



universität
wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Exploring the Physiology of the Ammonia Oxidizing
Archaeon *Nitrososphaera viennensis*“

verfasst von / submitted by

Logan Henry Hodgskiss

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on the student
record sheet:

A 794 685 437

Dissertationsgebiet lt. Studienblatt /
field of study as it appears on the student record sheet:

Biologie/Biology

Betreut von / Supervisor:

Univ.-Prof. Dr. Christa Schleper

Acknowledgements

The people that I would like to thank are numerous. This was a large undertaking and not one that I expected to take, and it would not have been possible without many of the people listed here.

First, I would like to thank my advisor Dr. Christa Schleper for all of the support and encouragement that she has offered over the years. Your enthusiasm for science makes a huge difference and I am grateful to have learned so much during my time spent in this lab. I am incredibly thankful for all of the opportunities that I have had and for your willingness to support and listen to ideas and concerns from not only me, but everyone in the group.

I also want to thank the other three group leaders within our division, Dr. Simon Rittmann, Dr. Silvia Bulgheresi, and Dr. Filipa Sousa. All of you have contributed ideas and support to the projects within these pages, and I thank you for your collaborative effort.

Additionally, a thank you goes to Dr. Thomas Nägele and his group, particularly Maria and Lisa, for welcoming and teaching me during my time with them in Munich.

I would also like to thank the neighboring Molecular Systems Biology Division for the support and assistance over the years.

A heartfelt thank you goes to the entire Archaea Biology and Ecogenomics Division and all of the people who have been a part of it during my time here. I have made many unforgettable friends and memories.

When I started, the division saw an influx of new PhD students and as such we all went on this journey together. So to Sinje, Kevin, Philipp, Gabriela, Michael, Lisa, and İpek, it has been rewarding to experience this with you.

I especially want to thank my original office mates, Ricardo, Isa, and İpek. The many long conversations, support, and friendships made truly mean a lot to me. Thank you for challenging me personally and scientifically and becoming friends along the way.

The same can be said for the people I share an office with at the end of this journey, Erika, Max, and Thomas. Although we met as you started your masters degrees, all three of you have grown to be good friends and I am excited to see how your own scientific adventures evolve.

There have been many post docs that helped shape my work and knowledge during my PhD for which I am grateful. In particular, I would like to thank Lokesh for always being willing to assist me with bioinformatics, and more importantly, for the friendship that came of it.

Within the AOA group, a special thank you goes to Melina for all of her advice, support, and expertise that she has shared over the years. Thank you for your support, friendship, and many valuable conversations along the way.

To Barbara, İpek, Lisa, and Chris, thank you for the many late nights, trips, and memories during the PhD adventure. Finding a tightknit and supportive group is no easy task,

and I feel lucky to have found such great people. Our adventures hold a special place in my time here in Vienna. The times that we shared are irreplaceable and I cannot thank you enough. This extends to the many friends I have been fortunate enough to make in Vienna outside of the department as well.

My final acknowledgements and recognition go to all of my dear friends and family both near and so very far away. Leaving Montana was never in my plan. However, your unwavering support and encouragement has always been something that I can rely on and will continue to remind me of my roots.

In particular, I want to thank my parents, Lyle and Beth Ann Hodgskiss. Words are simply not enough. I will always be appreciative of the sacrifices and choices that you made to support my siblings and me. I also want to thank my brothers, Taylor and Connor Hodgskiss, my sister and brother-in-law, Sarah and Dan Bock, and a special thank you to my cousin, Charlene Johnson. I am so proud of all of you. Thank you for inspiring me and supporting me over the last few years and beyond.

At last, I want to recognize my grandparents, Papa and Grandma (Wilson and Maryetta Hodgskiss) and Grandad and Grammy (Bill Brown Jr. and Emily Ann Alderson Brown). Although they are not all with us now, I want to thank them for investing in their neighbors and communities, valuing education, and instilling the beliefs and values that have been passed down to my parents and now to me. All of these actions helped set the stage for what I have now been able to accomplish.

Therefore, I would like to dedicate this thesis to them and my entire family.

Thank you.



Table of Contents

Abstract	1
Zusammenfassung	2
Chapter I - Introduction	3
Background	5
Major Goals and Contributions	15
Chapter II	
Unexpected Complexity of the Ammonia	
Monooxygenase in Archaea	29
Chapter III	
Genome Wide Transcriptomic Analysis of the Soil	
Ammonia Oxidizing Archaeon <i>Nitrososphaera</i>	
<i>viennensis</i> upon Exposure to Copper Limitation	79
Chapter IV	
Carbon Limitation of the Ammonia Oxidizing Archaeon	
<i>Nitrososphaera viennensis</i> Indicates Routes of Cellular	
Carbon Flow and Cellular Stress Responses	131
Chapter V – Discussion	185

Abstract

Ammonia oxidizing archaea (AOA) have become a focal point of nitrification research due to their recent discovery and their high abundance in many terrestrial and marine environments. While the substrates (ammonia/carbon dioxide/oxygen) and products (nitrite/biomass) of their metabolism may be the same as those of ammonia oxidizing bacteria, the pathways to achieve this metabolism are distinct and mostly unknown. Comparative genomics have shown a lack of iron-based cytochromes that could be replaced by the many copper-based plastocyanins in AOA. This switch in electron carriers positions AOA to be more dependent on copper than they already are due to the high copper demands from their primary energetic complex, the ammonia monooxygenase (AMO). Genomics have also indicated the possibility of a different subunit composition of the AMO complex, as well as a finely tuned 3-hydroxypropionate/4-hydroxybutyrate carbon fixation pathway that is so far unique to this group of archaea. The work within this thesis expands the current knowledge of AOA by experimentally testing these genomic hypotheses. This is achieved through a variety of biochemical and functional genomic methods with a focus on the soil AOA *Nitrososphaera viennensis* while encompassing the known information about various AOA species. The thesis is composed of three primary sections: (i) an examination of the gene expression of *N. viennensis* via transcriptomics under copper limiting conditions, which highlighted strategies of the organism to maximize copper acquisition and alleviate energy depletion (ii) characterization of the subunit composition of the archaeal AMO through biochemical and complexome analysis in *N. viennensis* and *Nitrosocaldus cavascurensis*, which led to the identification of three additional polypeptides in the complex and (iii) an investigation of the growth response of *N. viennensis* to carbon limitation using a proteomic and metabolomic systems biology approach guided by a thermodynamic model of AOA metabolism. The combined results from these studies have helped to elucidate the roles of proteins in the energy and carbon metabolism and stress physiology of AOA.

Zusammenfassung

Ammoniak oxidierende Archaea (AOA) sind aufgrund ihrer kürzlichen Entdeckung sowie ihres hohen Vorkommens in terrestrischen und marinen Ökosystemen zum Zentrum der Forschung im Feld der Nitrifikation geworden. Obwohl die Substrate (Ammoniak/Kohlendioxid/Sauerstoff) und Produkte (Nitrit/Biomasse) des Metabolismus gleich derer von Ammoniak oxidierenden Bakterien sind, unterscheiden sich die metabolischen Wege und sind bei AOA meist noch unbekannt. Durch vergleichende Genomik wurde das Fehlen von Eisen-basierten Cytochromen festgestellt, welche in AOA jedoch durch Kupfer-basiertes „Plastocyanin“ ersetzt werden könnten. Die durch den primären Energiekomplex, die „Ammoniak Monooxygenase“ (AMO) bereits bestehende Kupferabhängigkeit, wird durch diesen Wechsel an Elektronenüberträgern noch verstärkt. Genomische Studien haben weiters die Möglichkeit eines veränderten Aufbaus der AMO, sowie einen für diese Archaea-Gruppe einzigartigen, fein abgestimmten 3-Hydroxypropionate/4-Hydroxybutyrate Kohlenstofffixierungsweg, aufgezeigt. In dieser Arbeit wurden diese aus der vergleichenden Genomik entstandenen Hypothesen experimentell überprüft und somit die aktuellen Kenntnisse über AOA erweitert. Dies wurde durch eine Vielzahl an biochemischen und funktionellen genomischen Methoden, unter Einbezug aller verfügbaren Informationen von verschiedenen AOA Spezies, erreicht. Der Fokus liegt dabei auf dem Boden-Archaeon *Nitrososphaera viennensis*. Diese Dissertation unterteilt sich in drei primäre Kapitel: (i) die Transkriptom-Analyse von *N. viennensis* unter Kupferlimitation, welche die Kompensations-Strategien des Organismus, den Kupfererwerb zu maximieren und die Energienutzung zu maximieren, hervorhebt; (ii) die Charakterisierung der archaealen AMO mit biochemischen und Complexom-Analysen in *N. viennensis* und *Nitrosocaldus cavascurensis*, welche zur Identifikation von 3 zusätzlichen Untereinheiten im Komplex führte, und (iii) die systembiologische Analyse von *N. viennensis* unter Kohlenstoff-Limitierung durch Verwendung von Proteomik und Metabolomik, sowie geleitet durch ein thermodynamisches Modell des AOA Metabolismus. Insgesamt geben die kombinierten Forschungsergebnisse genauere Einblicke in die Rolle von Proteinen des Energie- und Kohlenstoff-Metabolismus und die Physiologie von AOA.

Chapter I

Introduction

Background

Microorganisms have fascinated biologists ever since their discovery. Their prevalence and metabolic diversity have ensured a profound effect on both human health as well as the environment that we live in. This extends to the many biogeochemical cycles that shape the planet including the nitrogen cycle, which is particularly diverse and dynamic. Within this cycle, nitrogen is converted across a wide range of oxidation states acting as both an electron donor and an electron acceptor ^{1,2}. The sub process of nitrification, the conversion of ammonium to nitrate, has attracted particular attention due to its agricultural and environmental relevance. While the nitrogen cycle is crucial for life to exist, it has become imbalanced as a consequence of anthropogenic activity, primarily through the addition of fertilizer to crops ³⁻⁵. Fertilization has allowed agricultural output to keep pace with an increasing population. However, large amounts of the added fertilizer are lost to the environment via nitrification as the conversion of ammonium to nitrite and/or nitrate makes the nitrogen more mobile in soil environments allowing it to be washed away or lost to other microbial reactions such as denitrification ^{6,7}. Nitrification is a microbially driven process and is also a key source of energy for nitrifiers, making it an alluring field to study from an energetics point of view.

Bacteria have been known to participate in the process of nitrification for over the last 125 years ⁸. Ammonia oxidizing bacteria (AOB) and nitrite oxidizing bacteria (NOB) fulfill separate niches of nitrification along with the recently discovered comammox bacteria which can complete the entire nitrification process (ammonium to nitrate) ^{9,10}. AOB have been routinely used to better understand the dynamics of ammonia oxidation. From these studies, the integral membrane protein ammonia monooxygenase (AMO) was identified as the starting point of the ammonia oxidation pathway ^{11,12}. Ammonia oxidizing archaea (AOA), however, represent a much newer addition to the growing list of nitrifying species.

The first hints that nitrifiers existed outside of the bacterial domain came with the advent of metagenomics. A genomic survey of the Sargosso Sea was the first to report the presence of ammonia monooxygenase genes within archaea ¹³. At the time, two primary phyla of archaea were identified: Euryarchaeota and Crenarchaeota. Based on metagenomic results the newly found AMO genes were assigned to unknown crenarchaeal species. Shortly after, fosmid clones from soil samples identified two primary AMO genes, *amoA* and *amoB*, on the same clone as crenarchaeal 16S and 23S rRNA genes ¹⁴. Definitive proof of ammonia oxidation by archaea came later when the first AOA, *Nitrosopumilus maritimus*, was successfully

isolated from a marine aquarium ¹⁵. The difference in environmental lifestyles from other known Crenarchaeota, which were primarily thermophiles and hyperthermophiles, and the genome information from the previously studied sponge symbiont *Candidatus Cenarchaeum symbiosum* ¹⁶, which was phylogenetically within the mesophilic Crenarchaeota, eventually led to AOA being moved to their own archaeal phylum, Thaumarchaeota ¹⁷.

Proof of ammonia oxidation and identification of archaeal AMO genes led to many comparative studies on the abundance of AOB and AOA. Initial studies in bacteria revealed that acetylene inactivates the complex and leads to labeling of the AmoA subunit ^{12,18–20}. This led to initial speculation that the active site could be housed within this subunit and the use of *amoA* as a marker for this process in the environment ²¹. The distant relationship between archaeal and bacterial *amoA* allowed it to be used along with 16S rRNA to investigate the abundance of AOA vs AOB in an environmental context. In many cases, AOA were found to be much more environmentally abundant than AOB ^{22–26}. This opened the possibility that many nitrogen fluxes could be either wrongly attributed to bacteria or underestimated entirely due to the vast but previously unknown presence of AOA. It also led to invigorated efforts to understand how AOA function and co-exist with AOB. The sequencing of the *N. maritimus* genome began to give some insight into the uniqueness of AOA ²⁷. This was complemented by the isolation of the first AOA from soil, *Nitrosospheara viennensis* ^{28,29}, sequencing, enrichment, and isolation of other AOA genomes ^{30–35} and comparative analyses across the known phylum ³⁶. AOA are now known to exist in a large array of environments ^{37–43} including hot springs ^{44–47}, soil ^{28,30,33,48–51}, marine ^{52,53}, freshwater ⁵⁴, wastewater treatment systems ^{55,56}, and even the human skin ^{57,58}. The wide range of available genomic information for archaea also led to another recent reclassification of AOA, that were formerly termed Thaumarchaeota. These species now collectively make up the order Nitrososphaerales in the class Nitrososphaeria within the phylum of Thermoproteota ⁵⁹. Although a wealth of genomic information exists, AOA prove difficult to grow making biomass very difficult to obtain. Additionally, no genetic system has been developed for AOA. Therefore, much of what was initially known about their metabolism was based off of genomic reconstructions and comparison to AOB.

Since the discovery of ammonia oxidation in both bacteria and archaea many questions have arisen concerning their co-existence in many environments. This has led to some distinctions between the two groups based on their affinity for ammonia ^{60–62}, suggesting that AOA are better adapted to low ammonia environments, as e.g. in the mesopelagic ocean and

pristine soils, while AOB respond better in higher levels of ammonia, as e.g. in fertilized soils. Regardless of their affinities, all energy derived for their metabolism comes from the oxidation of ammonia to nitrite. The process begins with the oxidation of ammonia to the intermediate hydroxylamine via ammonia monooxygenase (AMO). Hydroxylamine has been shown to be an intermediate in both archaea⁶³ and bacteria⁶⁴. Its discovery in bacteria coupled with tracer studies of labeled oxygen led to the first suggestion that a monooxygenase was involved in ammonia oxidizers' metabolism¹¹. Unfortunately, very little functional or structural information is available for any AMO complex, bacterial or archaeal. Most of what is known about AMO is based off the related particulate methane monooxygenase (PMO) complex found in methanotrophs that oxidizes methane to methanol. Together with some hydrocarbon monooxygenases, both AMO and PMO belong to the family of copper-dependent membrane monooxygenases (CuMMO)⁶⁵.

The first crystal structure for PMO came in 2005⁶⁶ and revealed the first look at any copper membrane monooxygenase. The structure revealed a homotrimeric complex – three protomers each containing a subunit of PmoA, PmoB, and PmoC (Figure 1). The structure also revealed two metal centers in the PmoB subunit and one metal center coordinate by residues from PmoA. Subsequent crystal structures^{67–70} and a cryo-EM structure⁷¹ led to a refinement of these initial results with 1-2 metal centers in PmoB and one metal center within PmoC. These results, along with expression and activity studies of PmoB⁷², shifted the consensus towards PmoB as housing the active site for the complex. This later came into contention with new data pointing towards the PmoC metal center⁷³ supported by mutagenesis experiment in a related hydrocarbon monooxygenase (HMO)⁷⁴. The debate between which metal center houses the active site, and even how many metal centers exist, has yet to be definitively solved. Regardless, residues coordinating the most likely active site metal centers are conserved within AOA. Even so, differences still arise as the AmoC and AmoB subunits of archaea are significantly shorter than their bacterial counterparts and appear to be missing regions containing transmembrane helices and a soluble domain^{75,76}, all assumed to be important for the structure of methane monooxygenases based on current PMO data.

Aside from AMO, archaea, like bacteria, rely on oxygen as a terminal electron acceptor via cytochrome oxidase (also known as Complex IV)^{27,36}. However, the similarity ends there. Although the beginning and end of the ammonia oxidation pathway appears to be the same, what happens in between is much more ambiguous. In bacteria, earlier studies showed that hydroxylamine is further oxidized via hydroxylamine oxidoreductase (HAO)⁷⁷. HAO was

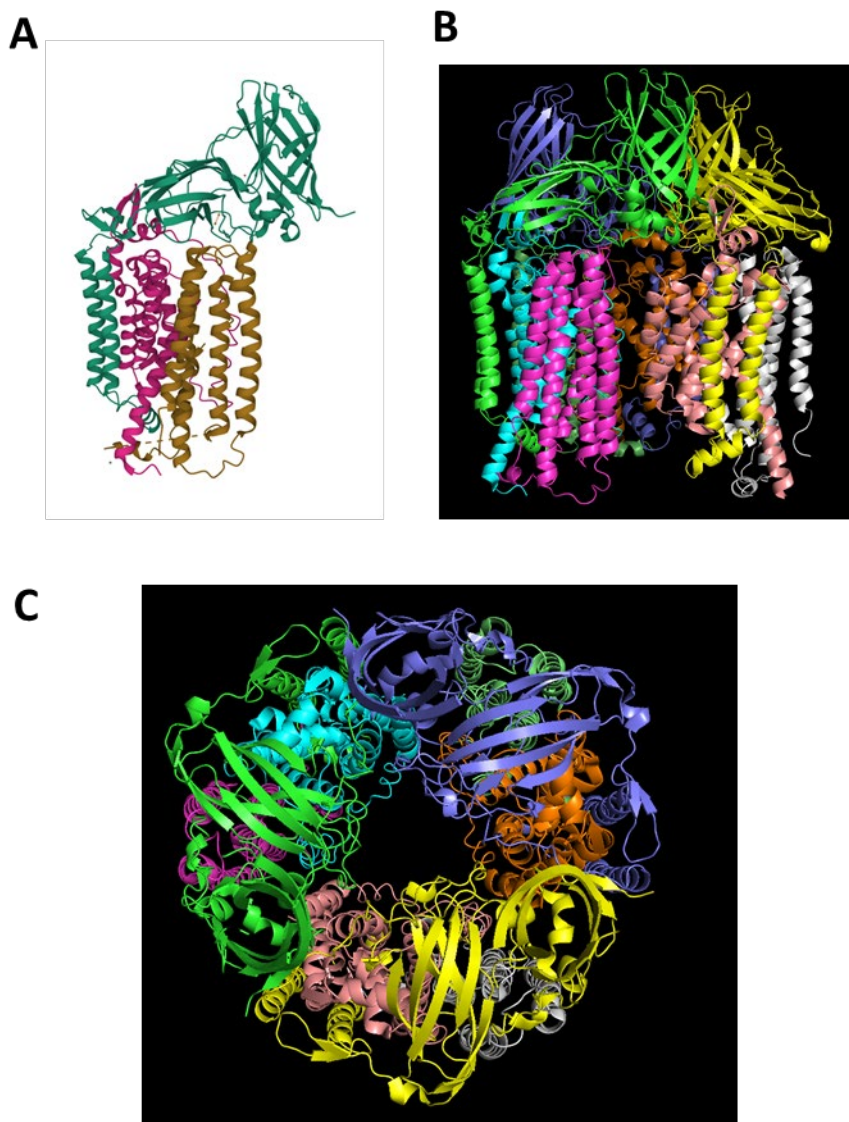


Figure 1: Crystal structure of *Methylococcus capsulatus* str. Bath. Images are shown of the first crystal structure of a particulate methane monooxygenase ⁶⁶. Structure (code: 1YEW) was downloaded from the Protein Data Bank (PDB) ⁷⁸. **A.)** Representation of a PMO protomer showing the configuration of PmoA (magenta), PmoB (green), and PmoC (brown). Image generated by Mol* ⁷⁹. **B.)** Representation of the fully assembled structure as a trimer. Subunits colored green, purple, and yellow at the top of the structure represent the soluble PmoB subunit from the three different protomers. One protomer can be visualized on the left side of the image with PmoA (cyan), PmoB (green), and PmoC (magenta) prominently shown. Image generated using PyMol ⁸⁰. **C.)** View of the trimeric PMO complex from the top. Soluble regions of PmoB are colored green, purple, and yellow. Image visualized using PyMol ⁸⁰.

long thought to produce primarily nitrite from this reaction (although NO production was also observed⁸¹), but recent evidence suggests that HAO actually produces nitric oxide (NO) and an unknown protein further converts NO to the end product of nitrite⁸². While the existence of HAO is well established in AOB, and even in some methanotrophs that can oxidize ammonia, no homolog for this crucial enzyme can be found in any of the AOA genomes^{27,36}. This implies that a novel enzyme oxidizes hydroxylamine produced by archaea. To complicate matters, a nitrite reductase (NirK), which classically converts nitrite to nitric oxide, has been observed to not only be present in most AOA genomes, but also abundantly transcribed within the environment^{83,84}. Although NirK from *N. viennensis* has shown the capacity for this reduction reaction (albeit with small turnover rates)⁸⁵, this seems counterintuitive as reduction of nitrite to NO would consume electrons produced by the energetics of ammonia oxidation. However, AOA are known to be sensitive to the availability of NO. The NO scavenger PTIO has been observed to arrest growth and metabolic activity of AOA while having little effect on bacteria^{86,87}. Other NO scavengers also have a greater effect on AOA than on AOB⁸⁸. These data together with the high expression of *nirK* suggests that NO plays a critical role in the metabolism of AOA.

Taken all together, two primary hypotheses have been proposed for the ammonia oxidation pathway in AOA: 1.) the ammonia oxidation pathway proceeds from hydroxylamine to nitric oxide via an unknown enzyme and nitric oxide is then oxidized to nitrite via NirK⁸⁹ (Figure 2A); 2.) the ammonia oxidation proceeds from hydroxylamine and nitric oxide via an unknown protein to directly produce nitrite. Nitrite is then reduced by NirK to provide the necessary NO for the reaction to take place⁸⁶ (Figure 2B).

