductive glycine pathway consumes less energy than the

rTCA, as low-potential electron carriers (i.e. reduced ferredoxin) are not needed for CO_2 fixation, and in total four electrons less are required per pyruvate molecule formed from formate instead of CO_2 (Figure 4.7). In this scenario, approximately 6% of the formate would be directly assimilated via the reductive glycine pathway, rather than oxidizing all formate to CO_2 followed by reassimilation (Supplementary Dataset 2). Genes for the reductive glycine pathway are all present in the *N. moscoviensis* genome (Supplementary Dataset 1) and were also found in the previously published *N. moscoviensis* transcriptome above median gene expression[12].

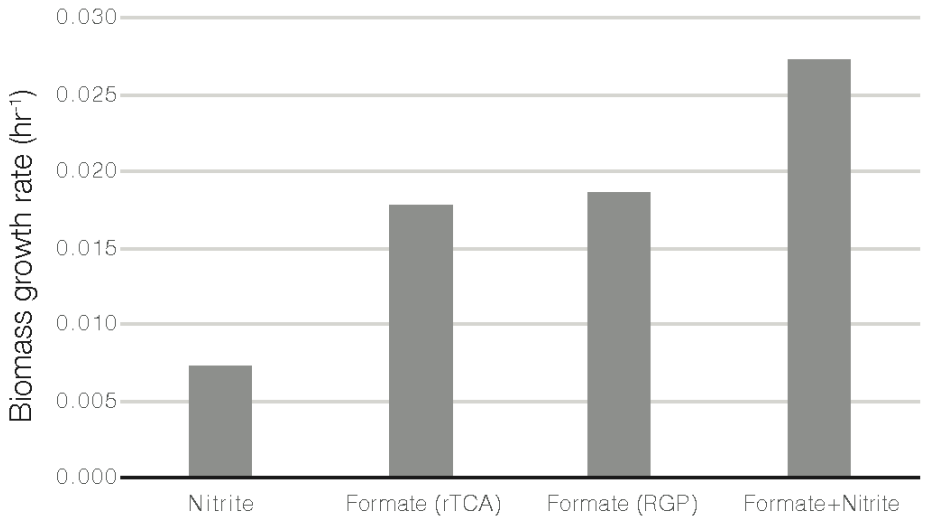

Figure 4.6 - Model predictions for growth rate based on different substrates (nitrite and formate) and different formate assimilation pathways. rTCA: reductive TCA cycle, RGP: reductive glycine pathway. In all scenarios, the uptake flux of each substrate was set equal to 8.5 mmol gDW^{-1} hr^{-1}.

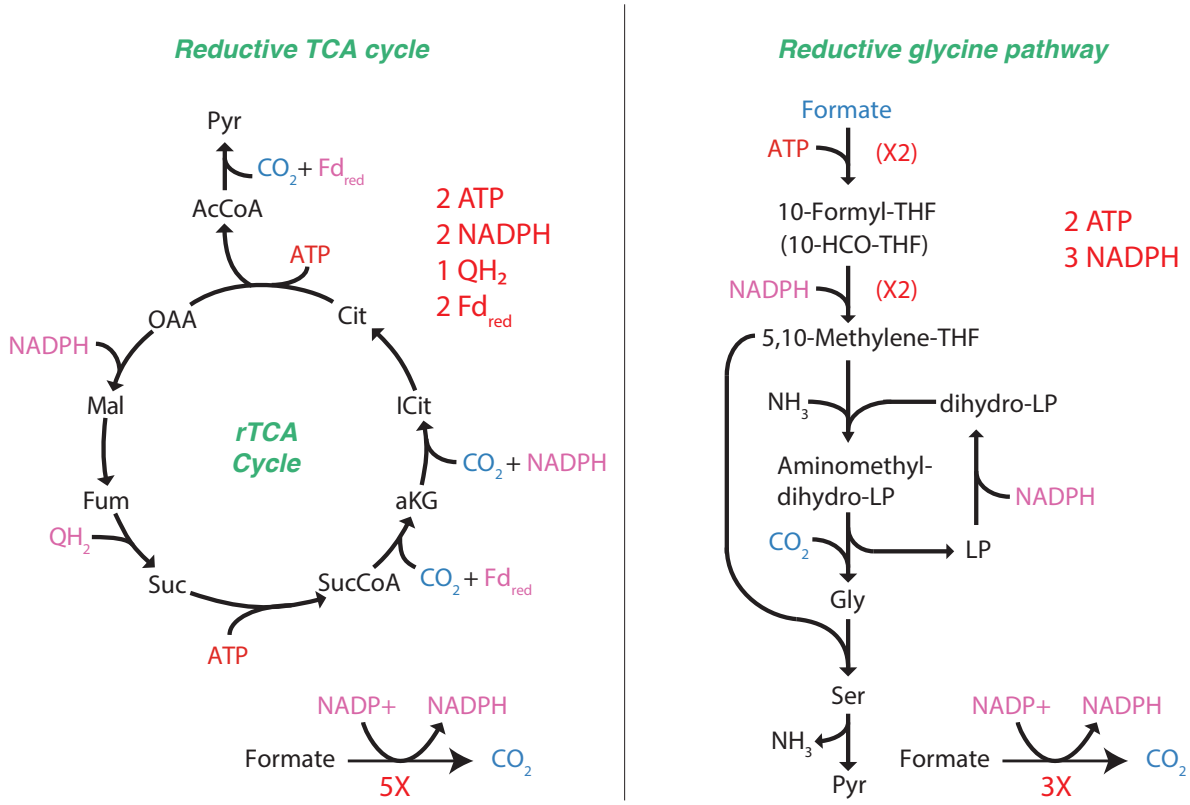

Figure 4.7 - Comparison of the reductive TCA cycle with the reductive glycine pathway.

4.2.5. ^{13}C-formate tracer experiments suggest formate is indirectly assimilated via CO_2 fixation

To experimentally determine the assimilation pathway of formate in *N. moscoviensis,* isotopic tracer experiments with ^{13}C-formate were performed with batch cultures in sealed serum bottles. Cultures were first acclimated to growth on 1 mM unlabelled formate without the presence of nitrite for 24 hours. Subsequently 0.5 mM ^{13}C-formate was added to the cultures and intracellular metabolome and gas headspace samples were collected before addition of ^{13}C-formate and at 15, 30, 60, 180, and 300 mins for isotopic analysis. Consistent with ^{13}C-formate oxidation to CO_2, continuous production of ^{13}C-CO_2 in the headspace gas was observed during the experiment (Figure 4.8A). Intracellular formate was also measured to be ~40% ^{13}C-enriched from 15 minutes onwards, consistent with formate transport into the cell (Figure 4.8C). However, the majority of

measured metabolites had no detectable ^{13}C enrichment, except for succinate, glutamate, glutamine, and fructose 6-phosphate (Supplementary Dataset 3, Figure 4.8C). Moreover, mass isotopomer distributions for these metabolites showed a consistent increase in labelled carbons overtime (Figure 4.8C), similar to results with ^{13}C-bicarbonate experiments (Figure 4.5). This suggests that formate was being oxidized to CO_2 and then assimilated via the rTCA cycle.

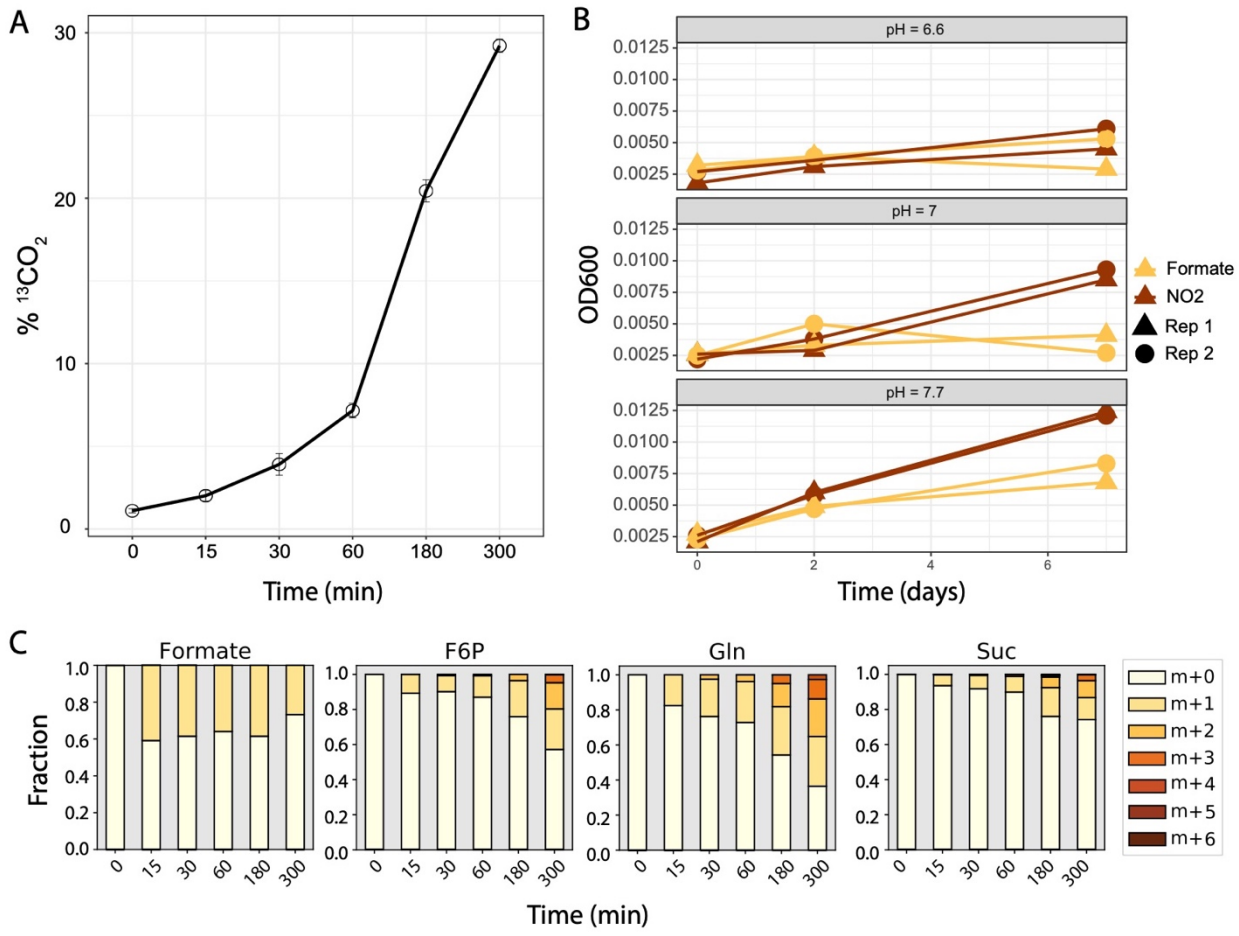

Figure 4.8 - Oxidation and assimilation of ^{13}C-formate by *N. moscoviensis.* (A) Production of ^{13}C-CO_2 from ^{13}C-formate oxidation. (B) Growth of *N. moscoviensis* using formate (yellow) or nitrite (red) as electron donor with oxygen as the terminal electron acceptor at different pH. (C) Mass isotopomer distributions for selected metabolites during batch ^{13}C-formate tracer experiments.

Batch experiments with *N. moscoviensis* also observed slower growth on formate than nitrite (Figure 4.8B). This was tested over a range of pH conditions, as previous studies indicated that formate oxidation may have a different pH optimum than nitrite oxidation[25,26]. These results reinforce the previous findings of poor growth of *N. moscoviensis* on formate[5] and do not support the model predictions of an improved growth rate on formate.

4.3 Discussion

Previous studies have shown the utility of constraint-based reconstruction and analysis for exploring the metabolic capabilities of microorganisms that have an important environmental role[27,28,29]. *iNmo691* provides a framework for examining *Nitrospira* metabolism at a systems-level and will serve as a knowledge base that can be continually refined to drive understanding and improved prediction of *Nitrospira* ecophysiology. Our analysis shows that *iNmo691* quantitatively predicts *N. moscoviensis* chemolithoautotrophic growth on nitrite. Flux balance analysis provided estimates for the amount of nitrite used for energy conservation and CO_2 fixation into biomass, offering a quantitative model for linking catabolism and anabolism. The model also accurately provides quantitative estimates of carbon and electron flux distribution, reinforcing the metabolic network predicted from recently reported *N. moscoviensis* gene expression levels[12].

^{13}C metabolomic analysis revealed a complex picture for central carbon metabolism in *Nitrospira*. While our results support that *Nitrospira* uses the rTCA cycle for CO_2 fixation under chemolithoautotrophic growth, ^{13}C enrichment values revealed extensive unlabelled pools of central carbon intermediates in the metabolome (Figure 4.5). Because all ^{13}C enrichment values increased monotonically, we reasoned that turnover of carbon storage reserves or macromolecules was not responsible for this observation. However, compartmentalization of metabolism or

substrate channeling may offer a possible explanation. In cells that contain multiple pools of the same metabolite (e.g. cells with mitochondria) differences in their labelling will result in an aggregated measurement of the mixed pool because they are extracted together. If only one of those pools is actively participating in metabolism, the aggregation of both labelled and unlabelled pools will serve to dilute the overall ^{13}C enrichment. While electron micrograph images suggest that no intracytoplasmic membranes or carboxosomes are present in *N. moscoviensis* cells, they do have an enlarged periplasmic space (40-80 nm wide), typical of *Nitrospira* [6,30,31]. The proposed role of this enlarged space has been suggested to accommodate the nitrite-oxidizing enzyme systems[32], however, it could also potentially harbour enzymes that directly participate in carbon metabolism.

An alternative mechanism to explain the low ^{13}C enrichment levels of some metabolites is substrate channeling; that is, the direct passing of pathway intermediates between enzyme active sites without escaping into the cytoplasm, facilitated by noncovalent dynamic enzyme complexes[33,34]. This can result in lower than expected ^{13}C enrichment values due to "leaked" metabolites that create a separate cytoplasmic pool with a much slower turnover rate[33,34]. Given that several metabolites immediately downstream of the rTCA cycle (aspartate, glutamate, glutamine) were more labelled than their precursor TCA intermediates, we suspect that rTCA cycle enzymes may interact to form a supramolecular complex, or metabolon, to efficiently transport reactants between enzyme active sites. The advantage of this is that high local substrate concentrations enable higher pathway fluxes and intermediates can be protected from the bulk phase, limiting competition between competing pathways and protecting the cell from toxicity[33,35]. Given the significant thermodynamic barriers encountered in the rTCA by the consecutive operation of alpha-ketoglutarate synthase and isocitrate dehydrogenase ($\Delta G > 40$ kJ/mol), as well

as pyruvate synthase ($\Delta G > 14$ kJ/mol)[36], substrate channeling maybe be important for overcoming these unfavorable reactions in addition to their indirect coupling to ATP hydrolysis via succinyl-CoA synthetase and ATP citrate lyase. Further confirmation of this channeling could be obtained through *in vivo* crosslinking coupled to proteomic analysis, as has been done to investigate the malate dehydrogenase–citrate synthase–aconitase complex of the oxidative TCA cycle[21,22].

In addition to chemolithoautotropic growth on nitrite, genome-scale modeling allowed us to assess different hypotheses regarding organotrophic and mixotrophic growth of *N. moscoviensis* on formate. *Nitrospira* have been observed to oxidize formate to CO_2 for energy conservation under laboratory conditions[5] and have also been observed to assimilate formate *in situ* during wastewater treatment[9]. However, while our analysis predicted that *N. moscoviensis* should grow 2-3 times faster on formate over nitrite, batch experiments showed lower biomass yields of *N. moscoviensis* on formate, consistent with previous reports[5]. This suggests that formate or formate metabolism may be toxic to *N. moscoviensis* via an unknown mechanism. Given that different lineages of *Nitrospria* may exhibit different formate utilization efficiencies *in situ*[9], it remains to be tested whether all *Nitrospria* grow slower using formate as an electron donor or whether other environmental conditions boost growth on formate.

Genome-scale modeling further allowed us to generate hypotheses on the use of formate as a carbon source for *N. moscoviensis* anabolism. While modeling predicted that small growth improvements would be favoured by directly assimilating formate via the reductive glycine pathway, *in vivo* ^{13}C-formate tracer experiments demonstrated that *N. moscoviensis* adhere to their autotrophic lifestyle, reducing CO_2 derived from formate via the rTCA cycle. This may be an evolutionary adaptation to avoid energetic costs associated with remodeling their proteome in

environments where formate may only become transiently available, for example at oxic-anoxic interfaces common to the habitats of *Nitrospira*[1].

In conclusion, our work provides the first genome-scale reconstruction and analysis of *Nitrospira* metabolism, offering unique insights on their versatile ecophysiology. The *iNmo691* GEM provides quantitative estimates of chemolithoautotrophic growth and pathway fluxes and serves as an invaluable tool for hypothesis-driven discovery. Our ^{13}C metabolomic results also provide the first insights on *Nitrospira's in vivo* central carbon metabolism, confirming the high activity of the rTCA cycle for CO_2 fixation. Future efforts to combine ^{13}C metabolomics with genome-scale modeling should provide a valuable approach for quantitatively understanding metabolic regulation of *Nitrospira* metabolism under different environmental conditions and microbe-microbe interactions. This will further expand the systems biology framework developed in this study, ultimately leading to the systematic prediction and control of *Nitrospira* metabolism in natural and engineered ecosystems.

4.4 Materials and Methods

4.4.1 Cultivation of N. moscoviensis cells

N. moscoviensis M-1 was grown in NOB mineral salts medium for lithoautotrophic growth as described in Spieck and Lipski (2011)[37] with the following modifications: $CaCO_3$ was replaced with $CaCl_2 \cdot 2\ H_2O$ in the same concentration and the following trace element composition was used per liter of medium: 34.4 µg of $MnSO_4 \cdot 1\ H_2O$, 50 µg of H_3BO_3, 70 µg of $ZnCl_2$, 72.6 µg of $Na_2MoO_4 \cdot 2\ H_2O$, 20 µg of $CuCl_2 \cdot 2\ H_2O$, 24 µg of $NiCl_2 \cdot 6\ H_2O$, 80 µg of $CoCl_2 \cdot 6\ H_2O$, and 2,000 µg of $FeSO_4 \cdot 7\ H_2O$. Nitrilotriacetic acid was added equimolar to all trace elements as a complexing agent.

N. moscoviensis was cultivated in a 7L membrane bioreactor (MBR) inoculated with an active batch culture and operated as described in Mundinger et al., (2019)[12]. The working volume of the reactor was 3 L and included pH, dissolved oxygen, temperature, and level controls (all by Applikon Biotechnology B.V., Schiedam, The Netherlands). The bioreactor was continuously sparged with Ar/CO_2 (95%/5% v/v) and air at a rate of 10 ml/min to maintain a dissolved oxygen concentration of ~30%. pH was controlled at 7.7 using a 1 M $KHCO_3$ buffer and the reactor temperature was maintained at 39ºC using a loop-type heat exchanger. The reactor was continuously stirred at 150 rpm. The reactor and all cultivation media and solutions were sterilized by autoclaving or sterile filtration prior to use and the reactor was operated aseptically to maintain culture purity. Nitrite and nitrate concentrations were check daily to ensure all nitrite was consumed stoichiometrically to nitrate (Nitrite test strips MQuant®, Merck, Darmstadt, Germany) and that nitrite was always limiting. Cultures were maintained in steady-state growth at a dilution rate of 0.006 hrs^{-1} and a substrate feeding rate of 2 – 2.5 mmol NO_2^- L^{-1} day^{-1}.

4.4.2 Genome-scale model reconstruction and analysis

The genome-scale metabolic model of *N. moscoviensis* (*iNmo691)* was reconstructed from the NCBI's genome sequence for *N. moscoviensis* (accession number NZ_CP011801.1) using the Model SEED pipeline[38] implemented in KBase[14], followed by manual curation using the MetaCyc database[13] and available literature[1,3,5]. The model was gap-filled manually through the addition of reactions not annotated in the genome to ensure that all biomass components could be produced on NOB minimal media. Growth and non-growth associated maintenance energy requirements were determined by plotting experimentally measured nitrite uptake rates as a function of the growth rate set during steady-state bioreactor cultivation[39]. The biomass equation was derived from biomass composition measurements of *N. moscoviensis.* The model was formulated in

Systems Biology Markup Language (SBML) level 3 version 1.0 and is available in the supplementary materials. Flux balance analysis was used to simulate *in silico* growth by solving the linear program:

$$\text{Max } v_{biomass}$$

$$s.t.$$

$$S \cdot v = 0$$

$$v_{min} \leq v \leq v_{max}$$

where $v_{biomass}$ is the flux through the biomass objective function, S is the stoichiometric matrix generated from the reconstruction with rows representing metabolites, columns representing reactions, and entries representing metabolite stoichiometric coefficients, v is the vector of steady-state reaction fluxes, and v_{min} and v_{max} are the minimum and maximum allowable reaction fluxes. Flux balance analysis was performed in Python version 3.7.2 using the COBRApy package[40].

The transcriptomic dataset reanalyzed here can be found under the NCBI Gene Expression Omnibus DataSet accession number GSE123406.

4.4.3 Biomass composition analysis

Cultures were centrifuged (10,000 rpm, 15 mins, 4°C) to obtain cell pellets, which were subsequently freeze-dried prior to analysis. Total protein concentration was determined using the Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific) and amino acid composition was determined according to Carnicer et al. (2009)[41] using a Varian 920-LC high performance liquid chromatography amino acid analyzer. Total carbohydrates were determined using the phenol-sulphuric acid method[42]. Total lipid content was determined via the sulfo-phospho-vanillin

reaction[43] and lipid composition for *N. moscoviensis* was taken from Lipski et al. (2001)[18]. Total RNA and DNA content was determined according to Benthin et al. (1991)[44]. Total inorganic content was determined by combustion of freeze-dried biomass in an oven at 550°C for 12 hours.

4.4.4 ^{13}C isotopic tracer experiments

^{13}C-labelled sodium bicarbonate was rapidly introduced (within 1 minute) into the bioreactor containing *N. moscoviensis* cells growing under steady-state conditions to a final concentration of approximately 30 mM. Following ^{13}C-label introduction, samples were rapidly withdrawn from the reactor at timepoints 0, 1.5, 3, 5, 8, 11, 15, 20, 30, 45, 60, 90, and 120 minutes. Samples were immediately filtered (Millipore 0.45 µm hydrophilic nylon filter HNWPO4700) using a vacuum pump to remove extracellular media, and filters were placed face down in 1.5 ml of -80°C extraction solvent (40:40:20 acetonitrile:methanol:water) for cell quenching and metabolite extraction. Samples were then centrifuged (10,000 rpm, 4°C, 5 mins) and 1 ml of cell-free supernatant was collected and stored at -80°C for metabolomic analysis. The time 0 min sample corresponded to the period directly before ^{13}C-label addition. The ratio of $^{13}C/^{12}C$ dissolved inorganic carbon remained constant during the course of the 2-hour experimental period as determined by GC-MS analysis.

4.4.5 Formate batch experiments

N. moscoviensis cells were harvested from a membrane bioreactor. The biomass was centrifuged 8,000xg, 15 min at 25°C and washed by resuspending the cells in fresh mineral NOB media. This was repeated until no nitrite or nitrate was detectable via nitrite/nitrate test strips (Nitrite test strips MQuant®, Merck) in the culture. Subsequently, the cells were transferred to sterile 120 ml serum bottles (triplicate experiments, 1 bottle per timepoint, 6 timepoints total) containing 50 ml NOB

mineral salts medium with 0.5 mM sodium formate and no nitrite but 0.187 mM NH_4Cl as nitrogen source. Bottles were crimp sealed with a rubber stopper to allow monitoring of the gas headspace and were incubated at 39°C in the dark. Following 24 hours of acclimation, 1 mM ^{13}C-sodium formate (Cambridge Isotopes Laboratories, MA, USA) was added to all incubations. At each timepoint (before addition, 15, 30, 60, 180, 300 minutes) the isotopic composition of the gas headspace was measured using GC-MS and subsequently bottles corresponding to a given timepoint were sacrificed for metabolomics analysis. Bottle contents were filtered at a given timepoint using a vacuum pump and metabolites were extracted using -80ºC extraction solvent (40:40:20 acetonitrile:methanol:water) as described above.

For growth experiments, *N. moscoviensis* cells from a membrane bioreactor were harvested and washed until no nitrite or nitrate was detectable via nitrite/nitrate test strips (Nitrite test strips MQuant®, Merck) (see above). The cells were transferred to sterile 120 ml serum bottles with 50 ml NOB mineral media containing 0.187 mM NH_4Cl as nitrogen source and a pH of 6.6, 7 or 7.7 adjusted using 1 M $KHCO_3$ and were sealed with a rubber stopper. The experiment was performed in duplicates (per pH value and substrate) and 5 mM sodium nitrite or 5 mM sodium formate were added to the incubations as sole energy source. Growth was monitored for 7 days using optical density measurements at wavelength 600 nm.

4.4.6 Metabolomic analysis

Samples were analysed using a high-performance HPLC–MS system consisting of a Vanquish™ UHPLC system (Thermo Scientific) coupled by electrospray ionization (ESI; negative polarity) to a hybrid quadrupole high-resolution mass spectrometer (Q Exactive Orbitrap, Thermo Scientific) operated in full scan mode for detection of targeted compounds based on their accurate masses. Properties of Full MS–SIM included a resolution of 140,000, AGC target of 1E6, maximum IT of

40 ms and scan range from 70 to 1,000 m/z. LC separation was achieved using an ACQUITY UPLC BEH C18 column (2.1 × 100 mm column, 1.7 μm particle size; part no. 186002352; serial no. 02623521115711, Waters). Solvent A was 97:3 water:methanol with 10 mM tributylamine (TBA) adjusted to pH 8.1–8.2 with 9 mM acetic acid. Solvent B was 100% methanol. Total run time was 25 min with the following gradient: 0 min, 5% B; 2.5 min, 5% B; 5 min, 20% B; 7.5 min, 20% B; 13 min, 55% B; 15.5 min, 95% B; 18.5 min, 95% B; 19 min, 5% B; 25 min, 5% B. Flow rate was 200 μl min^{-1}. The autosampler and column temperatures were 4 °C and 25 °C, respectively. Mass isotopomer distributions were corrected for natural abundance using the method of Su et al., (2017)[23] and ^{13}C enrichment values were calculated using the formula $(1/N)\ \sum_{i=1}^{N} Mi\ \times\ i$, where N is the number of carbon atoms in the metabolite and Mi is the fractional abundance of the i^{th} mass isotopomer.

4.4.7 GC-MS analysis of dissolved inorganic carbon isotopic fractions

Isotopic fractions of dissolved inorganic carbon in the liquid media were measured based on a modified headspace method[45]. 3 ml of liquid culture were collected from the bioreactor with a syringe and directly filtered through a 0.45 μm filter and 26G needle into a 120ml bottle containing 1 ml 6 M HCl (strong acid) crimp sealed with a rubber stopper. Prior to adding the liquid sample, bottles and HCl were flushed with either 100% N_2 or Ar gas to void the headspace of background CO_2. Samples were equilibrated with the acid in the bottles for at least 1 hour at room temperature to drive all dissolved inorganic carbon into the gas phase. 50 ul gas samples were then collected in a gas tight syringe with a needle (Hamilton) from the bottle's headspace and the isotopic fractions of $^{12}CO_2$ and $^{13}CO_2$ were determined using a gas chromatograph (Agilent 6890 equipped with 6 ft Porapak Q and molecular sieve columns) coupled with a mass spectrometer (GC-MS) (Agilent 5975C GC MSD; Agilent, Santa Clara, CA).

4.4.8 Supplementary Materials

Supplementary Dataset 1. *iNmo691* model reactions, metabolites, and biomass equation.

Supplementary Dataset 2. Flux balance analysis solutions for the model simulations performed in this study.

Supplementary Dataset 3. Metabolite mass isotopomer distributions and ^{13}C enrichment values from the ^{13}C-bicarbonate and ^{13}C-formate isotope tracer experiments.

Supplementary materials can be found online at:

https://github.com/celawson87/phdthesis/tree/master/chapter4

4.5 Acknowledgements

The authors would like to acknowledge Coty Weathersby for assistance with biomass composition analysis, Joshua Hamilton for helpful feedback on genome-scale model reconstruction, and David Stevenson for assistance with metabolomic analysis. Funding was provided by the National Science Foundation (CBET-1435661, CBET-1803055 and MCB-1518130), the Netherlands Organization for Scientific Research (Grants 016.Vidi.189.050 and SIAM Gravitation Grant 024.002.002), the European Research Council (ERC Advanced Grant ECO-MOM 339880), a Wisconsin Distinguished Graduate Fellowship, a Postgraduate Scholarship-Doctoral (PGS-D) by the National Sciences and Engineering Research Council of Canada (NSERC), and the UW-Madison Office of the Vice Chancellor for Research and Graduate Education through the Microbiome Initiative.

4.6 Reference

1. Daims, H., Lücker, S. & Wagner, M. A New Perspective on Microbes Formerly Known as Nitrite-Oxidizing Bacteria. *Trends Microbiol.* **24**, 699–712 (2016).
2. Daims, H., Nielsen, J. L., Nielsen, P. H., Schleifer, K.-H. & Wagner, M. In Situ Characterization of Nitrospira-Like Nitrite-Oxidizing Bacteria Active in Wastewater Treatment Plants. *Appl. Environ. Microbiol.* **67**, 5273 LP – 5284 (2001).
3. Lücker, S. *et al.* A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13479–84 (2010).
4. Koch, H. *et al.* Growth of nitrite-oxidizing bacteria by aerobic hydrogen oxidation. *Science* **345**, 1052–4 (2014).
5. Koch, H. *et al.* Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus Nitrospira. *Proc. Natl. Acad. Sci.* **112**, 201506533 (2015).
6. Ehrich, S., Behrens, D., Lebedeva, E. V, Ludwig, W. & Bock, E. A New Obligately Chemolithoautotrophic, Nitrite-Oxidizing Bacterium, Nitrospira-Moscoviensis Sp-Nov and Its Phylogenetic Relationship. *Arch. Microbiol.* **164**, 16–23 (1995).
7. van Kessel, M. A. H. J. *et al.* Complete nitrification by a single microorganism. *Nature* **528**, 555–559 (2015).
8. Daims, H. *et al.* Complete nitrification by Nitrospira bacteria. *Nature* **528**, 504–509 (2015).
9. Gruber-Dorninger, C. *et al.* Functionally relevant diversity of closely related Nitrospira in activated sludge. *ISME J.* **9**, 643–55 (2014).
10. Orth, J. D., Thiele, I. & Palsson, B. Ø. What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–8 (2010).
11. Jang, C., Chen, L. & Rabinowitz, J. D. Metabolomics and Isotope Tracing. *Cell* **173**, 822–837 (2018).
12. Mundinger, A. B., Lawson, C. E., Jetten, M. S. M., Koch, H. & Lücker, S. Cultivation and Transcriptional Analysis of a Canonical Nitrospira Under Stable Growth Conditions. **10**, 1–15 (2019).
13. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res.* **46**, D633–D639 (2017).

14. Arkin, A. P. *et al.* KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat. Biotechnol.* **36**, 566–569 (2018).
15. Bar-even, A., Noor, E., Flamholz, A. & Milo, R. Design and analysis of metabolic pathways supporting formatotrophic growth for electricity-dependent cultivation of microbes. *Biochim. Biophys. Acta* **1827**, 1039–1047 (2013).
16. Chadwick, G. L., Hemp, J., Fischer, W. W. & Orphan, V. J. Convergent evolution of unusual complex I homologs with increased proton pumping capacity : energetic and ecological implications. *ISME J.* **12**, 2668–2680 (2018).
17. Feist, A. M. & Palsson, B. O. The biomass objective function. *Curr. Opin. Microbiol.* **13**, 344–349 (2010).
18. Lipski, A., Spieck, E., Makolla, A. & Altendorf, K. Fatty Acid Profiles of Nitrite-oxidizing Bacteria Reflect theirPhylogenetic Heterogeneity. *Syst. Appl. Microbiol.* **24**, 377–384 (2001).
19. Neidhardt, F., Ingraham, J. & Schaechter, M. Physiology of the bacterial cell — A molecular approach. *Trends Genet.* **7**, 341 (1991).
20. Sumegi, B., Sherry, A. D. & Malloy, C. R. Channeling of TCA cycle intermediates in cultured Saccharomyces cerevisiae. *Biochemistry* **29**, 9106–9110 (1990).
21. Wu, F. & Minteer, S. Krebs Cycle Metabolon : Structural Evidence of Substrate Channeling Revealed by Cross-Linking and Mass Spectrometry **. *Angew. Chem. Int. Ed.* **54**, 1851–1854 (2015).
22. Zhang, Y. *et al.* Protein-protein interactions and metabolite channelling in the plant tricarboxylic acid cycle. *Nat. Commun.* **8**, 15212 (2017).
23. Su, X., Lu, W. & Rabinowitz, J. D. Metabolite Spectral Accuracy on Orbitraps. *Anal. Chem.* **89**, 5940–5948 (2017).
24. Oh, J. & Bowien, B. Structural Analysis of the fds Operon Encoding the NAD+-linked Formate Dehydrogenase of Ralstonia eutropha. *J. Biol. Chem.* **273**, 26349–26360 (1998).
25. Van Gool, A. & Laudelout, H. Formate utilization by Nitrobacter wibogradskyi. *Biochim. Biophys. Acta - Gen. Subj.* **127**, 295–301 (1966).
26. O'Kelley, J. C. & Nason, A. Particulate formate oxidase from Nitrobacter agilis. *Biochim. Biophys. Acta - Bioenerg.* **205**, 426–436 (1970).
27. Stolyar, S. *et al.* Metabolic modeling of a mutualistic microbial community. *Mol. Syst.*

Biol. **3**, 92 (2007).