Other differences between AOA and AOB also exist outside of the primary ammonia oxidation pathway. Bacteria rely on ubiquinone^{90,91} and iron containing cytochrome *c*^{92,93} to transport electrons, while AOA lack cytochrome *c*^{27,36} and have been shown to produce menaquinone in place of ubiquinone⁹⁴. The usage of menaquinone rather than ubiquinone likely also poses unique challenges and/or advantages due to its higher reduction potential when compared to ubiquinone. Although no cytochrome *c* proteins are identified, AOA genomes typically encode for several copper containing plastocyanins and blue copper proteins^{27,31,36}. It is assumed that AOA rely on copper containing proteins to transport electrons between complexes, effectively replacing cytochrome *c*. Copper is essential for both AOA and AOB as it is the primary metal in both the AMO complex and in the terminal oxidase (Complex IV). However, the use of plastocyanins and blue copper proteins in AOA would suggest a

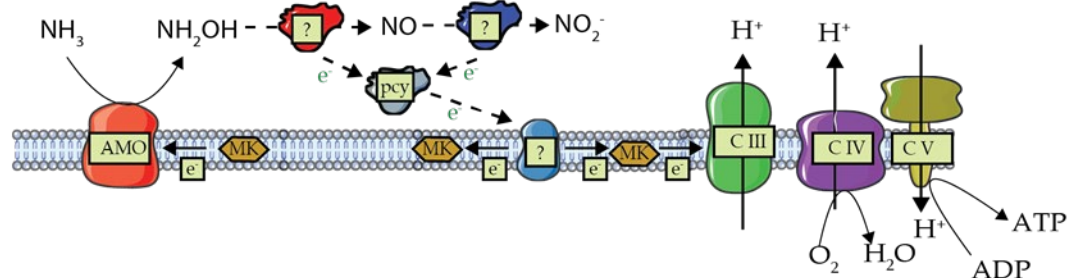
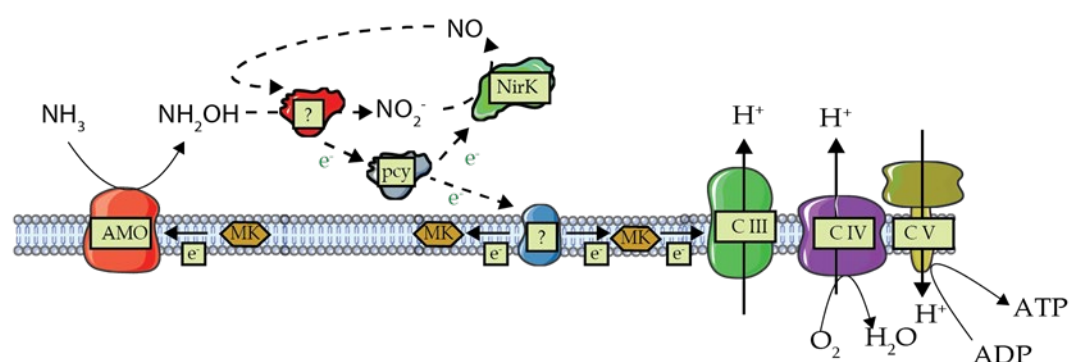
A**B**

Figure 2: Schematic representation of theorized ammonia oxidation pathway in archaea. **A.)** Ammonia is oxidized to hydroxylamine and follows two successive transformations to nitric oxide and nitrite by two unidentified proteins. Electrons are passed to plastocyanins and are transferred to the quinone pool via an unknown mechanism. Adapted from suggestions in Lancaster et al. (2018)⁸⁹. **B.)** Ammonia is oxidized to hydroxylamine. An unknown protein then converts hydroxylamine and nitric oxide to two molecules of nitrite. Nitric oxide is supplied by the reduction of nitrite via nitrite reductase (NirK). Electrons are passed to plastocyanins and are transferred to the quinone pool via an unknown mechanism. Adapted from Kozłowski et al. (2016)⁸⁶. Abbreviations: AMO, ammonia monooxygenase; e^- , electrons; MK, menaquinone; pcy, plastocyanin; C III, complex III (cytochrome *c* reductase); C IV, complex IV (cytochrome *c* oxidase); C V, complex V (ATP synthase); NirK, nitrite reductase. Note: Although CIII and CIV are named by their interaction with cytochrome *c*, AOA do not possess cytochrome *c*. Cytochrome *c* is hypothesized to be replaced by plastocyanins or copper containing proteins^{27,31,36}. Image produced using images from bioicons (<https://bioicons.com/>) with permissions from Servier (<https://smart.servier.com/>). Images are licensed under CC-BY 3.0 Unported <https://creativecommons.org/licenses/by/3.0/>.

higher reliance on copper in AOA and could indicate an important environmental difference between AOA and AOB. Although it is needed for many important proteins in AOA, too much copper is toxic for many organisms. While copper limiting thresholds have been found for two AOA^{95,96}, how AOA obtain this copper from the environment and regulate its abundance within the cell is not well understood.

While the archaeal ammonia oxidation pathway is ambiguous, there are many additional anomalies in its physiology. Regardless of how AOA obtain electrons, it is not clear as to how these electrons enter the cell as reducing equivalents to facilitate cellular processes. Based off of reverse electron flow in iron oxidizing bacteria⁹⁷, it has been suggested that electrons enter the cell via Complex I running in a reverse direction^{27,31}. However, even if this were the case it would not be clear what the electrons are transferred to as all archaea are missing genes for the subunits NuoEFG that house the interaction site of NADH⁹⁸. Electrons would therefore need to be put onto an intermediate reducing equivalent before being transferred to NAD(P)H for cellular activity. One possible candidate could be F420, which is predicted to occur in all AOA³⁶ and has been observed in high abundance in *Nitrososphaera gargensis* based off of its autofluorescence³¹.

Reducing equivalents within the cell would be needed for the buildup of cell biomass, but they would also play a crucial role in the AOA carbon fixation cycle, the 3-hydroxypropionate/4-hydroxybutyrate cycle. The cycle relies on acetyl-CoA to begin and produces succinyl-CoA in the first half of the cycle through subsequent reactions including two carboxylations using bicarbonate and a 3-hydroxypropionate intermediate⁹⁹. Acetyl-CoA is then replenished in the second half of the cycle using 4-hydroxybutyrate as an intermediate⁹⁹. This cycle is conserved across AOA³⁶ and has been shown to be unique as being the most efficient carbon fixation cycle¹⁰⁰. This is due to proteins playing multiple roles within the cycle as well as usage of ATP to ADP rather than the more energy intensive conversion to AMP. Most of what is known about this cycle is based off work in *Metallosphaera sedula*. While individual proteins have also been well characterized in *N. maritimus*^{100,101}, how the cycle connects to the rest of metabolism within AOA has not been thoroughly explored. However, based on isotope studies from *M. sedula*, it is assumed that the carbon fixation cycle connects with the central metabolism via succinate and/or succinyl-CoA rather than acetyl-CoA¹⁰². The TCA cycle is then assumed to run in the classical forward (oxidation) direction with malic enzyme and/or phosphoenolpyruvate carboxykinase facilitating the main connection point between the C-3 and C-4 carbon pools. Carbon would then proceed through

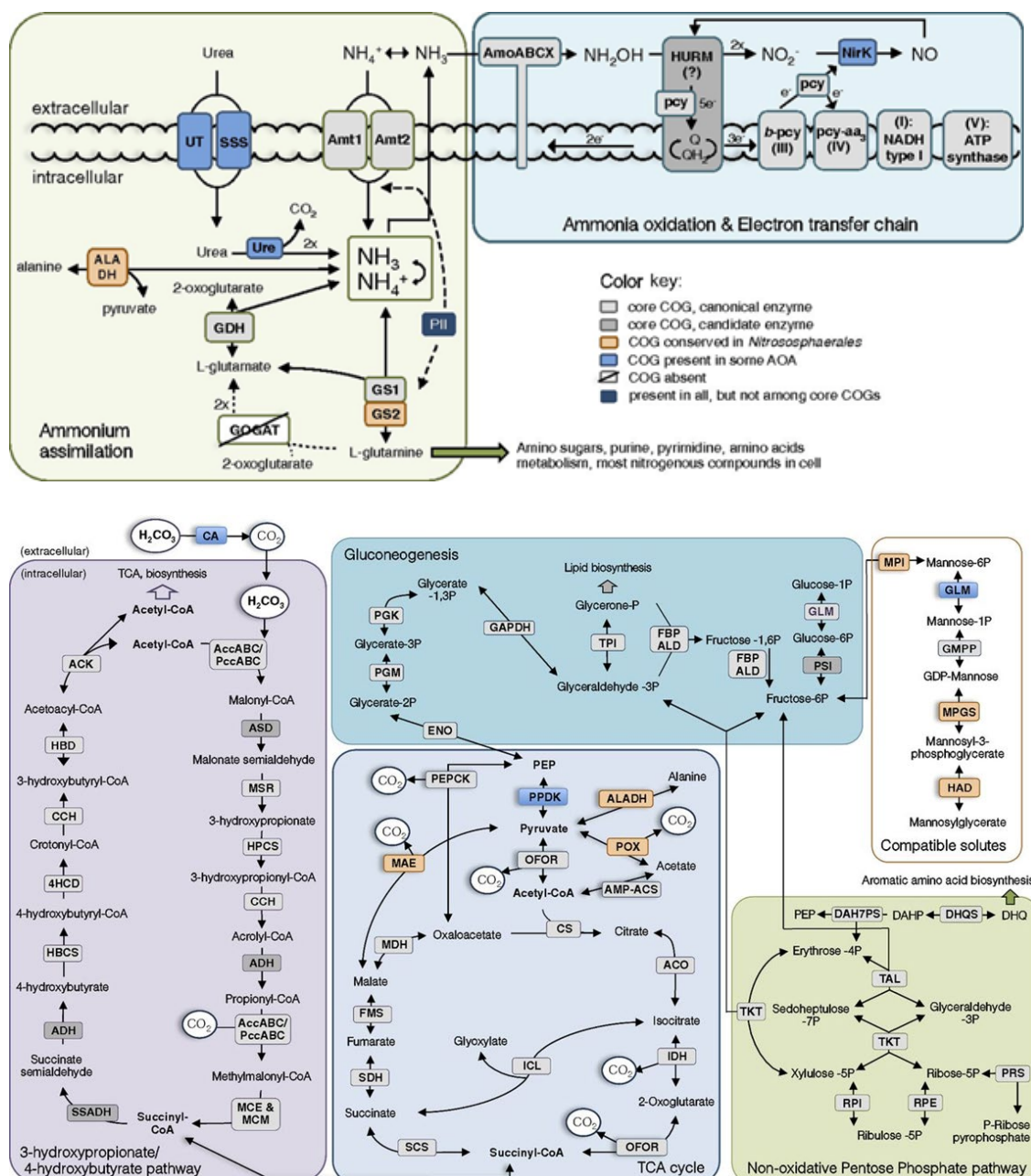


Figure 3: General overview of the metabolism of *N. viennensis*. Figures from Kerou et al. (2016) ³⁶. All energy for biomass production and growth comes from the oxidation of ammonia to nitrite ³⁶. Energy and reducing equivalents from the ammonia oxidation pathway are used to fix inorganic carbon in the form of bicarbonate via the 3-hydroxypropionate/4-hydroxybutyrate carbon fixation pathway ^{36,100}. Carbon is theorized to enter the central carbon metabolism via succinyl-CoA where it can then flow through the oxidizing tricarboxylic acid (TCA) cycle and up through gluconeogenesis via the C-3/C-4 connections from malate to pyruvate or oxaloacetate to phosphoenolpyruvate. Carbon from gluconeogenesis can be used to make nucleotides after passing through the non-oxidative pentose phosphate pathway ³⁶. Details on proteins and abbreviations can be found in Kerou et al. (2016).

gluconeogenesis and on to the non-oxidative pentose phosphate pathway (PPP). Connections in the TCA cycle would then be able to supply all necessary intermediates for amino acid synthesis while carbon funneled to the PPP would be used for synthesis of nucleotides ³⁶. Carbon for the isoprenoid lipids unique to archaea would theoretically pull directly from the carbon fixation pathway itself. In addition to these carbon sinks, AOA have also been shown to be able to store fixed carbon in the form of poly-hydroxy-butyrate (PHB) ³¹, which would also be pulled directly from the carbon fixation cycle. As a strict autotroph utilizing a restrictive lithotrophic lifestyle, the coordination of carbon resources within the cell is likely tightly controlled (Figure 3).

Although the carbon fixation cycle of AOA is now well characterized, there were also initial suggestions of a mixotrophic metabolism in AOA. After isolation of the first soil AOA, *N. viennensis*, it was discovered that the addition of α -keto organic acids, such as pyruvate, significantly improved growth ^{28,29}. Mixotrophy was suggested as a possible explanation. However, this was later debunked when the addition of pyruvate was shown to dramatically reduce the amount of reactive oxygen species (ROS) within the culture through an abiotic interaction with hydrogen peroxide ¹⁰³. This was particularly advantageous to *N. viennensis* as it lacks a catalase to detoxify hydrogen peroxide. Although Nano-SIMS studies with labeled pyruvate did reveal some incorporation ²⁸, it was concluded that this was solely due to the fixation of labeled CO₂ released by the abiotic pyruvate mechanism. This is also supported by a lack of heterotrophic genes found within the core metabolism of AOA. Later comparative analyses found that the lack of a catalase was not universal in soil AOA but was common in marine strains ¹⁰⁴. Therefore, addition of pyruvate or catalase to AOA cultures that do not encode for a catalase has been observed to help growth, suggesting a large amount of ROS production ^{52,53}. Where exactly the ROS is produced is not clear, but it is a known problem in aerobic organisms as a by-product of general electron transport chain function.

A lot has been learned about AOA since their discovery but many of the topics discussed above have yet to be fully resolved. The many open questions in AOA physiology are difficult to explore due to three primary reasons: 1.) cultured representatives are difficult to grow and maintain in pure culture and representatives are missing from abundant environmental clades ³⁹, 2.) biomass is difficult to obtain in large amounts for biochemical work, and 3.) the lack of a genetic system prevents identification and functional studies of crucial genes. However, the increasing availability of omics technologies and the vast amount

of knowledge collected from metagenomic studies allows for new and creative ways to explore these topics.

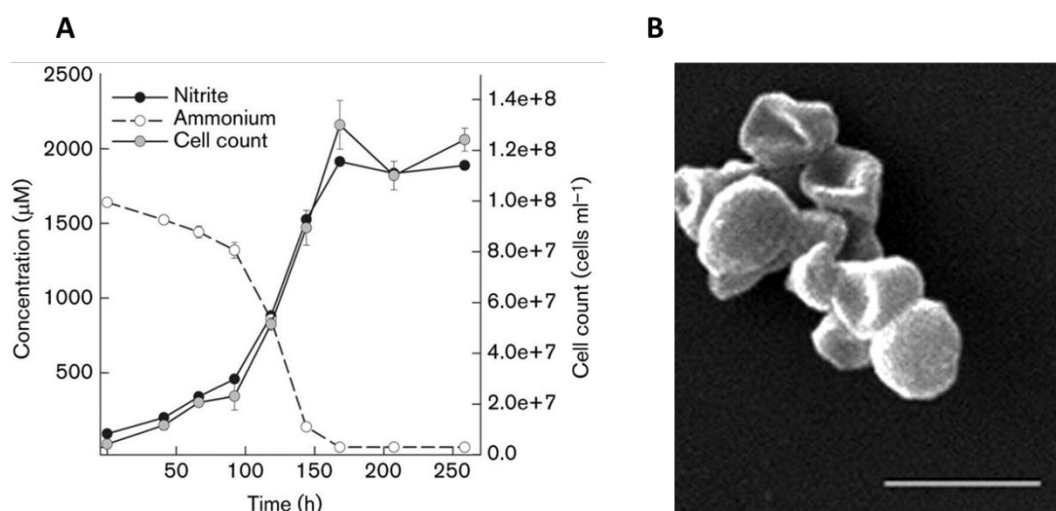


Figure 4: Typical growth curve (A) and scanning electron microscope image (B) of *N. viennensis*. (A) Growth curve from Stieglmeier et al. (2014)²⁹. (B) Image is from Tournai et al. (2011)²⁸. Scale bar represents 1 μm .

This thesis is aimed at expanding the knowledge of AOA physiology while using *N. viennensis* as a model organism (Figure 4). *N. viennensis* represents one of the few pure cultures of AOA and contains a larger genome that helps to account for the diversity of genomic potential that is seen in many AOA species. *N. viennensis* grows via an almost stoichiometric conversion of ammonia to nitrite. Therefore, production of nitrite is commonly used as a proxy for growth as cultures are not dense enough to obtain optical density measurements. Since its isolation, growth conditions have been optimized^{28,29} and are reproducible making it ideal to experiment with in the laboratory. Reactor growth in continuous conditions is also possible (Melcher et al., unpublished). While this is still a tedious process that can produce at best 500-700 mg in two weeks from two 2-liter-bioreactors, it allows for the harvest of biomass that can be used for biochemical work. In particular, methods can be utilized to investigate the membrane protein components as shown for other archaea¹⁰⁵, bacteria involved in the nitrogen cycle¹⁰⁶, and copper monooxygenase containing organisms¹⁰⁷. The work presented here was only possible due to the stable laboratory growth of *N. viennensis* and the production of

biomass. In all cases methods were tested and developed to take full advantage of the limited amount of biomass that can be reasonably obtained. In lieu of a genetic system, this work was also greatly complemented and guided by comparisons of gene presence both in and out of ammonia oxidizing archaea. The data produced during the course of this thesis in combination with known environmental and genomic patterns has helped reach a better understanding of how these intriguing microorganisms function.

Major Goals and Contributions

Chapter II

Unexpected Complexity of the Ammonia Monooxygenase in Archaea

Logan H. Hodgskiss, Michael Melcher, Melina Kerou, Weiqiang Chen, Rafael I. Ponce-Toledo, Savvas N. Savvides, Stefanie Wienkoop, Markus Hartl, & Christa Schleper

Status: Prepared for submission.

Major Goals

The primary goal of this chapter was to gain a better understanding of the archaeal AMO complex through biochemical methods, in particular to identify subunits of the protomer. This included showing the participation of AmoA, AmoB, and AmoC (predicted from homology to bacterial counterparts) and verifying the inclusion of AmoX (predicted based off of metagenomic studies). Methods were developed using the soil model organism *Nitrososphaera viennensis* and were then applied to the distantly related *Nitrosocaldus cavascurensis* with the hope of gaining knowledge that would be applicable to all ammonia oxidizing archaea. Membrane proteins were extracted and analyzed in their native form on Blue Native PAGE gels. Proteomics of cut bands were used to identify presence of the AMO complex. Correlation in BN-PAGE gels and syntenic analysis of genomes were utilized to confirm archaeal AMO subunits. Structural information was also gathered from in-gel cross-linking, gene analysis, and predictive modelling.

Contributions

I was responsible for most of the experimental design, laboratory work, and writing the first draft of this chapter. This included optimizing the membrane protein extraction from harvested biomass, optimizing Blue Native PAGE methods that could be applied to *N. viennensis*, and optimization of SDS-Tricine-PAGE from BN-PAGE cut-outs. The application of these methods to the only (based on biomass constraints) *N. cavascurensis* gel was also performed by me. Additionally, I was also responsible for the analysis and interpretation of all proteomic data obtained from BN-PAGE cut-outs (with the exception of cross-linked samples), genome collection and analysis of selected bacterial and archaea species, re-analysis of previous transcriptomic data, a structural search of the protein coding genes of *N. viennensis*, and creation of a phylogenetic tree of *amoC* genes within ammonia oxidizing archaea.

Chapter III

Genome Wide Transcriptomic Analysis of the Soil Ammonia Oxidizing Archaeon *Nitrososphaera viennensis* upon Exposure to Copper Limitation

Carolina Reyes, Logan H. Hodgskiss, Melina Kerou, Thomas Pribasnig, Sophie S. Abby,
Barbara Bayer, Stephan M. Kraemer, Christa Schleper

Status: Published in the ISME Journal (2020) 14:2659–2674.

Major Goals

The third chapter of this thesis focused on gaining a better understanding of how *N. viennensis* responds to copper limitation. *N. viennensis* was subjected to growth under copper limiting and copper replete conditions during which samples were taken for transcriptomic analysis. Information gained from this study was intended to further the understanding of how ammonia oxidizing archaea accumulate and regulate the required copper needed for their primary metabolism. Results from this chapter were also intended to be compared against a

similar study focusing on the marine AOA *Nitrosopumilus maritimus*. The comparison would be used to highlight differences within AOA based on the very different environments that they inhabit and the differences that can be seen based on comparative genomics. This study also represented the first transcriptomic study using RNA-seq for a soil AOA and was used to gain a better understanding of highly transcribed, and therefore likely crucial, genes within soil AOA.

Contributions

For this chapter, I was responsible for the transcriptomic data analysis of the RNA-seq data. This included filtering and sorting for high quality mRNA reads and eventual differential expression analysis between the copper replete and copper limited conditions. I was therefore also responsible for the data organization. In addition, I contributed the physiological interpretation of the transcriptomic data with respect to the electron transport chain and primary carbon pathways and identified differences between *N. viennensis* (soil) and *N. maritimus* (marine) responses. After identifying top up and down expressed genes, I utilized this information to carry out the promoter analysis. Combined extraction of RNA and proteins from frozen samples and subsequent proteomic analysis to verify trends seen in the transcriptome were also carried out by me.

Chapter IV

Carbon Limitation of the Ammonia Oxidizing Archaeon *Nitrososphaera viennensis* Indicates Routes of Cellular Carbon Flow and Cellular Stress Response

Logan H. Hodgskiss, Melina Kerou, Barbara Bayer, Andreas Maier, Wolfram Weckwerth,
Thomas Naegele, Christa Schleper

Status: Draft manuscript.

Major Goals

The final study in this thesis aimed to uncover patterns within the cell that could further the understanding of the general carbon flow in AOA and how these organisms respond to stress in the environment.

To achieve this, an accurate thermodynamic model for *N. viennensis* (and AOA as a whole) was to be constructed and used to predict minimal carbon requirements. Once predicted the organism was to be grown under different carbon concentrations and subjected to a combined proteomic and metabolomic approach to reveal responses in the primary carbon metabolism and major metabolisms.