28. Zhuang, K. *et al.* Genome-scale dynamic modeling of the competition between Rhodoferax and Geobacter in anoxic subsurface environments. *ISME J.* **5**, 305–316 (2011).
29. Mellbye, B. L. *et al.* Genome-Scale, Constraint-Based Modeling of Nitrogen Oxide Fluxes during Coculture of Nitrosomonas europaea and Nitrobacter winogradskyi. *mSystems* **3**, e00170-17 (2018).
30. Off, S., Alawi, M. & Spieck, E. Enrichment and Physiological Characterization of a Novel Nitrospira-Like Bacterium Obtained from a Marine Sponge. *Appl. Environ. Microbiol.* **76**, 4640 LP – 4646 (2010).
31. Nowka, B., Daims, H. & Spieck, E. Comparison of oxidation kinetics of nitrite-oxidizing bacteria: Nitrite availability as a key factor in niche differentiation. *Appl. Environ. Microbiol.* **81**, 745–753 (2015).
32. Spieck, E., Ehrich, S., Aamand, J. & Bock, E. Isolation and immunocytochemical location of the nitrite-oxidizing system in nitrospira moscoviensis. *Arch. Microbiol.* **169**, 225–30 (1998).
33. W, F. P., Williams, T. C. R., Sweetlove, L. J. & Ratcliffe, R. G. Capturing Metabolite Channeling in Metabolic. *Plant Physiol.* **157**, 981–984 (2011).
34. Sweetlove, L. J. & Fernie, A. R. substrate channelling in metabolic regulation. *Nat. Commun.* **9**, 2136 (2018).
35. Bulutoglu, B., Garcia, K. E., Wu, F., Minteer, S. D. & Banta, S. Direct Evidence for Metabolon Formation and Substrate Channeling in Recombinant TCA Cycle Enzymes. *ACS Chem. Biol.* **11**, 2847–2853 (2016).
36. Bar-even, A., Flamholz, A., Noor, E. & Milo, R. Thermodynamic constraints shape the structure of carbon fi xation pathways. *Biochim. Biophys. Acta* **1817**, 1646–1659 (2012).
37. Spieck, E. & Lipski, A. Cultivation, Growth Physiology, and Chemotaxonomy of Nitrite-Oxidizing Bacteria. in *Research on Nitrification and Related Processes, Part A* (ed. Klotz, M. G. B. T.-M. in E.) **486**, 109–130 (Academic Press, 2011).
38. Henry, C. S. *et al.* High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* **28**, 977–82 (2010).
39. Varma, A. & Palsson, B. O. Stoichiometric flux balance models quantitatively predict

growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *Appl. Environ. Microbiol.* **60**, 3724–3731 (1994).

40. Ebrahim, A., Lerman, J. A., Palsson, B. Ø. & Hyduke, D. R. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* **7**, 74 (2013).
41. Carnicer, M. *et al.* Macromolecular and elemental composition analysis and extracellular metabolite balances of Pichia pastoris growing at different oxygen levels. *Microb. Cell Fact.* **8**, 65 (2009).
42. Herbert, D., Phipps, P. J. & Strange, R. E. Chemical Analysis of Microbial Cells. in *Methods in Microbiology* (eds. Norris, J. R. & Ribbons, D. W. B. T.-M. in M.) **5**, 209–344 (Academic Press, 1971).
43. Izard, J. & Limberger, R. J. Rapid screening method for quantitation of bacterial cell lipids from whole cells. *J. Microbiol. Methods* **55**, 411–418 (2003).
44. Benthin, S., Nielsen, J. & Villadsen, J. A Simple and reliable method for the determination of cellular RNA content. *Biotechnol. Tech.* **5**, 39–42 (1991).
45. Åberg, J. & Wallin, B. Evaluating a fast headspace method for measuring DIC and subsequent calculation of pCO2 in freshwater systems. *Inl. Waters* **4**, 157–166 (2014).

5. Oxygen uptake flux predicted to be key selection parameter controlling the interaction between comammox and anammox bacteria

This chapter has been formatted as a manuscript for submission with the following authors:

Christopher E. Lawson, Maartje A. H. J. van Kessel, Martin Pabst, Mike S.M. Jetten, Daniel R. Noguera, Katherine D. McMahon, Sebastian Lücker

All supplementary materials are available online at:
https://github.com/celawson87/phdthesis/tree/master/chapter5

Author Contributions

C.E.L., S.L., K.D.M., and D.R.N. designed the study. C.E.L. performed the genome-scale modeling and analysis and performed the metagenomic analysis. M.A.H.J.K. operated the comammox-anammox reactor. M.P. and C.E.L. performed the metaproteomic analysis. C.E.L. wrote the manuscript. S.L., K.D.M., D.R.N., and M.P. M.A.H.J.K. provided valuable feedback and edits on the manuscript.

5.0 Abstract

Anaerobic ammonium oxidation (anammox) represents one of the most energy-efficient processes for the removal of nitrogen from wastewater. The performance of the anammox process in engineered ecosystems depends on the metabolic interactions between several functional guilds, including ammonia and nitrite-oxidizing bacteria (AOB and NOB, respectively), denitrifying bacteria, and anammox bacteria. Members of the genus *Nitrospira* were traditionally perceived to be strict NOB, but some species have recently been shown to be capable of complete ammonia oxidation to nitrate (or "comammox"). While the role of comammox *Nitrospira* in wastewater treatment remains largely unknown, they are expected to compete with anammox bacteria for ammonia or nitrite, based on their ability to use both substrates for energy conservation. However, recent findings demonstrate that both anammox and comammox organisms can co-occur in close association under low dissolved oxygen conditions, presumably as cooperators. Here, we use genome-scale dynamic modeling to investigate the metabolic interactions driving cooperation and competition between comammox *Nitrospira* and anammox bacteria affiliated with *Brocadia* under various oxygen conditions. We also use metagenomic and metaproteomic analysis to generate hypotheses on their interactions with heterotrophs. Our simulations show that when comammox *Nitrospira*'s growth is limited by oxygen uptake they perform incomplete nitrification and can cooperate with anammox bacteria to effectively cycle ammonia to dinitrogen gas. Moreover, nitrate produced by anammox bacteria can be used by comammox *Nitrospira* as a terminal electron acceptor resulting in nitrite comproportionation, further reducing oxygen requirements for ammonia removal. We also show that co-occurring heterotrophs encode genes for denitrification and express enzymes involved in the uptake of amino acids and peptides, which may be produced by *Nitrospira* and *Brocadia*. Our models provide valuable tools for exploring the metabolism of

nitrogen cycling microbial communities and suggest that microbiomes containing comammox *Nitrospira* and *Brocadia* could form efficient partial nitritation-anammox processes for mainstream wastewater treatment.

5.1 Introduction

Anaerobic ammonium oxidation (anammox) represents one of the most energy-efficient and cost-effective processes for the removal of nitrogen from wastewater. In practice, the process is commonly coupled to nitritation, where aerobic ammonia-oxidizing bacteria (AOB) first convert approximately 50% of the ammonium to nitrite using oxygen as a terminal electron acceptor, followed by anammox bacteria that anaerobically oxidize the remaining ammonium to nitrogen gas using nitrite as a terminal electron acceptor (termed partial nitritation-anammox, PNA)[1,2]. A diverse group of other nitrogen cycling organisms, including nitrite-oxidizing bacteria (NOB) and denitrifying heterotrophic bacteria, are also present in PNA processes[3,4]. These organisms often compete with anammox bacteria for nitrite, but in some cases may cooperate with them; for instance, by creating a "nitrite-nitrate loop"[3,4]. Thus, operation of PNA processes requires balanced activity between these functional guilds to maintain stable performance.

A particular operational challenge for PNA systems has been suppressing the activity of *Nitrospira*-affiliated NOB that compete for nitrite with anammox bacteria[3,5,6,7]. Common strategies for NOB suppression include biomass segregation, as NOB prefer smaller granules/flocs with lower mass transfer limitations compared to anammox bacteria[8,9], and/or intermittent aeration with low dissolved oxygen control[10]. While these strategies have had reasonable success in controlling NOB activity in systems treating high-strength ammonia wastewaters at higher temperatures (e.g. anaerobic digestion centrate), control of NOB activity under mainstream

treatment conditions remains a major challenge[7,10]).

Recently, bacteria from the genus *Nitrospira* have been discovered that contain both pathways for ammonia and nitrite oxidation and are capable of complete nitrification[11,12]. These organisms were predicted to exist prior to their discovery based on kinetic theory of optimal pathway length, which postulated them to have a lower growth rate and higher growth yield than canonical AOB[13], a theory recently confirmed with a pure isolate of *Nitrospira inopinata*[14]. To date, comammox *Nitrospria* have been detected in a range of mainstream nitrogen removal systems[11,15,16], but so far have not been detected in sidestream PNA systems treating high strength wastewaters[17], consistent with their preference for lower nitrogen loaded systems[13,14]. While it is expected that comammox *Nitrospira* would compete with anammox bacteria for ammonium under mainstream conditions, van Kessel et al. (2015)[12] have observed tight clustering and coexistence between comammox *Nitrospira* and *Brocadia*-affiliated anammox bacteria in oxygen-limited bioreactors[12]. This suggests that comammox bacteria may cooperate, rather than compete with anammox bacteria under certain conditions, although the mechanism behind this interaction remains unknown.

Here, we use genome-scale dynamic modeling to evaluate environmental controls and molecular mechanism underlying cooperation and competition between anammox and comammox bacteria. Possible interaction mechanisms between anammox, comammox, and heterotrophic organisms are also presented based on metagenomic and metaproteomic analyses. Our results show that when ammonia uptake flux is limiting, comammox *Nitrospira* perform complete nitrification and no nitrite is produced to support anammox bacterial growth. However, when oxygen uptake flux is limiting, comammox bacteria only oxidize ammonia to nitrite, which can be coupled with anammox for the complete removal of ammonium to nitrogen gas. We also show

that heterotrophic bacteria in the community are likely denitrifiers that express transporters for amino acid and peptide substrates possibly produced by anammox and comammox bacteria. We conclude by discussing how these interactions could be used to design mainstream PNA applications.

5.2 Results

5.2.1 Model reconstruction

The genome-scale metabolic network of *Nitrospira nitrosa* (*iNsa694)* and *Brocadia sinica* (*Bsi788*) were reconstructed from the most recent *N. nitrosa* and *B. sinica* genome annotations (NCBI WGS Project numbers NZ_CZQA01 and BAFN01, respectively), aided by reaction annotations in the MetaCyc[18] and ModelSEED databases available through the Department of Energy Systems Biology Knowledgebase[19]. The final reconstruction for *Bsi788* contained 751 reactions, 710 metabolites, and 788 genes, whereas the final reconstruction for *iNsa694* contained 672 reactions, 636 metabolites, and 694 genes. Reconstruction of *N. nitrosa's* electron transport chain was based on the recently reconstructed metabolism of *N. moscoviensis* (Chapter 4) for nitrite oxidation pathways, together with models of aerobic ammonia oxidation pathways from *Nitrosomonas europaea* for ammonia monooxygenase and hydroxylamine dehydrogenase[20] (Figure 5.1; Supplementary Dataset 1). Reconstruction of *B. sinica's* energy metabolism was based on current electron transport chain models for anammox bacteria (Figure 5.1; Supplementary Dataset 1)[21,22], whereas central carbon metabolism was reconstructed from genomic information and recent ^{13}C tracer studies (Chapter 3).

The biomass objective function for *N. nitrosa* was assumed to be the same as that determined for *N. moscoviensis* (Chapter 4). Reconstruction of the biomass objective function for

B. sinica was based on the measured biomass composition of the anammox bacterium *Kuenenia stuttgartiensis* (Table 5.1 and Table 5.2)[23], as a *B. sinica* enrichment culture was not available, and published fatty acid composition data[24].

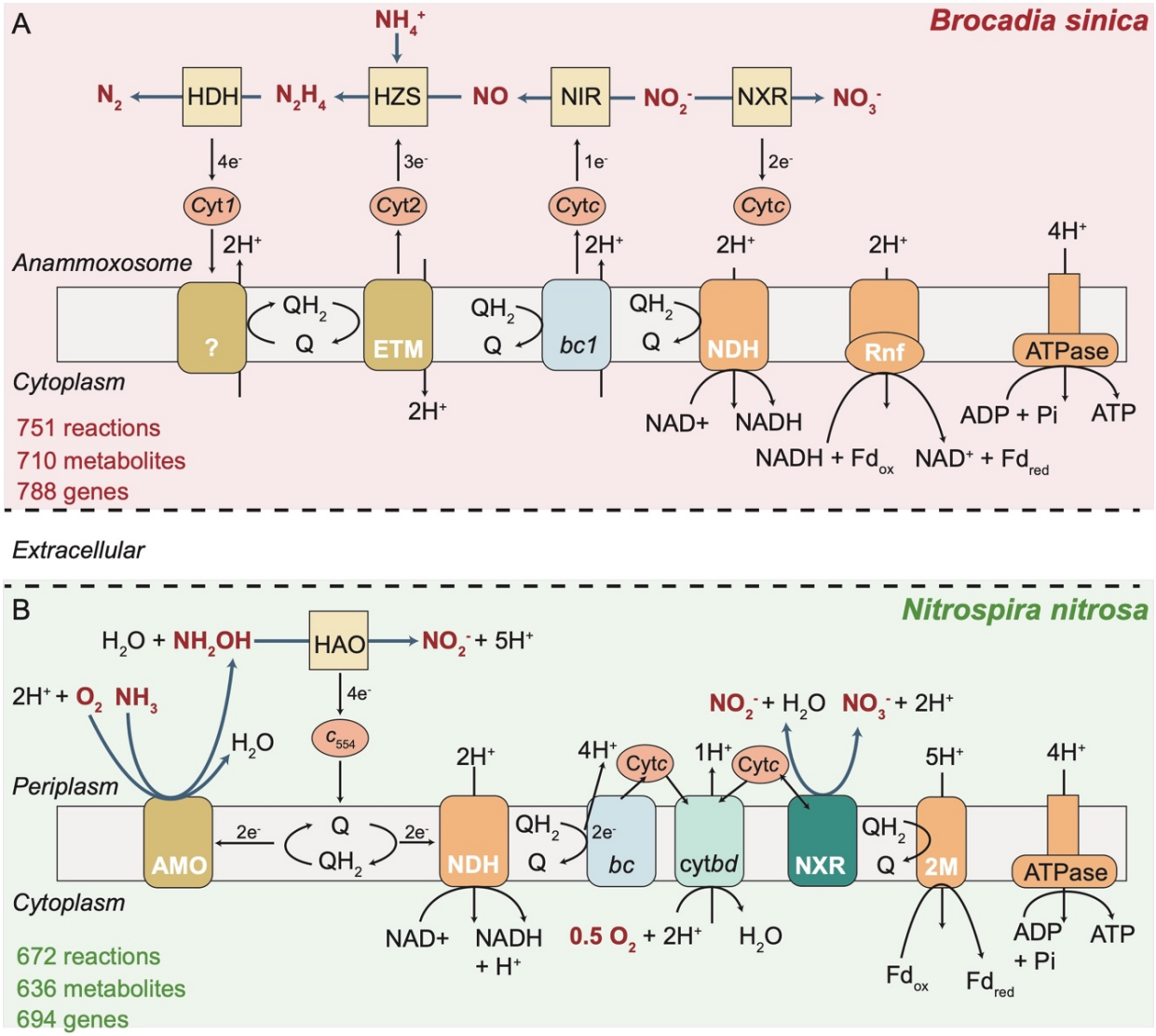

Figure 5.1 - Electron transport chain models for *Brocadia sinica* and *Nitrospira nitrosa*. HDH, hydrazine dehydrogenase; HZS, hydrazine synthase; NIR, nitrite reductase; NXR, nitrite oxidoreductase; Cyt, cytochromes (red); ETM, electron transfer module; *bc*, cytochrome *bc* complex (complex III); NDH, NADH:quinone dehydrogenase (complex I), Rnf, Rnf complex; cytbd, terminal cytochrome *c* oxidase (complex IV); ATPase, ATP synthase; AMO, ammonia monooxygenase; HAO, hydroxylamine dehydrogenase.

Table 5.1 - Anammox biomass composition

Molecule	Average (% DW)	Std Dev (% DW)
RNA	2	0.04
DNA	1	n.d.
Proteins	75	2
Carbohydrates	10	2
Lipids	15	2
Inorganics	5	n.d.

Table 5.2 - Anammox biomass amino acid composition

Amino Acid	Mass (µmol/mgDW)	Std Dev (µmol/mgDW)
Ala	375	39
Arg	279	57
Asp	130	10
Glu	107	8
Gly	779	117
His	42	8
Ile	268	43
Leu	393	55
Lys	155	44
Met	219	58
Phe	205	43
Pro	460	96
Ser	436	94
Thr	381	87
Tyr	19	8
Val	438	76

Notes: Biomass and amino acid composition based on measurements of the anammox bacterium Kuenenia stuttgartiensis assumed to be comparable to B. sinica. DW = biomass dry weight.

5.2.2 Oxygen-limited growth promotes incomplete nitrification by comammox bacteria

We first examined the growth of *N. nitrosa* in isolation to understand the environmental parameters controlling complete nitrification (ammonia oxidation to nitrate) versus incomplete nitrification

(ammonia oxidation to nitrite). We hypothesized that oxygen-limiting conditions would result in nitrite production over nitrate, as these conditions appeared to favor anammox-comammox cooperation *in situ*[12]. Therefore, three scenarios were examined using flux balance analysis to explore *N. nitrosa* growth and end-product formation: Case 1) ammonia-limited growth, Case 2) oxygen-limited growth, and Case 3) oxygen-limited growth with nitrate present in the media. FBA simulations were performed on minimal media, where either the ammonia (Case 1) or oxygen (Case 2 and 3) uptake flux was constrained to be limiting (see Figure 5.2 notes for details).

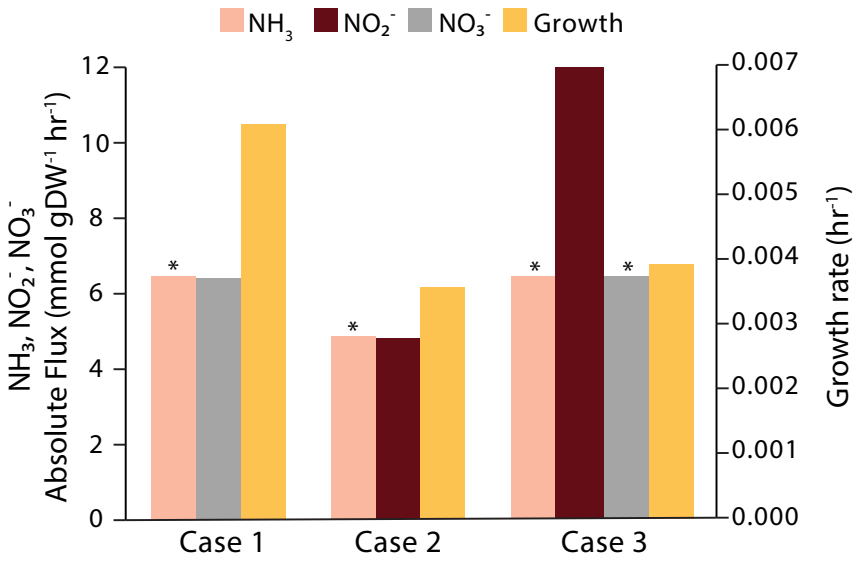

Figure 5.2 - Complete versus incomplete nitrification FBA predictions for *N. nitrosa*. Model constraints on substrate uptake were: Case 1 – ammonia limiting) NH_3 = -6.4 mmol gDW^{-1} hr^{-1} (V_{max}), O_2 = -50 mmol gDW^{-1} hr^{-1}. Case 2 – oxygen limiting) NH_3 = -6.4 mmol gDW^{-1} hr^{-1} (V_{max}), O_2 = -6 mmol gDW^{-1} hr^{-1}, Case 3 – oxygen limiting + nitrate) NH_3 = -6.4 mmol gDW^{-1} hr^{-1} (V_{max}), O_2 = -6 mmol gDW^{-1} hr^{-1}, NO_3^- = -50 mmol gDW^{-1} hr^{-1}. Maximum specific ammonium uptake rate (V_{max}) for *N. nitrosa* was assumed to be equal to *N. inopinata*[14]. * indicates negative flux values.

During ammonia-limited growth (Case 1), FBA simulations predicted that *N. nitrosa* would oxidize ammonia completely to nitrate (Figure 5.2). This serves to maximize *N. nitrosa*'s growth

rate because more energy can be generated through proton translocation from both the cytochrome *bc* complex (net 4 H^+) and the terminal oxidase (net 2 H^+) (Figure 5.3A). However, during oxygen-limited growth (Case 2), FBA simulations predicted that *N. nitrosa* would oxidize ammonia incompletely to nitrite (Figure 5.2). Analysis of the flux distribution through the electron transport chain allowed us to infer a mechanism for this incomplete nitrification phenotype (Figure 5.3B). When oxygen is limiting, *N. nitrosa* must efficiently use oxygen for two reactions of the electron transport chain: at the ammonia monooxygenase (AMO) and the terminal oxidase. Our flux analysis showed that using oxygen for these reactions to oxidize only ammonia is more favorable than using additional oxygen at the terminal oxidase to convert nitrite to nitrate. This is because more net energy is conserved per mol oxygen consumed when ammonia is oxidized to nitrite (net 3.3 H^+ per mol O_2) than complete oxidation to nitrate (net 3.0 H+ per mol O_2) (Figure 5.3). Thus, under oxygen-limited growth, incomplete nitrification is favoured because it maximizes *N. nitrosa's* growth rate and energy yield per oxygen consumed, generating nitrite that could be used for anammox bacterial growth.

Since anammox bacteria produce nitrate during growth[25], a final case evaluated with FBA was oxygen-limited growth of *N. nitrosa* with nitrate available in the media as an additional terminal electron acceptor (Case 3). Under these conditions, FBA predicted higher growth rates for *N. nitrosa* compared to oxygen-limited growth without nitrate available (Figure 5.2). Flux analysis revealed that under this scenario a new phenotype emerges, where all the oxygen is used by AMO to produce hydroxylamine, and nitrate is used as the sole terminal electron acceptor via reversal of the nitrite oxidoreductase (NXR) (Figure 5.3). This phenotype is predicted to maximize nitrite production via nitrite comproportionation, which could further benefit the growth of anammox bacteria though a "nitrite-nitrate" loop mechanism if grown together in co-culture.

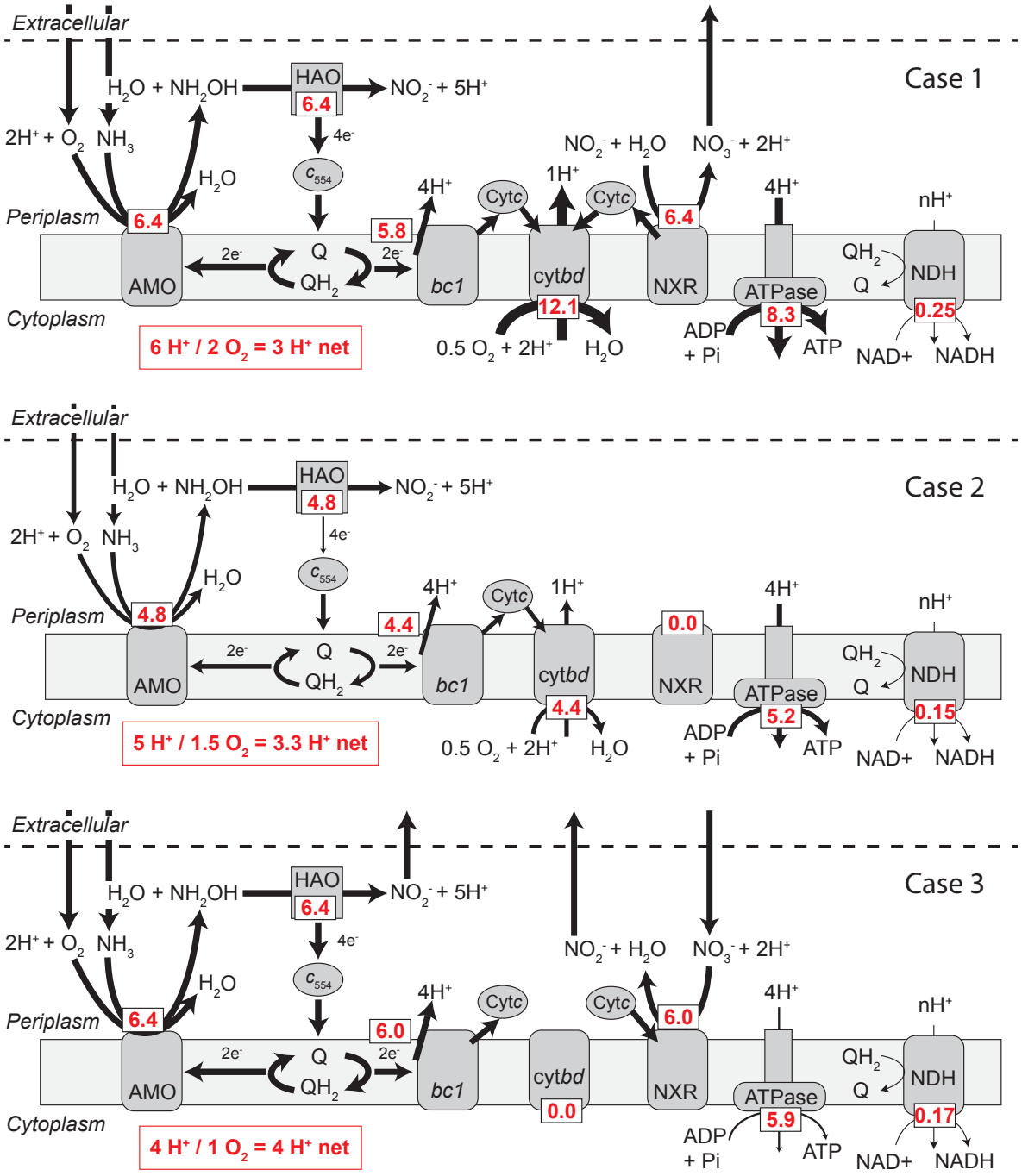

Figure 5.3 - Flux distributions through electron transport chain during complete nitrification (Case 1, Top), incomplete nitrification (Case 2, Middle), or nitrite comproportionation (Case 3, Bottom) by *N. nitrosa*.

5.2.3 Oxygen uptake rate balances activity between comammox and anammox bacteria

To evaluate how oxygen and ammonia availability influence the interactions between anammox and comammox bacteria, we next simulated their growth in co-culture using dFBA. Batch experiments were simulated using the reconstructed genome-scale models (*iBsi788* and *iNsa694)* with Monod kinetic terms, where individual genome-scale models were integrated into a dynamic multi-species framework that enabled simulation of batch reactor nutrient profiles, metabolic fluxes, and interspecies interactions (see methods). Consistent with predictions of *N. nitrosa* growth in isolation, when ammonia was limiting, *N. nitrosa* performed completed nitrification resulting in nitrate production over nitrite (Figure 5.4A). This prevented the growth of *B. sinica* because no nitrite was available for anammox. However, in simulations where oxygen was limiting growth, dFBA predicted that ammonia could be completely removed from the system as nitrogen gas (Figure 5.4B). Flux distributions predicted that when in co-culture with *B. sinica* under oxygen limited growth, *N. nitrosa* performed nitrite comproportionation, oxidizing ammonia to nitrite and also reducing the nitrate produced by *B. sinica* to nitrite (Figure 5.4B). Because nitrate production by *B. sinica* was not sufficient to support all ammonia oxidation, some oxygen was also used as a terminal electron acceptor by *N. nitrosa* in addition to nitrate.

The magnitude of the oxygen uptake rate by *N. nitrosa* was also predicted to impact interactions with *B. sinica*. Because the maximum specific ammonia uptake rate of *N. nitrosa* is higher than *B. sinica* (6.40 vs 2.95 mmol gDW^{-1} hr^{-1}), model simulations showed that tapering the oxygen uptake rate of *N. nitrosa* such that nitrite production was balanced with consumption by *B. sinica* was critical. If the oxygen uptake rate was too high, simulations showed that nitrite would accumulate in the media, which is known to inhibit anammox growth[26] (Figure 5.4C). Moreover, simulations showed that if ammonia was consumed too quickly, *N. nitrosa* would subsequently

compete with anammox bacteria for the accumulated nitrite (Figure 5.4C). Thus, our analysis suggests that precise control of *N. nitrosa's* oxygen uptake rate is important for balancing nitrite production for *B. sinica* and maximizing the conversion of ammonia to dinitrogen gas.

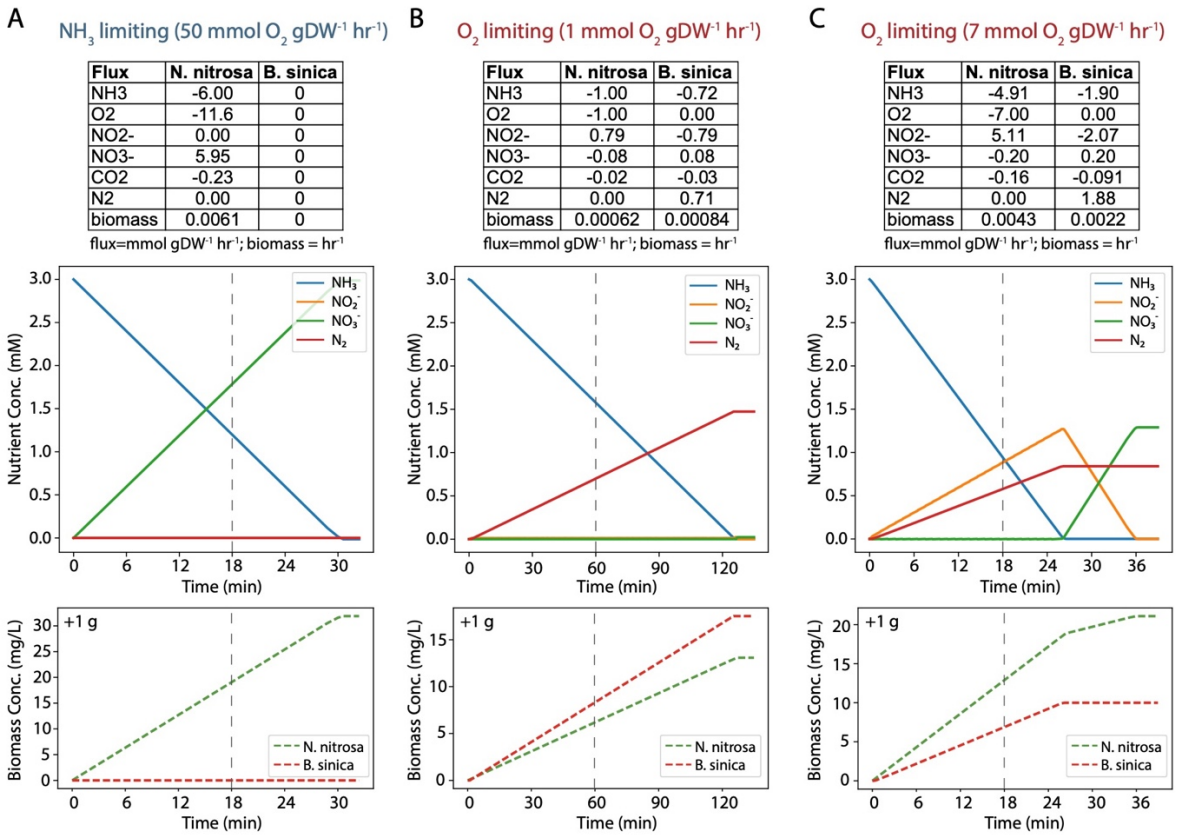

Figure 5.4 - Impact of oxygen uptake flux on comammox-anammox interaction determined via dFBA. Plots show (top) model exchanges fluxes, (middle) batch reactor nutrient profiles, and (bottom) batch reactor biomass concentrations. *N. nitrosa* maximum specific uptake rates for ammonia and nitrite were 6.40 and 8.45 mmol gDW^{-1} hr^{-1}, respectively. *N. nitrosa* half saturation constants for ammonia and nitrite were 0.049 and 372 µM respectively. *B. sinica* maximum specific uptake rates for ammonia and nitrite were 2.95 and 4.00 mmol gDW^{-1} hr^{-1}, respectively. *B. sinica* half saturation constants for ammonia and nitrite were 28 and 34 µM respectively. Allowable O_2 uptake flux (A) 50 mmol gDW^{-1} hr^{-1}, (B) 1 mmol gDW^{-1} hr^{-1}, (C) 7 mmol gDW^{-1} hr^{-1}.