Contributions

In the final chapter, I was responsible for calculation of the thermodynamic model, a large part of the experimental design, and analysis of results. This included growing the cultures of *N. viennensis* under varying carbon concentrations and executing a combined protein and metabolite extraction. At the time of harvest I was also responsible for taking samples for inorganic carbon analysis from the liquid phase, gas samples for gas chromatography, and samples for DNA extraction. After extracting proteins and metabolites, I prepped samples for subsequent LC-MS/MS and GC-MS respectively. Raw metabolomic files were manually curated by me and all analyses of proteomic, metabolomic, and combined approaches were also carried out by me. In addition, I was the primary writer for the draft manuscript.

References

1. Stein, L. Y. & Klotz, M. G. The nitrogen cycle. *Curr. Biol.* **26**, R94–R98 (2016).
2. M Kuypers, M. M., Marchant, H. K. & Kartal, B. The microbial nitrogen-cycling network. *Nat. Publ. Gr.* (2018) doi:10.1038/nrmicro.2018.9.
3. Erisman, J. W., Sutton, M. A., Galloway, J., Klimont, Z. & Winiwarter, W. How a century of ammonia synthesis changed the world. *Nat. Geosci.* **1**, (2008).

4. Gruber, A. N. & Galloway, J. N. An earth-system perspective of the global nitrogen cycle : with humans having an increasing impact on the planet , the interactions between the nitrogen cycle , the carbon cycle and climate are expected to become an increasingly important determinant of t. (2008).
5. Steffen, W. *et al.* Planetary Boundaries : Exploring the Safe Operating Space for Humanity. (2009).
6. Robertson, G. P. *et al.* Nitrogen-climate interactions in US agriculture. *Biogeochemistry* **114**, 41–70 (2013).
7. Norton, J. & Ouyang, Y. Controls and adaptive management of nitrification in agricultural soils. *Front. Microbiol.* **10**, 1–18 (2019).
8. Winogradsky, S. Recherches sur les organismes de la nitrification. *Ann Inst Pateur* **4:213-231.**, (1890).
9. Daims, H. *et al.* Complete nitrification by *Nitrospira* bacteria. *Nature* **528**, 504 (2015).
10. van Kessel, M. A. H. J. *et al.* Complete nitrification by a single microorganism. *Nature* (2015) doi:10.1038/nature16459.
11. Hollocher, T. C., Tate, M. E. & Nicholas, D. J. Oxidation of ammonia by *Nitrosomonas europaea*. Definite ¹⁸O-tracer evidence that hydroxylamine formation involves a monooxygenase. *J. Biol. Chem.* **256**, 10834–10836 (1981).
12. Hyman, M. R. & Wood, P. M. Suicidal inactivation and labelling of ammonia monooxygenase by acetylene. *Biochem. J.* **227**, 719–725 (1985).
13. Venter, J. C. *et al.* Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science* (80-.). **304**, 66–74 (2004).
14. Treusch, A. H. *et al.* Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ. Microbiol.* **7**, 1985–1995 (2005).
15. Könneke, M. *et al.* Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**, 543–546 (2005).
16. Schleper, C. *et al.* Genomic analysis reveals chromosomal variation in natural populations of the uncultured psychrophilic archaeon *Cenarchaeum symbiosum*. *J.*

- Bacteriol.* **180**, 5003–5009 (1998).
17. Brochier-Armanet, C., Boussau, B., Gribaldo, S. & Forterre, P. Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* **6**, 245–252 (2008).
 18. Prior, S. D. & Dalton, H. Acetylene as a suicide substrate and active site probe for methane monooxygenase from *Methylococcus capsulatus* (Bath). **29**, 105–109 (1985).
 19. Semrau, J. D. *et al.* Particulate Methane Monooxygenase Genes in Methanotrophs. **177**, 3071–3079 (1995).
 20. Holmes, A. J., Costello, A., Lidstrom, M. E. & Murrell, J. C. Evidence that particulate methane monooxygenase and ammonia monooxygenase may be evolutionarily related. **132**, 203–208 (1995).
 21. Rotthauwe, J. & Witzel, K. The Ammonia Monooxygenase Structural Gene *amoA* as a Functional Marker : Molecular Fine-Scale Analysis of Natural Ammonia-Oxidizing Populations. **63**, 4704–4712 (1997).
 22. Adair, K. L. & Schwartz, E. Evidence that ammonia-oxidizing archaea are more abundant than ammonia-oxidizing bacteria in semiarid soils of northern Arizona, USA. *Microb. Ecol.* **56**, 420–426 (2008).
 23. Hollibaugh, J. T., Gifford, S., Sharma, S., Bano, N. & Moran, M. A. Metatranscriptomic analysis of ammonia-oxidizing organisms in an estuarine bacterioplankton assemblage. *ISME J.* **5**, 866–878 (2011).
 24. Karner, M. B., Delong, E. F. & Karl, D. M. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**, 507–510 (2001).
 25. Leininger, S. *et al.* Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**, 806–809 (2006).
 26. Nicol, G. W., Leininger, S., Schleper, C. & Prosser, J. I. The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. *Environ. Microbiol.* **10**, 2966–2978 (2008).
 27. Walker, C. B. *et al.* *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc. Natl.*

- Acad. Sci.* **107**, (2010).
28. Tournai, M. *et al.* Nitrososphaera viennensis, an ammonia oxidizing archaeon from soil. *Proc. Natl. Acad. Sci. USA* **108**, 8420–8425 (2011).
 29. Stieglmeier, M. *et al.* Nitrososphaera viennensis gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* **64**, 2738–2752 (2014).
 30. Kim, B. K. *et al.* Genome sequence of an ammonia-oxidizing soil archaeon, ‘Candidatus Nitrosoarchaeum koreensis’ MY1. *J. Bacteriol.* **193**, 5539–5540 (2011).
 31. Spang, A. *et al.* The genome of the ammonia-oxidizing candidatus nitrososphaera gargensis: Insights into metabolic versatility and environmental adaptations. *Environ. Microbiol.* **14**, 3122–3145 (2012).
 32. Lehtovirta-Morley, L. E. *et al.* Identifying Potential Mechanisms Enabling Acidophily in the Ammonia-Oxidizing Archaeon “Candidatus Nitrosotalea devanaterrea”. *Appl. Environ. Microbiol.* **82**, 2608–2619 (2016).
 33. Zhalnina, K. V. *et al.* Genome sequence of Candidatus nitrososphaera evergladensis from group I.1b enriched from everglades soil reveals novel genomic features of the ammonia-oxidizing archaea. *PLoS One* **9**, 21–23 (2014).
 34. Santoro, A. E. *et al.* Genomic and proteomic characterization of “Candidatus Nitrosopelagicus brevis”: an ammonia-oxidizing archaeon from the open ocean. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 1173–8 (2015).
 35. Hatzenpichler, R. *et al.* A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proc. Natl. Acad. Sci.* **105**, 2134–2139 (2008).
 36. Kerou, M. *et al.* Proteomics and comparative genomics of Nitrososphaera viennensis reveal the core genome and adaptations of archaeal ammonia oxidizers. *Proc. Natl. Acad. Sci.* **113**, E7937–E7946 (2016).
 37. Pester, M., Schleper, C. & Wagner, M. The Thaumarchaeota: An emerging view of their phylogeny and ecophysiology. *Curr. Opin. Microbiol.* **14**, 300–306 (2011).
 38. Nicol, G. W. & Schleper, C. Ammonia-oxidising Crenarchaeota: important players in the nitrogen cycle? *Trends Microbiol.* **14**, 207–212 (2006).

39. Alves, R. J. E., Minh, B. Q., Urich, T., Von Haeseler, A. & Schleper, C. Unifying the global phylogeny and environmental distribution of ammonia-oxidising archaea based on amoA genes. (2018) doi:10.1038/s41467-018-03861-1.
40. Pester, M. *et al.* AmoA-based consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of amoA genes from soils of four different geographic regions. *Environ. Microbiol.* **14**, 525–539 (2012).
41. Hatzenpichler, R. Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl. Environ. Microbiol.* **78**, 7501–7510 (2012).
42. Prosser, J. I. & Nicol, G. W. Relative contributions of archaea and bacteria to aerobic ammonia oxidation in the environment. *Environ. Microbiol.* **10**, 2931–2941 (2008).
43. Stahl, D. A. & De La Torre, J. R. Physiology and diversity of ammonia-oxidizing archaea. *Annu. Rev. Microbiol.* **66**, 83–101 (2012).
44. Abby, S. S. *et al.* Candidatus Nitrosocaldus cavascurensis, an ammonia oxidizing, extremely thermophilic archaeon with a highly mobile genome. *Front. Microbiol.* **9**, 1–19 (2018).
45. Daebeler, A. *et al.* Cultivation and genomic analysis of ‘Candidatus Nitrosocaldus islandicus,’ an obligately thermophilic, ammonia-oxidizing thaumarchaeon from a hot spring biofilm in Graendalur valley, Iceland. *Front. Microbiol.* **9**, 1–16 (2018).
46. Luo, Z.-H. *et al.* Genomic Insights of “Candidatus Nitrosocaldaceae” Based on Nine New Metagenome-Assembled Genomes, Including “Candidatus Nitrosothermus” Gen Nov. and Two New Species of “Candidatus Nitrosocaldus”. *Front. Microbiol.* **11**, (2021).
47. Reigstad, L. J. *et al.* Nitrification in terrestrial hot springs of Iceland and Kamchatka. *FEMS Microbiol. Ecol.* **64**, 167–174 (2008).
48. Lehtovirta-Morley, L. E. *et al.* Isolation of ‘Candidatus Nitrosocosmicus franklandus’, a novel ureolytic soil archaeal ammonia oxidiser with tolerance to high ammonia concentration. *FEMS Microbiol. Ecol.* **92**, 1–10 (2016).
49. Alves, R. J. E. *et al.* Ammonia oxidation by the arctic terrestrial thaumarchaeote candidatus nitrosocosmicus arcticus is stimulated by increasing temperatures. *Front. Microbiol.* **10**, (2019).

50. Lehtovirta-Morley, L. E., Stoecker, K., Vilcinskas, A., Prosser, J. I. & Nicol, G. W. Cultivation of an obligate acidophilic ammonia oxidizer from a nitrifying acid soil. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15892–15897 (2011).
51. Jung, M. Y. *et al.* A mesophilic, autotrophic, ammonia-oxidizing archaeon of thaumarchaeal group I.1a cultivated from a deep oligotrophic soil horizon. *Appl. Environ. Microbiol.* **80**, 3645–3655 (2014).
52. Bayer, B. *et al.* *Nitrosopumilus adriaticus* sp. nov. and *Nitrosopumilus piranensis* sp. nov., two ammonia-oxidizing archaea from the Adriatic Sea and members of the class Nitrososphaeria. *Int. J. Syst. Evol. Microbiol.* **69**, 1892–1902 (2019).
53. Qin, W. *et al.* *Nitrosopumilus maritimus* gen. nov., sp. nov., *Nitrosopumilus cobalaminigenes* sp. nov., *Nitrosopumilus oxycinae* sp. nov., and *Nitrosopumilus ureiphilus* sp. nov., four marine ammonia-oxidizing archaea of the phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* (2017) doi:10.1099/ijsem.0.002416.
54. Sauder, L. A., Engel, K., Lo, C.-C., Chain, P. & Neufeld, J. D. “Candidatus Nitrosotenuis aquarius,” an Ammonia-Oxidizing Archaeon from a Freshwater Aquarium Biofilter. *Appl. Environ. Microbiol.* **84**, e01430-18 (2018).
55. 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).
56. Li, Y. *et al.* A novel ammonia-oxidizing archaeon from wastewater treatment plant: Its enrichment, physiological and genomic characteristics. *Sci. Rep.* **6**, 1–11 (2016).
57. Probst, A. J., Auerbach, A. K. & Moissl-Eichinger, C. Archaea on Human Skin. *PLoS One* **8**, (2013).
58. Moissl-Eichinger, C. *et al.* Human age and skin physiology shape diversity and abundance of Archaea on skin. *Sci. Rep.* **7**, 1–10 (2017).
59. Rinke, C. *et al.* A standardized archaeal taxonomy for the Genome Taxonomy Database. *Nat. Microbiol.* **6**, 946–959 (2021).
60. Dimitri Kits, K. *et al.* Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature* **549**, 269–272 (2017).

61. Jung, M. Y. *et al.* Ammonia-oxidizing archaea possess a wide range of cellular ammonia affinities. *ISME J.* **16**, 272–283 (2022).
62. Martens-Habben, 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).
63. Vajrala, N. *et al.* Hydroxylamine as an intermediate in ammonia oxidation by globally abundant marine archaea. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1006–1011 (2013).
64. Dua, R. D., Bhandari, B. & Nicholas, D. J. D. Stable isotope studies on the oxidation of ammonia to hydroxylamine by nitrosomonas europaea. *FEBS Lett.* **106**, 401–404 (1979).
65. Khadka, R. *et al.* Evolutionary History of Copper Membrane Monooxygenases. (2018) doi:10.3389/fmicb.2018.02493.
66. Lieberman, R. L. & Rosenzweig, A. C. Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. *Nature* **434**, 177–182 (2005).
67. Hakemian, A. S. *et al.* The metal centres of particulate methane mono-oxygenase from *Methylosinus trichosporium* OB3b. *Biochem. Soc. Trans.* **36**, 1134–1137 (2008).
68. Ro, S. Y. *et al.* From micelles to bicelles: Effect of the membrane on particulate methane monooxygenase activity. *J. Biol. Chem.* **293**, 10457–10465 (2018).
69. Sirajuddin, S. *et al.* Effects of zinc on particulate methane monooxygenase activity and structure. *J. Biol. Chem.* **289**, 21782–21794 (2014).
70. Smith, S. M. *et al.* Crystal structure and characterization of particulate methane monooxygenase from *Methylocystis* species strain M. *Biochemistry* **50**, 10231–10240 (2011).
71. Chang, W. H. *et al.* Copper Centers in the Cryo-EM Structure of Particulate Methane Monooxygenase Reveal the Catalytic Machinery of Methane Oxidation. *J. Am. Chem. Soc.* **143**, 9922–9932 (2021).
72. Culpepper, M. A. & Rosenzweig, A. C. Architecture and active site of particulate methane monooxygenase. *Crit. Rev. Biochem. Mol. Biol.* **47**, 483–492 (2012).

73. Ross, M. O. *et al.* Particulate methane monooxygenase contains only mononuclear copper centers. *Science* (80-.). **364**, 566–570 (2019).
74. Liew, E. F., Tong, D., Coleman, N. V. & Holmes, A. J. Mutagenesis of the hydrocarbon monooxygenase indicates a metal centre in subunit-C, and not subunit-B, is essential for copper-containing membrane monooxygenase activity. *Microbiol. (United Kingdom)* **160**, 1267–1277 (2014).
75. Tolar, B. B. *et al.* Integrated structural biology and molecular ecology of N-cycling enzymes from ammonia-oxidizing archaea. *Environ. Microbiol. Rep.* **9**, 484–491 (2017).
76. Lawton, T. J., Ham, J., Sun, T. & Rosenzweig, A. C. Structural conservation of the B subunit in the ammonia monooxygenase/particulate methane monooxygenase superfamily. *Bone* **23**, 1–7 (2014).
77. Andersson, K. K. & Hooper, A. B. O₂ and H₂O are each the source of one O in NO₂ produced from NH₃ by Nitrosomonas: 15N-NMR evidence. *FEBS Lett.* **164**, 236–240 (1983).
78. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
79. Sehnal, D. *et al.* Mol*Viewer: Modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Res.* **49**, W431–W437 (2021).
80. Schrodinger LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).
81. Hooper, A. B. & Terry, K. R. Hydroxylamine oxidoreductase of Nitrosomonas. Production of nitric oxide from hydroxylamine. *BBA - Enzymol.* **571**, 12–20 (1979).
82. Caranto, J. D. & Lancaster, K. M. Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase. *Proc. Natl. Acad. Sci.* **114**, 8217–8222 (2017).
83. Carini, P., Dupont, C. L. & Santoro, A. E. Patterns of thaumarchaeal gene expression in culture and diverse marine environments. *Environ. Microbiol.* **20**, 2112–2124 (2018).
84. Lund, M. B., Smith, J. M. & Francis, C. A. Diversity, abundance and expression of nitrite reductase (nirK)-like genes in marine thaumarchaea. *ISME J.* **6**, 1966–1977

- (2012).
85. Kobayashi, S. *et al.* Nitric oxide production from nitrite reduction and hydroxylamine oxidation by copper-containing dissimilatory nitrite reductase (Nirk) from the aerobic ammonia-oxidizing archaeon, *Nitrososphaera viennensis*. *Microbes Environ.* **33**, 428–434 (2018).
 86. Kozłowski, 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.* 1–10 (2016) doi:10.1038/ismej.2016.2.
 87. Martens-Habbena, W. *et al.* The production of nitric oxide by marine ammonia-oxidizing archaea and inhibition of archaeal ammonia oxidation by a nitric oxide scavenger. *Environ. Microbiol.* **17**, 2261–2274 (2015).
 88. Sauder, L. A., Ross, A. A. & Neufeld, J. D. Nitric oxide scavengers differentially inhibit ammonia oxidation in ammonia-oxidizing archaea and bacteria. *FEMS Microbiol. Lett.* **363**, fnw052 (2016).
 89. Lancaster, K. M., Caranto, J. D., Majer, S. H. & Smith, M. A. Alternative Bioenergy: Updates to and Challenges in Nitrification Metalloenzymology. *Joule* (2018) doi:10.1016/j.joule.2018.01.018.
 90. Hooper, A. B., Erickson, R. H. & Terry, K. R. Electron transport systems of *Nitrosomonas*: isolation of a membrane-envelope fraction. *J. Bacteriol.* **110**, 430–438 (1972).
 91. Whittaker, M., Bergmann, D., Arciero, D. & Hooper, A. B. Electron transfer during the oxidation of ammonia by the chemolithotrophic bacterium *Nitrosomonas europaea*. *Biochim. Biophys. Acta - Bioenerg.* **1459**, 346–355 (2000).
 92. Yamanaka, T. & Shinra, M. Cytochrome c-552 and Cytochrome c-554 Derived from *Nitrosomonas europaea*. *J. Biochem.* **75**, 1265–1273 (1974).
 93. Arciero, D. M., Hooper, A. B. & Balny, C. Spectroscopic and Rapid Kinetic Studies of Reduction of Cytochrome c554 by Hydroxylamine Oxidoreductase from *Nitrosomonas europaea*. *Biochemistry* **30**, 11466–11472 (1991).
 94. Elling, F. J. *et al.* Respiratory quinones in Archaea: Phylogenetic distribution and

- application as biomarkers in the marine environment. *Environ. Microbiol.* **18**, 692–707 (2016).
95. Reyes, C. *et al.* Copper limiting threshold in the terrestrial ammonia oxidizing archaeon *Nitrososphaera viennensis*. *Res. Microbiol.* **171**, 134–142 (2020).
 96. Qin, W. *et al.* Stress response of a marine ammonia-oxidizing archaeon informs physiological status of environmental populations. *Nat. Publ. Gr.* **doi**, (2017).
 97. Elbehti, A., Brasseur, G. & Lemesle-Meunier, D. First evidence for existence of an uphill electron transfer through the bc1 and NADH-Q oxidoreductase complexes of the acidophilic obligate chemolithotrophic ferrous ion-oxidizing bacterium *Thiobacillus ferrooxidans*. *J. Bacteriol.* **182**, 3602–3606 (2000).
 98. Friedrich, T. & Scheide, D. The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *Fed. Eur. Biochem. Soc.* (2000).
 99. Berg, I. A., Kockelkorn, D., Buckel, W. & Fuchs, G. A 3-Hydroxypropionate 4-Hydroxybutyrate Pathway in Archaea. *Science (80-.)*. **218**, (2007).
 100. Könneke, M. *et al.* Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 8239–44 (2014).
 101. Otte, J., Mall, A., Schubert, D. M., Könneke, M. & Berg, I. A. Malonic Semialdehyde Reductase from the Archaeon *Nitrosopumilus maritimus* Is Involved in the Autotrophic 3-Hydroxypropionate/4-Hydroxybutyrate Cycle. (2015)
doi:10.1128/AEM.03390-14.
 102. Estelmann, S. *et al.* Labeling and enzyme studies of the central carbon metabolism in *Metallosphaera sedula*. *J. Bacteriol.* **193**, 1191–200 (2011).
 103. Kim, J.-G. *et al.* Hydrogen peroxide detoxification is a key mechanism for growth of ammonia-oxidizing archaea. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7888–93 (2016).
 104. Abby, S. S., Kerou, M. & Schleper, C. Ancestral Reconstructions Decipher Major Adaptations of Ammonia-Oxidizing Archaea upon Radiation into Moderate. *MBio* **11**, (2020).
 105. Berger, S., Cabrera-oreface, A., Jetten, M. S. M., Brandt, U. & Welte, C. U.

- Investigation of central energy metabolism-related protein complexes of ANME-2d methanotrophic archaea by complexome profiling. *BBA - Bioenerg.* **1862**, 148308 (2021).
106. de Almeida, N. M. *et al.* Membrane-bound electron transport systems of an anammox bacterium: A complexome analysis. *Biochim. Biophys. Acta - Bioenerg.* **1857**, 1694–1704 (2016).
107. Versantvoort, W. *et al.* Complexome analysis of the nitrite-dependent methanotroph *Methylomirabilis lanthanidiphila*. *Biochim. Biophys. Acta - Bioenerg.* **1860**, 734–744 (2019).

Chapter II

Unexpected Complexity of the Ammonia Monooxygenase in Archaea

Logan H. Hodgskiss¹, Michael Melcher¹, Melina Kerou¹, Weiqiang Chen², Rafael I. Ponce-Toledo¹, Savvas N. Savvides³, Stefanie Wienkoop⁴, Markus Hartl^{2,5}, & Christa Schleper¹

Author affiliations:

¹Archaea Biology and Ecogenomics Unit, Dep. of Functional and Evolutionary Ecology, University of Vienna

²Mass Spectrometry Facility, Max Perutz Labs, Vienna BioCenter (VBC), Austria

³Unit for Structural Biology, Department of Biochemistry and Microbiology, Ghent University, Ghent, Belgium

⁴Molecular Systems Biology Unit, Dep. of Functional and Evolutionary Ecology, University of Vienna

⁵Department of Biochemistry and Cell Biology, Max Perutz Labs, University of Vienna, Austria

Prepared for submission.

Abstract

Ammonia monooxygenase belongs to the family of copper-dependent monooxygenases and catalyzes the first step in nitrification. Although it represents a key enzyme in the global nitrogen cycle fundamental knowledge of this protein, especially of the abundant ammonia oxidizing archaea (AOA), is missing. Here the structure of the enzyme is investigated in two species representing two separate lineages of AOA by a suite of techniques, including complexome analysis of native membrane complexes in blue-native gel electrophoresis, chemical cross linking, genomic linkage analyses, comparative genomics, and gene expression patterns. The well known AmoABC subunits are confirmed, as well as the earlier predicted subunit AmoX while two additional protein subunits which are conserved among all AOA were detected. Protein modelling suggests that the additional subunits likely serve an important structural and functional role in this key nitrification enzyme. These data open avenues for further structure-function studies of this important enzyme complex.

Introduction

Nitrification, the conversion of ammonium to nitrate, is a crucial step in the global nitrogen cycle solely performed by microorganisms. The process has attracted particular attention due to its agricultural and environmental relevance. The first and rate limiting ¹ step of nitrification is the oxidation of ammonia via the integral membrane protein complex ammonia monooxygenase (AMO) ^{2,3}. While ammonia oxidizing bacteria (AOB) were first discovered over 125 years ago ⁴ and have been extensively studied, this biological process was also detected in the archaeal domain in the last 20 years ⁵⁻⁷. Ammonia oxidizing archaea (AOA) have gained wide attention as they are widespread in nature and are more abundant than their bacterial counterparts in most terrestrial and marine environments, indicating important roles in nitrogen cycling ⁸⁻¹⁴. Their central nitrogen and carbon metabolism, however, is distinct from that of AOB ¹⁵⁻¹⁸. In particular, subunits of the AMO complex show only about 40% identity to those of bacteria ¹⁹ and archaeal proteins catalyzing the second step in ammonia oxidation, i.e. the conversion of hydroxylamine to nitrite, are still unknown ¹⁹⁻²¹.

Due to the difficulty of growing nitrifying organisms and the inherent problems with isolating membrane proteins, no structural studies have been successfully carried out for any AMO complex, bacterial or archaeal. This holds true for most of the diverse enzymes of the CuMMO (copper-dependent membrane monooxygenase) protein family, with a few notable exceptions. Crystal structures ²²⁻²⁶ and one cryo-EM structure ²⁷ of particulate methane

monooxygenase (pMMO) from five methanotrophs have consistently confirmed a three-polypeptide protomer (subunits-A, -B and -C) arranged in a trimer of $\alpha_3\beta_3\gamma_3$ configuration with at least two conserved metal sites in each protomer. Even so, the elucidation of the active site has remained ambiguous. It was first proposed to reside in the PmoB subunit of pMMO²⁸ as recently supported by cryo-EM analysis²⁷, while differing amino acid conservation in *Verrucomicrobia*²⁹, a recent spectroscopic analysis³⁰, and mutagenesis of a hydrocarbon monooxygenase³¹ give indications for its localization in the PmoC subunit.

Although no AMO structure has been solved, homology modelling for the AMO of the bacterium *Nitrosomonas europaea* using pMMO as a template confirmed a homotrimeric structure as well as conservation of the Cu_B and Cu_C copper sites³². The archaeal AMO complex is the most distantly related of all CuMMO proteins^{33,34} and very little is known so far about its structure or function. Based on comparative metagenomics alone, it has been suggested that an additional subunit might be present in the complex, termed AmoX^{15,35}.

To gain insights into the overall structure of the archaeal AMO complex and to investigate the role of the multiple AmoC subunits found in some AOA, membrane protein fractions from the well characterized soil AOA, *Nitrososphaera viennensis*, were analyzed biochemically using native gel electrophoreses, mass spectrometry, and chemical cross linking. Beside the three known AmoABC proteins, three additional potential subunits were identified and one of the six predicted AmoC proteins in *N. viennensis* was recognized as the primary homolog in the protein complex. In addition, the overall subunit composition of the AMO complex was confirmed in the distantly related thermophilic AOA *Nitrosocaldus cavascurensis*.

Results

Complexome analysis of native membrane complexes displays the AMO composition of Nitrososphaera viennensis

N. viennensis was grown in continuous culture for several weeks under optimal growth conditions in order to obtain enough biomass for biochemical analyses (Melcher et al. in preparation). Between 805.5 to 2042 μ g of membrane proteins were obtained from 450-550 mg of biomass per preparation, of which approximately 40-50 μ g were loaded per lane on blue-native PAGE gels³⁶. After optimization of conditions, 22 bands were cut out and subjected to mass spectrometry (see Materials and Methods). AMO subunits were among the most abundant proteins detected overall in these membrane fractions (Supplementary Figure S5). The relative

intensity profiles of AmoA, AmoB, and AmoC showed three distinct peaks corresponding to bands 4, 7, and 12, with the most prominent peak occurring at band 7 (Figure 1A). The subunits AmoA, AmoB, and AmoC made up 10%, 5%, and 14%, respectively, of the total protein found in band 7 based on iBAQ normalized intensities. AmoX was also present in band 7 representing 10% based on iBAQ normalized intensities. The AmoC subunit was predominantly represented by two of the six AmoC homologs, AmoC6 and AmoC4. These two homologs could not be distinguished based on the peptides identified in the BN-PAGE gel.

To identify proteins that might be part of the archaeal AMO complex a correlation analysis was conducted to identify proteins with a similar migration pattern as all three primary AMO subunits AmoA, AmoB, and AmoC4/C6 in the BN-PAGE gel. Patterns of the 50% most abundant proteins across the gel were compared to each other using a Kendall correlation to determine the likelihood of dependence between various proteins. Additional criteria were (i) their presence in fully sequenced ammonia oxidizing archaea, and (ii) their absence in species that do not oxidize ammonia³⁷. Two proteins initially met these criteria: the putative AMO subunit, AmoX and a hypothetical protein, NVIE_004540 (Table 1). The migration patterns for these proteins can be seen in Figure 1A. While this unbiased selection process produced intriguing additional AMO candidates, further analysis was needed to verify the presence of new subunits. Therefore, a multifaceted approach using genomics, proteomics, and transcriptomics was used to investigate this possibility.

Linkage analysis in AOA genomes supports proposed and additional AMO subunits

Earlier analyses of known subunits within the soil strains, or the family *Nitrososphaeraceae* (as defined by the Genome Taxonomy Database³⁸; used throughout), has shown a general lack of spatial clustering of all subunit genes. However, within the families *Nitrosopumilaceae* and *Nitrosocaldaceae*, the genes for the canonical AMO subunits, AmoABC, and the proposed subunit AmoX are syntenic^{35,39,40}. To investigate the conservation of the genomic neighborhood and putative additional subunits in the vicinity of the AMO cluster in those genomes, the syntenic status and conservation across AOA of the 5 genes upstream and downstream of the AMO cluster in *Nitrosocaldaceae* and *Nitrosopumilaceae* were analyzed. Of these genes, 19 were conserved in AOA with 5 being found exclusively in AOA (Supplementary Data). The 5 genes of interest included two canonical AMO genes (*amoA* and *amoB*) and the genes *amoX*, NVIE_004540, and NVIE_004550. The *amoX* gene was previously identified in metagenomic studies^{15,35} and NVIE_004540 was already a candidate identified from the BN-PAGE correlation analysis. The

additional conserved protein, NVIE_004550, was newly identified from this analysis and found to be located directly upstream of NVIE_004540. The close association of NVIE_004540 and NVIE_004550 homologs in many of the genomes investigated suggests that they could be co-transcribed (Figure 2). The two candidates encode for short polypeptides of 9.6 kDa and 12.8 kDa respectively, and – like the candidate subunit AmoX - their predicted secondary structure is predominantly helical and their subcellular localization transmembrane.

A closer analysis in *Nitrosocaldaceae*, the earliest diverging lineage in evolutionary reconstructions of AOA ³⁷, revealed that the genes for the three candidate subunits for AMO (AmoX, homolog of NVIE_004540, and homolog of NVIE_004550) clustered with the canonical subunits (AmoABC) and were syntenic in *Nitrosocaldus cavascurensis* and *Ca. Nitrosocaldus islandicus*. Clustering of all six subunit genes is also found in recently obtained MAGs ⁴¹ within the genus *Nitrosocaldus*. In the case of the newly proposed genus *Nitrosothermus* ⁴¹, AMO genes were split on multiple contigs and synteny could not be definitively determined. Additionally, all six *amo* genes belong to the fraction of genes that were predicted to have been newly acquired by the last common ancestor of all AOA ³⁷.

The emergence of *Nitrosopumilaceae* was accompanied by a separation of this genomic region into a primary cluster containing *amoABCX* and a secondary cluster containing the homologs of NVIE_004540 and NVIE_004550 (Figure 2). Within *Nitrosotalea* sp., these clusters are 11-12 genes apart, while the rest of *Nitrosopumilaceae* species have these clusters only 1-2 genes apart (with the exception of the sponge symbiont *Ca. Cenarchaeum symbiosum*). Apparently, the emergence of the family *Nitrososphaeraceae* led to a scattering of all subunit genes across the genome with the exception of *amoA* and *amoX*, which are typically linked.

A closer look was needed to account for the lack of association of NVIE_004550 with AmoABC in BN-PAGE of *N. viennensis*. When examining the relative abundance profile for NVIE_004550, the general pattern of AMO peptide peaks was followed. However, this remained undetected in the correlation analysis due to a high relative abundance peak occurring at the bottom of the gel peaking at the last band taken at approx. 66kD (Figure 1A). This suggests that NVIE_004550 could also be part of the AMO but possibly a weaker association lead to its dissociation from the complex and migration together with other complexed proteins to the bottom of the gel.

BN-PAGE protein gel indicates same AMO composition in the thermophilic archaeon Nitrosocaldus cavascurensis

To test the composition of the AMO complex outside of the context of *N. viennensis*, the BN-PAGE approach was applied to membrane protein fractions of *N. cavascurensis*, a distantly related thermophilic AOA species of the *Nitrosocaldaceae* family³⁹ that was recently obtained in pure culture (Melcher et al. in preparation).

Although a slightly different pattern of complexes was obtained (Figure 1B) a clear correlation of the additional subunits was also observed with AmoA, AmoB, and AmoC in this thermophilic organism (Kendall correlation of proteins, as performed for *N. viennensis*). The three proteins AmoX, NCAV_0486 (homolog for NVIE_004540), and NCAV_0488 (homolog for NVIE_04550) all had migration patterns within the gel that strongly correlated with AmoABC (Table 1). This analysis confirmed that the proposed subunits were translated in *N. cavascurensis*, and potentially had a physical connection within the AMO complex (Table 1).

SDS-PAGE verifies components of the AMO complex

To further explore the content of the AMO complex extracted from BN-PAGE, cut-outs from band 7 were subjected to SDS-Tricine-PAGE analysis^{42,43}. This band contained the highest relative intensities for both AmoA and AmoC. SDS bands containing AmoA, AmoB, AmoC, and AmoX were all identified using trypsin as the digestive enzyme (Figure 3). Apparent sizes of AmoB, AmoC, and AmoX proteins matched the expected molecular weight, while the band representing AmoA had a lower apparent molecular weight than expected (Figure 3), likely due to the high hydrophobicity of this subunit. Proteins with multiple transmembrane helices have been observed to run faster than expected in SDS gels⁴⁴. Bands containing SlaA (S-layer protein), NVIE_028570 (exported protein of unknown function), and NVIE_024150 (ABC uptake transporter) were also identified. The band containing AmoB also contained a large amount of NVIE_021780 (exported protein of unknown function) and the one containing the highest amount for AmoX was dominated by the two hypothetical proteins identified from the syntenic analysis, NVIE_004540 and NVIE_004550, which are proposed here to be part of the archaeal AMO complex. The band representing AmoC could be separated into two protein groups: one representing amoC4/C6 that made up 96% of the total AmoC intensities, and one representing amoC1/C2 at 4%.

In an attempt to obtain unique AmoC peptides to distinguish AmoC4 and C6 which are 96 % identical in their amino acid sequence (BLASTp)^{45,46}, additional replicate cut-outs from BN-PAGE gels corresponding to band 7 were subjected to SDS-Tricine-PAGE. Bands

corresponding to AmoC were digested using two different proteases (either gluC endoprotease or chymotrypsin). After chymotrypsin digestion it was possible to identify unique peptides from AmoC6, although the majority of AmoC peptides came from a shared region of both peptides. A unique peptide was not found for AmoC4 after correction of its predicted translational start site based on available transcriptomic and proteomic data (see details in Materials and Methods). Therefore, it was not possible to definitively distinguish in these experiments between AmoC4 and AmoC6 as the dominant homolog in the AMO complex.

Chemical cross-linking supports physical interaction of additional subunits

To estimate the physical proximity of the proposed subunits to known subunits and other proteins within the BN-PAGE gel, in-gel cross-linking⁴⁷ was performed using the DSSO cross-linker on additional BN-PAGE cut-outs from band 7. Mass spectrometry and cross-linking analysis showed multiple cross-links among AmoA, AmoB, AmoC, and AmoX as well as with the two newly proposed subunits NVIE_004540 and NVIE_04550 (Figure 4). Many cross-links were also connected to NVIE_016740, a putative S-layer protein that probably represents a highly abundant surface layer protein as known from other archaea (SlaA)^{48,49}. As this protein forms an outer sheet helping establish the pseudo-periplasm in AOA, it is not surprising to find it heavily cross-linked to membrane proteins.

AmoX also had individual cross-links to proteins of unknown function as well as AtpI (subunit of ATP synthase), Amt1 (ammonium transporter), NVIE_020280 (putative oxygen-sensitive ribonucleoside-triphosphate reductase) and NVIE_025270 (oligosaccharyl transferase STT3). As only single connections were found, and these proteins did not appear in any other syntenic or correlative analyses, they were not taken to represent a structural role in the AMO complex; cross-links can rather be attributed to their high abundance in the cell membrane.

Although this data seemed convincing, it has to be noted that in gel chemical cross-linking is a relatively new technique⁴⁷, and putative artifacts cannot be excluded due to the lack of a strong control.

*Expression patterns of AMO subunits in *N. viennensis**

Available transcriptomic studies of AOA were inspected to explore whether the expression patterns of the newly predicted subunits would corroborate their involvement in the AMO. A recent study on copper limitation in *N. viennensis*⁵⁰, confirmed that the genes *amoA*, *amoB*, and *amoC* have some of the highest transcription levels in the cell, as also shown in

previous studies ^{51–53}. A clustering analysis of the same dataset revealed that *amoA*, *amoB*, *amoC*, *amoX*, NVIE_004540, and NVIE_004550 all appear to be co-regulated, and fell into the clusters containing the most highly expressed genes. (Figure 5).

A re-evaluation of these transcriptomic data also revealed *amoC6* as the primarily transcribed *amoC* homolog. The original analysis showed high expression of both *amoC4* and *amoC6*. However, this was found to be an artefact of data processing. When taking into account only transcripts that can uniquely be assigned to a homolog, *amoC6* becomes the dominantly transcribed gene (Figure 5). This, in combination with the identification of a unique AmoC6 peptide from an SDS band digested with chymotrypsin indicates that AmoC6 is the primary structural AmoC homolog in the AMO complex of *N. viennensis*, at least under the applied growth conditions.

Transcriptomics of the marine strain, *N. maritimus*, also showed high expression of *amoA*, *amoB*, *amoC*, *amoX*, and Nmar_1506 (homolog of NVIE_004540). Nmar_1507 (homolog of NVIE_004550), albeit syntenic with Nmar_1506, exhibited lower expression levels ⁵¹.

The three newly proposed AMO subunits were also inspected in proteomic datasets that were generated with methods allowing for the improved recovery of membrane proteins. All six of the known and proposed subunits were found in membrane fractions from *N. viennensis* from a previous study ¹⁵ as well as in the proteome of *N. maritimus* ⁵¹. In other proteomic studies of AOA ^{54,55}, the three new subunits were not always present likely due to their small size and limited number of trypsin cleavage sites. Homologs of NVIE_004550 are particular hard to identify. Shot-gun proteomics of *N. viennensis* not focused on membrane fractions also reveals a low or lower identification of AmoX, NVIE_004540, and NVIE_004550 ⁵⁰ (Hodgskiss et al., unpublished results).

Structural search for missing components in the archaeal AMO complex

As previously observed, ⁵⁶ comparisons of the amino acid sequences of the three subunits AmoA,B,C from archaea with those of bacteria indicate that the primary differences between the archaeal AMO subunits and the bacterial AMO subunits are missing transmembrane helices, at least one in AmoB and two in AmoC (Supp. Figure S2 and S3), and a C-terminal soluble portion found in AmoB/PmoB. These observations also hold true for the new clade of archaeal AMO recently discovered in the Thermoplasmata phylum ⁵⁷. A HMMER search using the extended regions of the bacterial homologs against the genomes of collected AOA did not reveal any significant similarities. Therefore, a general structural search using

Phobius was carried out with the *N. viennensis* genome to search for genes that could encode a protein with the following criteria: (i) 1-3 transmembrane helices, (ii) conservation across all AOA³⁷, and (iii) present in the top 100 transcribed genes⁵⁰ (similar levels as the primary AMO subunits) (Table 2). This revealed 6 possible candidates: two blue (Type 2) copper domain containing proteins, a putative subunit for a heme/copper type cytochrome/quinol oxidase, amoX, NVIE_004540, and NVIE_004550. The only candidates to meet the structural requirements while maintaining syntenic and similar patterns of migration in BN-PAGE are amoX, NVIE_004540, and NVIE_004550.

The addition of the three proposed subunits in archaea increases the number of transmembrane helices from 10-11 to approximately 14 per protomer making it comparable to the number found in bacterial crystal structures of pMMO where each protomer of the trimer (i.e. one unit of PmoABC), contains 14-15 transmembrane helices^{23,58}.

Modelling of the archaeal AMO complex supports the integration of new subunits

The recently developed artificial intelligence system AlphaFold^{59,60} was used to model how the newly proposed subunits could interface with the known canonical subunits. The model supported all predicted transmembrane helices from AmoX, NVIE_004540, and NVIE_004550 as playing a role in anchoring the complex in the membrane along with the transmembrane helices from AmoA, AmoB, and AmoC (Figure 4D). Additionally, the N-terminal end of NVIE_004550 was predicted to contain two alpha helices that interact with the N-terminal domain of AmoB, thereby possibly replacing the missing C-terminal soluble domain found in PmoB and bacterial AmoB and offering the final piece of the missing complex in archaea when comparing to these bacterial counterparts (Figure 4D). A disulfide bond was also predicted to form within the soluble domain of NVIE_004550 (Figure 4B).

Discussion

The archaeal AMO complex is a key enzyme of AOA energy metabolism that is highly expressed in all ammonia oxidizing organisms investigated and has large implications on the environment due to its overwhelming presence in many ecosystems^{8-14,61,62}. In the domain of Archaea, the only confirmed structural information stems from the crystal structure of a heterologously expressed AmoB originating from *Candidatus Nitrosocaldus yellowstonensis*⁶³. This structure confirmed the lack of the C-terminal cupredoxin domain and revealed an extended amino acid region not found in bacteria made up of two helices and two loops. It was

proposed that this additional region could help stabilize the existing cupredoxin domain as supportive interactions are lacking due to the absence of the C-terminal domain. However, this amino acid extension is only found within the proposed genus of *Nitrosocaldus* (Supplementary Figure S2). Although more structural information on archaeal AMO is lacking, the work here profits from the recent improvements for cultivation of AOA in continuous cultures (Melcher et al. in preparation) and presents novel biochemical and comparative genomic evidence on the composition of the AMO complex in *Nitrososphaera viennensis* and other AOA.

The present analysis has verified that AmoX, NVIE_004540, and NVIE_004550 are all likely present within the archaeal AMO complex and proposes the naming of NVIE_004540 and NVIE_004550 as AmoY and AmoZ respectively. This is based on a host of analyses including BN-PAGE correlations with AmoABC in two distantly related AOA species, SDS-Tricine-PAGE from BN-PAGE cut-outs of the *N. viennensis* AMO complex, synteny with *amoABC* in the family *Nitrosocaldaceae*, predicted acquisition of genes in the last common ancestor of AOA, extant conservation in AOA, and transcriptional abundance and regulation patterns in *N. viennensis*. This proposal is additionally supported by a structural comparison and search for the missing parts of archaeal AMO (as compared to bacterial counterparts), the successful cross-linking of the newly proposed subunits with AmoA,B,C and X in BN-PAGE cut-outs from *N. viennensis*, and a predicted model of the full archaeal structure.

In both *N. viennensis* and *N. cavascurensis*, the AMO complex migrated well above the predicted height of a homotrimeric complex, even when considering the additional subunits (predicted molecular weight of a homotrimeric complex with 6 subunits per protomer: 296.94 kDa *N. viennensis*; 305.101 kDa *N. cavascurensis*). This is in contrast to the PMO complex from a *Methyloirablis* species that was also extracted using n-dodecyl- β -D-maltoside (DDM)⁶⁴, which could be explained by their differences in membrane composition. AOA contain unique ether-linked lipids (i.e. crenarchaeols)^{65–70} and rely on a proteinaceous S-layer rather than a lipid outer membrane to create a pseudo-periplasmic space^{48,49}. These differences likely have a large impact on the extraction of proteins from the membrane using non-ionic detergents (i.e. DDM). The most likely explanation for the presence of three distinct peaks of AMO is the co-migration with other proteins or complexes that it could be physically interacting with.

Previous work on bacteria that rely on CuMMOs have identified other putative proteins involved with the complex. Notably, monocistronic transcripts containing *amoABC* from the AOB *Nitrosococcus oceanus* ATCC 19707 contained two additional genes assigned as *amoR* and *amoD*⁷¹. *amoR* was found to be only present in *Nitrosococcus* and was therefore not

thought to be a conserved part of bacterial AMO. A recent study indicated that *AmoD/PmoD* (and the duplication *amoE*) play crucial roles in copper homeostasis, but they are not suspected to be a structural part of any CuMMO complex ⁷².