5.2.4 Potential functions and interactions of denitrifying heterotrophic bacteria

The initial bioreactor containing comammox and anammox bacteria was also observed to contained diverse heterotrophic bacteria, as many of the metagenomic contigs from the reactor belonged to taxonomic groups known to contain heterotrophs, such as the Proteobacteria, Bacteroidetes, Chloroflexi, Firmicutes, and Planctomycetes[12]. Therefore, we recovered metagenome-assembled genomes (MAGs) of five abundant putative heterotrophic bacteria from the comammox-anammox reactor and reconstructed their metabolism to explore their potential function and interactions in the microbiome. We also generated a metaproteome of this microbiome growing on defined autotrophic media to examine functions that were highly expressed by the community.

Table 5.3 shows five of the most abundant heterotrophic MAGs that were recovered via metagenomic binning. These MAGs spanned diverse taxonomic groups, including the orders Rhizobiales (PRO1), Planctomycetales (PLA1 and PLA2), Deferribacterales (DEF1), and the phylum Cyanobacteria (CYA1) and represented high-quality draft genomes[27]. Metabolic reconstruction showed that none of these organisms contained known CO_2 fixation pathways (Supplementary Data 2), suggesting that they acquired organic substrates from autotrophic bacteria, such as *Nitrospira* or *Brocadia*, in the reactor. All genomes encoded enzymes involved in denitrification and two (PLA2 and DEF1) encoded genes for dissimilatory nitrite reduction to ammonia mediated by cytochrome c nitrite reductase (Figure 5.5; Supplementary Data 2). This indicated that these organisms likely performed denitrification in the reactor. Interestingly, only one genome (PRO1) encoded the full denitrification pathway (nitrate to dinitrogen gas), while others encoded specific steps in denitrification. For example, PLA1 only encoded nitrous oxide reductase, whereas CYA1 encoded respiratory nitrate reductase and nitric oxide reductase (Figure

5.5). This is consistent with other nitrogen cycling reactors that found denitrification steps distributed across multiple organisms [3,4].

The MAGs encoded genomic potential to use formate, acetate, and H_2 as electron donors that could fuel denitrification (Figure 5.5). PRO1, PLA1, and DEF1 encoded formate dehydrogenase that oxidizes formate to CO_2 reducing NAD^+, whereas PLA1, DEF1, and CYA1 encoded NiFe hydrogenase that catalyzes NAD+-dependent H_2 oxidation to H^+ (Figure 5.5). PLA1 also encoded a bifurcating Ech hydrogenase that may allow ferredoxin-dependent energy conservation or be used to reduce ferredoxin[28,29]. Genes for acetate assimilation to acetyl-CoA mediated by acetyl-coA synthetase were found in PLA2, DEF1, and CYA1, whereas genes encoding acetate kinase and phosphate acetyltransferase were found in PRO1 and DEF1. PRO1, DEF1, and CYA1 also encoded malate synthase and isocitrate lyase of the glyoxylate cycle, which would allow these organisms to use acetate as a sole carbon source.

In addition to these simple substrates, all MAGs encoded a large number of transporters for amino acids, peptides, oligopeptides, dicarboxylic acids, and sugars (Supplementary Dataset 2). In particular, the most abundant heterotroph (PRO1) encoded a multitude (>75) of transporters involved in amino acid and peptide transport, suggesting that these were key substrates for these organisms. Consistent with this, PRO1 peptides assigned to enzymes involved in amino acid and peptide transport (e.g. oligopeptide transporters, branched chain amino acid transporters) were among the most abundant in the metaproteome (Table 5.4; Supplementary Dataset 3). This suggests that scavenging of amino acids and peptides derived from anammox and comammox bacteria present in the extracellular polymeric matrix and/or from cell turnover may be a key carbon and energy source for these organisms, as proposed for heterotrophs in other anammox systems[4].

Table 5.3 – Metagenome-assembled genome statistics

Bin ID	Complete (%)	Redundant (%)	size (Mbp)	# Contigs	GC (%)	Cov	Taxonomy
PRO1	99	1	5.34	34	64.5	76	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales
PLA1	95	1	5.07	171	66.4	73	Bacteria; Planctomycetes; Planctomycetia; Planctomycetales; Planctomycetaceae
PLA2	99	1	5.86	118	62	58	Bacteria; Planctomycetes; Planctomycetia; Planctomycetales; Planctomycetaceae
DEF1	96	2	6.70	45	51.5	56	Bacteria; Deferribacteres; Deferribacteres; Deferribacterales; Caldithrix
CYA1	94	6	6.76	51	50.6	37	Bacteria; Cyanobacteria

Notes: Cov = Coverage in metagenome.

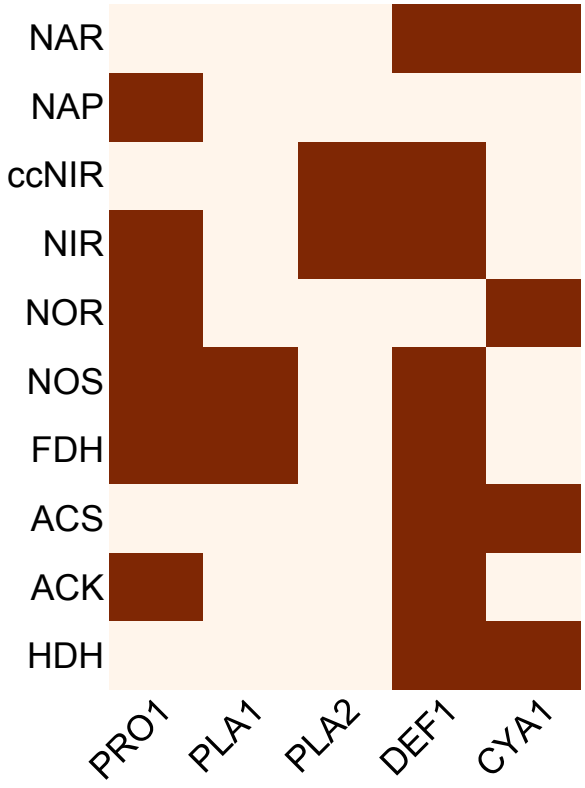

Figure 5.5 - Presence of denitrification genes and genes involved in formate, acetate, and hydrogen metabolism encoded across the analyzed MAGs. NAR, respiratory nitrate reductase; NAP, periplasmic nitrate reductase; ccNIR, ammonia-forming cytochrome c nitrite reductase; NIR, nitric oxide-forming nitrite reductase; NOR, nitric oxide reductase; NOS, nitrous oxide reductase; FDH, formate dehydrogenase; ACS, acetyl-CoA synthetase; ACK, acetate kinase; HDH, hydrogenase. All annotated open reading frames for each MAG can be found in Supplementary Dataset 2.

Table 5.4 - Top 15 most abundant peptides in the metaproteome

Organism	# Unique Peptides	Avg. Mass (Da.)	Annotation
PRO1	14	59,606	Oligopeptide ABC transporter
PRO1	12	48,700	Alpha-glucoside ABC transporter
PRO1	12	58,294	ABC transporter
PRO1	11	36,652	Sulfate transporter subunit
Brocadia	10	59,787	Hydroxylamine oxidoreductase
PRO1	10	26,632	Electron transfer flavoprotein subunit beta
PRO1	10	36,575	Amino acid ABC transporter
PRO1	9	33,873	Serine dehydratase
PRO1	9	38,080	Branched chain amino acid ABC transporter
N. nitrosa	9	35,201	Uncharacterized protein
PRO1	9	58,208	Uncharacterized protein
PRO1	9	35,695	Thiamine ABC transporter
PRO1	8	36,846	Serine protease
PRO1	8	40,738	Uncharacterized protein
PRO1	8	32,878	Amino acid ABC transporter

Note: full list of peptides can be found in Supplementary Dataset 3.

5.3 Discussion

Although comammox bacteria have only recently been discovered, understanding their contribution to and potential benefits for wastewater treatment has received considerable interest[16,17,30]. Our results show that microbiomes containing comammox *Nitrospira* and anammox bacteria could result in high nitrogen removal via PNA and that strict control of oxygen uptake by comammox *Nitrospira* is a key parameter controlling cooperation between these functional guilds. The coupling of comammox and anammox to achieve PNA would require even less oxygen than current PNA systems containing AOB, as the nitrate produced by anammox could be used as an electron acceptor by comammox, resulting in nitrite comproportionation. This may encourage the close spatial association between comammox and anammox bacteria in granules or flocs, as

observed by van Kessel et al. (2015)[12], limiting the opportunity for canonical *Nitrospira* or other NOB to compete for nitrite. However, this requires further experimentation.

Recent studies have suggested that comammox bacteria, in particular *N. nitrosa*, may be well adapted to low DO conditions[16,31,32] although other studies have been unable to draw such correlations[17]. Moreover, comammox *Nitrospira* are predicted to be adapted to lower ammonia concentrations based on their recently estimated ammonia affinities[14] and have also been observed to be more abundant in reactors with long solids retention times (SRT)[17,32]. Therefore, in addition to maintaining the oxygen-limited growth conditions predicted from our modeling analysis, we expect that designing systems with a low ammonia-to-microorganism ratio associated with a high SRT will also be important for achieving mainstream PNA with comammox-anammox microbiomes.

Overall, our model simulations highlight key mechanisms that may be involved in promoting cooperation between comammox and anammox bacteria. These models represent novel tools for the hypothesis-driven exploration of interactions between microorganisms in nitrogen cycling communities that may also be used to improve wastewater treatment bioprocess design. Further efforts to improve the models could include experimental validation of the reconstructed metabolic networks using genetic or fluxomic tools, as well as additional parameter estimation for Monod kinetic terms, such as oxygen half-saturation constants. We anticipate that incorporation of genome-scale models for heterotrophic denitrifiers identified here and elsewhere[3,4], as well as previously published models for AOB and NOB[20] (Chapter 4), into a dynamic modeling framework could predict the complete nitrogen-cycling transformations occurring in engineered ecosystems and the environment.

5.4 Materials and Methods

5.4.1 Cultivation of comammox-anammox enrichment co-culture on defined media

The comammox-anammox enrichment culture was inoculated with biomass from the comammox-anammox microbiome originally derived from an aquaculture trickling filter as described in van Kessel et al. (2015)[12]. The enrichment was cultivated using a sequencing batch reactor with the following 12-hour cycle: 11 hours influent filling, 15 minutes settling, 35 minutes effluent withdrawal, 10 minutes idling. The 5 L reactor (working volume) was feed 1.5 L of autoclaved media per day with the following composition (per 10 L): 580 mg NaCl, 2000 mg $MgSO_4 \cdot 7H_2O$, 2940 mg $CaCl_2 \cdot 2H_2O$, 285 mg KH_2PO_4, 500 µg H_3BO_3, 700 µg $ZnCl_2$, 726 µg $Na_2MoO_4 \cdot 2H_2O$, 200 µg $CuCl_2 \cdot 2H_2O$, 240 µg $NiCl_2 \cdot 6H_2O$, 800 µg $CoCl_2 \cdot 6H_2O$, 20 mg $FeSO_4 \cdot 7H_2O$, 4 ml NH_4Cl (1 M), 3.6 ml $NaNO_2$ (1 M), and 2.5 ml $NaNO_3$ (1M). The reactor and media were continuously sparged with 10 ml/min Argon/CO_2 (95/5%) to minimize oxygen inflow and was maintained at room temperature at pH 7 using a $KHCO_3$ buffer.

5.4.2 Biomass composition analysis

Kuenenia stuttgartiensis biomass was centrifuged (10,000 rpm, 15 mins, 4°C) to obtain cell pellets, which were subsequently freeze-dried prior to analysis. Total protein concentration was determined using the Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific) and amino acid composition was determined according to Carnicer et al. (2009)[33] using a Varian 920-LC high performance liquid chromatography amino acid analyzer. Total carbohydrates were determined using the phenol-sulphuric acid method[34]. Total lipid content was determined via the sulfo-phospho-vanillin reaction[35] and lipid composition for *B. sinica* was taken from Rattray et al., (2008)[24]. Total RNA and DNA content was determined according to Benthin et al. (1991)[36]. Total

inorganic content was determined by combustion of freeze-dried biomass in an oven at 550°C for 12 hours.

5.4.3 Genome-scale model reconstruction and analysis

The genome-scale metabolic models of *N. nitrosa* (*iNsa694)* and *B. sinica* (*Bsi788*) were reconstructed from the NCBI's genome sequences for *N. nitrosa* and *B. sinica* (NCBI WGS Project number NZ_CZQA01 and BAFN01, respectively) using the Model SEED pipeline[37] implemented in KBase[19], followed by manual curation using the MetaCyc database[18] and available literature[21,22,38,39] (Chapters 3 & 4). The model was gap-filled manually through the addition of reactions not annotated in the genome to ensure that all biomass components could be produced on minimal media with supplementary ammonia and nitrite. Growth and non-growth associate maintenance energy requirements for *B. sinica* were estimated by fitting these parameters to the measured maximum specific growth rate and specific ammonia uptake rate[40]. The biomass equation for *N. nitrosa* was assumed to be the same as *N. moscoviensis* (Chapter 4), whereas *B. sinica's* biomass equation was derived from biomass composition measurements of the anammox bacterium *K. stuttgartiensis* (Table 5.1 and Table 5.2), as no enrichment culture of *B. sinica* was available. The models were formulated in Systems Biology Markup Language (SBML) level 3 version 1.0 and are available in the supplementary materials. Parsimonious flux balance analysis (pFBA)[41] was used to simulate *in silico* growth by solving the following linear program:

$$Min \sum_{j \in reactions} \lceil v_j \rceil$$

$$s.t.$$

$$\text{Max } v_{biomass}$$

$$s.t.$$

$$S_{ij} \cdot v_j = 0$$

$$v_{min} \leq v_j \leq v_{max}$$

where $v_{biomass}$ is the flux through the biomass objective function, S is the stoichiometric matrix generated from the reconstruction with rows representing metabolites (i), columns representing reactions (j), and entries representing metabolite stoichiometric coefficients, v is the vector of steady-state reaction fluxes, and v_{min} and v_{max} are the minimum and maximum allowable reaction fluxes.

Dynamic flux balance analysis (dFBA)[42] was used to simulate growth of *N. nitrosa* and *B. sinica* in co-culture, based on a dynamic multi-species metabolic modeling framework[43]. Differential equations for the growth rate (dX/dt) of each organism in the community was given by:

$$\frac{dX_j}{dt} = \mu_j X_j$$

Where X_j and μ_j are the biomass concentration and specific growth rate, respectively, of the j^{th} organism in the community. The consumption or production rate (dS/dt) of each metabolite (i.e. ammonia, nitrite, nitrate, and dinitrogen gas) in the environment was given by:

$$\frac{dS_i}{dt} = \sum_{j=1}^{N} V_i^j X_j$$

Where V_i^j is the specific consumption/production rate of the i^{th} metabolite consumed/produced by the j^{th} organism in the community. μ_j and V_i^j are calculate using FBA based on the linear program presented in Equation 1. A Monod kinetic model was used to calculate the FBA constraint on allowable substrate uptake given by:

$$V_i^J \leq V_{i,max}^j \cdot \frac{[S_i]}{K_s^j + [S_i]}$$

Where $V_{i,max}$ is the maximum specific substrate uptake rate, K_s is the substrate half-saturation coefficient, and $[S_i]$ is the substrate concentration. FBA Flux balance analysis was performed in Python version 3.7.2 using the COBRApy package[44]. All code used for dFBA simulations can be found in the supplementary materials.

5.4.4 Metagenomic analysis

Metagenomic DNA used for analysis was extracted and sequenced as described in van Kessel et al. (2015). Paired-end Illumina reads were quality filtered using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and assembled using metaSpades version 3.10.1[45]. Reads were mapped to resulting contigs using bowtie2[46] to determine coverage. Automated binning with Metabat[47], Concoct[48], MaxBin2[49], BinSanity[50], and COCACOLA[51] using differential coverage was then performed and the best bins were selected using DAStool[52]. CheckM version 1.0.8 was used to determine bin completeness, redundancy, and taxonomy[53].

Metabolic reconstruction of the analyzed bins was performed using MetaPathways v2.5[54]. Briefly, ORFs were predicted using Prodigal v2.0[55], based on a minimum nucleotide length of 60, and queried against the SEED subsystems (accessed March 2013), Clusters of Orthologous Groups

(COG, accessed December 2013), RefSeq (accessed January 2017) and MetaCyc (accessed October 2011) protein databases using the optimized LAST algorithm (E value, 1E-6) for functional annotation[56]. Contig nucleotide sequences were also queried against the SILVA small subunit (SSU) 123 database to identify the taxonomy of recovered 16S rRNA genes.

5.4.5 Metaproteomic analysis

Proteins were extracted from bioreactor cell pellets by bead beating using glass beads (acid, washed, 0.1 mm diameter) in a suspension containing B-PER reagent (Thermo Scientific, Germany) and TEAB buffer (50 mM TEAB, 1% (w/w) NaDOC at pH 8). Following DTT reduction and alkylation using iodo acetamide (IAA) protein extracts were subject to proteolytic digestion using trypsin. Resulting peptides were purified by solid phase extraction using an Oasis HLB 96 well plate (Waters, UK), according to the manufacturer protocols. The purified peptide fraction was analysed via a one-dimensional reverse phase separation (Acclaim PepMap RSLC RP C18, 50 µm x 150 mm, 2µm, 100A) coupled to a Q-Exactive plus Orbitrap mass spectrometer (Thermo Scientific, Germany) operating in data dependent acquisition mode (DDA, shot-gun proteomics). The flow rate was maintained at 300 nL/min over a linear gradient from 5% to 30% over 90 minutes and finally to 75% solvent B over 25 minutes. Solvent A was H_2O containing 0.1% formic acid, and solvent B consisted of 80% acetonitrile in H_2O and 0.1% formic acid. The Orbitrap was operated in DDA mode acquiring peptide signals form 350-1400 m/z, where the top 10 signals (with a charge between 2-7) were isolated at a window of 2.0 m/z and fragmented using a NCE of 30. The AGC target was set to 1e5, at a max IT of 54 ms and 17.5 K resolution. Protein identification was determined from Tandem-MS data using PEAKS Studio X (BSI, Canada). All peptide spectra were matched against a protein database generated from predicted open reading frames from the metagenome assembly and all available draft genomes of anammox bacteria.

5.4.6 Supplementary Materials

Supplementary Dataset 1. Model reactions, compounds, and biomass equations.

Supplementary Dataset 2. ORF annotations for MAGs analyzed in this study.

Supplementary Dataset 3. Peptides identified in the metaproteomic dataset from comammox-anammox microbiome.

Supplementary materials can be found online at:

https://github.com/celawson87/phdthesis/tree/master/chapter5

5.5 Acknowledgements

The authors would like to acknowledge Coty Weathersby for assistance with biomass composition analysis, Joshua Hamilton for helpful feedback on genome-scale model reconstruction, and Jeroen Frank for assistance with metagenomic analysis. Funding was provided by the National Science Foundation (CBET-1435661, CBET-1803055 and MCB-1518130), the Netherlands Organization for Scientific Research (Grants 016.Vidi.189.050 and SIAM Gravitation Grant 024.002.002), the European Research Council (ERC Advanced Grant Ecomom 339880), a Wisconsin Distinguished Graduate Fellowship, a Postgraduate Scholarship-Doctoral (PGS-D) by the National Sciences and Engineering Research Council of Canada (NSERC), and the UW-Madison Office of the Vice Chancellor for Research and Graduate Education through the Microbiome Initiative.

5.6 References

1. Sliekers, A. O. *et al.* Completely autotrophic nitrogen removal over nitrite in one single reactor. *Water Res.* **36**, 2475–2482 (2002).
2. Van Dongen, U., Jetten, M. S. M. & Van Loosdrecht, M. C. M. The SHARON®-

Anammox® process for treatment of ammonium rich wastewater. *Water Sci. Technol.* **44**, 153–160 (2001).

3. Speth, D. R., in 't Zandt, M. H., Guerrero-Cruz, S., Dutilh, B. E. & Jetten, M. S. M. Genome-based microbial ecology of anammox granules in a full-scale wastewater treatment system. *Nat. Commun.* **7**, 11172 (2016).
4. Lawson, C. E. *et al.* Metabolic network analysis reveals microbial community interactions in anammox granules. *Nat. Commun.* **8**, 1–12 (2017).
5. Lackner, S. *et al.* Full-scale partial nitritation/anammox experiences – An application survey. *Water Res.* **55**, 292–303 (2014).
6. Laureni, M. *et al.* Biomass segregation between bio fi lm and fl ocs improves the control of nitrite-oxidizing bacteria in mainstream partial nitritation and anammox processes. *Water Res.* **154**, 104–116 (2019).
7. Cao, Y., van Loosdrecht, M. C. M. & Daigger, G. T. Mainstream partial nitritation–anammox in municipal wastewater treatment: status, bottlenecks, and further studies. *Appl. Microbiol. Biotechnol.* **101**, 1365–1383 (2017).
8. Vlaeminck, S. E. *et al.* Aggregate size and architecture determine microbial activity balance for one-stage partial nitritation and anammox. *Appl. Environ. Microbiol.* **76**, 900–909 (2010).
9. Winkler, M. K. H., Kleerebezem, R., Kuenen, J. G., Yang, J. & Loosdrecht, M. C. M. Van. Segregation of Biomass in Cyclic Anaerobic / Aerobic Granular Sludge Allows the Enrichment of Anaerobic Ammonium Oxidizing Bacteria at Low Temperatures. *Environ. Sci. Technol.* **45**, 7330–7337 (2011).
10. Ma, B. *et al.* Suppressing Nitrite-oxidizing Bacteria Growth to Achieve Nitrogen Removal from Domestic Wastewater via Anammox Using Intermittent Aeration with Low Dissolved Oxygen. *Sci. Rep.* **5**, 13048 (2015).
11. Daims, H. *et al.* Complete nitrification by Nitrospira bacteria. *Nature* **528**, 504–509 (2015).
12. van Kessel, M. A. H. J. *et al.* Complete nitrification by a single microorganism. *Nature* **528**, 555–559 (2015).
13. Costa, E., Pérez, J. & Kreft, J. U. Why is metabolic labour divided in nitrification? *Trends Microbiol.* **14**, 213–219 (2006).

14. Kits, K. D. *et al.* Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature* **549**, 269–272 (2017).
15. Chao, Y., Mao, Y., Yu, K. & Zhang, T. Novel nitrifiers and comammox in a full-scale hybrid biofilm and activated sludge reactor revealed by metagenomic approach. *Appl. Microbiol. Biotechnol.* **100**, 8239 (2016).
16. Roots, P. *et al.* Comammox Nitrospira are the dominant ammonia oxidizers in a mainstream low dissolved oxygen nitrification reactor. *Water Res.* **157**, 396–405 (2019).
17. Cotto, I. *et al.* Long solids retention times and attached growth phase favor prevalence of comammox bacteria in nitrogen removal systems. *Water Res.* **169**, 115268 (2020).
18. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res.* **46**, D633–D639 (2017).
19. Arkin, A. P. *et al.* KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat. Biotechnol.* **36**, 566–569 (2018).
20. Mellbye, B. L. *et al.* Genome-Scale, Constraint-Based Modeling of Nitrogen Oxide Fluxes during Coculture of Nitrosomonas europaea and Nitrobacter winogradskyi. *mSystems* **3**, e00170-17 (2018).
21. Kartal, B. *et al.* Molecular mechanism of anaerobic ammonium oxidation. *Nature* **479**, 127–30 (2011).
22. Almeida, N. M. De *et al.* Membrane-bound electron transport systems of an anammox bacterium : A complexome analysis. *BBA - Bioenerg.* **1857**, 1694–1704 (2016).
23. Thiele, I. & Palsson, B. Ø. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **5**, 93–121 (2010).
24. Rattray, J. E. *et al.* Ladderane lipid distribution in four genera of anammox bacteria. *Arch. Microbiol.* **190**, 51–66 (2008).
25. Strous, M., Heijnen, J. J., Kuenen, J. G. & Jetten, M. S. M. The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms. *Appl. Microbiol. Biotechnol.* **50**, 589–596 (1998).
26. Lotti, T., van der Star, W. R. L., Kleerebezem, R., Lubello, C. & van Loosdrecht, M. C. M. The effect of nitrite inhibition on the anammox process. *Water Res.* **46**, 2559–2569 (2012).
27. Bowers, R. M. *et al.* Minimum information about a single amplified genome (MISAG)

and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **35**, 725–731 (2017).

28. Meuer, J., Kuettner, H. C., Zhang, J. K., Hedderich, R. & Metcalf, W. W. Genetic analysis of the archaeon Methanosarcina barkeri Fusaro reveals a central role for Ech hydrogenase and ferredoxin in methanogenesis and carbon fixation. *Proc. Natl. Acad. Sci.* **99**, 5632 LP – 5637 (2002).
29. Welte, C. *et al.* Function of Ech Hydrogenase in Ferredoxin-Dependent, Membrane-Bound Electron Transport in Methanosarcina mazei *J. Bacteriol.* **192**, 674 LP – 678 (2010).
30. Lawson, C. E. & Lücker, S. Complete ammonia oxidation: an important control on nitrification in engineered ecosystems? *Curr. Opin. Biotechnol.* **50**, 158–165 (2018).
31. Camejo, P. Y., Domingo, J. S., Noguera, D. R. & McMahon, K. D. Genome-Enabled Insights into the Ecophysiology of the Comammox Bacterium "Candidatus Nitrospira nitrosa". *mSystems* **2**, e00059-17 (2017).
32. Beach, N. K. & Noguera, D. R. Design and Assessment of Species-Level qPCR Primers Targeting Comammox . *Frontiers in Microbiology* **10**, 36 (2019).
33. Carnicer, M. *et al.* Macromolecular and elemental composition analysis and extracellular metabolite balances of Pichia pastoris growing at different oxygen levels. *Microb. Cell Fact.* **8**, 65 (2009).
34. Herbert, D., Phipps, P. J. & Strange, R. E. Chemical Analysis of Microbial Cells. in *Methods in Microbiology* (eds. Norris, J. R. & Ribbons, D. W. B. T.-M. in M.) **5**, 209–344 (Academic Press, 1971).
35. Izard, J. & Limberger, R. J. Rapid screening method for quantitation of bacterial cell lipids from whole cells. *J. Microbiol. Methods* **55**, 411–418 (2003).
36. Benthin, S., Nielsen, J. & Villadsen, J. A Simple and reliable method for the determination of cellular RNA content. *Biotechnol. Tech.* **5**, 39–42 (1991).
37. Henry, C. S. *et al.* High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* **28**, 977–82 (2010).
38. Lücker, S. *et al.* A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13479–84 (2010).

39. Koch, H. *et al.* Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus Nitrospira. *Proc. Natl. Acad. Sci.* **112**, 201506533 (2015).
40. Oshiki, M., Satoh, H. & Okabe, S. Ecology and physiology of anaerobic ammonium oxidizing bacteria. *Environ. Microbiol.* **18**, 2784–2796 (2016).
41. Lewis, N. E. *et al.* Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. *Mol. Syst. Biol.* **6**, 390 (2010).
42. Mahadevan, R., Edwards, J. S. & Doyle III, F. J. Dynamic Flux Balance Analysis of Diauxic Growth in Escherichia coli. *Biophys. J.* **83**, 1331–1340 (2002).
43. Zhuang, K. *et al.* Genome-scale dynamic modeling of the competition between Rhodoferax and Geobacter in anoxic subsurface environments. *Isme J.* **5**, 305 (2010).
44. Ebrahim, A., Lerman, J. A., Palsson, B. Ø. & Hyduke, D. R. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.* **7**, 74 (2013).
45. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes : a new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).
46. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
47. Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).
48. Alneberg, J. *et al.* Binning metagenomic contigs by coverage and composition. *Nat. Methods* **11**, 1144–1146 (2014).
49. Wu, Y.-W., Tang, Y.-H., Tringe, S. G., Simmons, B. a & Singer, S. W. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome* **2**, 26 (2014).
50. Graham, E. D., Heidelberg, J. F. & Tully, B. J. BinSanity: unsupervised clustering of environmental microbial assemblies using coverage and affinity propagation. *PeerJ* **5**, e3035 (2017).
51. Lu, Y. Y., Chen, T., Fuhrman, J. A. & Sun, F. COCACOLA: binning metagenomic contigs using sequence COmposition, read CoverAge, CO-alignment and paired-end read LinkAge. *Bioinformatics* **33**, 791–798 (2017).
52. Sieber, C. M. K. *et al.* Recovery of genomes from metagenomes via a dereplication,

aggregation and scoring strategy. *Nat. Microbiol.* **3**, 836–843 (2018).

53. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–55 (2015).
54. Konwar, K. M. *et al.* MetaPathways v2.5: Quantitative functional, taxonomic and usability improvements. *Bioinformatics* **31**, 3345–3347 (2015).
55. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
56. Kielbasa, S. M., Wan, R., Sato, K., Horton, P. & Frith, M. C. Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21**, 487–493 (2011).

6. Common principles and best practices for engineering microbiomes

A version of this chapter has been published as:

Lawson, C.E., Harcombe, W.R., Hatzenpichler, R., Lindmann, S.R., Löffler, F., O'Malley, M.A., García Martín, H., Pfleger, B.F., Raskin, L., Venturelli, O.S., Weissbrodt, D.G., Noguera, D.R., McMahon, K.D. Common principles and best practices for engineering microbiomes. *Nat. Rev. Microbiol.* **17**, 725–741 (2019).

All supplementary data files are available online with the published article.

Terms in bold are defined in the glossary.

Author contributions

C.E.L. wrote the manuscript with direct input, edits, and critical feedback by all authors.

6.0 Abstract

Despite broad scientific interest in harnessing the power of Earth's microbiomes, knowledge gaps hinder their efficient use for addressing urgent societal and environmental challenges. We argue that structuring research and technology developments around a design-build-test-learn (DBTL) cycle will advance microbiome engineering and spur new discoveries on the basic scientific principles governing microbiome function. In this Review, we present key elements of an iterative DBTL cycle for microbiome engineering, focusing on generalizable approaches, including top-down and bottom-up design processes, synthetic and self-assembled construction methods, and emerging tools to analyze microbiome function. These approaches can be used to harness microbiomes for broad applications related to medicine, agriculture, energy, and the environment. We also discuss key challenges and opportunities of each approach and synthesize them into best practice guidelines for engineering microbiomes. We anticipate that adoption of a DBTL framework will rapidly advance microbiome-based biotechnologies aimed at improving human and animal health, agriculture, and enabling the bioeconomy.

6.1 Introduction

Microbial communities have seemingly limitless capabilities, driving Earth's biogeochemical cycles and occupying every environmental niche[1,2]. Engineers and scientists have tapped into this power for a long time; for example, by manipulating soil microbiomes to increase crop productivity[3], by stimulating naturally-occurring or introduced microbiomes to remediate contaminated groundwater[4], or by building reactor microbiomes to recover valuable resources from wastewater[5]. Although these accomplishments highlight the valuable functions of microbiomes, the vast majority of the microbial world's transformative capabilities have yet to be

unlocked and harnessed. Recent insights driven by DNA sequencing have shed light on the high genetic diversity of not-yet-cultured microorganisms and their crucial roles in diverse ecosystems[6,7], providing a window on potentially novel biotechnology applications.

In recognition of this unlocked potential, funding agencies and the international science community have called for a global effort to advance microbiome research[8,9]. These initiatives have recognized the need for **microbiome science** to move beyond descriptive studies, and embrace a systems approach that generates the mechanistic, predictive, and actionable understanding that enables rational **microbiome engineering**[8]. However, achieving this transition is hindered by the lack of tractable experimental systems that permit the detailed functional investigation of microbiomes, the large pool of microbiome gene and metabolite functions that remain unknown[10], the many uncharacterized interactions (for example, **syntrophy**) between microorganisms[11], inadequate tools to accurately measure and simulate microbiome functions across time and space, and the limited availability of approaches to precisely manipulate microbiome structure and function.