Although there is debate on whether *AmoC* harbors the primary active site in AMO, there is clear evidence that the metal site in *PmoC* plays a critical role in the complex of methanotrophs ^{27,30,31}. While the archaeal *AmoC* lacks a substantial section found in all bacteria that corresponds to two transmembrane helices, the metal site is conserved across all archaeal and bacterial species and its importance is supported by site directed mutagenesis studies in the genetically tractable Actinobacteria that contain the homologous hydrocarbon monooxygenase ³¹.

These results are intriguing as the soil model AOA, *N. viennensis*, like most other soil dwelling AOA from the family *Nitrososphaeraceae*, encodes multiple homologs of the *amoC* gene while retaining only single copies of *amoA* and *amoB* ¹⁵. Additional copies of *amoC* that are spatially disconnected from the AMO operon are encoded by some terrestrial AOB and were implicated in stress response based on transcriptional studies ^{73,74}. Within *Nitrososphaeraceae*, no conserved AMO operons exist (Figure 2). Duplications of the *amoC* gene (spatially distant from the other AMO genes) also occur in some species of the AOA marine associated family (*Nitrosopumilaceae*) and in two MAGs from AOA thermophiles (*Nitrosocaldaceae*), all discovered in sediments ^{41,75,76}. An *amoC* duplication is also found in an AOA sponge symbiont and copies of archaeal *amoC* are even found in marine viruses ⁷⁷. These findings together might indicate the metabolic importance of the *AmoC* subunit for ecophysiological adaptations in ammonia oxidation. While this work found one distinct copy of the six homologs, i.e. *AmoC6*, to be the primary homolog within the complex for *N. viennensis*, it is possible that (some of) the other *AmoC* subunits, which arose by gene duplications at the species level (Suppl. Figure S4), might be incorporated under certain environmental conditions and provide different activity profiles to the enzyme.

Conclusion

This study provides evidence through genomic, proteomic, and transcriptomic data for the presence of *AmoX* and the inclusion of *AmoY* and *AmoZ* as subunits within the archaeal AMO complex. A single protomer of the archaeal AMO would therefore consist of six subunits instead of three. As the anchoring of AMO in the membrane has previously been shown to be critical for its activity ²⁶, it seems plausible that the newly identified subunits play an important

role for the structural and functional integrity of AMO that allows it to properly function in archaea. The presence of a soluble domain within AmoZ that could fulfill the stabilizing function of the missing soluble domain in AmoB fulfills a potentially crucial missing piece of the AMO complex. Although the oligomerization of the complex cannot be interpreted from these data, the presence of two soluble domains in the archaea complex supports a multimeric organization similar to that seen in bacteria as these soluble domains appear to play an important stabilizing effect for the complex^{22,27}. Definitive proof of the oligomerization and organization of these subunits will not be possible until a crystal structure of archaeal AMO is realized.

In the absence of additional structure-function analyses it remains an open question, in how far the additional subunits in the archaeal complex rather reflect the vast evolutionary distance to all other known protein complexes of the CuMMO family³³, or if this difference in structure has also relevant functional implications. For instance, the bacterial AMO complexes are promiscuous enzymes able to oxidize methane and other compounds^{78–81}. Such investigations on alternative substrates have not yet been performed with the archaeal complex, but would be important for evaluating the functional role of archaea in the environment.

Since AmoXYZ appear to have important structural roles it will be crucial to include all subunits in future expression and structural studies of this environmentally relevant protein complex. Considering the wide distribution of AOA in virtually all ecosystems^{8–14,33} and their ecological relevance, it will be crucial to develop genetic tools for AOA and improve their biomass production to enable structure-function analysis and to elucidate the full pathway of ammonia oxidation in these archaea.

Data and Code Availability

All proteomic data were deposited to the ProteomeXchange Consortium via PRIDE⁸² partner repository. All data and scripts will be made available upon publication.

Methods

AMO Alignments for amoABC

50 archaeal species' and 29 bacterial species' genomes were collected and searched for known AMO/PMO subunits (amoA, amoB, and amoC). If a species was not annotated, genomic.fna files were collected from NCBI and coding sequences were searched for using

prodigal (version 2.6.3) ⁸³ using the parameter -p single. In the case of annotated species, RefSeq annotation files were given preference if available. When no RefSeq annotation was available, GenBank annotations were used. In the case of *Ca. Cenarchaeum symbiosum* A, the genome file was re-annotated using prodigal. This was done to search for coding regions that should theoretically be present that were not detected in the given annotation file.

Hidden Markov Models were made for archaea and bacteria separately based on amino acid sequences of well documented species with representatives from all major clades. In archaea, sequences from *Nitrososphaera viennensis* EN76 (Nitrososphaeraceae), *Nitrosocaldus cavascurensis* (Nitrosocaldaceae), *Nitrosopumilus maritimus* SCM1 (Nitrosopumilaceae, formerly Nitrosopumilales), and *Ca. Nitrosotalea devaneterra* (Nitrosopumilaceae, formerly Nitrosotaleales), were chosen to construct the model. In bacteria, sequences from *Nitrosococcus oceani* ATCC 19707 (γ -proteobacteria, ammonia oxidation), *Nitrospira multiformis* ATCC 25196 (β -proteobacteria, ammonia oxidation), *Ca. Nitrospira inopinata* (Nitrospira, comammox), *Methylosinus trichosporium* OB3b (α -proteobacteria, methanotroph), *Methylococcus capsulatus* str. Bath (γ -proteobacteria, methanotroph), *Methylacidiphilum kamchatkense* Kam1 (Verrucomicrobia, methanotroph), and *Mycolicibacterium chubuense* NBB4 (Actinobacteria, hydrocarbon oxidation), were chosen to construct the model.

Sequences from representative species of archaea and bacteria were aligned using Mafft (version 7.427) ^{84,85} and an HMM model was constructed using hmmbuild (HMMER 3.3, hmmer.org) for each subunit in archaea and bacteria separately. An HMM search using hmmsearch (HMMER 3.3) was performed on selected species and sequences were collected for archaea and bacteria. Archaeal and bacterial species were searched separately due to the distant phylogenetic relationship of the AMO/PMO complex within the two domains. A cut off value of 1e-20 was used for annotated species while a cut off value of 1e-10 was used for species analyzed with prodigal. A lower threshold was used for un-annotated species to account for the possibility of partial AMO/PMO genes on the edge of contigs. While a genome is not available for *Methylocystis* sp M., sequences were added to the appropriate bacteria files after the hmmsearch. Once collected, archaeal and bacterial sequences were combined and aligned using Mafft.

The archaeal amoC tree was constructed from a nucleotide BLAST (blastn, BLAST 2.12.0+) ⁸⁶ search using sequences from AOA species representing the dominant clades of AOA (see above). Mafft (version 7.427) ^{84,85} was used to align the nucleotide sequences. The alignment was then trimmed using BMGE (v1.12) ⁸⁷ and IQTree (version 2.1.2) ^{88,89} was used

to construct the phylogeny of archaeal amoC using the ultra-fast method and a bootstrap value of 1000. Nucleotide sequences were used for the tree construction as amino acid sequences were too similar to construct a reliable phylogeny.

Reactor Growth

N. viennensis was grown as a continuous culture in 2 L bioreactors (Eppendorf) filled with 1.5 L of fresh water medium (FWM) ^{68,90} with modified trace element solution ⁵, 7.5 μ M FeNaEDTA, 2 mM NH₄Cl and 1 mM pyruvate at 42 °C and pH 7.5. Carbonate was supplied by gassing the reactors with an 98 % Air 2 % CO₂ mixture, and the applied dilution rates ranged from 0.035 to 0.07 h⁻¹.

N. cavascurensis was grown as a batch culture in the same reactors, volume and medium as described for *N. viennensis*, but at 68 °C with 1 mM NH₄Cl and pH 7.0. Carbonate was also supplied by gassing, but with a mixture of Air/ N₂/ CO₂ to achieve an 10 % O₂ and 2 % CO₂ mixture. To increase the biomass, NH₄Cl was added stepwise with syringes via a septum to increase the final NO₂⁻ concentration to about 2.5 mM before harvesting the cultures.

Membrane Protein Extraction

Procedures for protein extraction and running a BN-PAGE gel were based off of Witting et al. 2006 ³⁶, Reisinger and Eichacker (2008) ⁹¹, and the NativePAGE™ Novex Bis-Tris Gel System manual from Life Technologies (MAN0000557). Study design and analysis for membrane extraction and BN-PAGE was also largely guided by de Almeida et al. (2016) ⁹² and Berger et al. (2021) ⁹³.

Frozen pellets of biomass were thawed on ice and resuspended in a sodium phosphate buffer solution (50 mM sodium phosphate, 200 mM NaCl, pH 7.0) to a concentration of ~ 20 mg/mL. Once resuspended, pepstatin and Complete Tablet EDTA-free inhibitor solution was added at concentrations of 1 μ g/mL and 40 μ L/mL (25x concentrated) respectively to inhibit protease activity. Cells were lysed using a One Shot machine set at 2.1 kbar of pressure. After lysis, samples were spun at 7000xg for 15 minutes at 4°C to remove cell debris. The supernatant was then taken for further processing.

Supernatant containing proteins and membrane were ultracentrifuged at 200,000xg (Beckman Coulter Ultracentrifuge; SW 41 Ti Swinging-Bucket Rotor, k_{max} =124) for 90 minutes at 4°C using 13.2 mL thin wall polypropylene tubes with a level of deceleration set to 7. The supernatant (containing cytoplasmic proteins) was removed and stored at -70°C with 10% glycerol. The remaining membrane pellet was resuspended/washed in a solution of 1M

NaCl, 50 mM Tris-HCl, pH 7.5 and then concentrated via ultracentrifugation at 200,000xg for 90 minutes at 4°C. The supernatant was removed and stored at -70°C with 10% glycerol. A final washing of the membrane fraction was performed by resuspending the pellet in 50 mM Tris-HCl, pH 7.5 and concentrated via ultracentrifugation at 200,000xg for 90 minutes at 4°C. (IMPORTANT: Washing solutions are made with Tris-Base and titrated with HCl to avoid accumulation of sodium ions that interfere with BN-PAGE gels. Do NOT make with Tris-HCl and titrate with NaOH.) After the final wash the supernatant was again removed and stored at -70°C with 10% glycerol.

The final membrane pellet was resuspended in 100-200 µL of NativePage Sample Buffer (Invitrogen BN2003) and 0.75 M of 6-aminocaproic acid. To aid in the resuspension of membrane fractions, samples were allowed to mix on a rotator set at 12 rpm at 4°C for 30 minutes. Following the incubation on the rotator, protein concentrations were measured using the Bradford assay. Based on the protein concentration, n-dodecyl-β-D-maltoside (DDM; Invitrogen BN2005) was added to the sample at a concentration of 0.5g DDM/g protein. Samples were again incubated on a rotator at 12 rpm and 4°C for 30 minutes. A final protein concentration was determined using the Bradford assay (Bio-Rad #5000006) with a bovine serum albumin (BSA) standard curve and controls to account for interference from DDM. Samples were then aliquoted into volumes containing approximately 40-50 µg of protein and frozen at -70°C for later analysis.

BN-PAGE

Frozen membrane samples (40-50 µg aliquots; ~10-20 µL) stored at -70°C were thawed on ice. Based on previously calculated protein concentrations, additional DDM was added to each sample to reach a DDM to protein ratio of 0.75g DDM / 1 g protein. Samples were then incubated in a shaker at 700 rpm for 15 minutes at 4°C. After incubation, samples were centrifuged at 9000xg for 60 minutes at 4°C to remove any cellular debris that was not solubilized by the addition of DDM. Supernatant from this centrifugation was transferred to a 1.5 mL LoBind protein tube. Coomassie (NativePage 5% G-250 Sample Additive; Invitrogen BN2004) was added to each sample to reach a Coomassie:DDM ratio of 1g:1g. Samples were then loaded on a 3-12% pre-cast BN-PAGE gel (Invitrogen BN1001). Approximately 5-7 µL of NativeMark Unstained Protein standard from Invitrogen (LC0725) was used as the ladder. The anode buffer was pre-ordered (Invitrogen BN2001) and consisted of a final concentration of 50 mM BisTris and 50 mM tricine at pH 6.8. The gel was run at 4°C in three stages. The first stage used dark blue cathode buffer (anode buffer with cathode buffer additive, Invitrogen

BN2002; 50 mM BisTris, 50 mM tricine, 0.02% Coomassie G-250, pH 6.8) and was run for 1 hour at 150V. For the second stage, the dark blue cathode buffer was replaced with light blue cathode buffer (50 mM BisTris, 50 mM tricine, 0.002% Coomassie G-250, pH 6.8) and the gel was run for an additional hour at 250V. For the third and final stage, the light blue cathode buffer was replaced with anode buffer and run for 45 minutes at 250V.

When finished, gels were stained using Simply Blue Safe Stain (maximum sensitivity protocol). Destaining was done with MilliQ (MilliporeSigma Milli-Q Reference A+ System) water and repeated until as much background could be removed as possible. Individual bands identified in each gel were cut and placed into LoBind protein Eppendorf tubes. Bands were stored at 4°C in 150 µL of MilliQ water until being processed for proteomic analysis, cross-linking, or SDS-Tricine-PAGE.

SDS-Tricine-PAGE

Procedures for running an SDS-Tricine-Page gel were primarily based off of Schägger and Jagow (1987)⁴² and Schägger (2006)⁴³. A 15% SDS-Tricine gel was made by mixing: 2.5 mL 30% acrylamide/Bis solution 37.5:1 (BioRad #1610158), 1.25 mL gel buffer (1.5 M Tris-HCl, 8.45 pH), 1.15 mL water, 50 µL 10% sodium-dodecyl sulfate (SDS), 50 µL 10% ammonium persulfate (APS), and 5 µL 1,2-bis(dimethylamino)ethan (TEMED). Once the gel solidified a stacking gel consisting of 340 µL 30% acrylamide/Bis solution 37.5:1, 250 µL gel buffer (1.5 M Tris-HCl, 8.45 pH), 1.36 mL water, 20 µL 10% sodium-dodecyl sulfate (SDS), 20 µL 10% ammonium persulfate (APS), and 2 µL 1,2-bis(dimethylamino)ethan (TEMED) was poured on top. Selected BN-PAGE bands were cut into 2-3 pieces and incubated together in 20 µL SDS loading buffer (0.2 M Tris-HCl, 0.3 M dithiothreitol (DTT), 277 mM SDS (8% w/v), 6 mM bromophenol blue, 4.3 M glycerol) at 65°C for 90 minutes in a shaker at 500 rpm. The loading dye solution from this incubation was used to load the SDS-Tricine gel. Either 3-5 µL of Color Prestained Protein Standard-Broad Range (11-245 kDa) (New England Biolabs, P7712) or 1-3 µL of PageRuler Prestained Protein Ladder (10-180 kDa) (Thermo Scientific, 26617) were used as a ladder. Stock solutions of 10x concentrated anode buffer (1 M Tris-base, adjusted with 6M HCl to a pH of 8.9) and 10x concentrated cathode buffer (1 M Tris-base, 1 M Tricine, 1% SDS, pH 8.3 (no pH adjustment necessary)) were previously made. Buffers were diluted to 1x when used for running the gel. The gel was run at 30V for 25 minutes to allow the proteins to leave the stacking gel followed by 200V for ~55 minutes or until the ladder reached the bottom of the gel. For clear visualization, SDS-Tricine gels were silver stained (see below). If the bands were to be used for proteomic analysis, gels were stained

with the less sensitive Simply Blue Safe Stain (maximum sensitivity protocol). After staining, bands were cut from the gel and placed in LoBind protein epis with 150 μ L of MilliQ water and stored at 4°C until being processed for proteomic analysis.

Silver Staining of SDS-Tricine Gels

All steps were done using a gentle shaker. Gels were soaked in fixing solution (50% methanol, 12% acetic acid) for at least 1 hour to overnight. Containers with gels fixed overnight were sealed with parafilm to prevent evaporation. Gels were washed for 20 minutes in 50% ethanol. Washing was repeated twice for a total of 3 times. The gel was then soaked for 1 minute in a freshly prepared solution of 1.2 g/L of sodium thiosulfate pentahydrate. Next the gel was washed for 30 seconds in MilliQ water. This was repeated twice for a total of 3 times. After washing, the gel was soaked in the dark in freshly prepared silver staining solution (2 g/L silver nitrate, 0.04% formaldehyde) for 25-30 minutes. Following staining, the gel was washed twice with water for 30 seconds. To develop bands, the gel was submerged in developer solution (60 g/L sodium carbonate, 0.04% formaldehyde, 0.036 g/L sodium thiosulfate pentahydrate). Bands developed within 1-3 minutes and development was stopped by adding destain solution (10% acetic acid, 1 % glycerol). The gel was then soaked in destain solution for approximately 5 minutes before being washed multiple times with water.

Silver Staining (Farmer's Reducer) of SDS-Tricine Gels

The gel was soaked in fixing solution as stated above. Fixing solution was removed and the gel was soaked in Farmer's Reducer (30 mM $K_3Fe(CN)_6$, 30 mM sodium thiosulfate pentahydrate) for 2 minutes. This will turn the gel yellow. The gel was then washed multiple times until the yellow background was completely removed (30-90 minutes). Once the background was removed, water was removed and replaced with 0.1% silver nitrate and was incubated in the dark for 15 minutes. After incubation, the gel was washed multiple times with MilliQ water. The gel was then submerged in a 2.5% sodium carbonate solution for 30 seconds. To develop bands, the solution was removed and replaced with a solution of 0.1% formaldehyde and 2.5% sodium carbonate. Development was stopped with destain solution as described above.

DSSO Cross-Linking Blue Native PAGE cut-outs

The protocol from Hevler et al. was followed⁴⁷. To summarize, BN-PAGE cut-outs were immersed in 90 μ L of sodium phosphate buffer solution (100 mM sodium phosphate,

0.15 M NaCl, pH 7.5: 42.6 mg $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$, 93.7 mg Na_2HPO_4 , 0.3 mL 0.5M NaCl in 10 mL MilliQ water; pH adjusted to 7.5). 1 mg of the mass spec cleavable cross-linker disuccinimidyl sulfoxide (DSSO; Thermo Scientific A33545) was resuspended in 51.5 μL of dimethyl sulfoxide (DMSO) to reach a final concentration of 50 mM DSSO. 10 μL of 50 mM DSSO was then added to each sample yielding a final DSSO concentration of 5 mM per sample. Samples were briefly vortexed and then incubated at room temperature for 30 minutes. The cross-linking reaction was then stopped with the addition of 2 μL of 1 M Tris-HCl. Samples were briefly vortexed and incubated for 15 minutes at room temperature. The solution was then removed and replaced with MilliQ water. Samples were then stored at 4°C until being processed for mass spectrometry analysis. To verify cross-linking effectiveness, some cross-linked samples were analyzed on SDS-Tricine gels and compared to non-cross-linked samples (examples seen with similar cross-linkers in Supplementary Figure S8).

Sample preparation for mass spectrometry (SDS or BN-gel)

The Coomassie-stained gel bands were destained with a mixture of acetonitrile (Chromasolv®, Sigma-Aldrich) and 50 mM ammonium bicarbonate (Sigma-Aldrich). The proteins were reduced using 10 mM dithiothreitol (Roche) and alkylated with 50 mM iodoacetamide. Trypsin (Promega; Trypsin Gold, Mass Spectrometry Grade) digestion was carried out at 37°C overnight in 50mM ammonium bicarbonate. Chymotrypsin (Roche) digestion was carried out at 25°C for 5 hours in 50mM ammonium bicarbonate. GluC (Roche) digestion was carried out at 37°C overnight in 50mM ammonium bicarbonate. 10% formic acid was used to stop the digestion and peptides were extracted twice with 5%FA for 10min in an cooled ultrasonic bath. Extracted peptides were pooled and desalted using C18 Stagetips⁹⁴.

Liquid chromatography separation coupled to mass spectrometry

Peptides were analyzed on an UltiMate 3000 HPLC RSLC nanosystem (Thermo Fisher Scientific) coupled to a Q Exactive HF-X, equipped with a nano-spray ion source using coated emitter tips (PepSep, MSWil). Samples were loaded on a trap column (Thermo Fisher Scientific, PepMap C18, 5 mm $\times$ 300 μm ID, 5 μm particles, 100 Å pore size) at a flow rate of 25 $\mu\text{L min}^{-1}$ using 0.1% TFA as mobile phase. After 10 min, the trap column was switched in-line with the analytical C18 column (Thermo Fisher Scientific, PepMap C18, 500 mm $\times$ 75 μm ID, 2 μm , 100 Å) and peptides were eluted applying a segmented linear gradient from 2% to 80% solvent B (80% acetonitrile, 0.1% formic acid; solvent A 0.1% formic acid) at a flow rate of 230 nL/min over 60 min. The mass spectrometer was operated in data-dependent mode,

survey scans were obtained in a mass range of 350-1600 m/z with lock mass activated, at a resolution of 120,000 at 200 m/z and an AGC target value of 1E6. The 15 most intense ions were selected with an isolation width of 1.2 Thomson for a max. of 150 ms, fragmented in the HCD cell at stepped normalized collision energy at 26%, 28%, and 30%. The spectra were recorded at an AGC target value of 1E5 and a resolution of 60,000. Peptides with a charge of +1, or >+7 were excluded from fragmentation, the peptide match feature was set to preferred, the exclude isotope feature was enabled, and selected precursors were dynamically excluded from repeated sampling for 20 seconds within a mass tolerance of 8 ppm.

Data analysis for identification of BN-PAGE and SDS-Tricine-Page bands

For peptide and protein identification raw data were processed using the MaxQuant software package ⁹⁵ (version 1.6.6.0) and spectra searched against *Nitrososphaera_Viennensis* reference proteome (Uniprot, downloaded fall 2021) with a starting site modification (see below) to AmoC4 (Uniprot accession: A0A060HLS1) and a database containing common contaminants. The search was performed with full trypsin specificity (or corresponding enzyme used in the digestion) and a maximum of 2 missed cleavages at a protein and peptide spectrum match false discovery rate of 1%. Carbamidomethylation of cysteine residues was set as fixed, oxidation of methionine, and N-terminal acetylation as variable modifications. The output option of iBAQ with log fit was selected. All other parameters were left at default.