Integrating basic scientific discovery with engineering can overcome these challenges and develop innovative solutions that support sustainable natural resources management and human and animal health. In particular, engineering approaches can be used to create experimental systems that permit the testing of conceptual knowledge and extraction of new knowledge that advances microbiome research. To accelerate both scientific discovery and translation into innovative solutions, we propose that microbiome engineering adopt an iterative design-build-test-learn (DBTL) cycle to structure research and the technology development process. This cycle involves developing an initial microbiome design or preliminary model system to achieve a defined engineering goal, building the microbiome, testing its function against a set of specified metrics to

determine whether the design-build solution(s) produced the design objective (i.e. establish causation), learning what worked, what did not (and why), and incorporating new knowledge into the decision making process of subsequent DBTL cycles (Figure 6.1). This approach has been used successfully in manufacturing[12], metabolic engineering[13], and entrepreneurship ('build, measure, learn')[14], and could rapidly advance our ability to develop much needed tools and design concepts for harnessing microbiomes, delivering innovative solutions and advancing scientific knowledge.

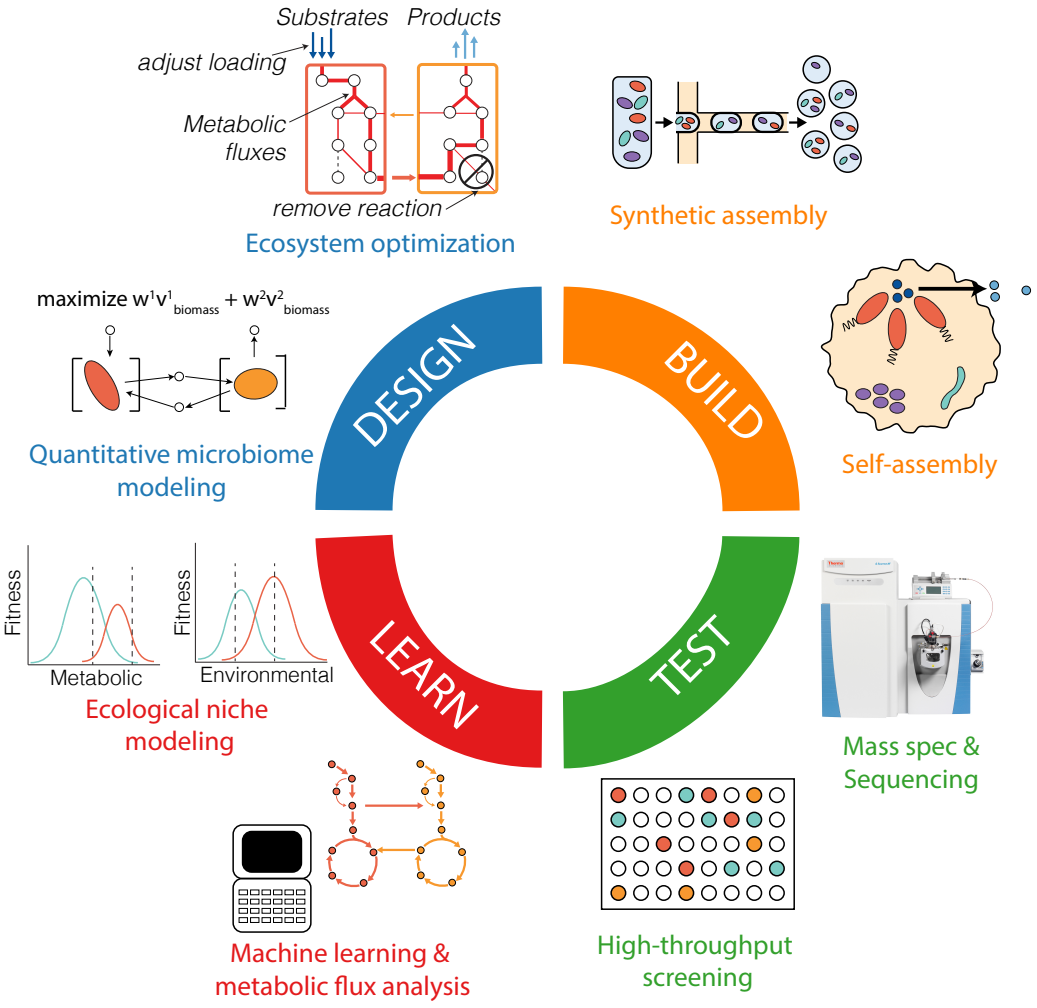

Figure 6.1 - The design-build-test-learn cycle for microbiome engineering. The figure presents key aspects and approaches of each phase of the design-build-test-learn (DBTL) cycle. The cycle starts with a defined engineering objective that determines the design and produces an engineered microbiome that performs the desired function(s).

In this Review, we present key elements of an iterative DBTL approach that can be implemented to advance the rational engineering of microbiomes for functions that benefit society. We review diverse approaches to harness microbiomes in medical, agricultural, energy, and environmental applications, and identify current challenges and opportunities associated with implementing each DBTL phase. Finally, we discuss how the DBTL cycle can be applied to build model systems to establish basic principles of microbial ecosystems and provide an outlook on the frontiers of microbiome engineering.

6.2 Designing microbiomes

Because of the high complexity and limited understanding of molecular-scale microbiome processes, microbiome design has conventionally followed a top-down approach. This approach tries to predict how ecosystem-level controls can create a microbiome with desired functions. However, recent advances in multi-omics have provided opportunities to design microbiomes from the bottom-up by predicting how the control of metabolic networks and their interactions can create a microbiome with desired functions. Combined, these approaches offer complementary strategies to design microbiomes for specific engineering goals, ranging from sustainable wastewater treatment to curing microbiome-associated human diseases.

6.2.1 Top-down design

Rather than deciding which organisms and detailed metabolic pathways to use *a priori*, the top-down approach uses carefully selected environmental variables (such as certain substrate loading rates, mean-cell retention times, and redox conditions) that force an existing microbiome (naturally occurring or inoculated) through ecological selection to perform the desired biological processes (or '**metaphenotypes**'[15]) (Figure 6.2). Here, 'top' refers to the ecosystem in which the desired

biological process occurs and top-down design denotes the methods used to predict how manipulation of the ecosystem's physical, chemical, and biological processes (that is, ecosystem processes) obtains the desired function. Predicting how to manipulate an ecosystem is informed by principles of **ecological engineering**[16] (also known as microbial resource management[17] or microbial community engineering[18]). This requires engineers to conceptualize the system as an ecosystem model that captures system inputs and outputs, physicochemical conditions (pH, temperature, redox potential, etc.), known abiotic and biotic processes, and environmental variables, and how their manipulation may promote or inhibit the biological process(es) being optimized[19,20]. Subsequently, mathematical modeling is used to perform mass balance analysis around chemicals and relevant microorganisms in the system and simulate chemical and biochemical transformation rates. These process-based models capture microbiome functions by representing key physiological or **functional guilds** of microorganisms (such as methanogens, fermenters, nitrifiers, or phototrophs) with specific stoichiometric (growth and product yields) and kinetic parameters (maximum specific growth rate, substrate uptake rate, and substrate affinity)[21,22,23]. The models can also integrate equations describing the three-dimensional physical transport processes (diffusion, advection, and dispersion) acting on chemicals and microorganisms, which are especially important in spatially structured systems such as biofilms[24,25].

6.2.2 Bottom-up design.

Although the conventional top-down design approach for microbiome engineering offers a framework for macro-scale processes and has been widely successful for wastewater treatment[21] and bioremediation[4], it often neglects the complex *in situ* metabolic networks driving microbial

and linked chemical transformations[26] and ignores processes that depend on intricate interactions between community members; for example, syntrophic interactions through direct interspecies electron transfer[27]. As a consequence, molecular-scale microbiome processes are often ignored during design, limiting system optimization through molecular-scale mechanistic insight. Recent advances in multi-omics and automation technology (for example, in metagenomics and microfluidics) have enabled researchers to develop bottom-up approaches and focus on engineering the microbiome's metabolic network and microbial interactions. Here, 'bottom' refers to the metabolic networks of individual organisms in the microbiome (expressed from their genomes) and 'bottom-up design' denotes the methods used to predict how metabolic flux through these interacting networks obtains the desired function. The general design process is to obtain the genomes of individual members of the microbiome[28] (especially **keystone species**[29], when known[30]), reconstruct their metabolic networks,[31,32] and use modeling[33] and/or network analysis tools[34] to guide design (Figure 6.2). Existing constraint-based methods such as **flux balance analysis (FBA)** provide a suitable framework for exploring which combinations of chemical transformations are possible using quantitative models, in which individual populations' reactions and metabolites can be compartmentalized and metabolic fluxes within and between populations can be simulated using optimality principles[35]. These models can also simulate steady-state flux distributions over time and space[36,37] and can be integrated into process-based and/or individual-based models[38] to predict metaphenotypes, self-organizing spatial patterns, and other emergent behaviours. Such bottom-up tools provide the engineer with a computational framework to systematically evaluate the metabolic networks driving biological processes and ecological interactions, and a platform for rationally designing microbiomes with specific properties, such as distributed pathways[39,40], modular species interactions[41], community resistance and resilience[42]

and spatiotemporal organization[43] that optimize ecosystem function and stability. However, the majority of these bottom-up design examples are based on simple communities with model organisms (such as *Escherichia coli* and *Saccharomyces cerevisiae*) that have engineered dependencies. Therefore, extending these designs to systems with non-model organisms of tens to hundreds of different species will require deeper insights into their metabolism and the principles governing their interactions and higher-order behavior.

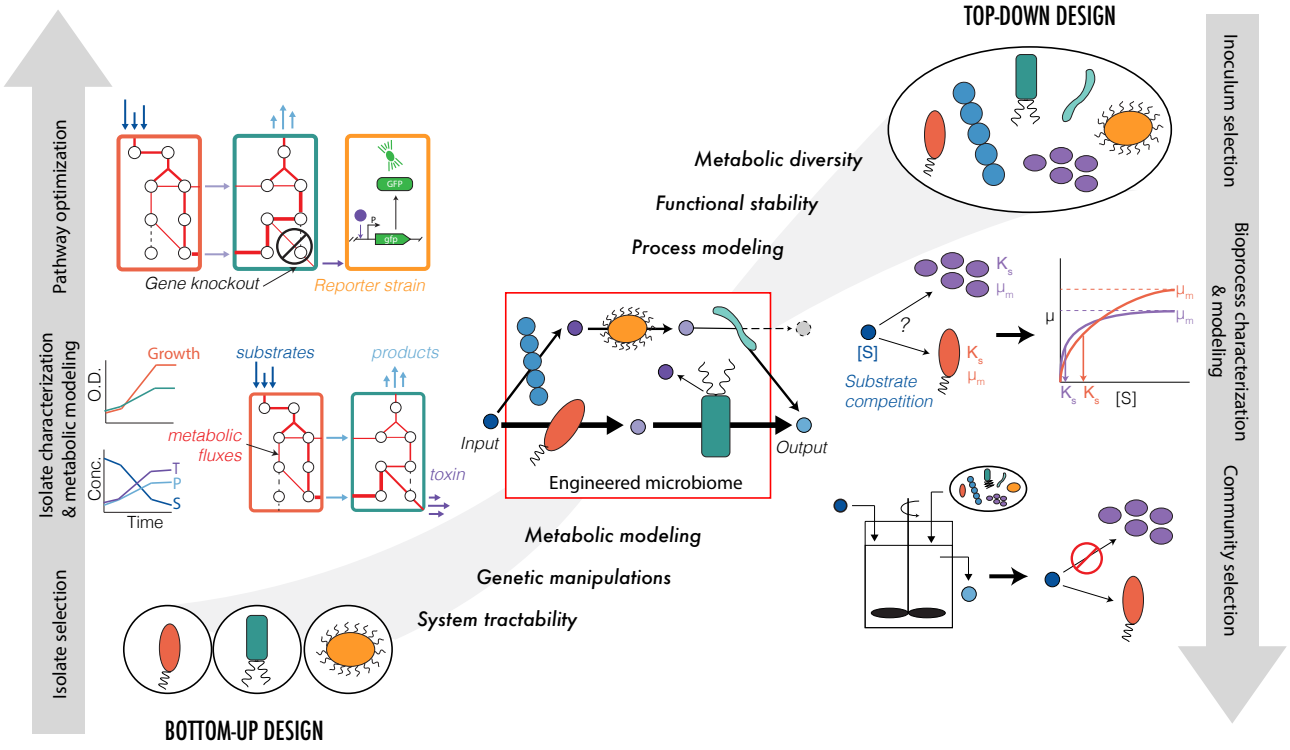

Figure 6.2 - Top-down and bottom-up approaches to design microbiomes. The left panel illustrates a bottom-up design workflow starting from pure isolates. Physiological characterization of individual organisms is performed, and metabolic modeling is used to design consortia for desired function (produce light blue compound from dark blue compound). Genetic engineering and synthetic biology strategies are used to optimize system function (identifying gene editing targets that re-route metabolic flux away from toxin (purple) and towards desired product; designing of toxin reporter strain). The right panel illustrates a top-down design starting with an inoculum containing uncultivated microorganisms from the environment. Community characterization of mixed microbiome is performed, and bioprocess modeling (mass balance analysis including kinetics and microbial growth) is used to develop selection strategies to achieve desired function (produce light blue compound from dark blue compound). Reactor engineering design is used to

optimize system function. The middle panel shows an integrated top-down bottom-up design. Combinations of uncultivated consortia and defined cultures are selected to achieve desired functions. Community characterization is performed and microbiome modeling that integrates process-based simulation with metabolic modeling is used to develop selection strategies and analyze microbiome metabolic fluxes. The shapes of the microorganisms represent different isolates or communities selected during design.

There are major challenges to implementing this bottom-up approach, including inaccurate and/or incomplete metabolic network reconstructions, unknown functions of many genes, proteins, and metabolites, poorly understood evolutionary pressures driving individual and community-level phenotypes, and limited understanding of gene, metabolic, and ecosystem regulatory schemes (for example, quorum sensing signal-response systems[44]). These limitations lead to high model uncertainty because key constraints on pathway stoichiometry and enzyme kinetics are either inappropriate or missing, and objective functions fail to capture the true evolutionary drivers of cell behavior[45], ultimately leading to poor predictions of *in situ* phenotypes.

As a starting point for bottom-up design, core metabolic models that capture central carbon and energy metabolism can be reconstructed from genome annotations and known physiological information. The predictive power of these models may be limited initially, as they ignore regulatory information, pathway kinetics, secondary metabolism, and evolution. However, when this knowledge is acquired and becomes incorporated into metabolic models through multiple cycles of testing and learning, accurate predictions of system function (for example, metabolic fluxes and metabolite exchange) may emerge. As a complementary approach, data-driven modeling techniques such as **ensemble modeling** and **machine learning** may offer more rapid methods to predict microbiome metabolic processes or obtain constraints and parameters required for microbiome modeling, without the need for detailed mechanistic understanding of metabolic regulation[46,47]. Such modeling frameworks have been used to predict pathway fluxes from

proteomic and metabolomic data[48], improve metabolite cross-feeding predictions through ensemble modeling-based FBA[49], and to obtain key catalytic turnover numbers needed for metabolic models[50]. Although these approaches are flexible and generalizable enough to be applied to microbial communities, they require substantial amounts of experimental data on the metabolism of individual strains and interacting communities. This information could be leveraged from prior test phases (for example, from high-throughput phenotypic screens and multi-omics) to enable data-driven design.

6.2.3 Integrated design

Moving forward, we envision that a judiciously balanced blend of top-down and bottom-up approaches will be needed for successful microbiome design, especially when working with complex microbiomes, such as human microbiota or activated sludge (Figure 6.2). A blended approach could involve selecting both undefined mixtures and defined consortia to achieve desired microbiome functions, merging process-based models with bottom-up metabolic models reconstructed from meta-omic information to simulate ecosystem processes, mass balances, and metabolite fluxes, and using genome-derived information to develop community selection strategies. Capturing higher-order properties in design, such as functional stability and dynamics, will likely also require top-down and bottom-up approaches to converge. In particular, new mathematical modeling approaches that quantify mechanisms of functional degeneracy, niche complementarity, and network buffering[51] using a metabolic framework may enable microbiome diversity to be optimized to sustain desired functions *in situ*. The need for a more comprehensive representation of microbiome metabolism will depend on the specific engineering objective and the degree of ecosystem tractability. For example, a more detailed representation of anaerobic microbiome metabolism is likely required for converting biomass into a specific commodity

chemical instead of methane because finer control over metabolism would be needed. In either case, the design phase encompasses defining the engineering problem, developing conceptual and quantitative models, identifying key biological processes to be manipulated, and evaluating multiple candidate design alternatives.

6.2.4 Practical design steps

There are five key steps when designing microbiomes, in particular complex microbiomes: defining the engineering problem, developing a conceptual ecosystem model, creating a quantitative model, identifying the microbiome process(es) to be engineered, and developing and evaluating candidate design strategies.

1. To drive the DBTL cycle, a clear ***definition of the problem*** with measurable design objectives must be established. These objectives could specify desired outcomes such as product titers, rates and yields, pollutant removal efficiency, crop productivity, or degree of functional stability and robustness. Design objectives should be complemented by **techno-economic assessments** and/or **life cycle analysis** to ensure that solutions are economically feasible and have positive environmental and societal impacts[52,53].
2. ***Conceptual ecosystem models*** can be used to contextualize the problem. Such models capture system boundaries, inputs and outputs, major pathways of carbon and nutrient flows, key organisms and interspecies interactions responsible for those transformations, and factors influencing their activity (for example, pH, temperature, redox potential, and residence times)[19]. They provide a concept map that describes current understanding of interactions between the microbiome and physical, chemical, and biological components of the ecosystem, helping to identify important gaps in system understanding and needs for data collection. At this stage, all relevant information should be collected from the literature, existing data (for

example, from the Human Microbiome Project[54]), and online databases (for example, MiDAS (microbial database for activated sludge)[55]) for ecosystem characterization. This includes reference genomes and physiological information for keystone organisms, previous multi-omic datasets, ecosystem physicochemical properties (such as pH, temperature and chemical concentrations) and processes (such as photochemical reactions and hydrogeological processes), site characteristics (such as nutrient loadings and dynamics, soil profiles and gut anatomy), and all other information needed to characterize the ecosystem. Missing information, such as unknown biochemical pathways and organisms that mediate them, can be targeted during the build-test-learn phases. This conceptual ecosystem model can be used by the scientific community for proposing and testing theories and serves as a roadmap for developing quantitative simulation tools.

3. ***Construction of quantitative modeling tools*** that enable the calculation and simulation of metabolic fluxes, microorganism abundances, mass balances, and ecosystem physicochemical parameters is critical for the systematic design of microbiomes. Several approaches could be used to create such models, including mechanistic metabolic modeling[33,] process-based modeling[21], data-driven modeling (for example, machine learning)[48], individual-based modeling[38] or their combination. Regardless of the approach, the simulation of complex microbiomes will likely require simplification based on experimentally valid assumptions. Simplification could include reducing the model to a set of core or keystone organisms that represent important functional guilds and control major carbon and energy flows, or reducing the metabolic network size of organisms to central carbon and energy metabolism. Moving forward, it will be important to ensure that models undergo rigorous experimental validation and iteration during build-test-learn cycles to increase their utility and widespread use in

microbiome engineering and to identify when modeling efforts fail, revealing gaps in conceptual understanding that can further facilitate model redesign and improvement.

4. ***Quantitative microbiome modeling*** (such as dynamic FBA) helps to identify the core and peripheral biochemical pathways that need to be directly manipulated, added, or removed to achieve the desired engineering objective. Objectives could include increasing butyrate production and non-digestible carbohydrate degradation by fermenting bacteria in the human gut, preventing toxin biosynthesis by cyanobacteria in freshwater ecosystems, or stimulating the degradation of toxic chloroorganics by bioaugmentation with organohalide-respiring bacteria.
5. ***Microbiome modeling*** can predict how environmental (such as substrate loading, pH, and solids retention time) or genetic manipulation (such as gene knockouts, pathway additions, and forced dependencies) could optimize microbiome functions towards the engineering objective. If necessary, synthetic microorganisms could be designed to improve microbiome function. Such synthetic microorganisms will need to be evaluated for their ability to cooperate and compete with existing microbiome members under relevant environmental conditions.

6.3 Building microbiomes

The build phase consists of physically assembling the designed microbiome by either top-down manipulation of a natural community (that is, a **self-assembled microbiome**) or bottom-up assembly using axenic or enrichment cultures of naturally-occurring or engineered microorganisms (that is, a **synthetic microbiome**). The build phase aims to bring the design specifications and predictions into reality.

6.3.1 Building by self-assembly

Self-assembled microbiomes may include those built as open mixed cultures using reactor engineering (for example, wastewater treatment bioreactor) or biostimulation (for example, additions to soils, sediments or groundwater aquifers), in which construction creates an environment that promotes the growth and desirable activity of resident microorganisms. Examples include manipulating reactor hydrodynamics to immobilize slow-growing microorganisms into compact granules that enable their retention and proliferation[56,57], use of non-human-digestible carbohydrates to stimulate fermentative production of short-chain fatty acids in the gut[58], or adding electron donors to drive the metabolism of organohalide-respiring bacteria during bioremediation of toxic chlorinated contaminants[4]. This approach is powerful when differences in physiological and physicochemical properties between functional guilds can be exploited for assembly through environmental manipulation (for example, differences in growth rates[59], main electron donors and acceptors[4,60], substrate affinities, cell and/or biofilm densities[61], and redox gradients). However, it can be limited when more fine-scale control over microbial metabolism and interactions is necessary (for example, controlling complex competitive interactions[62], producing valuable bioproducts at high yields and purity[63], or controlling organisms with versatile lifestyles[64]).

In addition, new strategies for evolutionary engineering have emerged as promising tools to build self-assembled microbiomes. Controlled exposure of an initial microbiome to multiple selection cycles and/or regimes results in the microbiome gaining or optimizing specific functions through adaptation or evolution. For example, successively transferring the microbiomes that maximize plant traits has generated microbiomes that improve plant biomass[65] and flowering time[66]. Response to community-level selection will often be driven by enrichment or adaptation of

single species[67,68]; however, selection for production of community biomass has also been shown to enhance desired species interactions in defined two and three species co-cultures[37,69]. Re-examining selection experiments to understand when and how mutations and/or adaptations altered microbiome phenotypes could elucidate the mechanisms underlying microbiome fitness optimization and inform design, as has been shown for *E. coli* in laboratory evolution experiments[70,71]. As similar evolutionary approaches (for example, adaptive laboratory evolution) have also been successfully applied to optimize strains for metabolic engineering[72], extension of experimental and computational protocols already developed for individual microorganisms to microbiomes could streamline the design phase and reduce the time required to complete evolution experiments.

6.3.2 Building synthetic microbiomes

Direct construction of microbiomes using axenic or enrichment cultures is also promising because of reduced complexity and the use of microorganisms that are genetically tractable and/or well-characterized. This bottom-up approach makes the growing suite of synthetic biology tools accessible for microbiome construction and optimization. An early approach for building microbiomes directly from cultured microorganisms is bioaugmentation. Here, defined laboratory consortia are added back to the environment to enhance the degradation rates of specific contaminants. A successful example has been the addition of consortia containing organohalide-respiring bacteria of the class *Dehalococcoidia* to contaminated groundwater aquifers and sediments to speed up the degradation of toxic chlorinated solvents. Crucial for the success of this approach was detailed knowledge of the physiology, nutritional requirements, and potential ecological interactions of the keystone dechlorinators with other microorganisms and the geochemical environment[4]. However, contrary to the success for chlorinated contaminants,

bioaugmentation approaches have largely failed for oil spills. Unlike organohalide-respiring *Dehalococcoidia* members that fill a unique ecological niche and cannot grow without the chlorinated contaminants, organisms capable of degrading oil hydrocarbons (especially aerobic bacteria) are ubiquitous, metabolically versatile, and do not depend on a specific substrate or redox couple for growth[64]. This metabolic versatility limits their utility for bioaugmentation given their unpredictable *in situ* activity. Other reasons why bioaugmentation can fail are that unrecognized mutualistic interactions and microorganisms performing critical functions are missing (for example, production of polysaccharide surfactants to increase hydrocarbon bioavailability[73]), or that consortia selected under laboratory conditions are no longer competitive enough under harsh and/or variable field conditions[74,75,76]. These examples highlight the need to better understand the interaction networks of synthetic consortia, especially the roles of supporting interactions (secondary functions), and the competitive landscape *in situ*, which are often difficult to predict in complex ecosystems.

Despite the appeal of building microbiomes bottom-up and the growing collection of cultured microorganisms from specific habitats[77,78], the majority of microorganisms relevant for human health, agriculture, and environmental applications remain uncultured, poorly characterized, genetically intractable, and difficult to maintain, making the construction of synthetic microbiomes challenging. To capture this uncharacterized metabolic diversity, innovative isolation and controlled microbiome assembly techniques are needed, such as single-cell sorting[79] coupled to high-throughput culturing (culturomics)[80,81] and phenotyping[82,83] across multiple conditions in parallel. Microfluidics[84,85], that is, creation and manipulation of microliter droplets, can facilitate this approach. Microfluidic chips can enable automated assembly and analysis of microbial communities from axenic or enrichment cultures through droplet

combination[86], elimination of specific species[87], sequencing, and multi-omics phenotyping of individual cells[88,89]. Combined with new gene editing techniques, such as CRISPR-based genomic tools[90] that improve the efficiency of homologous recombination-based gene editing[91,92], microfluidics could also automate synthetic biology techniques for the engineering of cells and microbiomes with novel capabilities[93].

Another challenge with synthetic microbiomes is maintaining their functional stability in the laboratory or in open systems (for example, human gut, soil, and wastewater treatment plants), which are susceptible to invasion by naturally-occurring microorganisms and dynamic heterogeneous environments. As mentioned above, the major reason for the success of bioaugmentation with organohalide-respiring *Dehalococcoidia* members is their highly specialized lifestyle that enables them to occupy an open ecological niche using chlorinated electron acceptors. However, the functional stability of organisms with versatile lifestyles in open systems is much less predictable. Few studies have examined the functional stability of synthetic consortia in open systems and the knowledge required to rationally engineer stable ecological interactions is limited. However, engineered bacteria have been successfully deployed as diagnostic sensors in the mammalian gut for up to 200 days maintaining robust function[94,95]. This feat, together with the bioaugmentation example of *Dehalococcoidia*[4], demonstrates that synthetic consortia can form stable microbiomes with previously established community members, provided key players can compete with resident microorganisms.

Observations from self-assembled microbiomes suggest that building communities with spatiotemporal organization will be important for achieving stable and multi-functional synthetic microbiomes. Highly diverse microbial communities, such as human microbiota or those used for wastewater treatment, self-assemble as biofilms, flocs, or granules comprised of multiple single-

species microcolonies attached together via species-specific extracellular polymeric substances (including polysaccharides, proteins, and DNA) and other poorly defined macromolecules (such as humics)[96,97]. These self-organizing microbial assemblages create diverse microenvironments and ecological niches that support the combination of seemingly incompatible functions (for example, both aerobic and anaerobic processes[98,99]) and functionally diverse population structures that can compensate for disturbances, such as changes in nutrients, physicochemical condition, or predation[100,101]. Although building such fine-scale and sophisticated architectures into synthetic microbiomes is nascent, microfluidic-based systems have been used to assemble simple communities with improved functional stability by controlling spatial structure and chemical communication[102]. Additionally, 3D bioprinting platforms could enable the construction of spatially organized systems, in which populations can be physically separated while remaining chemically interactive[103,104]. How to scale these spatially defined structures from experimental laboratory systems to real-world applications remains to be resolved, although knowledge gained from test and learn phases with model systems (such as synthetic polysaccharide particles [105,106]) should provide more insights. Until then, existing approaches based on top-down assembly and/or engineered biofilm carrier media[107] could be used to build self-organized synthetic microbiomes with better stability and functionality.

Designing synthetic genetic circuits in engineered hosts that can robustly perform sense-compute-respond programs in complex environments also remains a major challenge[108]. Therefore, it will be important to examine the molecular mechanisms that determine microbiome stability and adaptation to environmental perturbation in natural and engineered ecosystems, in order to extract design principles that can be used for rationally engineering robust functions. Given the potential utility of genetically engineered microorganisms and microbiomes in diverse

open environments, safeguards such as biocontainment systems (such as two-layered gene circuits and essential synthetic auxotrophies[109]) will also require further development and will be needed as integral components of constructed synthetic microbiomes that use genetically modified organisms in the future.

6.3.3 Integrating approaches

The ultimate goal for rational microbiome design is to develop tools that enable engineers to directly add, remove, or modify specific functions and phenotypes *in situ* over a range of desirable operational conditions. One emerging technique with promise to achieve such flexibility is *in situ* metagenomic engineering[110,111], which involves delivery of engineered mobile genetic elements to resident microorganisms. For example, donor strains engineered with **integrative and conjugative elements** have transferred DNA carrying a reporter and antibiotic resistance genes or multi-gene pathways (for example, nitrogen fixation (*nif*) gene cluster[112]) to bacteria in highly heterogeneous and diverse environments, such as soil[112] and the mammalian gut[111]. Further development of such tools in combination with existing CRISPR-Cas gene editing techniques would enable the precise manipulation of the microbiome's metabolic network *in situ*, effectively combining self-assembled and synthetic microbiomes (Figure 6.3; Box 1).

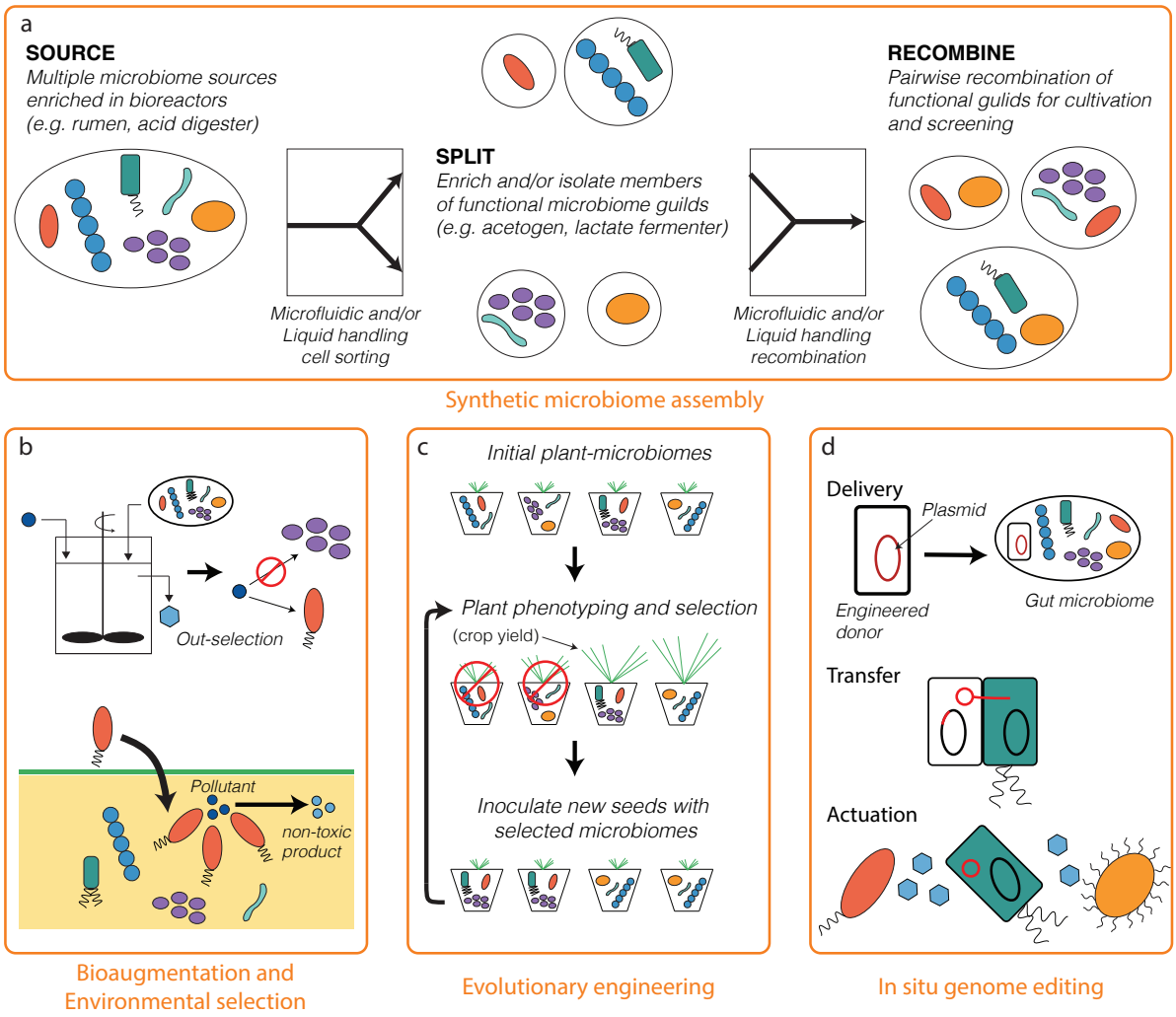

Figure 6.3 - Building self-assembled and synthetic microbiomes. (a) This example shows a protocol for assembling synthetic microbiomes from multiple microbiome sources. Complex microbiomes can be taken apart into key functional members using automated microfluidic cell sorting techniques. Isolated or enriched members can then be recombined into synthetic consortia using liquid handling robotics for downstream screening and/or cultivation. (b) Microbiome assembly can also be achieved through environmental selection via bioreactor manipulation or biostimulation (top) or using bioaugmentation with defined cultures (bottom). (c) Another option is microbiome assembly through directed adaptation and/or evolution of the microbiome to acquire or optimize a desired function. (d) *In situ* microbiome engineering can be used to add new functions to microbiomes residing in the environment.

6.4.1 Box 1 - A DBTL cycle to create synthetic microbiomes with desired functions

Here, we present a generalized DBTL cycle for creating synthetic microbiomes with desired functions, integrating both top-down and bottom-up approaches. We briefly describe two iterations of the cycle and identify opportunities for incorporating high-throughput approaches and automation to increase speed and reproducibility.

Top-down approach

Design: identify biological process(es)

An example of a process to harness or replicate is anaerobic conversion of complex lignocellulosic biomass into valuable commodity chemicals. The initial design step includes selection of different innocula that may contain microorganisms with desired functions (for example, acid phase anaerobic digester, herbivore rumen, or others). Conceptual ecosystem models that include environmental parameters (pH, temperature, nutrients, etc.) and expected functional guilds (hydrolytic bacteria, fermenting bacteria, methanogens, etc.) are used to select enrichment variables.

Build: enrich microbiomes from multiple sources

Source innocula are cultivated under different environmental conditions to select for desired function using real (for example, lignocellulosic hydrolysate or rumen fluid) and synthetic media. Modulation of environmental conditions and medium composition are done to improve desired function. For complex environments (such as soil) model laboratory ecosystems could be ideal platforms for microbiome enrichment[146].

Test: evaluate performance

Performance of enriched microbiomes are tested on real and synthetic media using high-throughput phenotypic screens. High-throughput screens could be developed using microfluidic or automated microbioreactor experiments. Deeper multi-omic measurements (such as metagenomics, metatranscriptomics, and metaproteomics) are collected from high performing microbiomes.

Learn: identify key functional roles of microbiome members

Besides key functions, bottlenecks for the desired function are identified using metabolic reconstruction and multi-omic analysis. This understanding helps to refine conceptual models of microbiome function and create quantitative models.

Bottom-up approach

Design: screen for new potential microbial partners

In silico metabolic modeling is used to screen for interacting microorganisms from high performing microbiome enrichments. Metagenome-assembled genomes (MAGs) can be used to reconstruct metabolic models of key microbiome members. Automated computational workflows (together with manual curation) will accelerate model building. FBA is used to predict each microorganism's requirements for optimal growth and activity, and unify individual metabolic models into a microbiome model to identify new potential partners that improve the design objective (for example, higher titers, rates, or yields of valuable product).

Build: recombine key microorganisms into new synthetic consortia

Following their isolation or enrichment, key microorganisms are assembled into new synthetic consortia based on *in silico* predictions at various ratios (for example, 1:1, 1:10). Microfluidic

devices and/or liquid handling robotics could be used for high-throughput isolation and recombination.

Test: test function and stability of consortia

High-throughput phenotypic screening coupled to multi-omic measurements can be used for testing. This step should also include validation of predicted metabolisms of individual isolates or enrichments.

Learn: identify microbial interactions that control function

Analyzing the metabolism of microorganisms growing in consortia versus in isolation using metabolic flux analysis (MFA) can identify important mechanisms and interactions. This understanding can be used to propose how microbiome function and stability could be optimized by environmental manipulation and/or *in situ* genome-engineering.

6.4 Testing microbiome function

The test phase involves measuring microbiome-associated phenotypes and properties to determine the efficacy of the design-build solution. The measurements should determine whether the design outcomes were achieved (for example, measuring the titer-rate-yield of a bioproduct, pollutant removal efficiency, or crop productivity) and whether the design-build solution was responsible for the observed outcome (establishing cause and effect). This typically requires readouts of ecosystem physicochemical properties (such as pH, temperature, and chemical concentrations), as well as the stoichiometry and kinetics of key ecosystem processes and microbiome functions (such as biomass growth, chemical transformations, nutrient assimilation, and metabolic fluxes). For example, acetate degradation rates and pathways to methane in an anaerobic digester microbiome could be tested using ^{13}C-labelled acetate and online biogas analysis that measures the flux through

acetoclastic methanogenesis versus syntrophic acetate oxidation coupled to hydrogenotrophic methanogenesis[113]. While the level of microbiome granularity measured during testing will depend on the specific design objectives and ecosystem complexity, the ability to quantify molecular microbial processes (for example, metabolic pathway rates and routes, enzyme activities, and individual organism growth rates) goes beyond bulk activity measurements and enables testing the specific mechanisms responsible for the observed microbiome functions. The challenge will be to develop tools that are high-throughput, quantitative, affordable, and easy to use, such that routine analyses of the microbiome over time, space, and under dynamic conditions can be accomplished.

Towards this goal, we envision a test phase comprised of high-throughput phenotypic screening of microbiome design-build solutions, followed by deeper investigation of promising solutions using multi-omic and metabolic flux analyses to obtain greater insights on underlying mechanisms (Figure 6.4). High-throughput phenotypic testing of constructed microbiomes could be achieved using droplet microfluidics, as has recently been demonstrated for screening ~100,000 synthetic communities[114]. Fully automated microbioreactor platforms that combine liquid handling and advanced sensing with microtiter plate or scaled-down bioreactor cultivation could also be used[82,83]. Combined with emerging methods to measure metabolic network activity and metabolic processes in heterogeneous environments (Box 2), rich information will be obtained to facilitate learning.

6.4.1 Microbiome metabolic network activity

To test predictions of microbiome function at a systems-level, measurement of the microbiome's *in situ* metabolic network structure and activity is critical. Multi-omic approaches (metagenomics, metatranscriptomics, metaproteomics, metabolomics) combined with bioinformatic tools have enabled the genome-centric analysis of individual species (or even strains[115]) within microbiomes

and global measurement of sequences, proteins, and metabolites[116,117,118]. These tools measure the microbiome's components on a spectrum from functional potential (for example, gene abundance) to expressed products (for example, protein and metabolite abundance), and through their combined activity produce microbiome metaphenotypes that drive system function. Currently, multi-omic approaches used to infer microbiome function have focused on correlating gene abundances or gene expression data across time and space with ecosystem geochemical data or process rates. This has included measurements of key functional genes and transcripts using qPCR assays (for example, ammonia monooxygenase[119]), microarrays (for example, GeoChip[120]), or untargeted high-throughput approaches (metatranscriptome and/or metaproteome). Although useful for overall system characterization and discovery, these approaches focus on measuring the components or "parts list" of the system, which are often limited predictors of emergent phenotypes due to metabolic network complexity, interactions, and regulation[121,122]. Therefore, new approaches and tools are needed to measure the *in situ* stoichiometry and fluxes of microbiome metabolic networks to permit the direct testing of design predictions and offer mechanistic insights into metabolic regulation.

MFA is the most authoritative method for measuring *in vivo* fluxes. This method calculates fluxes from metabolite stable isotope measurements obtained during isotopic labelling experiments using metabolic network modeling[123]. Although MFA has been used to measure fluxes in co-cultures[124], flux analysis in communities is challenging because metabolite pools cannot be easily assigned to individual cells and the number of possible reactions in a microbiome greatly exceed those of an individual organism. Nonetheless, isotopic tracers combined with **exometabolomics** and/or **off-gas analysis** have been used to determine process fluxes driving

important microbiome functions, such as syntrophic acetate oxidation and methanogenesis during anaerobic digestion[116].

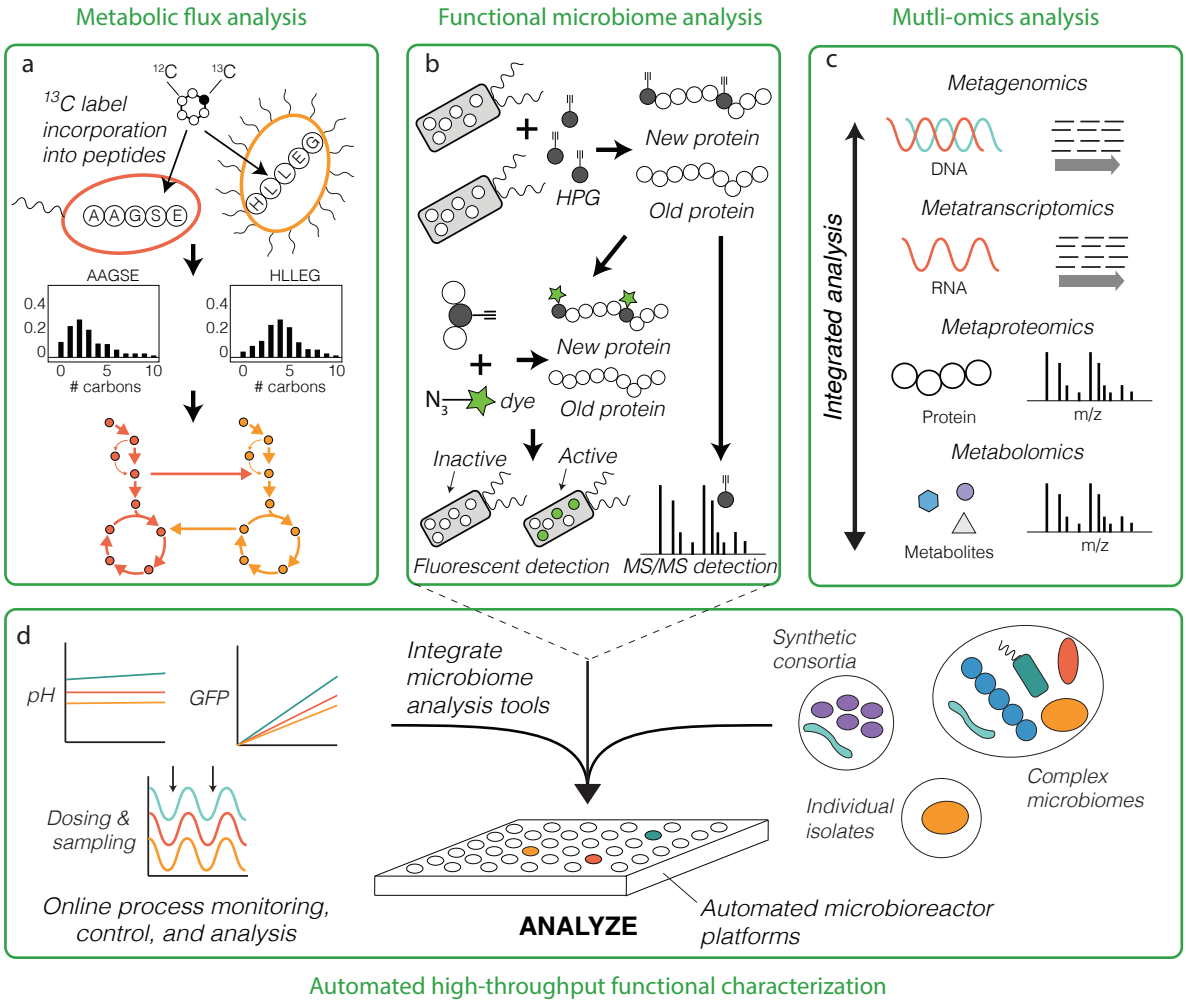

Figure 6.4 - Testing microbiome function. (a) Isotopic tracers combined with metaproteome can also be used to measure microbiome metabolic flux by analyzing isotopic labelling patterns of short peptides rather than amino acids (metabolome). (b) Biorthogonal non-canonical amino acid tagging (BONCAT) is a method for rapid profiling of the anabolic processes (growth) *in situ* using either fluorescent detection or metaproteomics. (c) Metagenomics, metatranscriptomics, metaproteomics, and metabolomics can be integrated to reconstruct and analysis metabolic network expression in microbiomes. (d) An automated microbioreactor platform enables high-throughput analysis of microbiome processes across diverse conditions (for example, with changing environmental or physiological variables). The platform can integrate tools for detailed functional analysis of individual microbiome members to complex communities. HPG: the amino acid homopropargylglycine.

To circumvent the challenges with metabolite measurements, a method analyzing labelling patterns from short peptides instead of amino acids for MFA was proposed[125]. Peptides can be assigned to individual species in a microbiome using high-throughput metaproteomic approaches, which opens the door to determining fluxes in microbial communities (that is, to 'metafluxomics'). Given that fluxes represent the final outcome of cellular regulation across all levels[126], further development and demonstration of metafluxomics will be essential for advancing microbiome engineering efforts and our understanding of metabolic regulation in microbiomes. This will also require new software packages for associated computational analyses, similar to existing ^{13}C-MFA software[127]. Such data may also allow metabolic modelers to infer, rather than assume, community and individual-level objective functions and to identify new constraints, enabling the accurate prediction and measurement of reaction rates driving microbiome function.

6.4.2 Measuring function in spatially heterogeneous environments

Most natural microbiomes, such as those associated with plants (for example, rhizosphere), humans (for example, oral microbiome), and industrial processes (for example, acid mine drainage), display highly-organized spatial organization across micro-scale physicochemical gradients that directly influences microbiome function. For example, the spatial proximity of microorganisms can control whether they interact through diffusible substrates or direct transfer[128], whereas variations in colony size can dramatically influence apparent substrate affinity constants and substrate competition between biofilm microorganisms[129]. Therefore, one of the biggest challenges will be to create tools that measure and report on microbiome spatial structure and function across all relevant scales (from μm to km). Current methods to measure **structure-function** relationships have focused on the μm to mm scale using approaches such as fluorescence *in situ* hybridization (FISH) combined with stable isotope labeling (SIP)[130], chemical

fingerprinting[131], mass spectrometry imaging[132], and/or fluorescence-based biorthogonal non-canonical amino acid tagging (BONCAT)[133] (*Box 2*). Although these techniques have successfully identified the substrate use and activity patterns of spatially distributed microorganisms in microbiomes, they are limited by throughput and can only examine and/or differentiate a limited number of organisms. The integrated application of labelling techniques (for example, SIP and BONCAT) with metaproteomics and cell sorting (for example, fluorescence-activated cell sorting (FACS)[133]) could be used to measure the metabolic activity of microorganisms in high throughput with spatial resolution. Combined with microsensor devices that profile microenvironmental chemical properties, for example, through microelectrodes[134] or engineered biosensors[95], microbiome structure, function, and ecosystem physicochemical parameters could be monitored in real-time.

6.4.3 Box 2 - A toolbox for measuring microbiome function

Multi-omics integration. The ability to assemble genomes from metagenomic data[28] has enabled the genome-resolved analysis of individual transcriptomes[63] and proteomes[118] from diverse communities and greatly increased the interpretive power of multi-omic datasets. A key challenge moving forward will be the integration of metabolomic information[163], both intracellular and extracellular, which cannot be readily assigned to individual members of the microbiome such as DNA, RNA, and proteins can be. The large amount of unknown or poorly characterized genes, enzymes and metabolites currently limits the interpretive power of multi-omic information. It does, however, create novel targets for further biochemical studies. Advances in bioinformatic tools, such as data-driven approaches (for example, statistical or machine learning methods) and knowledge-based approaches (for example, interaction networks or genome-scale metabolic

modeling)[164,165], will be key to the success of systematic measurements of microbiome function through coherent multi-omics data integration.

Isotopic tracers. Isotopic tracers have a long history in functional analysis in both pure cultures and communities, and have been combined with DNA[166], RNA[167], and protein[116] measurements to link individual populations to specific *in situ* functions. Moving forward, more efforts to incorporate isotopic tracers with multi-omics (especially metaproteomics and metabolomics) are needed for illuminating the complex metabolic networks within microbiomes. The combination of these techniques should also pave the way for measurement of intracellular and extracellular reaction rates ('metafluxomics')[124,125], which has been one of the most powerful tools for elucidating *in vivo* phenotypes, pathway constraints, and metabolic regulation in pure cultures used for engineering purposes.

Mass spectrometry imaging. Mass spectrometry imaging (MSI) techniques visualize the distribution of elements and their isotopes as well as biomolecules within complex samples. MSI is well suited for the analysis of spatially structured microbiomes and for the investigation of cellular interactions. When combined with FISH, MSI also enables the linking of microbiome structure with function[168,169]. The chemical coverage, spatial resolution, and sample preparation that can be obtained with different MSI techniques depends on the selected ionization method[132]. Although nanoscale secondary ion mass spectrometry (nanoSIMS) has superior lateral resolution compared to matrix-assisted laser desorption-ionization (MALDI) or desorption electrospray ionization (DESI; 50 nm, 3-50 mm and 100 mm, respectively), its relative chemical versatility is very low (elements and isotopes versus peptides, lipids, metabolites, and other molecules).

Therefore, nanoSIMS has generally been applied to study substrate use of single cells, whereas MALDI has been used to visualize chemical interactions between populations[132]. Although MALDI-MSI and DESI-MSI are more accessible than nanoSIMS[170] and could be well positioned to visualize the broad range of chemical interactions within microbiomes, they have very low throughput and their lateral resolution and sensitivity currently prohibit single-cell metabolic profiling[132]. A technique that combines the best of these two methods is nanostructure-initiator mass spectrometry (NIMS). NIMS is a matrix-free desorption-ionization technique that depends on initiator molecules trapped in 30 nm large pores to achieve the ionization of small molecules adsorbed to the pore surface. NIMS offers a lateral resolution of ~150 nm and is particularly well suited for the analyses of peptides and metabolites[171]. So far, NIMS has only seen limited application in microbiology[172],[173]. We expect advances that improve these issues will make MSI a useful and more widely applied tool for functional analysis of microbiomes in the near future[174].

Bioorthogonal chemistry. Metabolic labeling techniques, such as bioorthogonal non-canonical amino acid tagging (BONCAT), offer additional approaches to measure microbiome anabolic activity *in situ*. BONCAT is based on the *in vivo* translational incorporation of a non-canonical amino acid (for example, *L*-azidohomoalanine, a *L*-methionine surrogate), followed by fluorescent labelling of tagged cellular proteins by azide-alkyne click chemistry[175]. The technique can be used together with rRNA-targeted FISH to directly link taxonomy with *in situ* activity[175]. BONCAT has also been combined with FACS to separate active cells from complex samples and further characterize them by DNA sequencing[133]. In addition, tagged proteins can be selectively enriched through bead-capture and subjected to proteomic analysis[176]. The combined application of these methods could enable the high-throughput tracking of newly synthesized proteins from

uncultivated microorganisms under different physicochemical conditions. Although BONCAT can be limited due to differences in cellular amino acid uptake and metabolic perturbation, the technique offers a flexible tool for the comparatively simple, inexpensive, and high-throughput analysis of *in situ* activity on a single-cell level.

Microfluidics. Devices that enable the high-throughput analyses of microorganisms at single-cell resolution will be important for the rapid cultivation and functional analysis of microbiomes. Microfabricated devices such microfluidic 'lab-on-chip' technology could offer multiple applications, including isolation of individual cells and populations from complex microbiomes[177], creation of *in vitro* cell-based models that facilitate assembly of synthetic microbiomes and experimentation under heterogenous microenvironmental conditions[178], and online diagnostics for rapid monitoring and detection of desired phenotypes. These applications are still in early stages of development and several challenges remain, including reliable detection of microorganisms in droplets, precise control of gas concentrations, cross contamination, and technology accessibility[177,179].

Automation. To increase the reproducibility, throughput, efficiency, and standardization of microbiome engineering, advances in automation will be necessary. This includes incorporating liquid handling robotics, microfluidic devices, automated cultivation systems, online physicochemical measurement sensors, and software into data generation and analysis workflows. Emerging examples include the use of liquid handling robotics coupled to automated micro-fermentation platforms for high-throughput cultivation[82], or microfluidics to automate the analysis of thousands of droplet experiments that probe microbial community interactions[180,114]. Such

automated platforms could also integrate several functional tools (for example, single-cell analyses and multi-omics), resulting in rich reproducible data sets that could be leveraged for machine learning and other big data analytics.

6.5 Learning microbiome design principles

Progressing through the design-build-test phases of microbiome engineering presents a unique opportunity to learn from previous failures and successes, and to incorporate new knowledge into subsequent cycles. Indeed, the learn phase of the DBTL cycle is critical for success and for improving microbiome engineering efficacy. To date there are no general strategies, techniques, or approaches that guarantee success in translating information obtained from the test phase into new knowledge that informs the next design phase. Therefore, we stress the importance of devoting enough emphasis and resources to the learn phase early on, so as to avoid, for example, the difficulties encountered in metabolic engineering due to a relative lack of investment in the learn step[13]. Further development of computational methods to formalize the learn phase will be needed, including machine learning algorithms[48,135,136], metabolic flux analysis and constraint-based analysis[36,124,125,137], ecosystem modeling approaches[138], and regulatory network analysis[139]. Together, these analyses could isolate the principal drivers of microbiome interactions and function from large datasets to inform microbiome design. For example, **generalized Lotka-Volterra equations** could infer interacting species from temporal population dynamics data that become the starting point for bottom-up design[140] or constraint-based analysis could be applied to identify key metabolite exchange reactions from ^{13}C-metabolomic data that improve flux simulation accuracy and design of anaerobic consortia[137].

More broadly, we envision the learn phase to focus on translating data into generalizable principles for microbiome engineering, through the continuous refinement of conceptual knowledge and proposed theory (for example, from traditional macroecology[141,142,143,144,51]) with each DBTL cycle. We propose that model laboratory ecosystems should be utilized to drive microbiome engineering inquiry and learning. Model laboratory ecosystems are experimental platforms that can replicate the physicochemical conditions of a complex environment (natural or engineered) in a simplified and controlled manner and contain model microbial communities (for example, the model rhizosphere microbiome THOR[145]) that can be used as testing grounds for learning how to design, construct, and optimize engineered microbiomes. These ecosystems have reduced complexity, are accessible for experimentation, and can be established in a reproducible manner, which is often not possible when working in natural environments.

Recently, model laboratory ecosystems have been developed for studying plant-soil microbiome interactions[146]. These fabricated ecosystem (EcoFAB) use 3D printing, sensing, and analytical and imagining technologies to create an experimental device that replicates the native soil ecosystem, in which microorganism and host phenotypes can be monitored in response to changing variables, enabling the systematic dissection of microbial interactions and metabolite exchanges influencing plant health[146,147]. EcoFABs offer a middle ground between model organisms and complex natural microbiomes, and can be established collaboratively between expert investigators to create standardized and reproducible devices and protocols for dissemination to the broader research community. Such model systems offer the ability to experimentally develop engineered microbiomes with desired functions in a tractable manner, and permit results to be compared with results from natural settings. This cross-examination between model and natural ecosystems will be a valuable and necessary approach for learning engineering

principles and practices that are relevant to real-world systems (not laboratory artifacts), and for acquiring knowledge on scaling-up lab-based engineering strategies to full-scale applications (Figure 6.5). For example, microfluidic-based *in vitro* models of the human gut microbiome that contain co-cultures of human cells with different bacterial consortia are already producing physiological (including epithelial cell monolayer formation, cell growth and viability, cytokine levels, and metabolomic profiles) and environmental (including oxygen gradients and laminar flow) variables that are comparable to *in vivo* variables[148].

The combination of model ecosystems with the DBTL cycle may be particularly fruitful for understanding the mechanisms governing microbial interactions and functional stability. Substantial knowledge is available on specific microorganisms that co-aggregate and exchange metabolites, such as bacteria involved in nitrogen cycling[2], consortia of methane-oxidizing archaea and sulphate-reducing bacteria[149,150,128], and syntrophic bacteria partnered with hydrogenotrophic methanogens[151,152]. However, we are only beginning to understand the complex mechanisms (such as quorum sensing and secondary metabolites) involved in regulating the behavior, interactions, and kin discrimination of microorganisms in communities[153]. Although studies have established links between microbiome functional redundancy, diversity, and stability[154], a framework to predict or engineer functionally stable microbiomes has not been attained. Through the use of model laboratory ecosystems together with existing knowledge of microbial ecology and engineering design, it may be possible to decipher the chemical language of microbiomes and discover mechanisms of other important processes (including evolution, selection, dispersal limitation, and neutral processes[155]) that enable robust and stable microbiome function. Translating this theory into engineering design practice will require a quantitative framework that links these

mechanisms to metabolic interaction networks, and new approaches that enable ecological properties to emerge from metabolic models (Box 3).

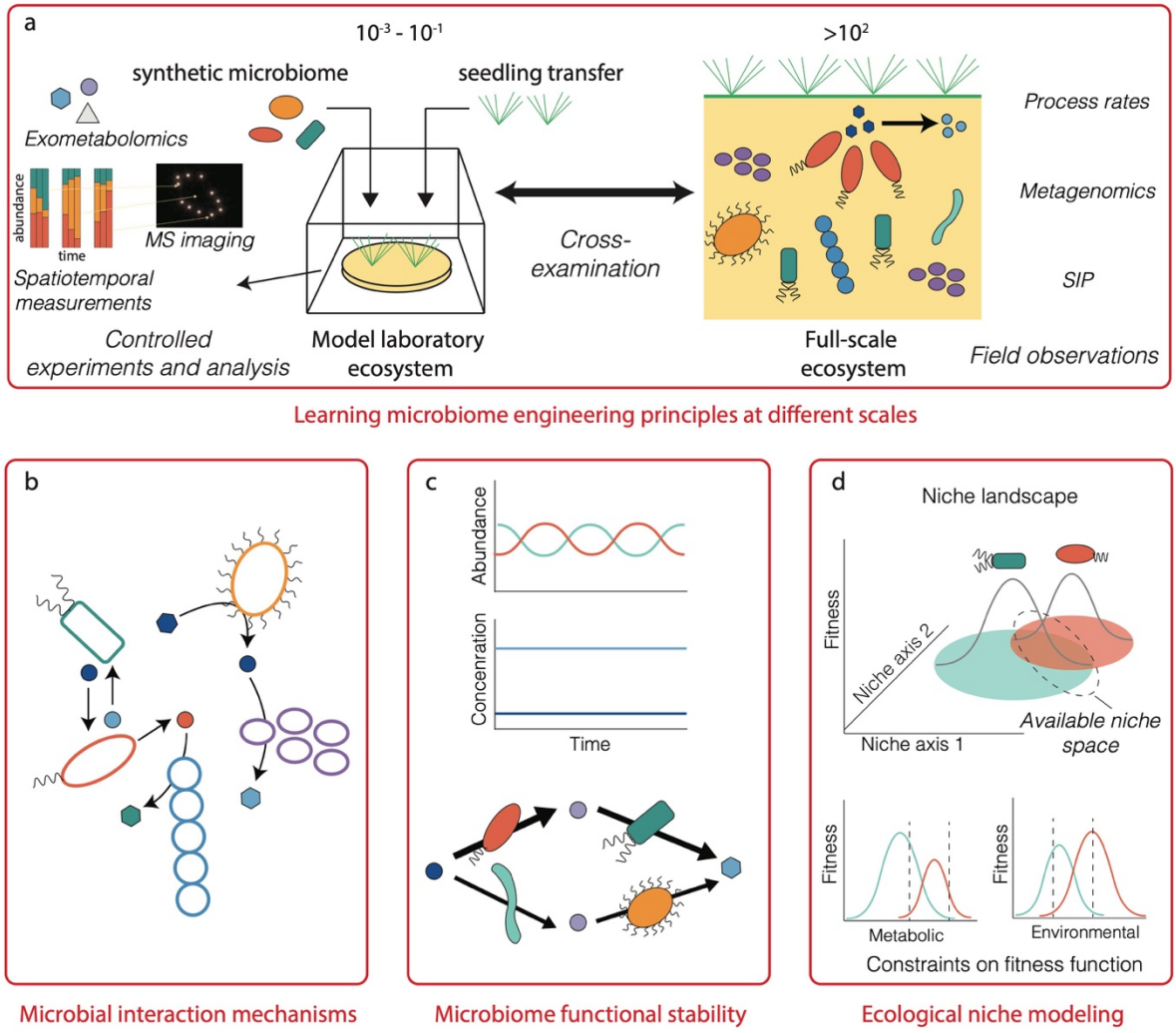

Figure 6.5 - Learning fundamental principles for microbiome engineering. (a) Model laboratory ecosystems can be used for controlled experiments with simplified microbiomes and environmental properties, representing an in-between of pure lab conditions (such as test tubes or flasks) and complex natural environments (such as soil or the ocean). Continuous cross-examination between laboratory-scale models and natural complex ecosystems will be needed for developing engineering principles and practices that are robust in real systems, while also tractable in the lab. This will require close collaboration between multiple stakeholders, including researchers and end-users (such as hospitals or treatment plants) that have expertise and experience with issues specific to each scale. Key principles that need to be learned to enable systematic microbiome engineering are microbial interaction mechanisms, mechanisms governing functional stability and degeneracy, and

frameworks for quantitatively mapping and simulating ecological niches in complex ecosystems.

6.5.1 Box 3 - Emerging principles for microbiome engineering: a case for niche modeling

Ecological niche modeling could be used to systematically design higher-order properties such as functional stability and robustness into engineered microbiomes. However, to develop such a framework, mechanistic understanding on how diversity is maintained within microbiomes and how it imparts properties such as functional stability is needed. Here we propose that this understanding could come from applying the DBTL cycle to answer key questions:

Does functional degeneracy lead to productivity and functional stability?

Diversity has been correlated with productivity and functional stability in communities of macro-organisms[143,181], yet the role that diversity has in improving microbiome function and functional stability remains open. For microbiome engineering, we propose that diversity be viewed, discussed, and defined through the lens of functional redundancy (as described previously[154]), or more specifically, functional degeneracy. This is the degree to which a set of organisms perform an identical role in ecosystem functionality (for example, methane oxidation, nitrogen fixation, or polymer hydrolysis), but exhibit degeneracy with respect to other physiological traits (for example, pH optima or biofilm formation), which enables them to achieve realized niche space and coexistence[51]. The DBTL cycle offers an excellent opportunity to understand the molecular basis of functional degeneracy and to examine how emergent community-level properties, such as resilience to perturbation or susceptibility to invasion by another species, are predictable from quantifying the fundamental and realized niche space in microbiomes. We propose that ecological niche modeling could be a particularly useful framework to achieve this goal.

How is diversity maintained in microbial ecosystems?

To create a framework for ecological niche modeling, it will be important to understand how diversity is maintained. Competitive exclusion suggests that two species with identical resource requirements cannot coexist in the same ecological niche[144]. Therefore, we need to understand the mechanisms that create niche space and enable diversity to develop and be maintained. For example, what role do the processes of spatiotemporal variability, dormancy, predation, nutrient loading, secondary metabolite production and resistance, cell motility, and biofilm formation have in niche differentiation? And how can these processes be manipulated to achieve and maintain a desired level of functional degeneracy in a microbiome? Answers to these questions will offer microbiome engineering mechanisms to design and control ecological niche space for desired microbiome properties.

How does ecological niche modeling underlie microbiome engineering?

To enable the systematic engineering of desirable higher-order microbiome properties, we propose that microbiome engineering develops a framework for ecological niche modeling. The goal of this framework would be to quantify community and individual **fundamental niche** and **realized niche** space by integrating multi-omic data, physiological information, nutrient availability, and environmental parameters, and use them to develop strategies for controlling cooperation and competition in microbiomes. To achieve this goal, new mathematical representations of the fundamental and realized niche of an organism or guild will need to be defined, together with fitness functions that describe responses to environmental variables. When incorporated into microbiome modeling, this framework will enable the ecological forecasting of higher-order properties, as well as quantification of cooperative and competitive microbiome landscapes.

Moreover, such frameworks will help guide important unresolved microbiome design questions, such as the trade-off between functional redundancy and minimal diversity.