Data analysis for identification of BN-PAGE cross-linked bands

Peptide and protein identification was described as above except with MaxQuant version 1.6.17.0 and with the a reference proteome that was not corrected for AmoC4. Two programs were used to identify cross-linked peptides: Annika and MeroX.

To identify cross-linked peptides, the raw data were searched with either MS Annika ⁹⁶ in Proteome Discoverer 2.3 or with MeroX 2.0 ⁹⁷ against the sequences of the top abundant protein hits (with at least 10 MS/MS counts) from the MaxQuant search. Although it had less than 10 MS/MS counts, the protein encoded by NVIE_004550 was also added based on other proteomic and syntenic analysis. DSSO was selected as the cross-linking chemistry. Carbamidomethyl on Cys was set as fixed, oxidation of Met, and protein N-terminal acetylation as variable modifications. Enzyme specificity was selected according to the protease used for digestion. Search results were filtered for 1% FDR on the PSM level limiting the precursor mass deviation to 10 ppm. Further filtering is done for only non-decoy and high confidence

PSMs in MS Annika and for a score higher than 40 in MeroX 2.0. <http://crosslinkviewer.org/> is used to draw the XL maps.

amoC6 as the Primary AmoC subunit in N. viennensis

Six homologs of *amoC* (*amoC1-6*) are present in the genome of *N. viennensis*. Predicted start sites for AmoC1-3, AmoC5, and AmoC6 are consistent across annotations (GenBank and RefSeq). However, the start site of AmoC4 is significantly different between the two annotations. Additionally, there is a third possibility based on a methionine that would resemble the start site of AmoC1-3,5-6. Based on AmoC alignments in all AOA and transcriptional information from Reyes et al. (2020), it was determined that the most accurate start site for this protein resides at the third option, resembling that of the other AmoC subunits in *N. viennensis*. The proteome downloaded from Uniprot was manually annotated to reflect this decision. This is the proteome used for BN-PAGE and SDS-Tricine reference files. With this correction, it is not possible to distinguish between AmoC4 and AmoC6 in any of the samples. However, if the original proteome from Uniprot is used, MaxQuant will identify unique peptides for AmoC4 and AmoC6. Based on the current analysis, this is an inaccurate interpretation of the data. With the corrected AmoC4 annotation, only unique peptides from AmoC6 are found. To further verify that AmoC4 is not playing a significant role, the raw data from the chymotrypsin digest was subjected to an unspecific and semi-specific closed search for unique peptides from AmoC4. An open search was also carried out using the FragPipe software (version 17.1) ⁹⁸. None of these searches revealed unique AmoC4 peptides therefore strengthening the argument that the primary structural homolog in *N. viennensis* is AmoC6.

BN-PAGE correlation analysis

A Kendall correlation using R was used to find proteins that had a similar migration pattern with AmoABC. Any protein expected to be a part of the AMO complex should be in high abundance. Therefore, only the top 50% proteins, based on iBAQ abundance, were used for the analysis. Each protein was correlated with every other protein using the function `cor.test` in R with `method="kendall"` and `use="complete.obs"`. P-values were adjusted using the Benjamini-Hochberg method. Results were filtered for proteins correlated with one of the AMO proteins (AmoA, AmoB, or AmoC) and a correlation (tau) greater than or equal to 0.7. All filtered results had an adjusted p-value higher than 0.001.

For genes of interest, conservation across AOA was checked according to the data set curated in Abby et al. (2020). A gene was considered exclusive to AOA if it was found in AOA but not found in any species outside of AOA according to the dataset in Abby et al. (2020). The full dataset from this study was provided by the authors.

Phylogenomic analysis

A total of 106 MAGs and completely sequenced genomes (98 AOA and 8 non-AOA genomes) were collected from NCBI, IMG or DDBJ databases, followed by protein prediction using Prodigal v2.6.3⁸³. The identification of phylogenetic markers to perform the phylogenomic tree reconstruction was based on the workflow proposed by⁹⁹ using the archaeal single-copy gene collection (e-value 10^{-10})¹⁰⁰. 32 ribosomal proteins detected in at least 90 out of the 106 genomes present in the collected genome database were selected. Protein families were aligned independently using the mafft-linsi algorithm implemented in MAFFT v7.427⁸⁵ followed by a trimming step in BMGE⁸⁷ with default parameters. Trimmed protein families were concatenated using a tailor-made python script and the concatenated alignment was used to reconstruct a maximum likelihood (ML) phylogenomic tree in IQTREE (v2.0-rc1)⁸⁸ under the LG+C20+F+G model with 1000 bootstrap replicates.

Transcriptomic Clustering

Transcriptomic reads from Reyes et al.⁵⁰ were re-processed taking into account strandedness and using hisat2 (version 2.1.0)¹⁰¹ rather than bowtie2 for read mapping. Count values were evaluated using featureCounts (v2.0.0)¹⁰². Counts were then normalized using the Transcripts per Kilobase Million (TPM) method. TPM is calculated by dividing each gene count by the total length of the gene in kilobases giving a reads per kilobase value. “reads per kilobase” was summed up for all genes in a sample and divided by 1,000,000 to give a scaling factor. Each genes’ “reads per kilobase” was divided by this scaling factor to give the final TPM. TPM values were converted to log₂ values and clustered using hierarchical clustering with the default values of heatmap.2¹⁰³ in R version 3.6.3¹⁰⁴. Clustered genes were split into 15 clusters. 3 clusters represented the genes with the highest abundance of transcripts in both the copper replete and copper limited cultures. Cluster data is visualized using Anvio 7.1¹⁰⁵.

Structural Search for Missing AMO Sections in Archaea

Extended pieces of bacterial AmoB and AmoC were trimmed from alignments of combined archaea and bacteria species using trimal (v1.4rev15) ¹⁰⁶. One sequence was removed from AmoB sequences as it was not actually the AmoB subunit. Trimmed sections were used to create HMMs and then searched against archaeal species to look for proteins that might supply these missing pieces. No hits were found. Therefore, a structural search for proteins containing transmembrane helices in *N. viennensis* was carried out for highly transcribed genes.

TPM reads from Reyes et al. 2020 were averaged for the five replete conditions and sorted by highest abundance to obtain the top 100 transcribed genes. The GenBank translated CDS region for *N. viennensis* was analyzed using Phobius ¹⁰⁷ to identify transmembrane helices and signal peptides in all proteins. The highest transcribed genes were then filtered to include genes with 1-3 predicted transmembrane helices. Candidate genes were then analyzed for conservation in AOA and correlation in BN-PAGE gels from *N. viennensis*.

Predicted model of archaeal AMO using AlphaFold

Sequences for AmoA, AmoB, AmoC, AmoX, NVIE_004540, and NVIE_004550 from *N. viennensis* were used for AlphaFold version 2 ^{59,60} predictions. In the case of AmoB and NVIE_004550, predicted signal peptides were removed based off of predictions using SignalP 5.0 (archaea) ¹⁰⁸. AmoC6 was used to represent the AmoC subunit. The crystal structure of *Methylocystis* sp. strain M ²⁴ was used as a template for the model prediction. Visualization of the predicted structure was performed using PyMol ¹⁰⁹.

Identification and Alignments of amoXYZ in Ammonia Oxidizing Archaea

Amino acid sequences for AmoX, AmoY, and AmoZ were obtained from the genes *amoX*, NVIE_004540, and NVIE_004550 respectively. Homologs in other AOA were initially searched for using blastp (v2.12.0+) ⁸⁶ with a threshold of 1e-4. Not all species had identified hits. A more sensitive analysis was performed by creating an HMM model using hmmbuild (HMMER 3.3, hmmer.org) for each new subunit from the top BLASTp hit from each species with a BLAST result. The HMMs for AmoX, AmoY, and AmoZ were then used with hmmsearch (HMMER 3.3, hmmer.org) in all collected AOA genomes. This produced hits in all species for all subunits except *amoY* and *amoZ* in Thaumarchaeota archaeon J079, a MAG that is also missing AmoB and is only 84% complete and *amoZ* in the GenBank protein file for *Ca. Cenarchaeum symbiosum* A. A re-analysis of the *Ca. Cenarchaeum symbiosum* A genome

using prodigal (see above) was able to identify a coding sequence for AmoZ while maintaining the coding sequences for all other AMO subunits.

References

1. Wong-Chong, G. M. & Loehr, R. C. The kinetics of microbial nitrification. *Water Res.* **9**, 1099–1106 (1975).
2. Hyman, M. R. & Wood, P. M. Suicidal inactivation and labelling of ammonia mono-oxygenase by acetylene. *Biochem. J.* **227**, 719–725 (1985).
3. Hollocher, T. C., Tate, M. E. & Nicholas, D. J. Oxidation of ammonia by *Nitrosomonas europaea*. Definite ¹⁸O-tracer evidence that hydroxylamine formation involves a monooxygenase. *J. Biol. Chem.* **256**, 10834–10836 (1981).
4. Winogradsky, S. Recherches sur les organismes de la nitrification. *Ann Inst Pateur* **4:213-231.**, (1890).
5. Könneke, M. *et al.* Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**, 543–546 (2005).
6. Treusch, A. H. *et al.* Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ. Microbiol.* **7**, 1985–1995 (2005).
7. Venter, J. C. *et al.* Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science* (80-.). **304**, 66–74 (2004).
8. Leininger, S. *et al.* Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**, 806–809 (2006).
9. Nicol, G. W., Leininger, S., Schleper, C. & Prosser, J. I. The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. *Environ. Microbiol.* **10**, 2966–2978 (2008).
10. Adair, K. L. & Schwartz, E. Evidence that ammonia-oxidizing archaea are more abundant than ammonia-oxidizing bacteria in semiarid soils of northern Arizona, USA. *Microb. Ecol.* **56**, 420–426 (2008).
11. Karner, M. B., Delong, E. F. & Karl, D. M. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**, 507–510 (2001).
12. Shi, Y., Tyson, G. W., Eppley, J. M. & Delong, E. F. Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean.

- ISME J.* **5**, 999–1013 (2011).
13. Baker, B. J., Lesniewski, R. A. & Dick, G. J. Genome-enabled transcriptomics reveals archaeal populations that drive nitrification in a deep-sea hydrothermal plume. *ISME J.* **6**, 2269–2279 (2012).
 14. Hollibaugh, J. T., Gifford, S., Sharma, S., Bano, N. & Moran, M. A. Metatranscriptomic analysis of ammonia-oxidizing organisms in an estuarine bacterioplankton assemblage. *ISME J.* **5**, 866–878 (2011).
 15. Kerou, M. *et al.* Proteomics and comparative genomics of *Nitrososphaera viennensis* reveal the core genome and adaptations of archaeal ammonia oxidizers. *Proc. Natl. Acad. Sci.* **113**, E7937–E7946 (2016).
 16. Walker, C. B. *et al.* *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc. Natl. Acad. Sci.* **107**, (2010).
 17. Berg, I. A., Kockelkorn, D., Buckel, W. & Fuchs, G. A 3-Hydroxypropionate/4-Hydroxybutyrate Autotrophic Carbon Dioxide Assimilation Pathway in Archaea. *Science* (80-.). **318**, (2007).
 18. Könneke, M. *et al.* Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 8239–44 (2014).
 19. Lancaster, K. M., Caranto, J. D., Majer, S. H. & Smith, M. A. Alternative Bioenergy: Updates to and Challenges in Nitrification Metalloenzymology. *Joule* (2018) doi:10.1016/j.joule.2018.01.018.
 20. Simon, J. & Klotz, M. G. Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. *Biochim. Biophys. Acta - Bioenerg.* **1827**, 114–135 (2013).
 21. 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.* 1–10 (2016) doi:10.1038/ismej.2016.2.
 22. Lieberman, R. L. & Rosenzweig, A. C. Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. *Nature* **434**, 177–182 (2005).
 23. Hakemian, A. S. *et al.* The metal centres of particulate methane mono-oxygenase from *Methylosinus trichosporium* OB3b. *Biochem. Soc. Trans.* **36**, 1134–1137 (2008).
 24. Smith, S. M. *et al.* Crystal structure and characterization of particulate methane
-

- monooxygenase from *Methylocystis* species strain M. *Biochemistry* **50**, 10231–10240 (2011).
25. Sirajuddin, S. *et al.* Effects of zinc on particulate methane monooxygenase activity and structure. *J. Biol. Chem.* **289**, 21782–21794 (2014).
 26. Ro, S. Y. *et al.* From micelles to bicelles: Effect of the membrane on particulate methane monooxygenase activity. *J. Biol. Chem.* **293**, 10457–10465 (2018).
 27. Chang, W. H. *et al.* Copper Centers in the Cryo-EM Structure of Particulate Methane Monooxygenase Reveal the Catalytic Machinery of Methane Oxidation. *J. Am. Chem. Soc.* **143**, 9922–9932 (2021).
 28. Balasubramanian, R. *et al.* Oxidation of methane by a biological dicopper centre. *Nature* **465**, 115–119 (2010).
 29. Op den Camp, H. J. M. *et al.* Environmental, genomic and taxonomic perspectives on methanotrophic Verrucomicrobia. *Environ. Microbiol. Rep.* **1**, 293–306 (2009).
 30. Ross, M. O. *et al.* Particulate methane monooxygenase contains only mononuclear copper centers. *Science (80-.).* **364**, 566–570 (2019).
 31. Liew, E. F., Tong, D., Coleman, N. V. & Holmes, A. J. Mutagenesis of the hydrocarbon monooxygenase indicates a metal centre in subunit-C, and not subunit-B, is essential for copper-containing membrane monooxygenase activity. *Microbiol. (United Kingdom)* **160**, 1267–1277 (2014).
 32. Musiani, F., Broll, V., Evangelisti, E. & Ciurli, S. The model structure of the copper-dependent ammonia monooxygenase. *JBIC J. Biol. Inorg. Chem.* **1**, 3 (2020).
 33. Alves, R. J. E., Minh, B. Q., Urich, T., Von Haeseler, A. & Schleper, C. Unifying the global phylogeny and environmental distribution of ammonia-oxidising archaea based on amoA genes. (2018) doi:10.1038/s41467-018-03861-1.
 34. Khadka, R. *et al.* Evolutionary History of Copper Membrane Monooxygenases. (2018) doi:10.3389/fmicb.2018.02493.
 35. Bartossek, R., Spang, A., Weidler, G., Lanzen, A. & Schleper, C. Metagenomic Analysis of Ammonia-Oxidizing Archaea Affiliated with the Soil Group. *Front. Microbiol.* **0**, 208 (2012).
 36. Wittig, I., Braun, H.-P. & Schägger, H. Blue native PAGE. (2006) doi:10.1038/nprot.2006.62.
 37. Abby, S. S., Kerou, M. & Schleper, C. Ancestral Reconstructions Decipher Major Adaptations of Ammonia-Oxidizing Archaea upon Radiation into Moderate. *MBio* **11**, (2020).
-

38. Rinke, C. *et al.* A standardized archaeal taxonomy for the Genome Taxonomy Database. *Nat. Microbiol.* **6**, 946–959 (2021).
 39. Abby, S. S. *et al.* Candidatus Nitrosocaldus cavascurensis, an ammonia oxidizing, extremely thermophilic archaeon with a highly mobile genome. *Front. Microbiol.* **9**, 1–19 (2018).
 40. Nicol, G. W. & Schleper, C. Ammonia-oxidising Crenarchaeota: important players in the nitrogen cycle? *Trends Microbiol.* **14**, 207–212 (2006).
 41. Luo, Z.-H. *et al.* Genomic Insights of “Candidatus Nitrosocaldaceae” Based on Nine New Metagenome-Assembled Genomes, Including “Candidatus Nitrosothermus” Gen Nov. and Two New Species of “Candidatus Nitrosocaldus”. *Front. Microbiol.* **11**, (2021).
 42. Schägger, H. & von Jagow, G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368–379 (1987).
 43. Schägger, H. Tricine-SDS-PAGE. *Nat. Protoc.* **1**, (2006).
 44. Rath, A. & Deber, C. M. Correction factors for membrane protein molecular weight readouts on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. (2012) doi:10.1016/j.ab.2012.11.007.
 45. Altschul, S. F. *et al.* Protein database searches using compositionally adjusted substitution matrices. *FEBS J.* **272**, 5101–5109 (2005).
 46. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
 47. Hevler, J. F. *et al.* Selective cross-linking of coinciding protein assemblies by in-gel cross-linking mass spectrometry. *EMBO J.* (2021) doi:10.15252/embj.2020106174.
 48. Stieglmeier, M. *et al.* Nitrososphaera viennensis gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* **64**, 2738–2752 (2014).
 49. Albers, S. V. & Meyer, B. H. The archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414–426 (2011).
 50. Reyes, C. *et al.* Genome wide transcriptomic analysis of the soil ammonia oxidizing archaeon Nitrososphaera viennensis upon exposure to copper limitation. *ISME J.* **14**, 2659–2674 (2020).
 51. Qin, W. *et al.* Stress response of a marine ammonia-oxidizing archaeon informs physiological status of environmental populations. *Nat. Publ. Gr.* doi, (2017).
-

52. Carini, P., Dupont, C. L. & Santoro, A. E. Patterns of thaumarchaeal gene expression in culture and diverse marine environments. *Environ. Microbiol.* **20**, 2112–2124 (2018).
53. Stewart, F. J., Ulloa, O. & Delong, E. F. Microbial metatranscriptomics in a permanent marine oxygen minimum zone. *Environ. Microbiol.* **14**, 23–40 (2012).
54. Bayer, B. *et al.* *Proteomic Response of Three Marine Ammonia-Oxidizing Archaea to Hydrogen Peroxide and Their Metabolic Interactions with a Heterotrophic Alphaproteobacterium* Downloaded from. *msystems.asm.org* vol. 4 <http://msystems.asm.org/> (2019).
55. Santoro, A. E. *et al.* Genomic and proteomic characterization of “Candidatus Nitrosopelagicus brevis”: an ammonia-oxidizing archaeon from the open ocean. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 1173–8 (2015).
56. Tolar, B. B. *et al.* Integrated structural biology and molecular ecology of N-cycling enzymes from ammonia-oxidizing archaea. *Environ. Microbiol. Rep.* **9**, 484–491 (2017).
57. Diamond, S. *et al.* Soils and sediments host Thermoplasmata archaea encoding novel copper membrane monooxygenases (CuMMOs). *ISME J.* 1–15 (2022) doi:10.1038/s41396-021-01177-5.
58. Hakemian, A. S. & Rosenzweig, A. C. The Biochemistry of Methane Oxidation. (2007) doi:10.1146/annurev.biochem.76.061505.175355.
59. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, (2021).
60. Varadi, M. *et al.* NAR Breakthrough Article AlphaFold Protein Structure Database : massively expanding the structural coverage of protein-sequence space with high-accuracy models. **50**, 439–444 (2022).
61. Ernst, J. *et al.* STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics* **7**, 191 (2006).
62. Klotz, M. G. & Stein, L. Y. Nitrifier genomics and evolution of the nitrogen cycle. (2007) doi:10.1111/j.1574-6968.2007.00970.x.
63. Lawton, T. J., Ham, J., Sun, T. & Rosenzweig, A. C. Structural conservation of the B subunit in the ammonia monooxygenase/particulate methane monooxygenase superfamily. *Bone* **23**, 1–7 (2014).
64. Versantvoort, W. *et al.* Complexome analysis of the nitrite-dependent methanotroph *Methyloiridis lanthanidiphila*. *Biochim. Biophys. Acta - Bioenerg.* **1860**, 734–744

- (2019).
65. Pitcher, A. *et al.* Crenarchaeol dominates the membrane lipids of Candidatus Nitrososphaera gargensis, a thermophilic Group I.1b Archaeon. *ISME J.* **4**, 542–552 (2010).
 66. Villanueva, L., Damsté, J. S. S. & Schouten, S. A re-evaluation of the archaeal membrane lipid biosynthetic pathway. *Nat. Rev. Microbiol.* **12**, 438–448 (2014).
 67. Sinninghe Damsté, J. S., Schouten, S., Hopmans, E. C., Van Duin, A. C. T. & Geenevasen, J. A. J. Crenarchaeol: The characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota. *J. Lipid Res.* **43**, 1641–1651 (2002).
 68. De La Torre, J. R., Walker, C. B., Ingalls, A. E., Könneke, M. & Stahl, D. A. Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ. Microbiol.* **10**, 810–818 (2008).
 69. Sinninghe Damsté, J. S. *et al.* Intact polar and core glycerol dibiphytanyl glycerol tetraether lipids of group I.1a and I.1b Thaumarchaeota in soil. *Appl. Environ. Microbiol.* **78**, 6866–6874 (2012).
 70. Elling, F. J., Könneke, M., Mußmann, M., Greve, A. & Hinrichs, K. U. Influence of temperature, pH, and salinity on membrane lipid composition and TEX86 of marine planktonic thaumarchaeal isolates. *Geochim. Cosmochim. Acta* **171**, 238–255 (2015).
 71. El Sheikh, A. F., Poret-Peterson, A. T. & Klotz, M. G. Characterization of two new genes, amoR and amoD, in the amo operon of the marine ammonia oxidizer Nitrosococcus oceanus ATCC 19707. *Appl. Environ. Microbiol.* **74**, 312–318 (2008).
 72. Fisher, O. S. *et al.* Characterization of a long overlooked copper protein from methane- and ammonia-oxidizing bacteria. *Nat. Commun.* **9**, 1–12 (2018).
 73. Berube, P. M., Samudrala, R. & Stahl, D. A. Transcription of All amoC Copies Is Associated with Recovery of Nitrosomonas europaea from Ammonia Starvation. *J. Bacteriol.* **189**, 3935–3944 (2007).
 74. Berube, P. M. & Stahl, D. A. The divergent AmoC3 subunit of ammonia monooxygenase functions as part of a stress response system in Nitrosomonas europaea. *J. Bacteriol.* **194**, 3448–56 (2012).
 75. Lebedeva, E. V. *et al.* Enrichment and genome sequence of the group I.1a ammonia-oxidizing archaeon ‘Ca. Nitrosotenuis uzonensis’ representing a clade globally distributed in thermal habitats. *PLoS One* **8**, 1–12 (2013).
 76. Qin, W. *et al.* Nitrosopumilus maritimus gen. nov., sp. nov., Nitrosopumilus
-