6.6 Outlook

True advancement in microbiome engineering will need multiple rounds of DBTL to capture the necessary ecological principles to manipulate microbiomes in a precise manner with predictable outcomes (Figure 6.1). For example, incorporating direct interspecies electron transfer discovered during previous DBTL cycles into metabolic models and bioreactor construction (for example, by adding conductive materials) could optimize the efficiency of biogas production from waste[27]; or designing engineered *E. coli* to control levels of previously discovered autoinducers could tailor gut microbiota under conditions of dysbiosis towards a healthier state[156]. However, developing new knowledge and tools with fast turnaround will require next-generation infrastructure for data collection, data sharing, and knowledge integration. To accelerate progress, developing the predictive capabilities needed for the learn phase is a priority. Model laboratory ecosystems combined with advances in automation, such as liquid-handling robots, microfluidics, and data analysis pipelines[157,158], will offer a starting point for the testing of multiple designs in a rigorous and reproducible manner. Capturing new knowledge from this process and integrating information into subsequent DBTL cycles will accelerate microbiome engineering developments, creating innovative biotechnologies and practices for the management of microbiomes across medicine, agriculture, manufacturing, and environmental stewardship. Examples that show particular promise for advancing microbiome engineering across these fields include illuminating the roles that phages and metabolite cross-feeding have in controlling ruminal carbon turnover[159], harnessing untapped anaerobic fungal-bacterial consortia to improve biomass conversion to

valuable bioproducts[160,161], creating microfluidic cell sorting techniques to automatically sort stable isotope-labelled cells from high diversity samples for subsequent multi-omic analysis or cultivation[162], and developing *in situ* metagenomic engineering tools to introduce new functions into microbiomes in their native environment[111].

To move the DBTL approach forward, interdisciplinary research teams with expertise in experimentation (for example, in culturing, molecular genetics, or biochemistry) computation (for example, metabolic modeling, machine learning, or bioinformatics), automation (for example, robotics, or microfluidics), and practice (for example, professional engineers, or medical doctors) are essential. The road ahead for microbiome engineering seems long, given our nascent understanding of microbial ecology; however, structuring research and technology developments around the DBTL cycle offers a promising approach for advancing microbiome engineering and providing innovative solutions for addressing pressing societal and environmental problems.

6.7 Acknowledgements

The authors would like to acknowledge the College of Engineering at the University of Wisconsin-Madison that provided financial support for a workshop during the Madison Microbiome Meeting on April 27th, 2018 in which all authors attended and participated in discussions that led to the creation of this article. CEL is supported by a Postgraduate Scholarship-Doctoral (PGS-D) award from the National Sciences and Engineering Research Council of Canada (NSERC) and a Wisconsin Distinguished Graduate Fellowship. KDM and DRN acknowledge support from the National Science Foundation (CBET-1803055 and MCB-1518130) and from the University of Wisconsin-Madison Wisconsin Alumni Research Foundation via the Microbiome Initiative. DRN and BFP acknowledge support from the U.S. Department of Energy (DOE) Great Lakes Bioenergy

Research Center grants (DOE Office of Science BER DE-SC0018409). BFP acknowledges support from the National Science Foundation (CBET-1703504; MCB-1716594). MAO and HGM are funded by the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley Laboratory and the U.S. Department of Energy. HGM is also funded by the DOE Agile BioFoundry (http://agilebiofoundry.org), supported by the U.S. Department of Energy, Energy Efficiency and Renewable Energy, Bioenergy Technologies Office, through the same contract DE-AC02-05CH11231. HGM is also supported by the Basque Government through the BERC 2018-2021 program and by Spanish Ministry of Economy and Competitiveness MINECO through BCAM Severo Ochoa excellence accreditation SEV-2017-071. FEL acknowledges support by DOD's Strategic Environmental Research and Development Program (SERDP) and the Governor's Chair program through the University of Tennessee and Oak Ridge National Laboratory.

6.8 Glossary

Microbiome science: discovery and testing of fundamental principles governing microbiome function and assembly.

Microbiome engineering: leveraging fundamental scientific principles and quantitative design to create microbiomes that perform desired functions.

Metaphenotypes: sets of emergent functions of a microbiome resulting from the interactions between individual microbial genomes (metagenome) and their interaction with the environment.

Ecological engineering: the process of designing and operating bioreactors and other engineered systems to foster the development of specific microbial communities that can perform desired functional processes.

Exometabolomics: an analytical technique to quantify extracellular small molecule metabolites from environmental and/or biological samples typically through gas/liquid chromatography-mass spectrometry or nuclear magnetic resonance.

Functional guilds: groups organisms that use similar resources (for example, electron donors, electron acceptors, or carbon source) and occupy a similar ecological niche.

Fundamental niche: the entire set of environmental conditions in which an organism can survive and reproduce (that is, an organism's niche in the absence of interspecific competition).

Generalized Lotka-Volterra equation: A set of ordinary differential equations used to represent population dynamics based on experimentally inferred species interaction parameters.

Off-gas analysis: the monitoring of gas flow rate and chemical composition (e.g. carbon dioxide, hydrogen, methane) produced from a biological system.

Realized niche: the set of environmental conditions used by a species after considering interspecific competition (competition, predation, and others).

Keystone species: An organism that has a disproportionately large effect on maintaining the microbiome's function and microbial interactions (both between micoorganisms and with the environment).

Flux balance analysis: a constraint-based mathematical modeling technique for simulating metabolic fluxes through a metabolic network reconstructed from genomic information.

Ensemble modeling: Use of multiple models to address uncertainty by simulating a set of possibilities and selecting those consistent with measured data.

Machine learning: A technique used to build predictive models through patterns and inferences obtained from sample data, rather than explicit or mechanistic relationships.

Self-assembled microbiome: a microbiome built through environmental manipulation that selects for desired functions.

Synthetic microbiome: a microbiome built using pre-defined axenic or enrichment cultures to achieve a desired function.

Syntrophy: an obligately mutualistic process that is mediated by metabolite cross-feeding between two or more organisms that cannot be catalyzed by one organism alone.

Techno-economic assessment: A tool used to evaluate the technical and economic viability of an integrated process through a combination of process design, modeling, and economic evaluation.

Life cycle analysis: a tool used to evaluate the environmental impacts associated with all stages of a product or processes life, such as energy and water consumption, and air pollutant and greenhouse gas emissions.

Integrative and conjugative elements (ICEs): ICEs are mobile genetic elements able to integrate into DNA sites via site-specific recombination that carry genes encoding the machinery necessary for conjugation.

Structure-function relationships: the influence of the microbiomes three-dimensional spatial organization on its function.

6.9 Reference

1. Falkowski, P. G., Fenchel, T. & Delong, E. F. The microbial engines that drive earth ’s

biogeochemical cycles. *Science.* **320**, 1034–1039 (2008).

2. Kuypers, M. M. M., Marchant, H. K. & Kartal, B. The microbial nitrogen-cycling network. *Nat. Rev. Microbiol.* **16**, 263–276 (2018).
3. O'Connell, K. P., Goodman, R. M. & Handelsman, J. Engineering the rhizosphere: Expressing a bias. *Trends Biotechnol.* **14**, 83–88 (1996).
4. Löffler, F. E. & Edwards, E. A. Harnessing microbial activities for environmental cleanup. *Curr. Opin. Biotechnol.* **17**, 274–284 (2006).
5. Mccarty, P. L., Bae, J. & Kim, J. Domestic wastewater treatment as a net energy producer à can this be achieved ? *Environ. Sci. Technol.* **45**, 7100–7106 (2011).
6. Rinke, C. *et al.* Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431–7 (2013).
7. Anantharaman, K. *et al.* Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat. Commun.* **7**, 1–11 (2016).
8. Alivisatos, A. P. *et al.* A unified initiative to harness Earth's microbiomes. *Science.* **350**, 507–508 (2015).
9. Dubilier, N., McFall-Ngai, M. & Zhao, L. Create a global microbiome effort. *Nature* **526**, 631–634 (2015).
10. Price, M. N. *et al.* Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature* **557**, 503–509 (2018).
11. Zengler, K. & Zaramela, L. S. The social network of microorganisms — how auxotrophies shape complex communities. *Nat. Rev. Microbiol.* **16**, 383–390 (2018).
12. Wheelwright, S. C. & Clark, K. B. *Revolutionizing Product Development: Quantum Leaps in Speed, Efficiency, and quality*. (The Free Press, 1992).
13. **Nielsen, J. & Keasling, J. D. Engineering cellular metabolism. *Cell* 164, 1185–1197 (2016).**

This review highlights experiences, success stories, and challenges associated with implementing the DBTL cycle for metabolic engineering.

14. Blank, S. & Dorf, B. *The Startup Owner's Manual: The Step-by-Step Guide for Building a Great Company*. **1**, (K&S Ranch, Inc., 2012).

15. Jansson, J. K. & Hofmockel, K. S. The soil microbiome — from metagenomics to metaphenomics. *Curr. Opin. Microbiol.* **43**, 162–168 (2018).
16. Briones, A. & Raskin, L. Diversity and dynamics of microbial communities in engineered environments and their implications for process stability. *Curr. Opin. Biotechnol.* **14**, 270–276 (2003).
17. Verstraete, W. *et al.* Microbial resource management: the road to go for environmental biotechnology. *Eng. Life Sci.* **7**, 117–126 (2007).
18. Moralejo-Gárate, H., Mar'Atusalihat, E., Kleerebezem, R. & Van Loosdrecht, M. C. M. Microbial community engineering for biopolymer production from glycerol. *Appl. Microbiol. Biotechnol.* **92**, 631–639 (2011).
19. Nielsen, P. H. *et al.* A conceptual ecosystem model of microbial communities in enhanced biological phosphorus removal plants. *Water Res.* **44**, 5070–5088 (2010).
20. Winkler, M.-K. H. *et al.* An integrative review of granular sludge for the biological removal of nutrients and recalcitrant organic matter from wastewater. *Chem. Eng. J.* **336**, 489–502 (2018).
21. Henze, M., Gujer, W., Mino, T. & Van Loosdrecht, M. C. M. *Activated sludge models ASM1, ASM2, ASM2d and ASM3*. (IWA Publishing, 2000).
22. Batstone, D. J., Puyol, D. & Rodrı, X. F. J. Mathematical modelling of anaerobic digestion processes : applications and future needs. 595–613 (2015).
23. Muñoz-Tamayo, R., Giger-Reverdin, S. & Sauvant, D. Mechanistic modelling of in vitro fermentation and methane production by rumen microbiota. *Anim. Feed Sci. Technol.* **220**, 1–21 (2016).
24. Picioreanu, C., Kreft, J. & Loosdrecht, M. C. M. Van. Particle-based multidimensional multispecies biofilm model. *Appl. Environ. Microbiol.* **70**, 3024–3040 (2004).
25. Yunjie Ma, Carlos Domingo-Félez, Benedek Gy. Plósz, and B. F. S. Intermittent aeration suppresses nitrite-oxidizing bacteria in membrane-aerated biofilms : A Model-Based Explanation. *Environ. Sci. Technol.* **51**, 6146–6155 (2017).
26. Nobu, M. K. *et al.* Microbial dark matter ecogenomics reveals complex synergistic networks in a methanogenic bioreactor. *ISME J.* **9**, 1710–1722 (2015).
27. Rotaru, A. E. *et al.* A new model for electron flow during anaerobic digestion: Direct interspecies electron transfer to Methanosaeta for the reduction of carbon dioxide to

methane. *Energy Environ. Sci.* **7**, 408–415 (2014).

28. Albertsen, M. *et al.* Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat. Biotechnol.* **31**, 533-538 (2013).
29. Banerjee, S., Schlaeppi, K. & Heijden, M. G. A. Keystone taxa as drivers of microbiome structure and functioning. *Nat. Rev. Microbiol.* **16**, 567–576 (2018).
30. Röttjers, L. & Faust, K. Can we predict keystones? *Nat. Rev. Microbiol.* **17**, 193 (2019).
31. Thiele, I. & Palsson, B. Ø. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **5**, 93–121 (2010).
32. Wang, L., Dash, S., Ng, C. Y. & Maranas, C. D. A review of computational tools for design and reconstruction of metabolic pathways. *Synth. Syst. Biotechnol.* **2**, 243–252 (2017).
33. Zomorrodi, A. R. & Segre, D. Synthetic ecology of microbes: mathematical models and applications. *j. mol. biol.* **428**, 837-86 (2015).
34. Borenstein, E., Kupiec, M., Feldman, M. W. & Ruppin, E. Large-scale reconstruction and phylogenetic analysis of metabolic environments. *Proc. Natl. Acad. Sci.* **105**, 14482–14487 (2008).
35. Orth, J. D., Thiele, I. & Palsson, B. Ø. What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–8 (2010).
36. **Zhuang, K. *et al.* Genome-scale dynamic modeling of the competition between Rhodoferax and Geobacter in anoxic subsurface environments. *ISME J.* 5, 305–316 (2011).**

This study integrates multiple genome-scale models for dynamic flux balance analysis of a microbial community.

37. Harcombe, W. R. *et al.* Ecosystem interactions and spatial dynamics. *Cell Rep.* **7**, 1104–1115 (2014).
38. Hellweger, F. L., Clegg, R. J., Clark, J. R., Plugge, C. M. & Kreft, J. Advancing microbial sciences by individual-based modelling. *Nat. Rev. Microbiol.* **14**, 461–471 (2016).
39. **Zhou, K., Qiao, K., Edgar, S. & Stephanopoulos, G. Distributing a metabolic**

pathway among a microbial consortium enhances production of natural products. *Nat. Biotechnol.* 33, 377–383 (2015).

This study shows that designing distributed metabolic pathways over multiple microbial taxa can optimize a desired function.

40. Lilja, E. E. & Johnson, D. R. Segregating metabolic processes into different microbial cells accelerates the consumption of inhibitory substrates. *ISME J.* **10**, 1568–78 (2016).
41. Mee, M. T., Collins, J. J., Church, G. M. & Wang, H. H. Syntrophic exchange in synthetic microbial communities. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E2149–E2156 (2014).
42. Shade, A. *et al.* Fundamentals of microbial community resistance and resilience. *Front. Microbiol.* **3**, 417 (2012).
43. Balagaddé, F. K. *et al.* A synthetic Escherichia coli predator-prey ecosystem. *Mol. Syst. Biol.* **4**, 1–8 (2008).
44. Papenfort, K. & Bassler, B. L. Quorum sensing signal-response systems in Gram-negative bacteria. *Nat. Rev. Microbiol.* **14**, 576–588 (2016).
45. Feist, A. M. & Palsson, B. O. What do cells actually want? *Genome Biol.* **17**, 110 (2016).
46. Oyetunde, T., Bao, F. S., Chen, J. W., Martin, H. G. & Tang, Y. J. Leveraging knowledge engineering and machine learning for microbial bio-manufacturing. *Biotechnol. Adv.* **36**, 1308–1315 (2018).
47. Tran, L. M., Rizk, M. L. & Liao, J. C. Ensemble modeling of metabolic networks. *Biophys. J.* **95**, 5606–5617 (2008).
48. Costello, Z. & Martin, H. G. A machine learning approach to predict metabolic pathway dynamics from time-series multiomics data. *npj Syst. Biol. Appl.* **4**, 1–14 (2018).
49. Medlock, G. L. *et al.* Inferring metabolic mechanisms of interaction within a defined gut microbiota. *Cell Syst.* **7**, 245-257.e7 (2018).
50. Heckmann, D. *et al.* Machine learning applied to enzyme turnover numbers reveals protein structural correlates and improves metabolic models. *Nat. Commun.* **9**, 5252 (2018).
51. Konopka, A., Lindemann, S. & Fredrickson, J. Dynamics in microbial communities: unraveling mechanisms to identify principles. *ISME J.* **9**, 1488 (2014).

52. Smith, A. L. *et al.* Navigating wastewater energy recovery strategies: A life cycle comparison of anaerobic membrane bioreactor and conventional treatment systems with anaerobic digestion. *Environ. Sci. Technol.* **48**, 5972–5981 (2014).
53. Balakrishnan, M. *et al.* Novel pathways for fuels and lubricants from biomass optimized using life-cycle greenhouse gas assessment. *Proc. Natl. Acad. Sci.* **112**, 7645-7649 (2015).
54. Turnbaugh, P. J. *et al.* The human microbiome project. *Nature* **449**, 804–810 (2007).
55. McIlroy, S. J. *et al.* MiDAS: The field guide to the microbes of activated sludge. *Database* **2015**, 1–8 (2015).
56. Arne Alphenaar, P., Visser, A. & Lettinga, G. The effect of liquid upward velocity and hydraulic retention time on granulation in UASB reactors treating wastewater with a high sulphate content. *Bioresour. Technol.* **43**, 249–258 (1993).
57. Liu, Y. & Tay, J.-H. The essential role of hydrodynamic shear force in the formation of biofilm and granular sludge. *Water Res.* **36**, 1653–1665 (2002).
58. L., Z. *et al.* Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. *Science.* 359, 1151–1156 (2018).

This study shows how biostimulation strategies can be used to create microbiomes with desired functions, such as enhanced short chain fatty acid production.

59. Van Dongen, U., Jetten, M. S. M. & Van Loosdrecht, M. C. M. The SHARON®-Anammox® process for treatment of ammonium rich wastewater. *Water Sci. Technol.* **44**, 153–160 (2001).
60. Mueller, U. G. & Sachs, J. L. Engineering microbiomes to improve plant and animal health. *Trends Microbiol.* **23**, 606–617 (2015).
61. Winkler, M. K. H., Kleerebezem, R., Kuenen, J. G., Yang, J. & Loosdrecht, M. C. M. Van. Segregation of biomass in cyclic anaerobic / aerobic granular sludge allows the enrichment of anaerobic ammonium oxidizing bacteria at low temperatures. *Environ. Sci. Technol.* **45**, 7330–7337 (2011).
62. Laureni, M. *et al.* Biomass segregation between bio film and flocs improves the control of nitrite-oxidizing bacteria in mainstream partial nitritation and anammox processes. *Water Res.* **154**, 104–116 (2019).

63. Scarborough, M. J., Lawson, C. E., Hamilton, J. J., Donohue, T. J. & Noguera, D. R. Metatranscriptomic and thermodynamic insights into medium-chain fatty acid production using an anaerobic microbiome. *mSystems* 3, 1–21 (2018).

This study provides a detailed reconstruction of the metabolism and interactions of an anaerobic microbiome using multi-omic and thermodynamic analyses.

64. Head, I. M., Jones, D. M. & Röling, W. F. M. Marine microorganisms make a meal of oil. *Nat. Rev. Microbiol.* **4**, 173–82 (2006).
65. Swenson, W., Wilson, D. S. & Elias, R. Artificial ecosystem selection. *Proc. Natl. Acad. Sci.* **97**, 9110 – 9114 (2000).
66. Panke-buisse, K., Poole, A. C., Goodrich, J. K., Ley, R. E. & Kao-kniffin, J. Selection on soil microbiomes reveals reproducible impacts on plant function. **9**, 980–989 (2014).
67. Williams, H. T. P. & Lenton, T. M. Artificial selection of simulated microbial ecosystems. **104**, (2007).
68. King, K. C. *et al.* Rapid evolution of microbe-mediated protection against pathogens in a worm host. *ISME J.* **10**, 1915–1924 (2016).
69. Hillesland, K. L. & Stahl, D. A. Rapid evolution of stability and productivity at the origin of a microbial mutualism. **107**, 1–6 (2010).
70. Barrick, J. E. *et al.* Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature* **461**, 1243–1247 (2009).
71. Utrilla, J. *et al.* Global rebalancing of cellular resources by pleiotropic point mutations illustrates a multi-scale mechanism of adaptive evolution. *Cell Syst.* **2**, 260–271 (2016).
72. LaCroix, R. A., Palsson, B. O. & Feist, A. M. A model for designing adaptive laboratory evolution experiments. *Appl. Environ. Microbiol.* **83**, 1–14 (2017).
73. Iwabuchi, N. *et al.* Extracellular polysaccharides of rhodococcus rhodochrous s-2 stimulate the degradation of aromatic components in crude oil by indigenous marine bacteria. *Appl. Environ. Microbiol.* **68**, 2337–2343 (2002).
74. Palková, Z. Multicellular microorganisms: laboratory versus nature. *EMBO Rep.* **5**, 470–476 (2004).
75. Eydallin, G., Ryall, B., Maharjan, R. & Ferenci, T. The nature of laboratory domestication

changes in freshly isolated Escherichia coli strains. *Environ. Microbiol.* **16**, 813–828 (2014).

76. Steensels, J., Gallone, B., Voordeckers, K. & Verstrepen, K. J. Domestication of industrial microbes. *Curr. Biol.* **29**, R381–R393 (2019).
77. Seshadri, R. *et al.* Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. *Nat. Biotechnol.* **36**, 359 (2018).
78. Forster, S. C. *et al.* A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat. Biotechnol.* **37**, 186–192 (2019).
79. Jiang, C.-Y. *et al.* High-throughput single-cell cultivation on microfluidic streak plates. *Appl. Environ. Microbiol.* **82**, 2210–8 (2016).
80. **Lagier, J. C. *et al.* Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol.* 1, (2016).**

This study, together with Browne et al.[81], describe approaches for high-throughput culturing of microorganisms that can be integrated with metagenomics and enable genome-sequencing, archiving, and phenotypic analysis.

81. Browne, H. P. *et al.* Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* **533**, 543 (2016).
82. Huber, R. *et al.* Robo-Lector - A novel platform for automated high-throughput cultivations in microtiter plates with high information content. *Microb. Cell Fact.* **8**, 1–15 (2009).
83. Clark, C. *et al.* Characterization of TAP Ambr 250 disposable bioreactors, as a reliable scale-down model for biologics process development. *Biotechnol. Prog.* **33**, 478–489 (2017).
84. Gach, P. C. *et al.* A droplet microfluidic platform for automating genetic engineering. *ACS Synth. Biol.* **5**, 426–433 (2016).
85. Prakadan, S. M., Shalek, A. K. & Weitz, D. A. Scaling by shrinking: Empowering single-cell 'omics' with microfluidic devices. *Nat. Rev. Genet.* **18**, 345–361 (2017).
86. Cole, R. H. *et al.* Printed droplet microfluidics for on demand dispensing of picoliter droplets and cells. *Proc. Natl. Acad. Sci.* **114**, 8728–8733 (2017).

87. Ando, H., Lemire, S., Pires, D. P. & Lu, T. K. Engineering modular viral scaffolds for targeted bacterial population editing. *Cell Syst.* **1**, 187–196 (2015).
88. Lan, F., Demaree, B., Ahmed, N. & Abate, A. R. Single-cell genome sequencing at ultra-high-throughput with microfluidic droplet barcoding. *Nat. Biotechnol.* **35**, 640–646 (2017).
89. Heinemann, J. *et al.* On-chip integration of droplet microfluidics and nanostructure-initiator mass spectrometry for enzyme screening. *Lab Chip* **17**, 323–331 (2017).
90. Shapiro, R. S., Chavez, A. & Collins, J. J. CRISPR-based genomic tools for the manipulation of genetically intractable microorganisms. *Nat. Rev. Microbiol.* **16**, 333–339 (2018).
91. Cobb, R. E., Wang, Y. & Zhao, H. High-efficiency multiplex genome editing of Streptomyces species using an engineered CRISPR/Cas system. *ACS Synth. Biol.* **4**, 723–728 (2015).
92. Nayak, D. D. & Metcalf, W. W. Cas9-mediated genome editing in the methanogenic archaeon *Methanosarcina acetivorans*. *Proc. Natl. Acad. Sci.* **114**, 2976–2981 (2017).
93. Shih, S. C. C. *et al.* A versatile microfluidic device for automating synthetic biology. *ACS Synth. Biol.* **4**, 1151–1164 (2015).
94. Kotula, J. W. *et al.* Programmable bacteria detect and record an environmental signal in the mammalian gut Jonathan. *Proc. Natl. Acad. Sci.* **111**, 4838–4843 (2014).
95. Riglar, D. T. *et al.* Engineered bacteria can function in the mammalian gut long-term as live diagnostics of inflammation. *Nat. Biotechnol.* **35**, 653–658 (2017).
96. Nielsen, P. H., Saunders, A. M., Hansen, A. A., Larsen, P. & Nielsen, J. L. Microbial communities involved in enhanced biological phosphorus removal from wastewater — a model system in environmental biotechnology. *Curr. Opin. Biotechnol.* **23**, 452–459 (2012).
97. Nadell, C. D., Drescher, K. & Foster, K. R. Spatial structure, cooperation and competition in biofilms. *Nat. Rev. Microbiol.* **14**, 589 (2016).
98. Vlaeminck, S. E., Cloetens, L. F. F., Carballa, M., Boon, N. & Verstraete, W. Granular biomass capable of partial nitritation and anammox. *Water Sci. Technol.* **58**, 1113–1120 (2008).
99. Mark Welch, J. L., Rossetti, B. J., Rieken, C. W., Dewhirst, F. E. & Borisy, G. G.

Biogeography of a human oral microbiome at the micron scale. *Proc. Natl. Acad. Sci.* **113**, E791 LP-E800 (2016).

100. Werner, J. J. *et al.* Bacterial community structures are unique and resilient in full-scale bioenergy systems. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4158–63 (2011).
101. Gruber-Dorninger, C. *et al.* Functionally relevant diversity of closely related Nitrospira in activated sludge. *ISME J.* **9**, 643-655 (2015).
102. Kim, H. J., Boedicker, J. Q., Choi, J. W. & Ismagilov, R. F. Defined spatial structure stabilizes a synthetic multispecies bacterial community. *Proc. Natl. Acad. Sci.* **105**, 18188 – 18193 (2008).
103. Connell, J. L., Ritschdorff, E. T., Whiteley, M. & Shear, J. B. 3D printing of microscopic bacterial communities. *Proc. Natl. Acad. Sci.* **110**, 18380 – 18385 (2013).
104. Schaffner, M., Rühs, P. A., Coulter, F., Kilcher, S. & Studart, A. R. 3D printing of bacteria into functional complex materials. *Sci. Adv.* **3**, eaao6804 (2017).
105. Datta, M. S., Sliwerska, E., Gore, J., Polz, M. & Cordero, O. X. Microbial interactions lead to rapid micro-scale successions on model marine particles. *Nat. Commun.* 7, 11965 (2016).

This study shows how synthetic polysaccharide particles can be used as a model system to study ecological processes and microbe-microbe interactions.

106. Enke, T. N. *et al.* Modular assembly of polysaccharide-degrading marine microbial communities. *Curr. Biol.* **29**, 1528-1535.e6 (2019).
107. Rusten, B., Eikebrokk, B., Ulgenes, Y. & Lygren, E. Design and operations of the Kaldnes moving bed biofilm reactors. *Aquac. Eng.* **34**, 322–331 (2006).
108. Venturelli, O. S., Egbert, R. G. & Arkin, A. P. Towards engineering biological systems in a broader context. *J. Mol. Biol.* **428**, 928–944 (2016).
109. Lee, J. W., Chan, C. T. Y., Slomovic, S. & Collins, J. J. Next-generation biocontainment systems for engineered organisms. *Nat. Chem. Biol.* **14**, 530–537 (2018).
110. Sheth, R. U., Cabral, V., Chen, S. P. & Wang, H. H. Manipulating bacterial communities by in situ microbiome engineering. *Trends Genet.* **32**, 189–200 (2016).
111. Ronda, C., Chen, S. P., Cabral, V., Yaung, S. J. & Wang, H. H. Metagenomic

engineering of the mammalian gut microbiome in situ. *Nat. Methods* 16, (2019).

This study, together with Brophy et al.[112], provide new techniques to transfer engineered mobile genetic elements into microorganisms living in their native environment.

112. Brophy, J. A. N. *et al.* Engineered integrative and conjugative elements for efficient and inducible DNA transfer to undomesticated bacteria. *Nat. Microbiol.* **3**, 1043–1053 (2018).
113. Mulat, D. G. *et al.* Quantifying contribution of synthrophic acetate oxidation to methane production in thermophilic anaerobic reactors by membrane inlet mass spectrometry. *Environ. Sci. Technol.* **8**, 2505–2511 (2014).
114. Kehe, J. *et al.* Massively parallel screening of synthetic microbial communities. *Proc. Natl. Acad. Sci.* 116, 12804-12809 (2019).

This study develops a high-throughput phenotypic screen using droplet-based microfluidics that can analyze ~100,000 multi-species synthetic communities per day against any optically assayable function.

115. Truong, D. T., Tett, A., Pasolli, E., Huttenhower, C. & Segata, N. Microbial strain-level population structure and genetic diversity diversity from metagenomes. *Nat. Methods* **12**, 626–638 (2017).
116. Mosbæk, F. *et al.* Identification of syntrophic acetate-oxidizing bacteria in anaerobic digesters by combined protein-based stable isotope probing and metagenomics. *ISME J.* **10**, 2405–2418 (2016).
117. Lawson, C. E. *et al.* Metabolic network analysis reveals microbial community interactions in anammox granules. *Nat. Commun.* **8**, 1–12 (2017).
118. Hawley, A. K., Brewer, H. M., Norbeck, A. D., Pasa-Toli, L. & Hallam, S. J. Metaproteomics reveals differential modes of metabolic coupling among ubiquitous oxygen minimum zone microbes. *Proc. Natl. Acad. Sci.* **111**, 11395–11400 (2014).
119. Bowen, J. L., Babbin, A. R., Kearns, P. J. & Ward, B. B. Connecting the dots : linking nitrogen cycle gene expression to nitrogen fluxes in marine sediment mesocosms. *Front. Microbiol.* **5**, 1–10 (2014).

120. He, Z. *et al.* GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J.* **1**, 67–77 (2007).
121. Hellerstein, M. K. In vivo measurement of fluxes through metabolic pathways: The missing link in functional genomics and pharmaceutical research. *Annu. Rev. Nutr.* **23**, 379–402 (2003).
122. Sauer, U. Metabolic networks in motion: ^{13}C-based flux analysis. *Mol. Syst. Biol.* **2**, 62 (2006).
123. Antoniewicz, M. R., Kelleher, J. K. & Stephanopoulos, G. Elementary Metabolite Units (EMU): a novel framework for modeling isotopic distributions. *Metab. Eng.* **9**, 68–86 (2007).
124. Gebreselassie, N. A. & Antoniewicz, M. R. (13)C-metabolic flux analysis of co-cultures: A novel approach. *Metab. Eng.* **31**, 132–9 (2015).
125. Ghosh, A. *et al.* A peptide-based method for 13C metabolic flux analysis in microbial communities. *PLoS Comput. Biol.* 10, (2014).

This study develops a novel method for calculating metabolic fluxes in microbial communities using ^{13}C-labelled peptides.

126. Nielsen, J. It Is All about Metabolic Fluxes. *J. Bacteriol.* **185**, 7031–7035 (2003).
127. Beyß, M., Azzouzi, S., Weitzel, M., Wiechert, W. & Nöh, K. The design of FluxML: A universal modeling language for ^{13}C metabolic flux analysis. *Front. Microbiol.* **10**, 1022 (2019).
128. McGlynn, S. E., Chadwick, G. L., Kempes, C. P. & Orphan, V. J. Single cell activity reveals direct electron transfer in methanotrophic consortia. *Nature* **526**, 531–535 (2015).
129. Picioreanu, C., Pérez, J. & van Loosdrecht, M. C. M. Impact of cell cluster size on apparent half-saturation coefficients for oxygen in nitrifying sludge and biofilms. *Water Res.* **106**, 371–382 (2016).
130. Nielsen, J. L. & Nielsen, P. H. Advances in microscopy: Microautoradiography of single cells. *Methods Enzymol.* **397**, 237–256 (2005).
131. Huang, W. E. *et al.* Raman-FISH: Combining stable-isotope Raman spectroscopy and fluorescence in situ hybridization for the single cell analysis of identity and function.

Environ. Microbiol. **9**, 1878–1889 (2007).

132. Dunham, S. J. B., Ellis, J. F., Li, B. & Sweedler, J. V. Mass spectrometry imaging of complex microbial communities. *Acc. Chem. Res.* **50**, 96–104 (2017).

133. Hatzenpichler, R. *et al.* Visualizing in situ translational activity for identifying and sorting slow-growing archaeal−bacterial consortia. *Proc. Natl. Acad. Sci.* 113, E4069–E4078 (2016).

This study develops a high-throughput approach for visualizing protein synthesis in individual cells within microbiomes by combining bioorthogonal non-canonical amino acid tagging with fluorescence-activated cell sorting.

134. Okabe, S., Satoh, H. & Watanabe, Y. Analysis of microbial structure and function of nitrifying biofilms. *Methods Ecol. Evol.* **337**, 213-224 (2001).

135. DiMucci, D., Kon, M. & Segrè, D. Machine learning reveals missing edges and putative interaction mechanisms in microbial ecosystem networks. *mSystems* **3**, e00181-18 (2018).

136. Qu, K., Guo, F., Liu, X., Lin, Y. & Zou, Q. Application of machine learning in microbiology. *Front. Microbiol.* **10**, 827 (2019).

137. Wang, P.-H. *et al.* An interspecies malate–pyruvate shuttle reconciles redox imbalance in an anaerobic microbial community. *ISME J.* 13, 1042–1055 (2019).

This study combines metabolic modeling with ^{13}C metabolomic experiments to resolve poorly understood metabolite exchange reactions driving ecosystem function in anaerobic microbiomes.

138. Faust, K. & Raes, J. Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* **10**, 538 (2012).

139. Imam, S., Noguera, D. R. & Donohue, T. J. An integrated approach to reconstructing genome-scale transcriptional regulatory networks. *PLoS Comput. Biol.* **11**, 1–35 (2015).

140. Venturelli, O. S. *et al.* Deciphering microbial interactions in synthetic human gut microbiome communities. *Mol. Syst. Biol.* **14**, e8157 (2018).

141. MacArthur, R. Fluctuations of animal populations and a measure of community stability.

Ecology **36**, 533–536 (1955).

142. Martin, H. G. & Goldenfeld, N. On the origin and robustness of power-law species – area relationships in ecology. *Proc. Natl. Acad. Sci.* **103**, 10310–10315 (2006).
143. Tilman, D. Biodiversity : Population versus ecosystem stability. *Ecology* **77**, 350–363 (1996).
144. Garrett Hardin. The competitive exclusion principle. *Science*. **131**, 1292–1297 (1960).
145. Lozano, G. L. *et al.* Introducing THOR, a model microbiome for genetic dissection of community behavior. *MBio* **10**, e02846-18 (2019).
146. Zengler, K. *et al.* EcoFABs: advancing microbiome science through standardized fabricated ecosystems. *Nat. Methods* 16, 567–571 (2019).

This paper describes the construction and use of standardized fabricated ecosystems for the development of theory and predictive models for microbiomes.

147. Zhalnina, K., Zengler, K., Newman, D. & Northen, T. R. Need for laboratory ecosystems to unravel the structures. *MBio* **9**, 1–8 (2018).
148. Shah, P. *et al.* A microfluidics-based in vitro model of the gastrointestinal human–microbe interface. *Nat. Commun.* **7***,* 11535 (2016).
149. Wegener, G., Krukenberg, V., Riedel, D., Tegetmeyer, H. E. & Boetius, A. Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria. *Nature* **526**, 587–590 (2015).
150. Scheller, S., Yu, H., Chadwick, G. L., McGlynn, S. E. & Orphan, V. J. Artificial electron acceptors decouple archaeal methane oxidation from sulfate reduction. *Science.* **351**, 703–707 (2016).
151. Schink, B. Energetics of syntrophic cooperation in methanogenic degradation. *Microbiol. Mol. Biol. Rev.* **61**, 262–280 (1997).
152. Jackson, B. E. & McInerney, M. J. Anaerobic microbial metabolism can proceed close to thermodynamic limits. *Nature* **415**, 454–456 (2002).
153. Phelan, V. V., Liu, W. T., Pogliano, K. & Dorrestein, P. C. Microbial metabolic exchange--the chemotype-to-phenotype link. *Nat. Chem. Biol.* **8**, 26–35 (2012).
154. Louca, S. *et al.* Function and functional redundancy in microbial systems. *Nat. Ecol. Evol.*

2, 936–943 (2018).

155. Ladau, J. & Eloe-Fadrosh, E. A. Spatial, temporal, and phylogenetic scales of microbial ecology. *Trends Microbiol.* **27**, 662–669 (2019).
156. Thompson, J. A., Oliveira, R. A., Djukovic, A., Ubeda, C. & Xavier, K. B. Manipulation of the quorum sensing signal AI-2 affects the antibiotic-treated gut microbiota. *Cell Rep.* **10**, 1861–1871 (2015).
157. Arkin, A. P. *et al.* KBase: The united states department of energy systems biology knowledgebase. *Nat. Biotechnol.* **36**, 566–569 (2018).
158. Morrell, W. C. *et al.* The experiment data depot: A web-based software tool for biological experimental data storage, sharing, and visualization. *ACS Synth. Biol*. **6**, 2248-2259 (2017).
159. Solden, L. M. *et al.* Interspecies cross-feeding orchestrates carbon degradation in the rumen ecosystem. *Nat. Microbiol.* **3**, (2018).
160. Podolsky, I. A. *et al.* Harnessing nature's anaerobes for biotechnology and bioprocessing. *Annu. Rev. Chem. Biomol. Eng.* **10**, 105–128 (2019).
161. Swift, C.L., Brown, J.L., Seppälä, S. & O'Malley, M.A. Co-cultivation of the anaerobic fungus Anaeromyces robustus with Methanobacterium bryantii enhances transcription of carbohydrate active enzymes. *J. Ind. Microbiol. Biotechnol.* (2019). doi: 10.1007/s10295-019-02188-0
162. Lee, K. S. *et al.* An automated Raman-based platform for the sorting of live cells by functional properties. *Nat. Microbiol.* **4**, 1035-1048 (2019).
163. Swenson, T. L., Karaoz, U., Swenson, J. M., Bowen, B. P. & Northen, T. R. Linking soil biology and chemistry in biological soil crust using isolate exometabolomics. *Nat. Commun.* **9**, (2018).
164. Noor, E., Cherkaoui, S. & Sauer, U. Biological insights through omics data integration. *Curr. Opin. Syst. Biol.* **15**, 39–47 (2019).
165. Guido Zampieri, Supreeta Vijayakumar, Elisabeth Yaneske, C. A. Machine and deep learning meet genome-scale metabolic modelling. *PLoS Comput. Biol.* (2019).
166. Ziels, R. M., Sousa, D. Z., Stensel, H. D. & Beck, D. A. C. DNA-SIP based genome-centric metagenomics identifies key long-chain fatty acid-degrading populations in anaerobic digesters with different feeding frequencies. *ISME J.* **12**, 112–123 (2018).

167. Fortunato, C. S. & Huber, J. A. Coupled RNA-SIP and metatranscriptomics of active chemolithoautotrophic communities at a deep-sea hydrothermal vent. *ISME J.* **10**, 1925–1938 (2016).

168. Orphan, V. J., Orphan, V. J., House, C. H. & Hinrichs, K. Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. **484**, 484–488 (2013).

169. Kaltenpoth, M., Strupat, K. & Svatoš, A. Linking metabolite production to taxonomic identity in environmental samples by (MA)LDI-FISH. *ISME J.* **10**, 527-531 (2015).

170. Nuñez, J., Renslow, R., Cliff, J. B. & Anderton, C. R. NanoSIMS for biological applications: Current practices and analyses. *Biointerphases* **13**, 03B301 (2018).

171. Northen, T. R. *et al.* Clathrate nanostructures for mass spectrometry. *Nature* **449**, 1033 (2007).

172. Louie, K. B. *et al.* "Replica-extraction-transfer" nanostructure-initiator mass spectrometry imaging of acoustically printed bacteria. *Anal. Chem.* **85**, 10856–10862 (2013).

173. Johnson, C. H. *et al.* Metabolism links bacterial biofilms and colon carcinogenesis. *Cell Metab.* **21**, 891–897 (2015).

174. Gilmore, I. S., Heiles, S. & Pieterse, C. L. Metabolic imaging at the single-cell scale: recent advances in mass spectrometry imaging. *Annu. Rev. Anal. Chem.* **12**, 201–224 (2019).

175. Hatzenpichler, R. *et al.* In situ visualization of newly synthesized proteins in environmental microbes using amino acid tagging and click chemistry. *Environ. Microbiol.* **16**, 2568–2590 (2014).

176. Ma, Y. & Yates, J. R. Proteomics and pulse azidohomoalanine labeling of newly synthesized proteins: what are the potential applications? *Expert Rev. Proteomics* **15**, 545–554 (2018).

177. Kaminski, T. S., Scheler, O. & Garstecki, P. Droplet microfluidics for microbiology: Techniques, applications and challenges. *Lab Chip* **16**, 2168–2187 (2016).

178. Bein, A. *et al.* Microfluidic organ-on-a-chip models of human intestine. *Cmgh* **5**, 659–668 (2018).

179. Aleklett, K. *et al.* Build your own soil: Exploring microfluidics to create microbial habitat structures. *ISME J.* **12**, 312–319 (2018).

180. Hsu, R. H., Clark, R. L., Tan, J. W., Romero, P. A. & Venturelli, O. S. Rapid microbial

interaction network inference in microfluidic droplets. *bioRxiv* 521823 (2019). doi:10.1101/521823

181. Tilman D. *et al.* Diversity and productivity in a long-term grassland experiment. *Science.* **294**, 843–846 (2001).

7. Conclusions and future directions

7.0 Summary of major findings

Our work provides important understanding on the metabolic versatility and interactions occurring in nitrogen cycling microbiomes and develops new computational models that enable their prediction in natural and engineered ecosystems. We show that heterotrophic bacteria in partial-nitritation anammox (PNA) systems are involved in different steps of denitrification based on metatranscriptomic analysis, likely using amino acid and peptide substrates produced by anammox bacteria as a carbon and energy source (Chapter 2). We also show that nitrate produced by anammox bacteria during growth may be recycled back to nitrite by partial denitrifying bacteria, such as certain *Chlorobi spp.*, forming a nitrite-nitrate loop that results in higher nitrogen removal efficiencies (Chapter 2). This highlights an important connection between carbon and nitrogen cycling in PNA systems.

To further understand the carbon metabolism of nitrogen cycling microbiomes, we applied cutting-edge metabolomic tools, ^{13}C fluxomics, and metabolic modeling to elucidate the autotrophic and mixotrophic metabolic networks operating in the anammox bacterium *"Candidatus"* Kuenenia stuttgartiensis (Chapter 3) and the NOB *Nitrospira moscoviensis* (Chapter 4). We find that *K. stuttgartiensis* operates an oxidative TCA to synthesize biomass precursors, despite not encoding a known citrate synthase (Chapter 3). We also elucidated the role of the Wood-Ljungdahl pathway for formate assimilation by *K. stuttgartiensis* and show that reversal of this pathway can additionally be used for acetate oxidation (Chapter 3). This provides mechanisms underlying their observed versatility, which may inform organic carbon feeding strategies to PNA systems. For *N. moscoviensis*, we experimentally verified their use of the reverse

tricarboxylic acid (rTCA) cycle and show this pathway is also used to assimilate formate indirectly. Together, these findings provide further insights on the metabolic versatility of *K. stuttgartiensis* and *N. moscoviensis* outside their chemolithoautotrophic lifestyles.

Insights from these metabolic networks were then used to build genome-scale models for the anammox bacterium *Brocadia sinica* and the comammox bacterium *Nitrospira nitrosa* to predict their interactions during co-culture (Chapter 5). Our simulations reveal that when growth is limited by oxygen uptake, comammox *Nitrospira* perform incomplete nitrification and can cooperate with anammox bacteria to effectively cycle ammonia to dinitrogen gas (Chapter 5). This novel interaction could be used to develop new comammox-anammox based PNA systems and highlights the value of our models as novel tools for hypothesis-driven exploration of interactions between microorganisms in nitrogen cycling microbiomes. We also found that heterotrophic bacteria co-occurring with anammox and comammox organisms encoded genes for denitrification and expressed amino acid and peptide transporters (Chapter 5). This reflected similar heterotrophic-anammox interactions observed in Chapter 2, suggesting this interaction mechanism may be widespread in nitrogen cycling microbiomes.

From reflecting on this work and the work of others, we provide outlook on the field of microbiome engineering and create a generalized framework for harnessing microbiomes using a design-build-test-learn (DBTL) cycle. (Chapter 6). The cycle presents common principles and best practices for engineering microbiomes to achieve desired functions, including top-down and bottom-up design processes, synthetic and self-assembled construction methods, and emerging tools to analyze microbiome function. We anticipate that adoption of this DBTL framework will rapidly advance microbiome-based biotechnologies aimed at improving human and animal health, agriculture and enabling the bioeconomy.

Overall, our results have opened new research areas to further advance the understanding and engineering of nitrogen cycling microbiomes as outlined below.

7.1 Measuring microbiome fluxes

Our ^{13}C metabolomic studies (Chapters 3 and 4) offer the first measurements of metabolic flux in any nitrogen cycling organism, offering quantitative insights on *in vivo* metabolic network activity. Extending these techniques to measure intracellular and extracellular metabolic fluxes in microbial communities, as opposed to individual organisms, would enable the experimental discovery and quantification of metabolite exchanges controlling microbe-microbe interactions. While computational frameworks to perform "metafluxomics" have been proposed, such as peptide-based metabolic flux analysis[1], they have yet to be demonstrated on complex microbiomes experimentally. Therefore, new techniques that combine metabolomics, metaproteomics, isotopic tracing and network modeling to perform metafluxomics should be explored.

7.2 Understanding interactions between denitrifying and anammox bacteria in the presence of organic substrates

Our metatranscriptomic study highlighted potential mutualist interactions between heterotrophic denitrifying bacteria and anammox bacteria under autotrophic conditions (Chapter 2). We also showed that anammox bacteria can use simple organic substrates, including acetate and formate, as electron donors and for growth (Chapter 3). However, our experiments did not examine competition between denitrifying bacteria and anammox bacteria for these substrates, which is expected to occur because they are suitable electron donors for denitrification. Indeed, preliminary chemostat studies conducted by us have shown that heterotrophic community members do increase

in abundance overtime during continuous feeding of organic substrates (*Christopher Lawson, Michele Laureni, Aina Jofra, unpublished*). Therefore, further experimental studies to identify these denitrifying bacteria and the factors influencing competition for organic substrates with anammox bacteria should be explored.

The nitrite-nitrate loop interaction between anammox and partial denitrifying bacteria we identified (Chapter 2) also points to a new control in PNA systems that could be more purposefully manipulated. For example, during nitrate build-up in PNA systems caused by unwanted NOB activity, partial denitrification could be stimulated with organic substrates. While this has recently been attempted for mainstream PNA[2], no strategy to rationally intervene and stimulate these organisms exists. Therefore, future studies should identify the specific substrates, operating conditions, and mechanisms that promote partial denitrification over complete denitrification to keep nitrite available for anammox bacteria.

7.3 Defining ecological niches of nitrogen cycling bacteria

Our studies on the carbon metabolism of *K. stuttgartiensis* and *N. moscoviensis* highlight the complexity and detail of their metabolism and metabolic versatility (Chapter 3 and 4). To achieve the ultimate goal of systematically engineering nitrogen cycling microbiomes for desired functions, detailed mapping of each organism's ecological niches will be required. This includes quantitatively mapping substrate utilization and tolerances, environmental optima (pH, temperature), growth and substrate kinetics (e.g. affinities, maximum rates), toxicities, interaction landscapes, structural properties (e.g. EPS formation), and more. Given the large number of variables that need to be characterized, high-throughput experimental pipelines, such as

microfluidic based assays[3] or automated microfermentation systems[4,5], that would enable the collection of this information in more realistic timeframes should be developed.

7.4 Quantitative ecosystem model for PNA systems

The genome-scale models we developed for *Brocadia* and *Nitrospira* (Chapters 4 and 5) represent a starting point for building a quantitative ecosystem model for the PNA process. To expand this, new models for key denitrifying bacteria identified here (Chapter 2) and elsewhere[6,7] will be needed. When integrated with the models for *Brocadia* and *Nitrospira* (Chapters 4 and 5), and the recently published model for *Nitrosomonas europaea*[8], it would be theoretically possible to simulation interactions between the major functional guilds present in PNA systems. At a minimum, this would also require determination of important kinetic parameters (affinity constants, maximum substrate utilization rates) required for dynamic simulation. Given the importance of spatial structure in granular based anammox systems[9], extension of the dFBA framework used in our study to consider diffusion and reaction of substrates within biofilms[10] should also be explored.

7.5 References

1. Ghosh, A. *et al.* A peptide-based method for 13C Metabolic Flux Analysis in microbial communities. *PLoS Comput. Biol.* **10**, e1003827 (2014).
2. Le, T. *et al.* Nitrate residual as a key parameter to efficiently control partial denitrification coupling with anammox. *Water Environ. Res.* **91**, 1455–1465 (2019).
3. Kehe, J. *et al.* Massively parallel screening of synthetic microbial communities. *Proc. Natl. Acad. Sci.* 201900102 (2019). doi:10.1073/pnas.1900102116
4. Huber, R. *et al.* Robo-Lector - A novel platform for automated high-throughput cultivations in microtiter plates with high information content. *Microb. Cell Fact.* **8**, 1–15

(2009).

5. Clark, C. *et al.* Characterization of TAP Ambr 250 Disposable Bioreactors , as a Reliable Scale-Down Model for Biologics Process Development. *Biotechnol. Prog.* **33**, 478–489 (2017).
6. Lawson, C. E. *et al.* Metabolic network analysis reveals microbial community interactions in anammox granules. *Nat. Commun.* **8**, 15416 (2017).
7. Speth, D. R., in 't Zandt, M. H., Guerrero-Cruz, S., Dutilh, B. E. & Jetten, M. S. M. Genome-based microbial ecology of anammox granules in a full-scale wastewater treatment system. *Nat. Commun.* **7**, 11172 (2016).
8. Mellbye, B. L. *et al.* Genome-Scale, Constraint-Based Modeling of Nitrogen Oxide Fluxes during Coculture of Nitrosomonas europaea and Nitrobacter winogradskyi. *mSystems* **3**, e00170-17 (2018).
9. Vlaeminck, S. E. *et al.* Aggregate size and architecture determine microbial activity balance for one-stage partial nitritation and anammox. *Appl. Environ. Microbiol.* **76**, 900–909 (2010).
10. Picioreanu, C., Kreft, J.-U. & van Loosdrecht, M. C. M. Particle-Based Multidimensional Multispecies Biofilm Model. *Appl. Environ. Microbiol.* **70**, 3024 LP – 3040 (2004).

8. Appendix A - Complete ammonia oxidation: an important control on nitrification in engineered ecosystems?

This chapter has been published as:

Lawson, C.E. & Lucker, S. Complete ammonia oxidation: an important control on nitrification in engineered ecosystems? *Curr. Opin. Biotechnol.* **50**, 158–165 (2018).

Author Contributions

C.E.L. and S.L. performed the analysis and wrote the paper.

8.0 Abstract

Nitrification has long been considered to be mediated by two distinct microbial guilds, the ammonium oxidizing bacteria and archaea, and the nitrite oxidizing bacteria. The process has been widely applied as an environmental biotechnology for ammonium removal during water and wastewater treatment. Recently, bacteria capable of complete nitrification of ammonia to nitrate (a process termed complete ammonia oxidation, or comammox) have been discovered. These novel nitrifiers have been identified in a range of engineered, natural freshwater and terrestrial ecosystems, challenging previously held knowledge on the key microorganisms and biochemical pathways controlling nitrification process performance. This paper discusses the distribution of comammox bacteria with a focus on engineered ecosystems, as well as emerging insights from recent genomic and experimental studies on their ecophysiology.

8.1 Introduction

Nitrification, the biological oxidation of ammonia (NH_3) to nitrate (NO_3^-), has been widely applied in engineered ecosystems for ammonia removal during drinking water and wastewater treatment. In drinking water systems, nitrification is employed to treat ground waters with elevated ammonium concentrations to produce biologically stable water for distribution[1]. Control of unwanted nitrification in water distribution systems has also been an engineering objective to avoid decay of chloramine disinfectants by nitrifying microorganisms[2]. In wastewater systems, nitrification is employed to prevent oxygen depletion and fish toxicity in receiving waters, and is often combined with denitrification (reduction of NO_2^-/ NO_3^- to nitrogen gas, N_2) and more recently anaerobic ammonium oxidation (anammox; oxidation of NH_4^+ with NO_2^- to N_2) for complete nitrogen removal to prevent eutrophication[3,4]. Furthermore, nitrification is used to

remove ammonium in recirculating aquaculture systems (RAS) and bioreactors treating landfill leachate[5,6].

Nitrification has traditionally been thought to occur as a two-step process, where ammonia-oxidizing bacteria and archaea (AOB and AOA) first oxidize NH_3 to nitrite (NO_2^-), and nitrite-oxidizing bacteria (NOB) subsequently convert NO_2^- to NO_3^-. However, this view has been radically altered recently when bacteria able to oxidize NH_3 all the way to NO_3^- were identified[7,8]. While the thermodynamic feasibility of this so-called comammox process (for complete ammonia oxidation) had been predicted[9,10], it took more than 120 years from the first description of nitrifying bacteria[11] to its discovery[7,8]. Since then, comammox organisms have been detected in a range of engineered and natural ecosystems, challenging previously held knowledge about the key microorganisms performing nitrification in them.

In this review, we summarize the identification and distribution of comammox *Nitrospira* in engineered ecosystems, highlight recent insights on their ecophysiology inferred from comparative genomic studies, and examine their interactions with other nitrogen-cycling microorganisms. We also discuss the importance of incorporating comammox bacteria into nitrification process control development and highlight future research directions.

8.2 Comammox in engineered ecosystems

All complete nitrifiers identified to date[12,7,13,14,8,15] belong to sublineage II of the genus *Nitrospira* (Figure 8.1A)[16,17]. This genus has been long known[18], but all members described were canonical NOB[19,17,20]. Consequently, despite several studies observing disproportionally high abundances of *Nitrospira* in response to ammonium amendment[21,22], their potential role in ammonia oxidation had previously been overlooked. Since their original discovery, comammox *Nitrospira* have been

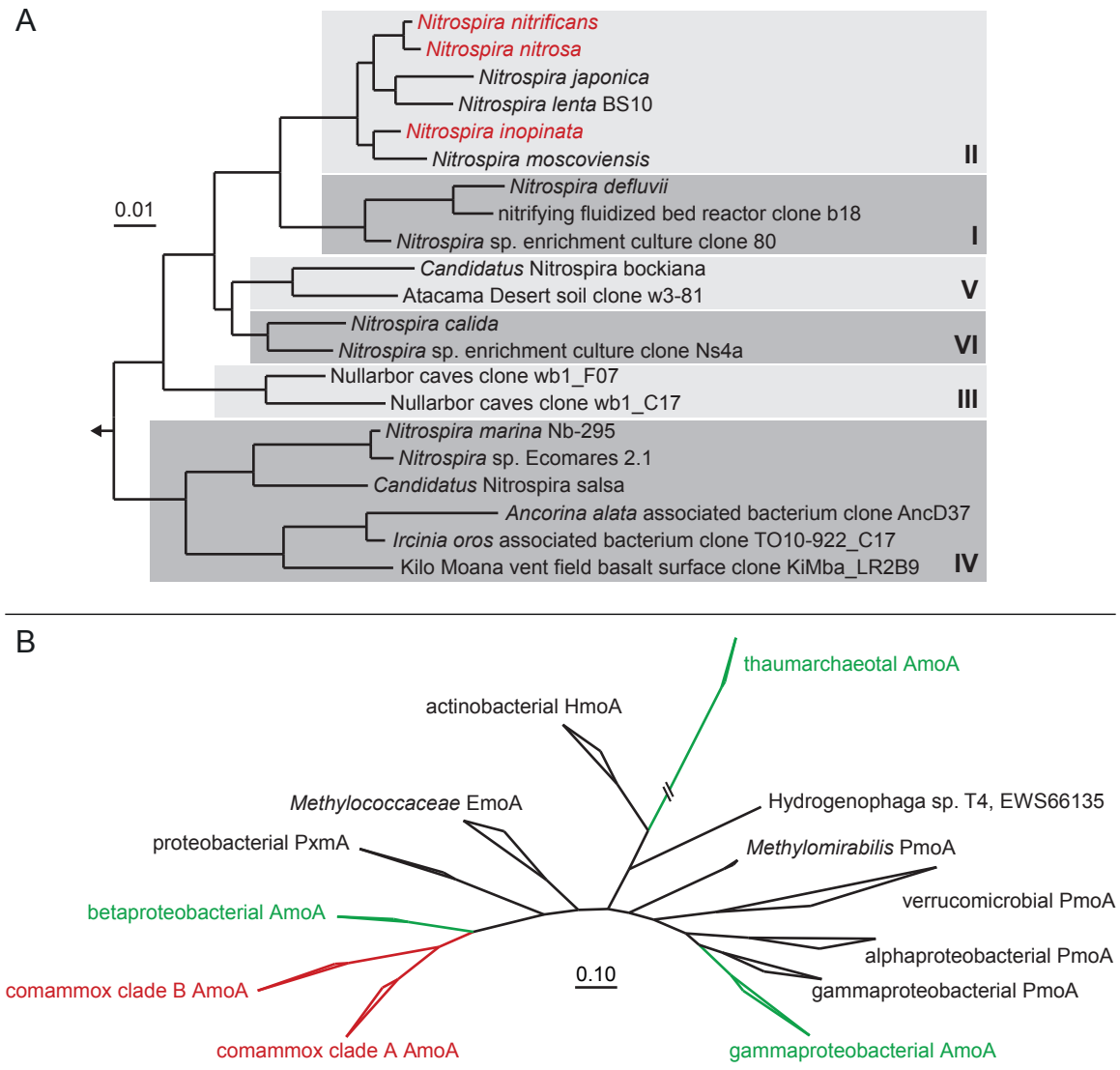

Figure 8.1 - Phylogenetic analyses of comammox *Nitrospira*. (A) 16S rRNA-based phylogenetic tree of the genus *Nitrospira*. Comammox *Nitrospira* are depicted in red. *Nitrospira* sublineages are indicated by roman numbers. (B) Phylogenetic analysis of the ammonia/methane monooxygenase family. AmoA protein clusters of comammox *Nitrospira* are shown in red, AmoA sequences of canonical ammonia oxidizers in green. Amo, ammonia monooxygenase; Emo, ethane monooxygenase; Hmo, hydrocarbon/butane monooxygenase; Pmo/Pxm, particulate methane monooxygenase.

identified in a range of engineered systems, including aquaculture biofiltration units[12,8], drinking water treatment[7,13,14] and distribution systems[15], and wastewater treatment plants [23,7,24]. They usually co-occur with AOB and/or AOA, but can greatly outnumber other ammonia-oxidizing

prokaryotes (AOP), highlighting their potential contribution to nitrification in these systems[13,24,25,15]. However, to date it is still unknown to which extend biogeography, process configuration and chemical composition of the treated (waste)water influences their distribution and abundance.

8.3 Identification of comammox

Nitrifying microorganisms can exhibit ecophysiological differences and it thus is important to identify and monitor their abundance and activity in engineered ecosystems to inform process control efforts. Initial identification of comammox *Nitrospira* resulted from screening of metagenomic datasets and subsequent assignment of ammonia monooxygenase (AMO), hydroxylamine oxidoreductase (HAO), and nitrite oxidoreductase (NXR) gene sequences to single *Nitrospira* genomes or population-level genomic bins[7,13,14,8,15]. These studies showed that comammox *Nitrospira* possess phylogentically distinct forms of AMO compared to canonical AOP (Figure 8.1B)[7,8], a key reason why previous surveys using canonical *amoA* PCR primers overlooked comammox *Nitrospira*[26-29]. Furthermore, it was observed that different comammox *Nitrospira* encode AMO orthologs that are as dissimilar to each other as to the betaproteobacterial AMO (with only *ca*. 60% amino acid identity between members of the different clades), which are referred to as comammox AMO clades A and B (Figure 8.1B)[7,30]. In addition to metagenomics, PCR primer sets targeting the distinct comammox *Nitrospira amoA* sequences have now been designed to quantify their abundance[12,24], with the primer sets reported by Pjevac *et al.*[24] having the broadest coverage (approximately 95% of clade A and 92% of clade B comammox *Nitrospira* sequences). This approach enables more rapid monitoring of comammox *Nitrospira* in engineered ecosystem, and could also be suitable for profiling potential activity via quantification of *amoA*

transcripts using RT-qPCR[24]. It is important to emphasize that comammox *Nitrospira* do not form a monophyletic group in 16S rRNA gene sequence or *nxrB*-based phylogenetic analyses, but are instead interspersed with strict nitrite-oxidizing *Nitrospira* (Figure 8.1)[7,14,8]. Consequently, it is impossible to distinguish comammox from canonical *Nitrospira* using 16S rRNA or *nxrB* gene sequences, despite published attempts suggesting otherwise [31].

8.4 Distribution of comammox

The high abundance of comammox *Nitrospira* in several engineered ecosystems highlights their potential importance for nitrification. In rapid sand filters used for groundwater treatment, *Nitrospira* were observed to be the most dominant guild (5-10% of the microbial community) in a survey of Danish drinking water treatment plants, having up to two and six orders of magnitude higher abundance than canonical AOB and AOA, respectively[32,25]. Given that these *Nitrospira* all belonged to sublineage II and a comammox *Nitrospira* population-level genome bin (CG24)[13] from one of these plants also had high abundance (30% of all mapped sequencing reads), it is probable that complete nitrification was occurring in these rapid sand filters[25]. This is consistent with recent metagenomic evidence that found comammox *Nitrospira* are broadly distributed in several drinking water systems in Asia and North America[14,15], and previous observations of disproportionally high *Nitrospira* abundances in other full-scale biological filtration units treating ground and surface waters[22,29]. Intriguingly, high abundance patterns of sublineage II *Nitrospira* have also been observed in chloraminated drinking water distributions systems[33], highlighting a need to evaluate the contribution of comammox *Nitrospira* to unwanted nitrification episodes during water supply.

In addition to drinking water systems, Bartelme et al.[12] showed that comammox *Nitrospira* were the most abundant ammonia oxidizers in a RAS biofilter, having 1.9 times the abundance of AOA, and outnumbering betaproteobacterial AOB by three orders of magnitude. Comammox *Nitrospira* have also been identified in WWTPs[23,7], consistent with the oftentimes high abundance and microdiversity of *Nitrospira* in activated sludge[21]. Although their prevalence in WWTPs remains largely unexplored, one study from a full-scale plant in Vienna, Austria, reported that comammox represented 43-71% of the total *Nitrospira* population, with greater abundances than canonical AOB[7]. Together, these studies indicate that comammox *Nitrospira* are important players performing nitrification in engineered ecosystems, interacting with a diverse community of AOA, AOB, and strictly nitrite-oxidizing *Nitrospira.*

8.5 Genomic insights on comammox *Nitrospira*

Despite their potential importance to engineered and natural ecosystems, the factors determining niche partitioning between comammox *Nitrospira*, canonical AOP, and NOB remain unknown. However, recent comparative genomic studies have shed light on the adaptation and metabolic potential of comammox *Nitrospira*[34,30] (Figure 8.2). Based on genomic evidence, comammox *Nitrospira* appear to be adapted to microaerophilic environments (Figure 8.2; Figure 8.3A). Like all *Nitrospira*, comammox encode genes for the reductive tricarboxylic acid (rTCA) cycle[34,7,35,36,8]. As discussed by Lücker *et al.*[36], the rTCA is typically employed by anaerobic or microaerophilic organisms due to oxygen-sensitive reactions with ferredoxin, and offers greater energy efficiency for CO_2 fixation compared to the oxygen-tolerant Calvin-Benson-Bassham (CBB) and thaumarchaeal hydroxypropionate-hydroxybutyrate (HP/HB) cycles employed by AOB and AOA, respectively[37,38] (Figure 3B). However, the mechanism *Nitrospira* uses to regenerate reduced

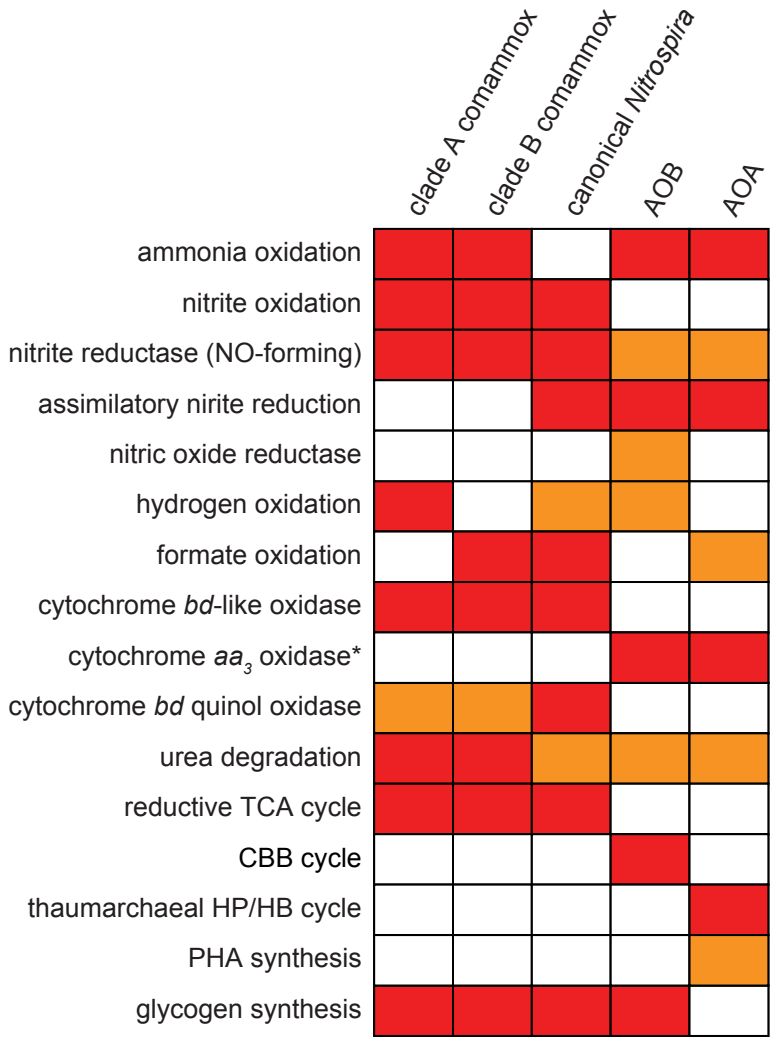

Figure 8.2 - Heatmap comparing the metabolic features encoded by comammox and canonical *Nitrospira*, AOB, and AOA. Red, function is encoded in all genomes; orange, function is encoded in some to most genomes; white, function is not encoded. *Note that the cytochrome *aa*$_3$ oxidase in AOB and AOA both belong to the heme-copper oxidase type A, but are phylogenetically distinct.

ferredoxin and how it is protected from autoxidation by O_2 are unknown and might constitute considerable energy costs. Adaptation to microaerophilic environments is also reflected in comammox *Nitrospira's* electron transport chain, which employs novel cytochrome *bd*-like terminal oxidases (described in detail for *Candidatus* N. defluvii)[36]. These cytochrome *c* oxidases are speculated to have higher oxygen affinities compared to the low-affinity *aa*$_3$-type heme-copper

oxidases employed by most canonical AOP and may confer comammox *Nitrospira* a selective advantage in engineered processes with low dissolved oxygen. Indeed, this agrees with observations from Camejo *et al.*[34] that enriched comammox *Nitrospira* in a low-dissolved oxygen bioreactor inoculated with activated sludge. However, it should be noted that these novel terminal oxidases have not been functionally characterized and both their ability to pump protons and their oxygen affinity remain to be determined experimentally.