- cobalaminigenes sp. nov., Nitrosopumilus oxyclinae sp. nov., and Nitrosopumilus ureiphilus sp. nov., four marine ammonia-oxidizing archaea of the phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* (2017) doi:10.1099/ijsem.0.002416.
77. Ahlgren, N. A., Fuchsman, C. A., Rocap, G. & Fuhrman, J. A. Discovery of several novel, widespread, and ecologically distinct marine Thaumarchaeota viruses that encode amoC nitrification genes. *ISME J.* (2018) doi:10.1038/s41396-018-0289-4.
78. Hyman, M. R. & Wood, P. M. Methane oxidation by Nitrosomonas europaea. *Biochem. J.* **212**, 31–37 (1983).
79. Hyman, M. R. & Wood, P. M. Ethylene oxidation by Nitrosomonas europaea. *Arch. Microbiol.* **137**, 155–158 (1984).
80. Hyman, M. R., Murton, I. B. & Arp, D. J. Interaction of Ammonia Monooxygenase from Nitrosomonas europaea with Alkanes, Alkenes, and Alkynes. *Appl. Environ. Microbiol.* **54**, 3187–3190 (1988).
81. Jones, R. D. & Morita, R. Y. Methane oxidation by Nitrosococcus oceanus and Nitrosomonas europaea. *Appl. Environ. Microbiol.* **45**, 401–410 (1983).
82. Perez-Riverol, Y. *et al.* The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).
83. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, (2010).
84. Katoh, K., Misawa, K., Kuma, K. I. & Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
85. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
86. Camacho, C. *et al.* BLAST+: Architecture and applications. *BMC Bioinformatics* **10**, 1–9 (2009).
87. Criscuolo, A. & Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): A new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, (2010).
88. Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
89. Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
90. Tournai, M. *et al.* Nitrososphaera viennensis, an ammonia oxidizing archaeon from ,
-

- soil. *Proc. Natl. Acad. Sci. USA* **108**, 8420–8425 (2011).
91. Reisinger, V. & Eichacker, L. A. Solubilization of membrane protein complexes for blue native PAGE. *J. Proteomics* **71**, 277–283 (2008).
 92. de Almeida, N. M. *et al.* Membrane-bound electron transport systems of an anammox bacterium: A complexome analysis. *Biochim. Biophys. Acta - Bioenerg.* **1857**, 1694–1704 (2016).
 93. Berger, S., Cabrera-oreface, A., Jetten, M. S. M., Brandt, U. & Welte, C. U. Investigation of central energy metabolism-related protein complexes of ANME-2d methanotrophic archaea by complexome profiling. *BBA - Bioenerg.* **1862**, 148308 (2021).
 94. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
 95. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry–based shotgun proteomics. *Nat. Protoc.* **11**, (2016).
 96. Pirklbauer, G. J. *et al.* MS Annika: A New Cross-Linking Search Engine. *J. Proteome Res.* **20**, 2560–2569 (2021).
 97. Iacobucci, C. *et al.* A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein–protein interactions. *Nat. Protoc.* **13**, 2864–2889 (2018).
 98. Kong, A. T., Leprevost, F. V., Avtonomov, D. M., Mellacheruvu, D. & Nesvizhskii, A. I. MSFragger: Ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* **14**, 513–520 (2017).
 99. Graham, E. D., Heidelberg, J. F. & Tully, B. J. Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J.* **12**, 1861–1866 (2018).
 100. Rinke, C. *et al.* Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431–437 (2013).
 101. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
 102. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
 103. Warnes, G. R. *et al.* gplots: Various R Programming Tools for Plotting Data. (2020).
- ,
-

104. R Core Team (2020). R: A language and environment for statistical computing. (2020).
105. Eren, A. M. *et al.* Community-led, integrated, reproducible multi-omics with anvi'o. *Nat. Microbiol.* **6**, 3–6 (2021).
106. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
107. Käll, L., Krogh, A. & Sonnhammer, E. L. L. A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* **338**, 1027–1036 (2004).
108. Almagro Armenteros, J. J. *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
109. Schrodinger LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).
110. Sehnal, D. *et al.* Mol*Viewer: Modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Res.* **49**, W431–W437 (2021).
111. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

Tables

Table 1: Correlations of proteins with occurrence of AmoA,B, and C in (A) *N. viennensis* and (B) *N. cavascurens* BN-PAGE gels.

A

AMO Correlation Results for *N. viennensis*

Locus Tag	Gene	Protein*	Conserved in Extant AOA [§]	Exclusive to AOA [§]	Correlations [†]		
					AmoA	AmoB	AmoC6
NVIE_016740	NVIE_016740	surface associated S-layer protein		X	X	X	X
NVIE_011620	nuoJ	Complex I	X		X	X	X
NVIE_011600	nuoM	Complex I	X		X	X	X
NVIE_027530	coxB	Complex IV			X	X	X
NVIE_027540	coxA1	Complex IV			X		
NVIE_013530	NVIE_013530	protein of unknown function		X	X	X	X
NVIE_027260	NVIE_027260	conserved protein of unknown function		X	X	X	X
NVIE_017130	NVIE_017130	protein of unknown function DUF373	X		X		X
NVIE_027280	amoX	potential AMO subunit	X	X	X	X	X
NVIE_004540	NVIE_004540	hypothetical protein	X	X	X	X	X

B

AMO Correlation Results for *N. cavascurens*

Locus Tag	Gene	Protein*	Conserved in Extant AOA [§]	Exclusive to AOA [§]	Correlations [†]		
					AmoA	AmoB	AmoC6
NCAV_1585	coxA	Cytochrome c oxidase polypeptide 1			X	X	
NCAV_1739	NCAV_1739	Uncharacterized protein		X	X	X	
NCAV_1586	coxB	Putative heme-copper oxidase subunit II			X	X	X
NCAV_0011	NCAV_0011	ABC-1 domain-containing protein			X	X	
NCAV_1743	amt	Ammonium transporter	X		X	X	
NCAV_0191	petB	Putative cytochrome b/b6			X		
NCAV_1587	NCAV_1587	Putative heme/copper-type cytochrome/quinol oxidase, subunit	X	X	X		
NCAV_0486	NCAV_0486	Uncharacterized protein (NVIE_004550 homolog)	X	X	X	X	X
NCAV_0488	NCAV_0488	Uncharacterized protein (NVIE_004540 homolog)	X	X	X	X	X
NCAV_0491	NCAV_0491	Uncharacterized protein (AmoX)	X	X	X	X	X

*Proteins with at least one AmoABC protein correlation.

[§]Conservation and exclusiveness to AOA based on results from Abby et al. (2020).

[†]Correlation ≥ 0.7 and adjusted p-value ≥ 0.001 .

Table 2: Structural search for missing AMO subunits.

Gene Expression & Structural Search				Previous Analyses		Protein
Gene	TPM log2*	TM†	AOA Conservation§	BN-PAGE Corr. ‡	AMO Synteny*	
NVIE_013530	13.82	1		X		protein of unknown function
NVIE_004540	13.55	1	X	X	X	hypothetical protein
NVIE_004550	13.39	1	X	X	X	hypothetical protein
amoX	12.14	2	X	X	X	putative ammonia monooxygenase - associated protein
coxB	11.87	1		X		putative heme-copper oxidase subunit II
NVIE_014370	11.75	1				putative Copper binding protein, plastocyanin/azurin family
NVIE_010490	11.75	1				putative phosphoesterase, DHHA1
NVIE_027520	11.40	1	X			putative heme/copper-type cytochrome/quinol oxidase, subunit
NVIE_027550	11.33	1	X			putative blue (Type 1) copper domain protein
atpK	11.30	2				archaeal A1A0-type ATP synthase, subunit K
NVIE_021780	10.90	1				exported protein of unknown function
NVIE_006540	10.90	1				exported protein of unknown function
NVIE_029580	10.20	1	X			blue (Type 1) copper domain-containing protein

*Based on transcriptomic counts averaged from replete copper conditions from Reyes et al. (2020).

†Number of predicted transmembrane helices.

§As predicted from Abby et al. (2020).

‡Represents proteins correlated with AmoA, AmoB, and AmoC in BN-PAGE bands from *N. viennensis* in this study.

* Genes syntenic with *amoA*, *amoB*, and *amoC* based on the syntenic analysis from this study.

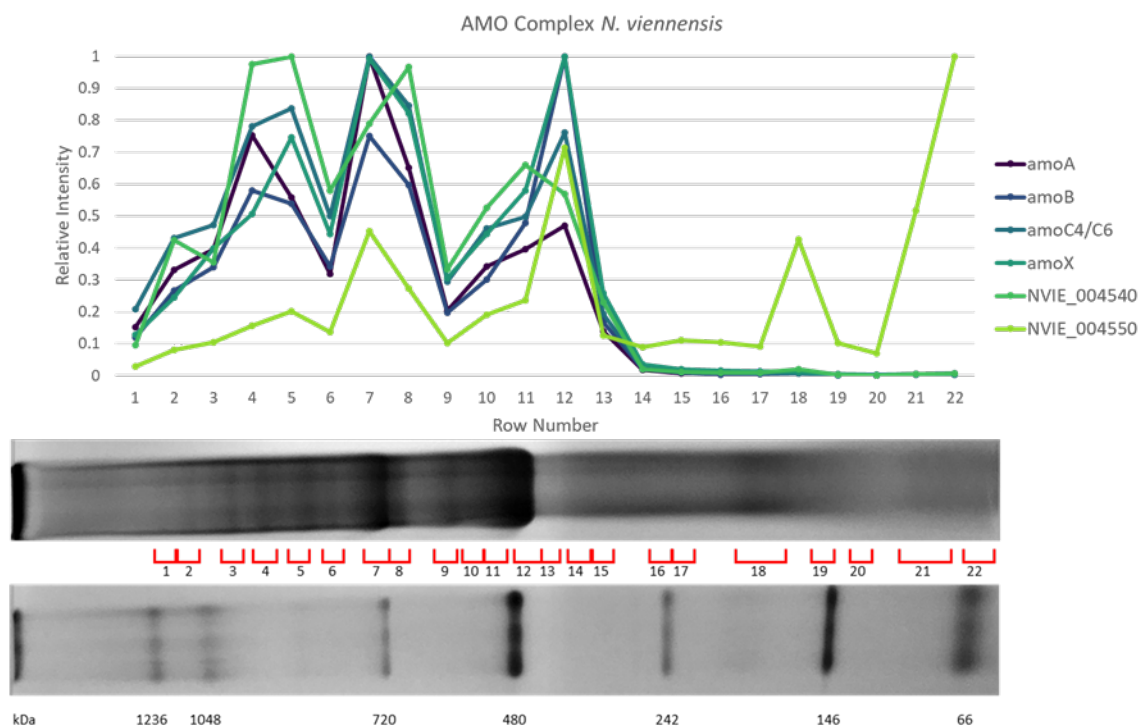
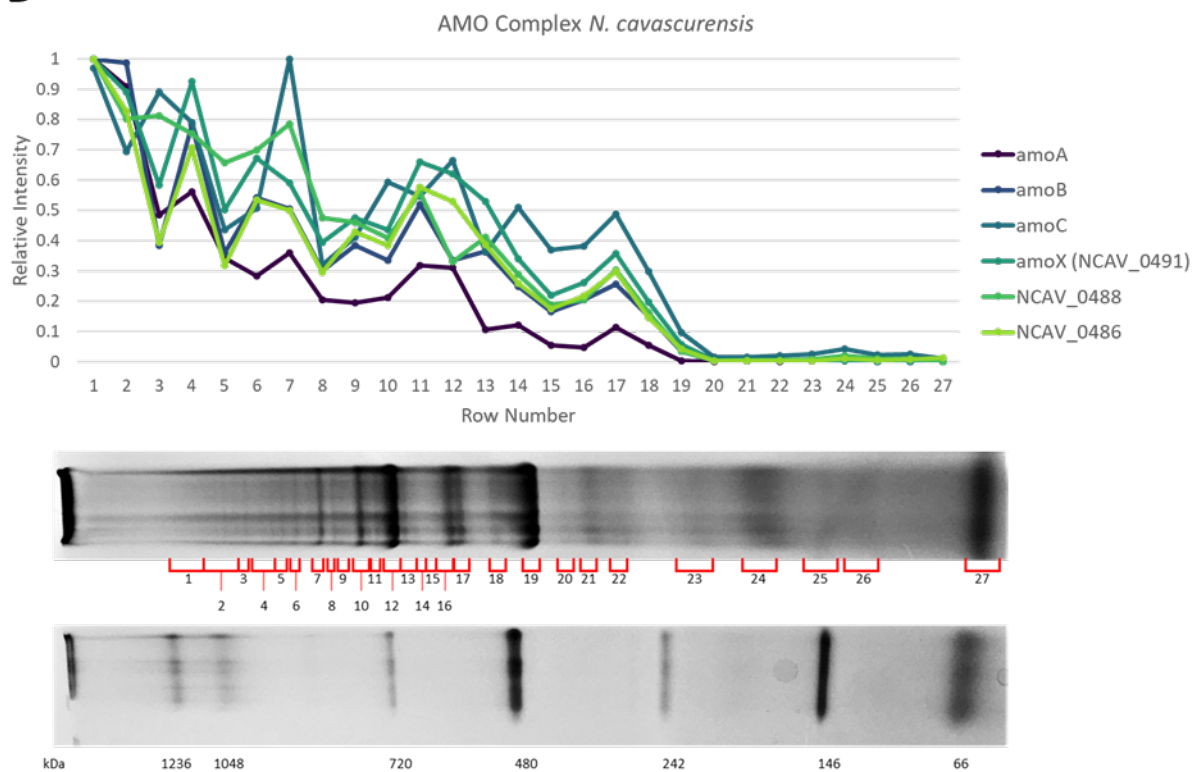
A**B**

Figure 2 (previous page): Genomic comparison of AMO subunit synteny in AOA. The left side of the figure represents a phylogenetic tree of AOA based on 32 conserved ribosomal proteins. Blue circles represent bootstrap values of 100%. Taxonomic labels are colored according to GTDB family identity: *Nitrosocaldaceae*-red, *Nitrosopumilaceae*-blue, *Nitrososphaeraceae*-orange. Labels in bold were included in syntenic analysis. Labels in red represent species with proteomic evidence from BN-PAGE gels. Separate contigs are represented by a double forward slash. Gaps between genes on the same contig are marked by a zig-zag line. Added information next to the phylogenetic tree includes: (A) all six AMO subunits found (full circle); (B) AMO subunits found on the same contig (full circle); (C) multiple *amoC* copies found within the genome (full circle); (D) different genera indicated by different colors: *Nitrosothermus* (bright red), *Nitrosocaldus* (pink), *Nitrosotenuis* (sea blue), *Nitrosopelagicus* (aqua), *Cenarchaeum* (light purple), *Nitrosoarchaeum* (light blue), *Nitrosospongia* (dark purple), *Nitrosopumilus* (dark blue), *Nitrosotalea* (green), *Nitrosocosmicus* (yellow), *Nitrososphaera* (orange). With the exception of *Nitrosothermus* sp., species with AMO genes on different contigs could still plausibly fit into the shown arrangements. The sponge species *Ca. Nitrosospongia ianthellae* and *Ca. Nitrosopumilus* sp. LS AOA, as well as *Nitrosoarchaeum limnia* SFB1 differ from the other arrangements of marine AOA by having 1 gene inserted between *AmoY* and *AmoZ*.

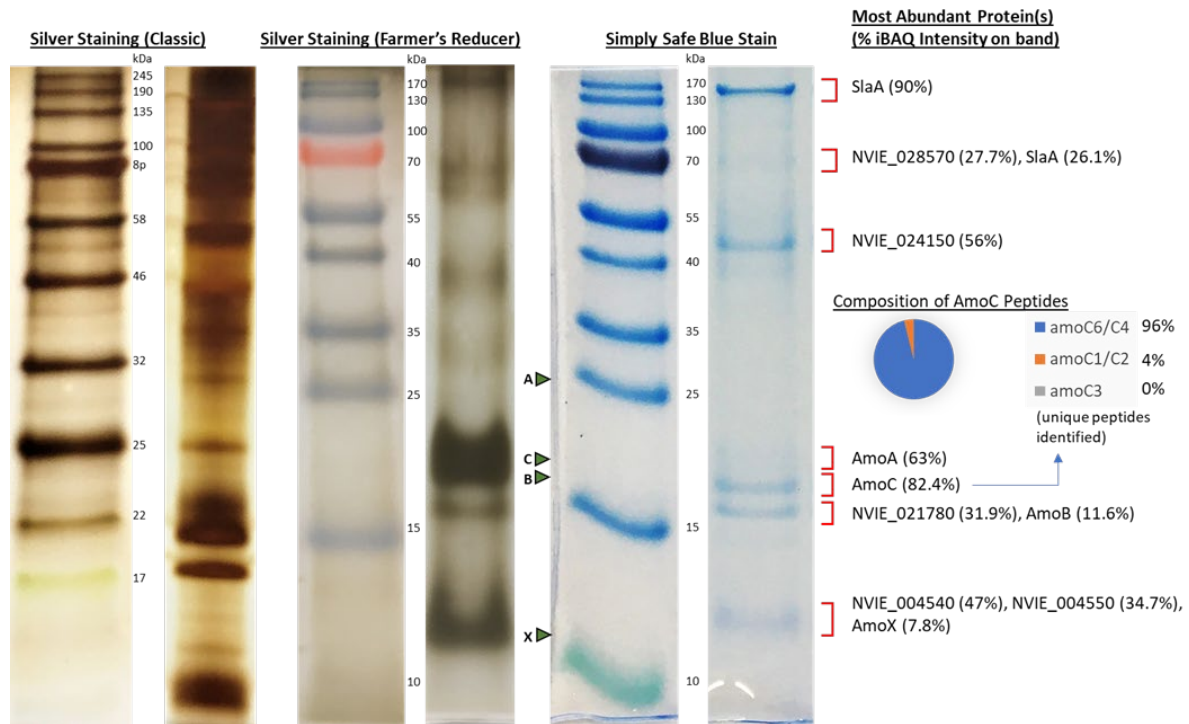


Figure 3: SDS-Tricine-PAGE of AMO bands from BN-PAGE gels. Comparison of three different staining methods for SDS-Tricine-PAGE gels. Bands cut for analysis from a gel stained with Simply Blue Safe Stain and digested using trypsin are indicated by red brackets. Percentages represent the percentage of iBAQ normalized protein intensities for each individual band. Green arrows marked A, B, C, and X represent expected heights of bands for AMO subunits AmoA, AmoB, AmoC, and AmoX respectively. A pie chart from the band with the highest amount of AmoC shows the percentage of AmoC bands coming from different distinguished AmoC homologs.

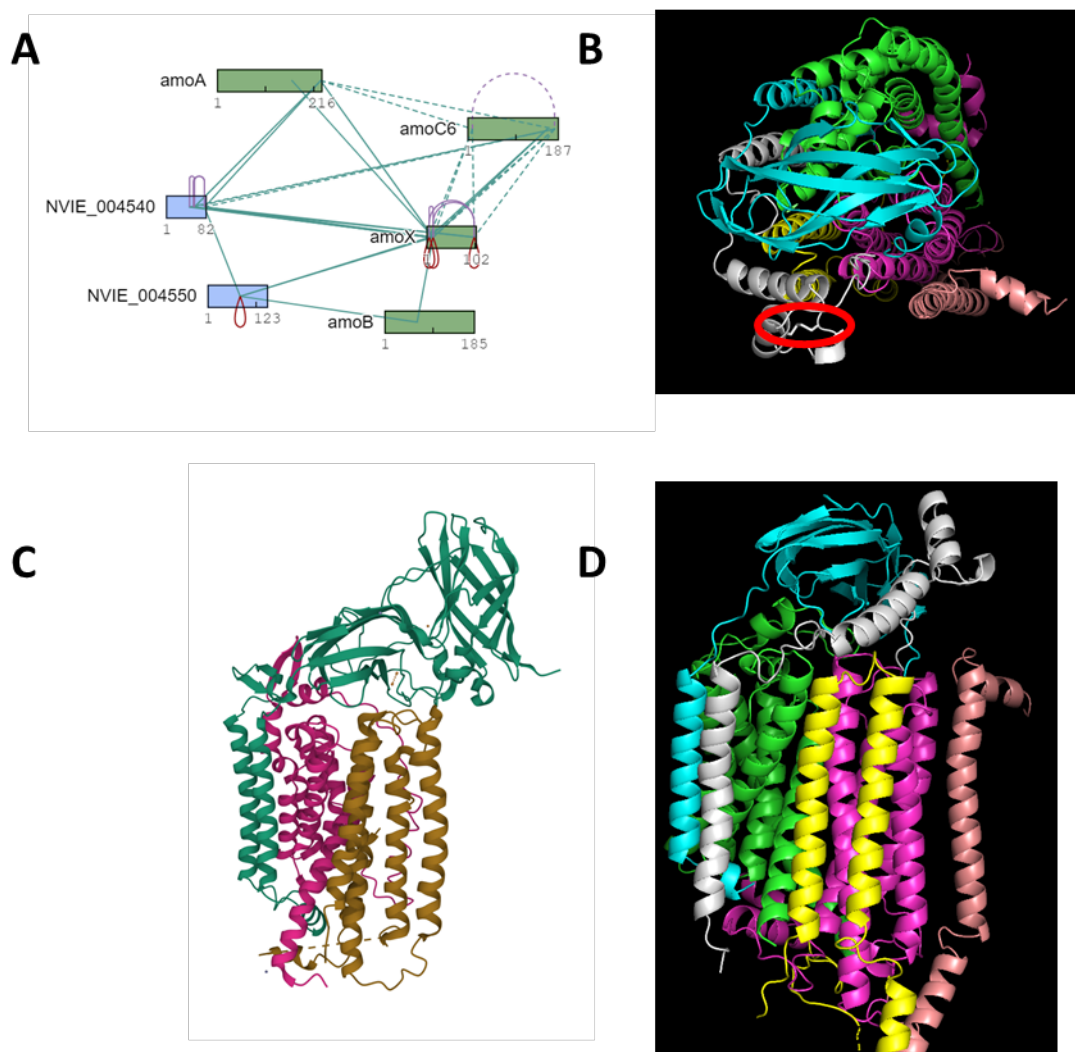


Figure 4: Structural support of proposed AMO subunits based on BN-PAGE cross-linking and Alpha Fold modelling. **A.)** Representation of identified cross-links among existing and proposed AMO subunits of an AMO band cut from a BN-PAGE gel of *N. viennensis*. Dotted lines are ambiguous, i.e. cannot be assigned to a specific AmoC homolog as the cross-linked peptide came from a shared region between AmoC homologs. Green: suspected subunits based on comparative genomics. Blue: newly proposed subunits based on BN-PAGE correlation and syntenic analysis. **B.)** Visualization of a predicted model of AMO in *N. viennensis* using Alpha fold version 2 in PyMol¹⁰⁹. Subunits are colored as: AmoA, green; AmoB, blue; AmoC, magenta; AmoX, yellow; AmoY (NVIE_004540), salmon; AmoZ (NVIE_004550) grey. Interactions in the soluble domain can be seen between AmoB and AmoZ. The red circle in D indicates a predicted disulfide bond in the soluble portion of AmoZ. **C.)** Representation of a PMO protomer from *Methylococcus capsulatus* str. Bath²² showing the configuration of PmoA (magenta), PmoB (green), and PmoC (brown). Image generated by Mol*¹¹⁰. Structure (code: 1YEW) was downloaded from the Protein Data Bank (PDB)¹¹¹. **D.)** Side view of the predicted protomer model of AMO in *N. viennensis*. Colors and visualization are the same as in (B). Comparison of (C) and (D) show strong similarities in the orientation of transmembrane helices and soluble domains while retaining some differences. Predicted interactions can be seen in (D) between soluble domains of AmoB and AmoZ.