Genomic information also indicates that comammox *Nitrospira* can use diverse substrates for energy conservation, similar to other *Nitrospira* species (Figure 8.2). They have the potential to uptake and degrade urea to ammonia via high-affinity urea ABC transporters and urease genes[4,8]. This activity has been functionally demonstrated in canonical *Nitrospira*[35] and could provide a selective advantage to comammox *Nitrospira* in ammonia-limited environments. Interestingly, all comammox species investigated to date lack genes for assimilatory nitrite reduction[7,30,8] and have been shown to depend on the presence of ammonium for growth[7,8]. Most comammox *Nitrospira* also encode group 3 [NiFe]-hydrogenase (sulfhydrogenase) and clade B representatives encode formate dehydrogenase, which could allow growth on H_2 and formate, respectively. These substrates are common fermentation end products and would enable anaerobic growth of comammox *Nitrospira* (via nitrate respiration), as has been shown for *N. moscoviensis*[39,35]. Indeed, this metabolic versatility would be particularly favorable for comammox *Nitrospira* at anoxic-aerobic interfaces typically present in biofilms or flocs in engineered ecosystems[40]. Given the known functional diversity of *Nitrospira* spp. in engineered ecosystems[21,24], efforts to recover and analyze new comammox genomes should continue to yield new ecophysiological insights. Moreover, confirmation of these predicted functions with other multi-omic approaches (transcriptomic, proteomic, metabolomics) and physiological activity

measurements are needed to elucidate how such metabolic flexibility is regulated by environmental cues and microbe-microbe interactions.

8.6 Controlling ecophysiology and microbial interactions

Environmental biotechnology seeks to control the physiology of individual microbes within a community to achieve ecosystem-level phenotypes and ensure stable process performance. This requires detailed understanding of each microbe's ecophysiology and of the interactions between community members. For nitrifying systems, comammox organisms represent a new control on nitrification for biotechnologists to manage. Kinetic theory of optimal pathway length predicted that comammox organisms would have low growth rates, high growth yields, and therefore be adapted to environments with low substrate fluxes[9]. This was recently confirmed based on kinetic analysis of a *Nitrospira inopinata* pure culture, which yielded apparent Km values 4-2500-fold lower than any kinetically characterized AOB and non-marine AOA ($Km_{(app)}$=0.84 μM), and growth yields at least 30% greater than any examined AOP[41] (Table 8.1). While these data represent yet only one comammox organism, uncultured comammox *Nitrospira* were found at high abundances in many of the engineered systems described above, where ammonium concentrations can range from 4-60 μM in bulk waters to even lower in biofilms due to diffusion gradients[41]. This suggests that ammonium availability is a key factor determining niche differentiation of comammox *Nitrospira* and other AOP (Figure 8.3A), which could be controlled by manipulating specific substrate loading rates in engineered processes.

Genomic analyses of comammox *Nitrospira* suggest that they are also adapted to microaerophilic environments and future studies should elucidate their oxygen affinity parameters to confirm this prediction. This might explain a previous report from low dissolved oxygen

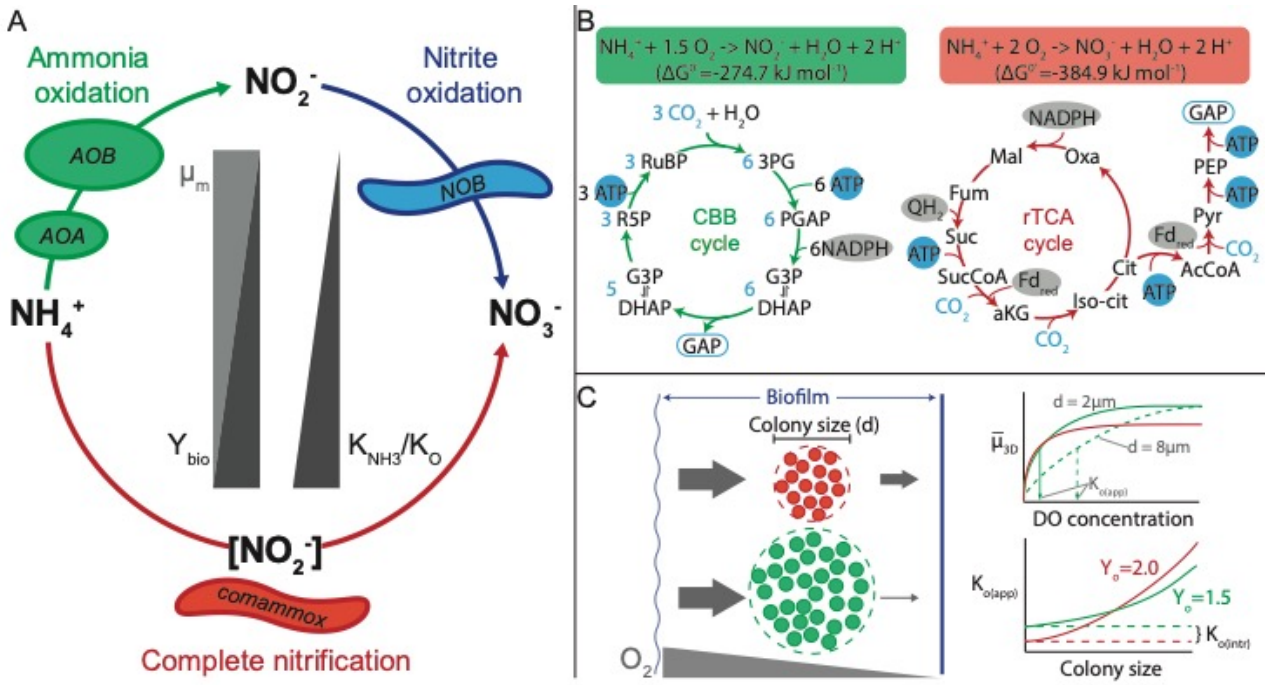

Figure 8.3 - Ecophysiological parameters impacting AOP niche differentiation. (A) Comammox *Nitrospira* have lower growth rates, higher biomass yields, and higher ammonium affinities compared to other kinetically characterized AOB and non-marine AOA, as demonstrated for *N. inopinata*[9,41]. Comammox bacteria are also predicted to have a higher oxygen affinity[9]. (B) Differences in energy yield and carbon fixation pathways between comammox *Nitrospira* and betaproteobacterial AOB. Comammox *Nitrospira* have a greater energy yield and more efficient carbon fixation pathway compared to AOB. Calvin-Benson-Bassham (CBB) cycle overall stoichiometry: 3 CO_2 + 9 ATP + 6 NADPH => $C_3H_7O_6P$ + 6 $NADP^+$ + 9 ADP + 9 P_i. Reductive TCA (rTCA) overall stoichiometry: 3 CO_2 + 4 ATP + 1 NADPH + 1 QH_2 + 2 Fd_{red} => $C_3H_7O_6P$ + 1 $NADP^+$ + 1 Q + 2Fd_{ox} + 4 ADP + 4 P_i. RuBP, ribulose 1,5-bisphosphate; 3PG, 3-phosphoglycerate; PGAP, 1,3-bisphosphoglycerate; GAP, glyceraldehyde-3-phosphate; DHAP, dihydroxyacetone phosphate; R5P, ribulose-5-phosphate; Oxa, oxaloacetate; Mal, malate; Suc, succinate; SucCoA, succinyl-CoA; aKG, alpha-ketoglutarate; Iso-cit, Isocitrate; Cit, citrate; AcCoA, acetyl-coenzyme A; Pyr, pyruvate; PEP, phosphoenolpyruvate. (C) Impact of cell cluster size and oxygen consumption stoichiometry on apparent oxygen half-saturation coefficients [$K_{O(app)}$] in nitrifying biofilms[48]. (Left) Impact of cell cluster size on local oxygen concentrations within the biofilm. Gray arrows indicate the amount of oxygen consumed by small (red) and large (green) microcolonies. (Top right) Larger cell cluster sizes create local oxygen limitations within microcolonies, reducing average specific growth rates in three-dimensional biofilms (μ_{3D}) and increasing $K_{O(app)}$. Solid lines denote uniform cell cluster size; dashed line denotes larger (green) cell cluster sizes. (Bottom right) Larger oxygen consumption stoichiometry (Y_O) can also create local oxygen limitations with microcolonies, influencing the rate that $K_{O(app)}$ increases with colony size. $K_{O(intr)}$, intrinsic oxygen half-saturation coefficient.

nitrifying reactors, where clade II *Nitrospira* composed 20% of the total community abundance with no detectable known AOP population[42]. These parameters are also important for understanding oxygen competition in nitrifying reactors and are often used in practice as the basis of control strategies. One example is partial nitritation-anammox, where *Nitrospira* have typically been viewed to have negative impacts on process performance and aeration control strategies have been developed for their out-selection[43-45]. While these strategies have had moderate success, robust control of partial nitritation is still difficult to achieve, in part because of the functional diversity among nitrifying microorganisms *in situ*, especially *Nitrospira*. Indeed, future monitoring of *Nitrospira* will need to distinguish between comammox and strict nitrite-oxidizing *Nitrospira*, which is now possible with the new *amoA*-targeting PCR primers developed[24]. Moreover, better understanding of how biofilm chemical gradients and microcolony formation influences microbial interactions is needed. For example, in nitrifying flocs and biofilms that have steep oxygen gradients[46,47], parameters such as cell cluster size and position in the biofilm can greatly impact apparent oxygen half-saturation coefficients [$K_{O(app)}$] and competition between nitrifying microorganisms[48] (Figure 8.3C). For biofilm and granules, a stratified structure of nitrifying microorganisms has been observed, with AOB dominating the outer and NOB occupying deeper layers[49,50]. This consequently increases the $K_{O(app)}$ of NOB through oxygen limitation deep in the biofilm, while causing a lower $K_{O(app)}$ for AOB on the surface[48]. The opposite effect is observed for activated sludge flocs, where due to the random distribution of nitrifying colonies the cell cluster size has the strongest influence on the observed $K_{O(app)}$. Here, the larger oxygen consumption stoichiometry combined with the typically larger colony size of AOB (Table 8.1) causes oxygen depletion inside the cell clusters, decreasing growth rates and increasing $K_{O(app)}$ of AOB more strongly than of NOB (Figure 8.3C) and resulting in reversion of apparent AOB and

NOB oxygen affinities versus the intrinsic values reported[48]. However, it should be noted that these intrinsic oxygen affinities have to date not been obtained for any *Nitrospira* pure cultures but are derived from *Nitrobacter* pure cultures or activated sludge enrichments, indicating a possible underestimation of the intrinsic K_O of *Nitrospira* in these systems. Comammox *Nitrospira* have a larger oxygen consumption stoichiometry than canonical AOB, but will still form the smaller cell clusters typical for *Nitrospira* (Table 8.1). As colony size is expected to have the strongest influence on $K_{O(app)}$[48], this would give them a competitive advantage in activated sludge flocs under oxygen-limited conditions, but not necessarily in highly structured biofilm systems. Here, they would need to outcompete AOB in the outermost biofilm layers, which might be achieved through a combination of ammonium and oxygen limitation due to their expected higher ammonia affinities[41] in combination with the formation of smaller colonies and thus reduced $K_{O(app)}$.

Table 8.1 - Physiological parameters of nitrifying microorganisms.

Parameter	clade A comammox	clade B comammox	canonical *Nitrospira*	AOB	AOA[1]
K_M (NH_3) [μM]	0.049	n.d.	n.a.	1.9-200	0.36-4.4
K_O [μM]	n.d.	n.d.	4.06-16.88	1-40	2-4
μ_{max} [h^{-1}]	0.0061	n.d.	0.019-0.058	0.028-0.068	0.010-0.013
Y_{bio} [mg protein/mol NH_3]	395	n.d.	120-213[2]	250	300
microcolony size [μm]	n.d.	n.d.	1-8	2-16	-
oxygen consumption [mol O_2:NH_3]	2	2	0.5[2]	1.5	1.5
references	[41]		[52-55]	[54,56-58]	[41,59-61]

[1]referenced for non-marine AOA
[2]per mol NO_2^-
n.d. = not determined; n.a. = not applicable

In addition to out-selecting *Nitrospira*, it is also intriguing to envision new biotechnologies that incorporate comammox *Nitrospira*. van Kessel *et al.*[8] have shown that comammox *Nitrospira* can form tight co-aggregates with anammox bacteria under hypoxia, suggesting these organisms

can cooperate, rather than compete. Presumably, comammox *Nitrospira* perform canonical ammonia oxidation under these conditions, supplying the anammox bacteria with nitrite. However, further studies on the physiological bases of this interaction are needed in order to clarify potential control strategies that could maintain this interaction in a stable nitrogen removal process.

8.7 Summary and future outlook

Since their initial discovery, high abundances of comammox *Nitrospira* have been detected in some recirculating aquaculture and drinking water systems, wastewater treatment plants, and a range of natural freshwater and terrestrial ecosystems. While these initial findings highlight their potential contribution to nitrification in engineered processes, it is important to recognize that *amoA* gene abundances, and even gene expression, does not always correlate with *in situ* nitrification activity, as demonstrated for AOA in a WWTP[51]. Accordingly, future efforts should focus on measuring the *in situ* activity of comammox *Nitrospira* to better define their functional contribution to nitrification and other nutrient cycling processes both in natural and engineered systems. In particular, understanding the ecophysiological parameters regulating comammox *Nitrospira* function will be critical for creating novel process control strategies for improved and robust management of nitrification. This includes not only understanding how environmental parameters, such as oxygen and ammonium availability, influence comammox *Nitrospira* behavior, but also how cooperative and competitive interactions between AOP, NOB, anammox bacteria, denitrifying bacteria, and comammox organisms determine ecosystem-level phenotypes. Integration of multi-omics information with physiological and reactor experiments will be needed to drive such mechanistic understanding of nitrogen cycling within engineered ecosystems. This, together with metabolic models that incorporate three-dimensional biofilm structure and kinetic

parameters, is ultimately needed to develop a predictive framework that enables the rational engineering of nitrifying microbiomes for sustainable water management.

8.8 Acknowledgements

This work was supported by funding from the Netherlands Organization for Scientific Research (VENI grant 863.14.019 and SIAM Gravitation grant 024 002 002), a Wisconsin distinguished graduate fellowship, and a Postgraduate Scholarship-Doctoral (PGS-D) by the National Sciences and Engineering Research Council of Canada (NSERC).

8.9 References

1. Rittmann BE, Snoeyink VL: Achieving biologically stable drinking water. *Journal - American Water Works Association* 1984, **76**:106-114.

2. Regan JM, Harrington GW, Noguera DR: Ammonia- and nitrite-oxidizing bacterial communities in a pilot-scale chloraminated drinking water distribution system. *Applied and Environmental Microbiology* 2002, **68**:73-81.

3. Kartal B, Kuenen JG, van Loosdrecht MCM: Sewage treatment with anammox. *Science* 2010, **328**:702-703.

4. Schmidt I, Sliekers O, Schmid M, Bock E, Fuerst J, Kuenen JG, Jetten MS, Strous M: New concepts of microbial treatment processes for the nitrogen removal in wastewater. *FEMS Microbiology Reviews* 2003, **27**:481-492.

5. Knox K: Leachate treatment with nitrification of ammonia. *Water Research* 1985, **19**:895-904.

6. Martins CIM, Eding EH, Verdegem MCJ, Heinsbroek LTN, Schneider O, Blancheton JP, d'Orbcastel ER, Verreth JAJ: New developments in recirculating aquaculture systems in

Europe: A perspective on environmental sustainability. *Aquacultural Engineering* 2010, **43**:83-93.

7. Daims H, Lebedeva EV, Pjevac P, Han P, Herbold C, Albertsen M, Jehmlich N, Palatinszky M, Vierheilig J, Bulaev A, et al.: Complete nitrification by *Nitrospira* bacteria. *Nature* 2015, **528**:504-509.

   ** This study, together with [8], describes the initial discovery of comammox *Nitrospira* and examines distribution of comammox *Nitrospira* in natrual and engineered environments.

8. van Kessel MA, Speth DR, Albertsen M, Nielsen PH, den Camp HJO, Kartal B, Jetten MS, Lücker S: Complete nitrification by a single microorganism. *Nature* 2015, **528**:555-559.

   ** This study, together with [7], describes the initial discovery of comammox *Nitrospira*, and demostrates that comammox and anammox form tight co-aggregates and are active under hypoxia.

9. Costa E, Pérez J, Kreft JU: Why is metabolic labour divided in nitrification? *Trends in Microbiology* 2006, **14**:213-219.

   ** This paper presented a theoretical prediction of the existence of comammox bacteria in environments with low substrate fluxes and first termed the name "comammox".

10. van de Leemput IA, Veraart AJ, Dakos V, de Klein JJ, Strous M, Scheffer M: Predicting microbial nitrogen pathways from basic principles. *Environmental Microbiology* 2011, **13**:1477-1487.

11. Winogradsky S: Recherches sur les organismes de la Nitrification. *Comptes rendus de l' Académie des Sciences* 1890, **110**:1013-1016.

12. Bartelme RP, McLellan SL, Newton RJ: Freshwater recirculating aquaculture system operations drive biofilter bacterial community shifts around a stable nitrifying consortium of

ammonia-oxidizing archaea and comammox *Nitrospira*. *Frontiers in Microbiology* 2017, **8**:101.

13. Palomo A, Jane Fowler S, Gulay A, Rasmussen S, Sicheritz-Ponten T, Smets BF: Metagenomic analysis of rapid gravity sand filter microbial communities suggests novel physiology of *Nitrospira* spp. *ISME Journal* 2016, **10**:2569-2581.

14. Pinto AJ, Marcus DN, Ijaz UZ, Bautista-de lose Santos QM, Dick GJ, Raskin L: Metagenomic evidence for the presence of comammox *Nitrospira*-like bacteria in a drinking water system. In *mSphere*; 2015.

15. Wang Y, Ma L, Mao Y, Jiang X, Xia Y, Yu K, Li B, Zhang T: Comammox in drinking water systems. *Water Research* 2017, **116**:332-341.

16. Daims H, Nielsen JL, Nielsen PH, Schleifer KH, Wagner M: *In situ* characterization of *Nitrospira*-like nitrite-oxidizing bacteria active in wastewater treatment plants. *Applied and Environmental Microbiology* 2001, **67**:5273-5284.

17. Lebedeva EV, Off S, Zumbrägel S, Kruse M, Shagzhina A, Lücker S, Maixner F, Lipski A, Daims H, Spieck E: Isolation and characterization of a moderately thermophilic nitrite-oxidizing bacterium from a geothermal spring. *FEMS Microbiology Ecology* 2011, **75**:195-204.

18. Watson SW, Bock E, Valois FW, Waterbury JB, Schlosser U: *Nitrospira marina* gen. nov. sp. nov.: a chemolithotrophic nitrite-oxidizing bacterium. *Archives of Microbiology* 1986, **144**:1-7.

19. Ehrich S, Behrens D, Lebedeva E, Ludwig W, Bock E: A new obligately chemolithoautotrophic, nitrite-oxidizing bacterium, *Nitrospira moscoviensis* sp. nov. and its phylogenetic relationship. *Archives of Microbiology* 1995, **164**:16-23.

20. Spieck E, Hartwig C, McCormack I, Maixner F, Wagner M, Lipski A, Daims H: Selective enrichment and molecular characterization of a previously uncultured *Nitrospira*-like bacterium from activated sludge. *Environmental Microbiology* 2006, **8**:405-415.

21. Gruber-Dorninger C, Pester M, Kitzinger K, Savio DF, Loy A, Rattei T, Wagner M, Daims H: Functionally relevant diversity of closely related *Nitrospira* in activated sludge. *ISME Journal* 2015, **9**:643-655.

22. LaPara TM, Hope Wilkinson K, Strait JM, Hozalski RM, Sadowksy MJ, Hamilton MJ: The bacterial communities of full-scale biologically active, granular activated carbon filters are stable and diverse and potentially contain novel ammonia-oxidizing microorganisms. *Applied and Environmental Microbiology* 2015, **81**:6864-6872.

23. Chao Y, Mao Y, Yu K, Zhang T: Novel nitrifiers and comammox in a full-scale hybrid biofilm and activated sludge reactor revealed by metagenomic approach. *Applied Microbiology and Biotechnology* 2016, **100**:8225-8237.

24. Pjevac P, Schauberger C, Poghosyan L, Herbold CW, van Kessel M, Daebeler A, Steinberger M, Jetten MSM, Lucker S, Wagner M, et al.: *amoA*-targeted polymerase chain reaction primers for the specific detection and quantification of comammox *Nitrospira* in the environment. *Frontiers in Microbiology* 2017, **8**:1508.

    * In this paper, *amoA*-targeted primer sets are presented, which allow the specific detection of comammox *Nitrospira* and are used to survey comammox distribution and abundance in a range of natural and engineered environments.

25. Tatari K, Musovic S, Gülay A, Dechesne A, Albrechtsen HJ, Smets BF: Density and distribution of nitrifying guilds in rapid sand filters for drinking water production: Dominance of *Nitrospira* spp. *Water Research* 2017, **127**:239-248.

26. Bagchi S, Vlaeminck SE, Sauder LA, Mosquera M, Neufeld JD, Boon N: Temporal and spatial stability of ammonia-oxidizing archaea and bacteria in aquarium biofilters. *PLOS ONE* 2014, **9**:e113515.

27. Park HD, Wells GF, Bae H, Criddle CS, Francis CA: Occurrence of ammonia-oxidizing archaea in wastewater treatment plant bioreactors. *Applied and Environmental Microbiology* 2006, **72**:5643-5647.

28. Rotthauwe JH, Witzel KP, Liesack W: The ammonia monooxygenase structural gene *amoA* as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing populations. *Applied and Environmental Microbiology* 1997, 63:4704-4712.

29. White CP, Debry RW, Lytle DA: Microbial survey of a full-scale, biologically active filter for treatment of drinking water. *Applied and Environmental Microbiology* 2012, **78**:6390-6394.

30. Palomo A, Pedersen AG, Fowler SJ, Dechesne A, Sicheritz-Pontén T, Smets BF: Comparative genomics sheds light on niche differentiation and the evolutionary history of comammox *Nitrospira*. *bioRxiv* 2017:doi:10.1101/138586.

    * This study presents the first extensive comparative genomic analysis between comammox and canonical *Nitrospira* and other AOP, highlighing differences in metabolic potential and niche adaptation.

31. Gonzalez-Martinez A, Rodriguez-Sanchez A, van Loosdrecht MCM, Gonzalez-Lopez J, Vahala R: Detection of comammox bacteria in full-scale wastewater treatment bioreactors using tag-454-pyrosequencing. *Environmental Science and Pollution Research* 2016, **23**:25501-25511.

32. Gülay A, Musovic S, Albrechtsen H-J, Al-Soud WA, Sørensen SJ, Smets BF: Ecological patterns, diversity and core taxa of microbial communities in groundwater-fed rapid gravity filters. *ISME Journal* 2016, **10**:2209.

33. Regan JM, Harrington GW, Baribeau H, Leon RD, Noguera DR: Diversity of nitrifying bacteria in full-scale chloraminated distribution systems. *Water Research* 2003, **37**:197-205.

34. Camejo PY, Santo Domingo J, McMahon KD, Noguera DR, Summers ZM: Genome-enabled insights into the ecophysiology of the comammox bacterium "*Candidatus* Nitrospira nitrosa". *mSystems* 2017, **2**:e00059-00017.

35. Koch H, Lücker S, Albertsen M, Kitzinger K, Herbold C, Spieck E, Nielsen PH, Wagner M, Daims H: Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the

genus *Nitrospira*. *Proceedings of the National Academy of Sciences of the USA* 2015, **112**:11371-11376.

* This study gave experimental proof for the metabolic versatility of *Nitrospira* and demonstrated aerobic and anaerobic formate oxidation and urea degradation.

36. Lücker S, Wagner M, Maixner F, Pelletier E, Koch H, Vacherie B, Rattei T, Damsté JSS, Spieck E, Le Paslier D, et al.: A *Nitrospira* metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *Proceedings of the National Academy of Sciences of the USA* 2010, **107**:13479-13484.

* In this study, the first *Nitrospia* genome was analyzed, yielding a detailed metabolic reconstruction of their electron transport chain and central carbon metabolism.

37. Berg IA: Ecological aspects of the distribution of different autotrophic CO_2 fixation pathways. *Applied and Environmental Microbiology* 2011, **77**:1925-1936.

38. Könneke M, Schubert DM, Brown PC, Hugler M, Standfest S, Schwander T, Schada von Borzyskowski L, Erb TJ, Stahl DA, Berg IA: Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO_2 fixation. *Proceedings of the National Academy of Sciences of the USA* 2014, **111**:8239-8244.

39. Koch H, Galushko A, Albertsen M, Schintlmeister A, Gruber-Dorninger C, Lücker S, Pelletier E, Le Paslier D, Spieck E, Richter A, et al.: Growth of nitrite-oxidizing bacteria by aerobic hydrogen oxidation. *Science* 2014, **345**:1052-1054.

40. Daims H, Lücker S, Wagner M: A new perspective on microbes formerly known as nitrite-oxidizing bacteria. *Trends in Microbiology* 2016, **24**:699-712.

* An exhaustive recent review summarizing the known metabolic diveristy and niche adaptations of *Nitrospira*.

41. Kits KD, Sedlacek CJ, Lebedeva EV, Han P, Bulaev A, Pjevac P, Daebeler A, Romano S, Albertsen M, Stein LY, et al.: Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature* 2017, **549**:269–272.

** This study described the first pure culture isolation and kinetic analysis of a complete nitrifier, demonstraing their high ammonia affinity.

42. Fitzgerald CM, Camejo P, Oshlag JZ, Noguera DR: Ammonia-oxidizing microbial communities in reactors with efficient nitrification at low-dissolved oxygen. *Water Research* 2015, **70**:38-51.

43. Pérez J, Lotti T, Kleerebezem R, Picioreanu C, van Loosdrecht MCM: Outcompeting nitrite-oxidizing bacteria in single-stage nitrogen removal in sewage treatment plants: A model-based study. *Water Research* 2014, **66**:208-218.

44. Regmi P, Bunce R, Miller MW, Park H, Chandran K, Wett B, Murthy S, Bott CB: Ammonia-based intermittent aeration control optimized for efficient nitrogen removal. *Biotechnology and Bioengineering* 2015, **112**:2060-2067.

45. Regmi P, Miller MW, Holgate B, Bunce R, Park H, Chandran K, Wett B, Murthy S, Bott CB: Control of aeration, aerobic SRT and COD input for mainstream nitritation/denitritation. *Water Research* 2014, **57**:162-171.

46. Downing LS, Nerenberg R: Effect of oxygen gradients on the activity and microbial community structure of a nitrifying, membrane-aerated biofilm. *Biotechnology and Bioengineering* 2008, **101**:1193-1204.

47. Okabe S, Satoh H, Watanabe Y: *In situ* analysis of nitrifying biofilms as determined by *in situ* hybridization and the use of microelectrodes. *Appl Environ Microbiol* 1999, **65**:3182-3191.

48. Picioreanu C, Pérez J, van Loosdrecht MCM: Impact of cell cluster size on apparent half-saturation coefficients for oxygen in nitrifying sludge and biofilms. *Water Research* 2016, **106**:371-382.

* A thorough biofilm modeling analysis highlighting the impact of cell cluster size and oxygen consumption stoichiometry on apparent oxygen affinities in nitrifying flocs and biofilms.

49. Matsumoto S, Katoku M, Saeki G, Terada A, Aoi Y, Tsuneda S, Picioreanu C, Van Loosdrecht MCM: Microbial community structure in autotrophic nitrifying granules characterized by experimental and simulation analyses. *Environmental Microbiology* 2010, **12**:192-206.

50. Vlaeminck SE, Terada A, Smets BF, De Clippeleir H, Schaubroeck T, Bolca S, Demeestere L, Mast J, Boon N, Carballa M, et al.: Aggregate size and architecture determine microbial activity balance for one-stage partial nitritation and anammox. *Applied and Environmental Microbiology* 2010, **76**:900-909.

51. Mußmann M, Brito I, Pitcher A, Sinninghe Damsté JS, Hatzenpichler R, Richter A, Nielsen JL, Nielsen PH, Müller A, Daims H, et al.: Thaumarchaeotes abundant in refinery nitrifying sludges express amoA but are not obligate autotrophic ammonia oxidizers. *Proceedings of the National Academy of Sciences of the USA* 2011, **108**:16771-16776.

52. Blackburne R, Vadivelu VM, Yuan Z, Keller J: Kinetic characterisation of an enriched *Nitrospira* culture with comparison to *Nitrobacter*. *Water Research* 2007, **41**:3033-3042.

53. Manser R, Gujer W, Siegrist H: Consequences of mass transfer effects on the kinetics of nitrifiers. *Water Research* 2005, **39**:4633-4642.

54. Manser R, Gujer W, Siegrist H: Membrane bioreactor versus conventional activated sludge system: population dynamics of nitrifiers. *Water Science and Technology* 2005, **52**:417-425.

55. Park M-R, Park H, Chandran K: Molecular and kinetic characterization of planktonic *Nitrospira* spp. selectively enriched from activated sludge. *Environmental Science & Technology* 2017, **51**:2720-2728.

56. Laanbroek HJ, Gerards S: Competition for limiting amounts of oxygen between *Nitrosomonas europaea* and *Nitrobacter winogradskyi* grown in mixed continuous cultures. *Archives of Microbiology* 1993, **159**:453-459.

57. Nowka B, Daims H, Spieck E: Comparison of oxidation kinetics of nitrite-oxidizing bacteria: nitrite availability as a key factor in niche differentiation. *Applied and Environmental Microbiology* 2015, **81**:745-753.

58. Park HD, Noguera DR: Characterization of two ammonia-oxidizing bacteria isolated from reactors operated with low dissolved oxygen concentrations. *Journal of Applied Microbiology* 2007, **102**:1401-1417.

59. Park B-J, Park S-J, Yoon D-N, Schouten S, Sinninghe Damste JS, Rhee S-K: Cultivation of autotrophic ammonia-oxidizing archaea from marine sediments in coculture with sulfur-oxidizing bacteria. *Applied and Environmental Microbiology* 2010, **76**:7575-7587.

60. Sauder LA, Albertsen M, Engel K, Schwarz J, Nielsen PH, Wagner M, Neufeld JD: Cultivation and characterization of *Candidatus* Nitrosocosmicus exaquare, an ammonia-oxidizing archaeon from a municipal wastewater treatment system. *ISME Journal* 2017, **11**:1142.

61. Stahl DA, de la Torre JR: Physiology and diversity of ammonia-oxidizing archaea. *Annual Review of Microbiology* 2012, **66**:83-1