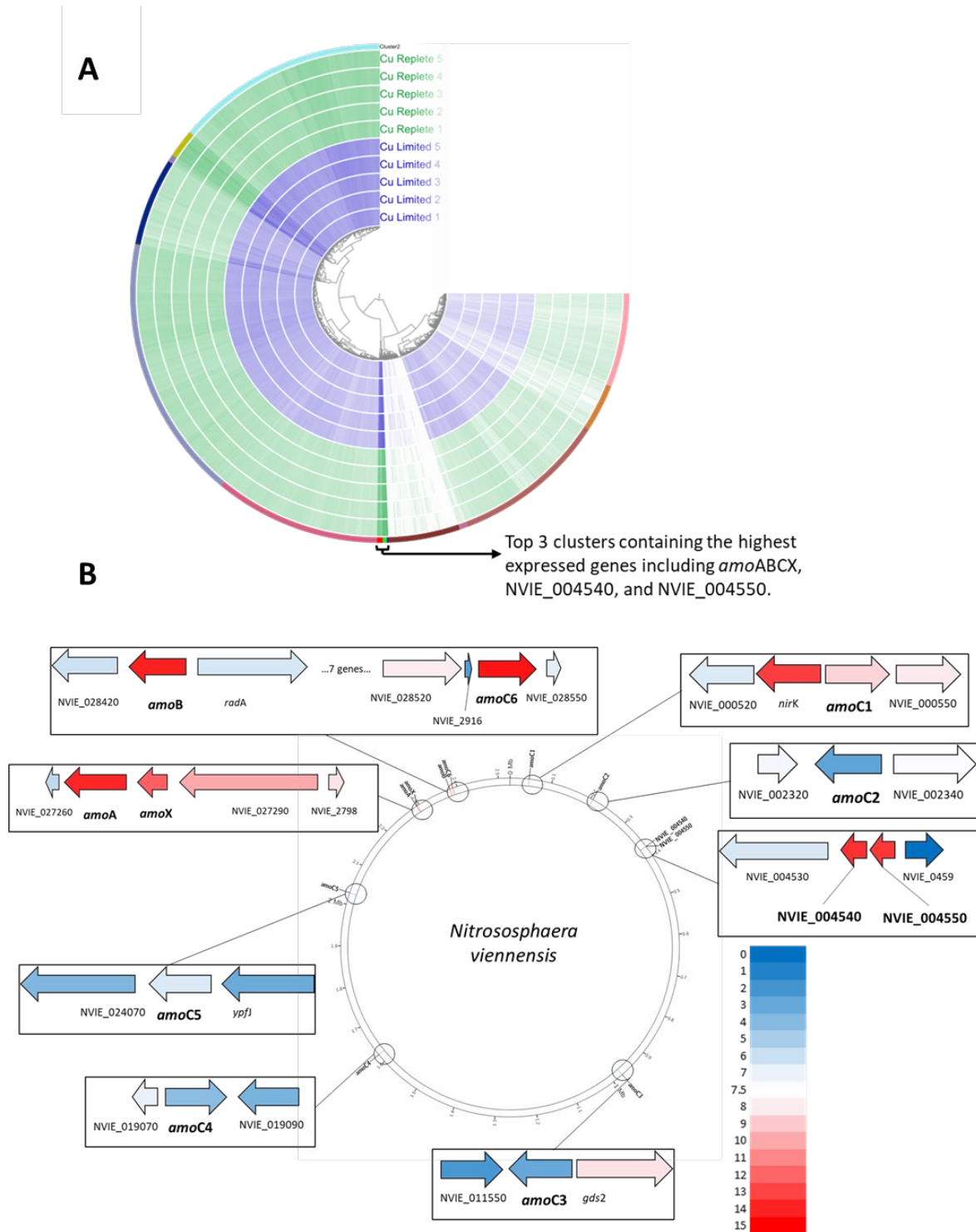


Figure 5: Transcriptomics analysis of AMO subunits in *N. viennensis*. A.) Clustering analysis of gene expression in *N. viennensis* under copper replete and limited conditions from Reyes et al. (2020). Genes were clustered into 15 separate clusters represented by different colors on the outer edge of the circular plot. B.) Genomic representation of *N. viennensis* showing location of AMO subunits. Boxes show AMO genes and immediate neighbors colored based on gene expression from copper replete cultures.

Supplementary Figures

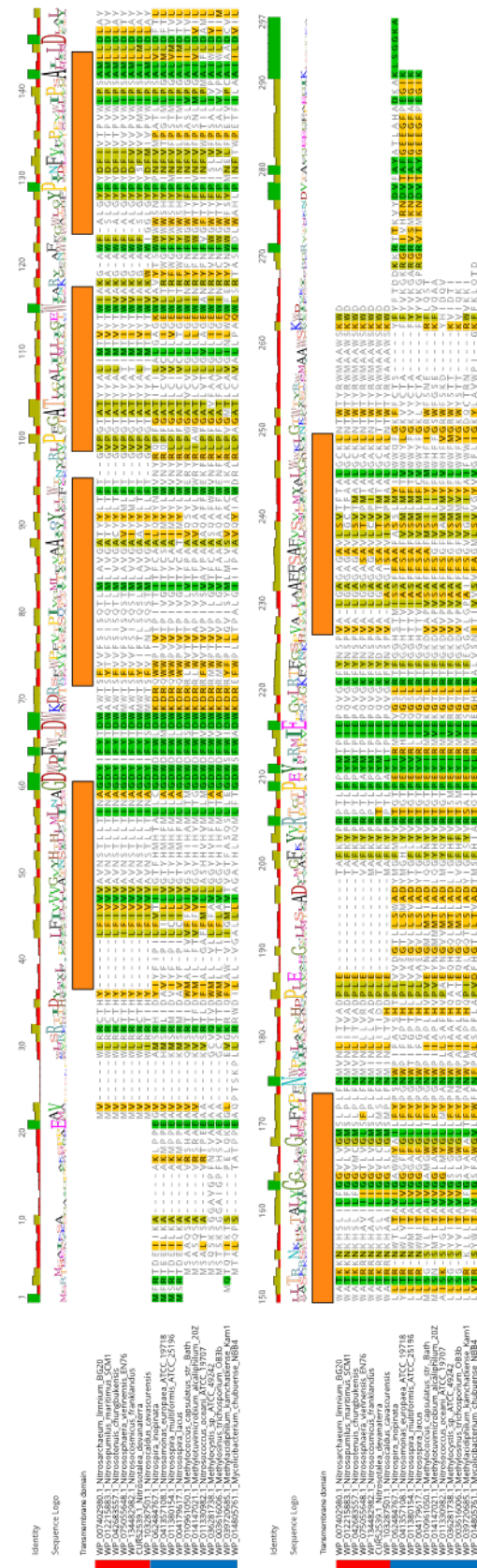
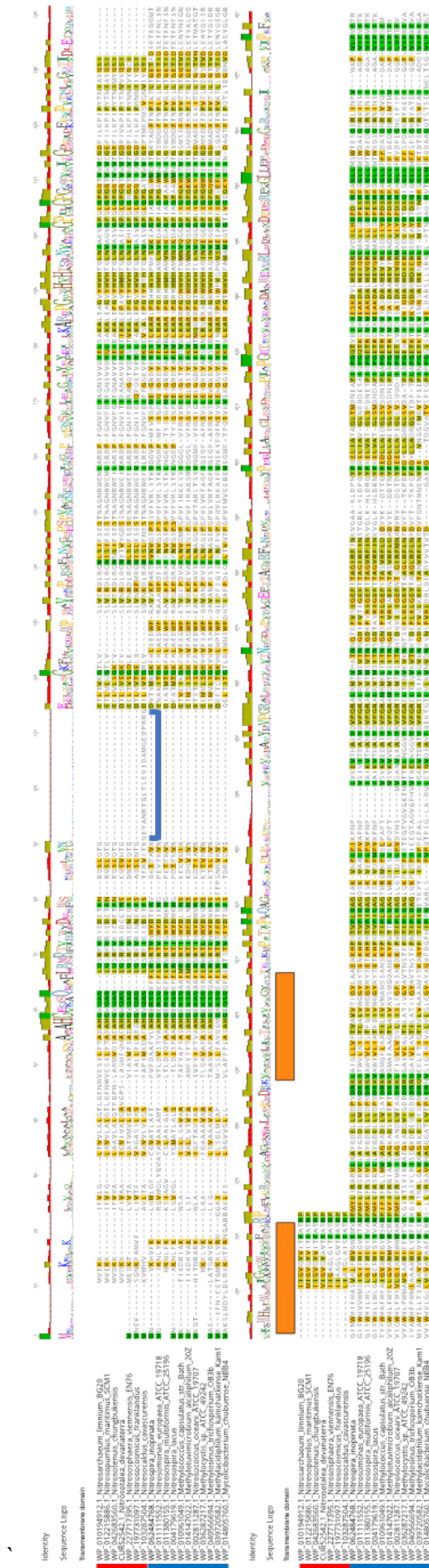


Figure S1: Protein alignments of PmoA/AmoA subunits from selected species. Percent identity of amino acids across species are indicated. Sequence logo visually represents the proportion of amino acids at each position. Orange bars represent transmembrane helices as found in the crystal structure of PMO from *Methylosinus trichosporium* OB3b (Hakemian et al. (2008)). Red bars represent archaeal species and blue bars represent bacterial species.



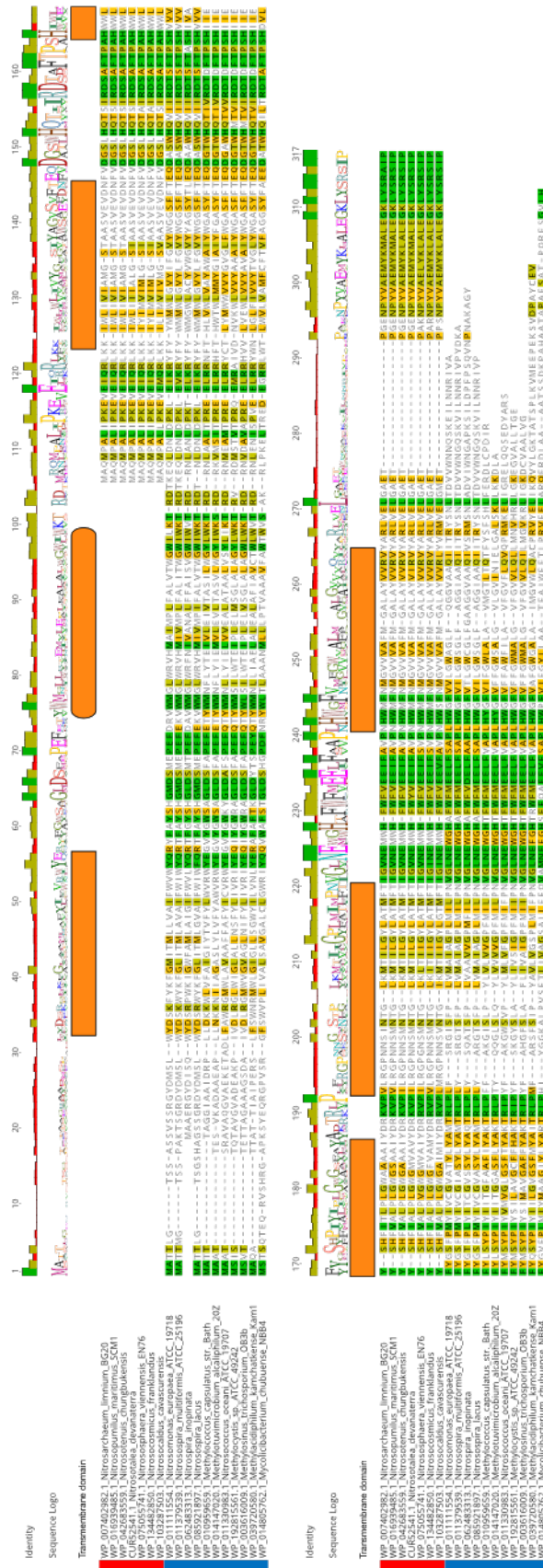


Figure S3: Protein alignments of PmoC/AmoC subunits from selected species. Percent identities of amino acids across species is indicated. Sequence logo visually represents the proportion of amino acids at each position. Orange bars represent transmembrane helices as found in the crystal structure of PMO from *Methylosinus trichosporium* OB3b (Hakemian et al. (2008)). Red bars represent archaeal species and blue bars represent bacterial species.

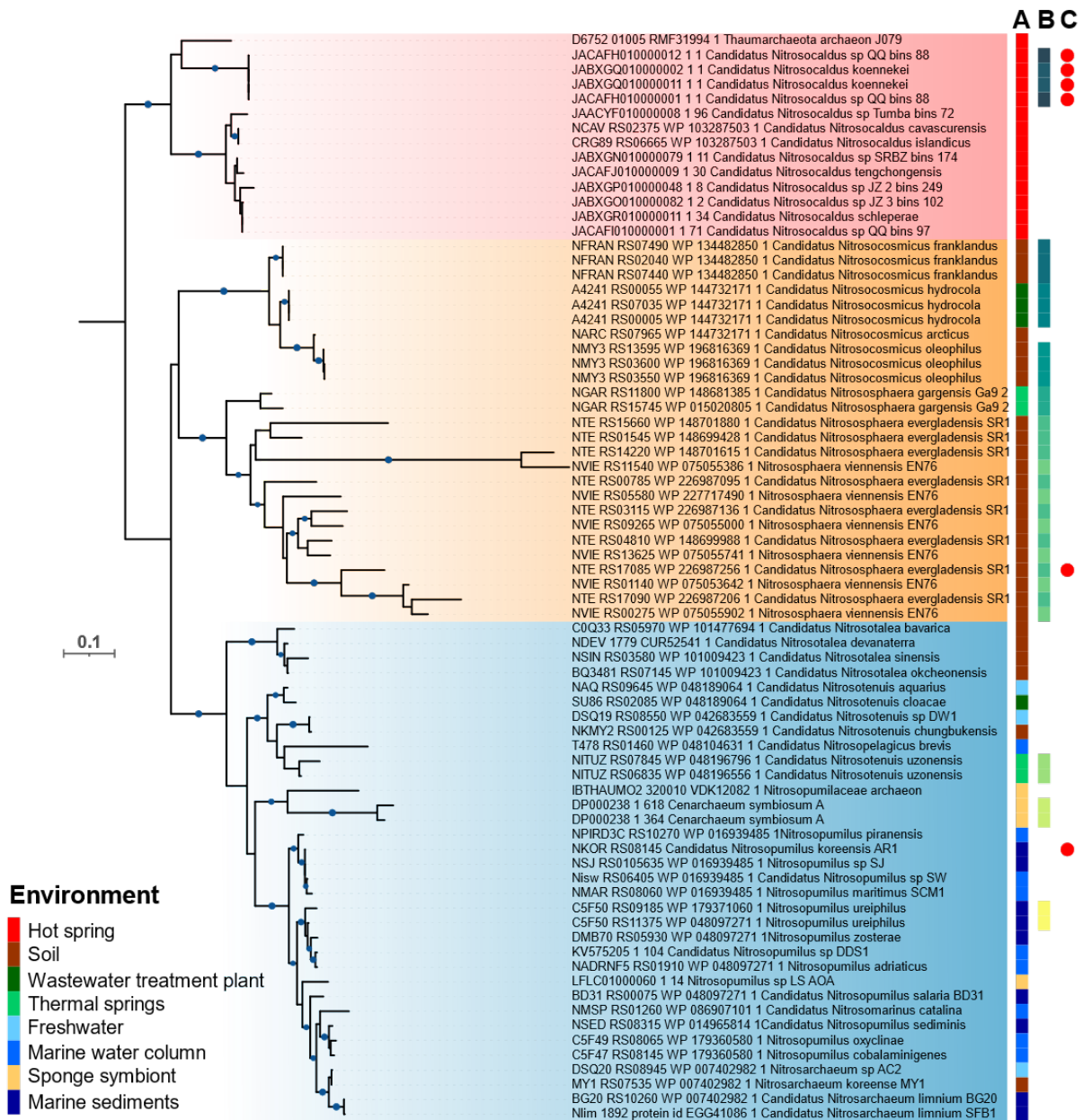


Figure S4: Phylogenetic tree of *amoC* genes in ammonia oxidizing archaea. Labels are colored according to GTDB families. Blue circles represent bootstrap values $\geq 85\%$. Duplications are colored according to species. Added information next to the phylogenetic tree includes: (A) environment; (B) *amoC* duplications represented by different colors for each specie; (C) the *amoC* shown is a fragment (full circle).

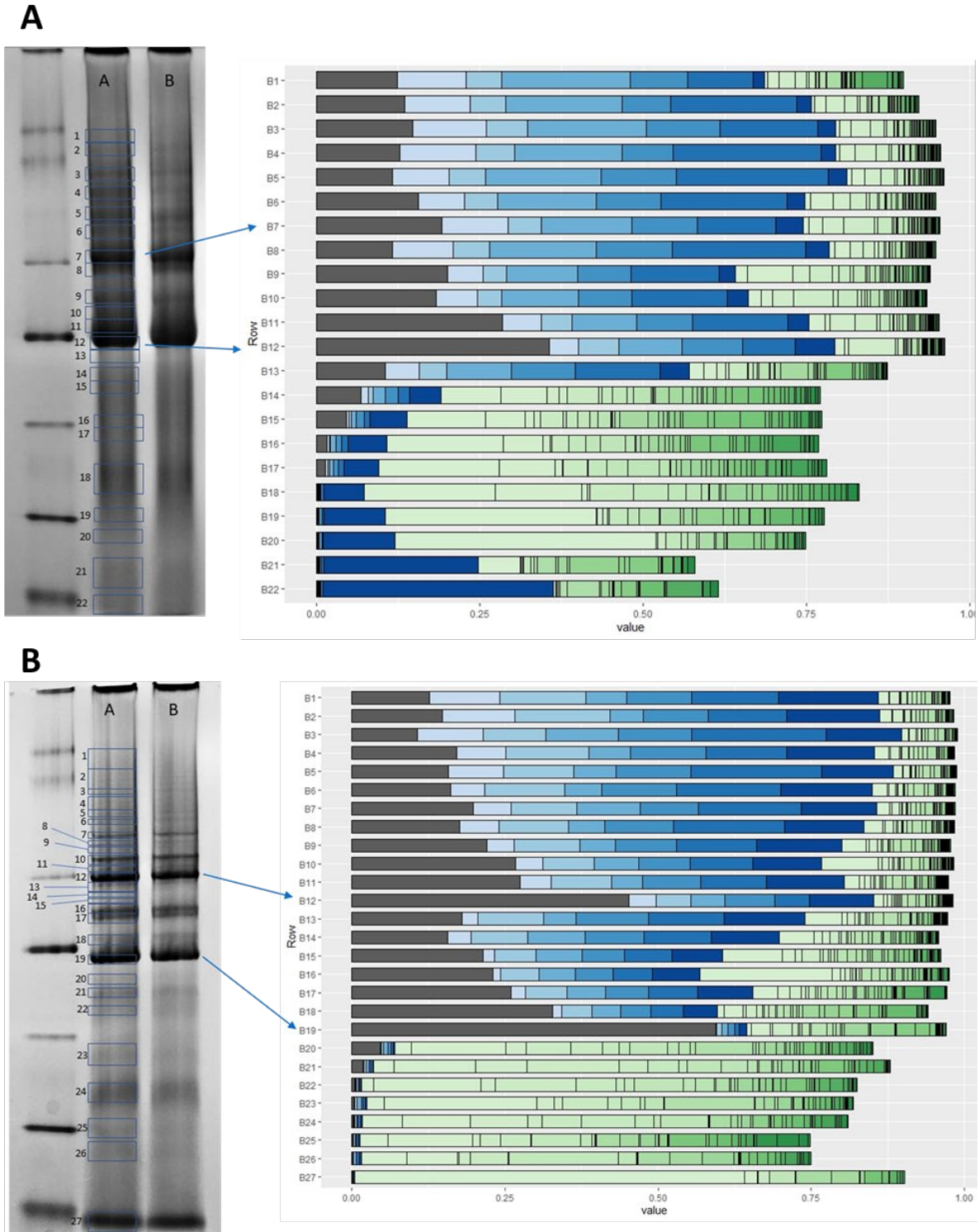


Figure S5: Bar charts showing the primary composition of normalized iBAQ intensities for each band of BN-PAGE based on the top 50 identified proteins across samples. Percent of total intensity is represented by the x-axis. Grey bars represent contribution from S-layer protein intensities (SlaA). Blue: intensities contributed from AMO subunits (AmoABCXYZ). Green: intensities from any other protein aside from SlaA and AmoABCXYZ. A.) *N. viennensis*. B.) *N. cavascurensis*.

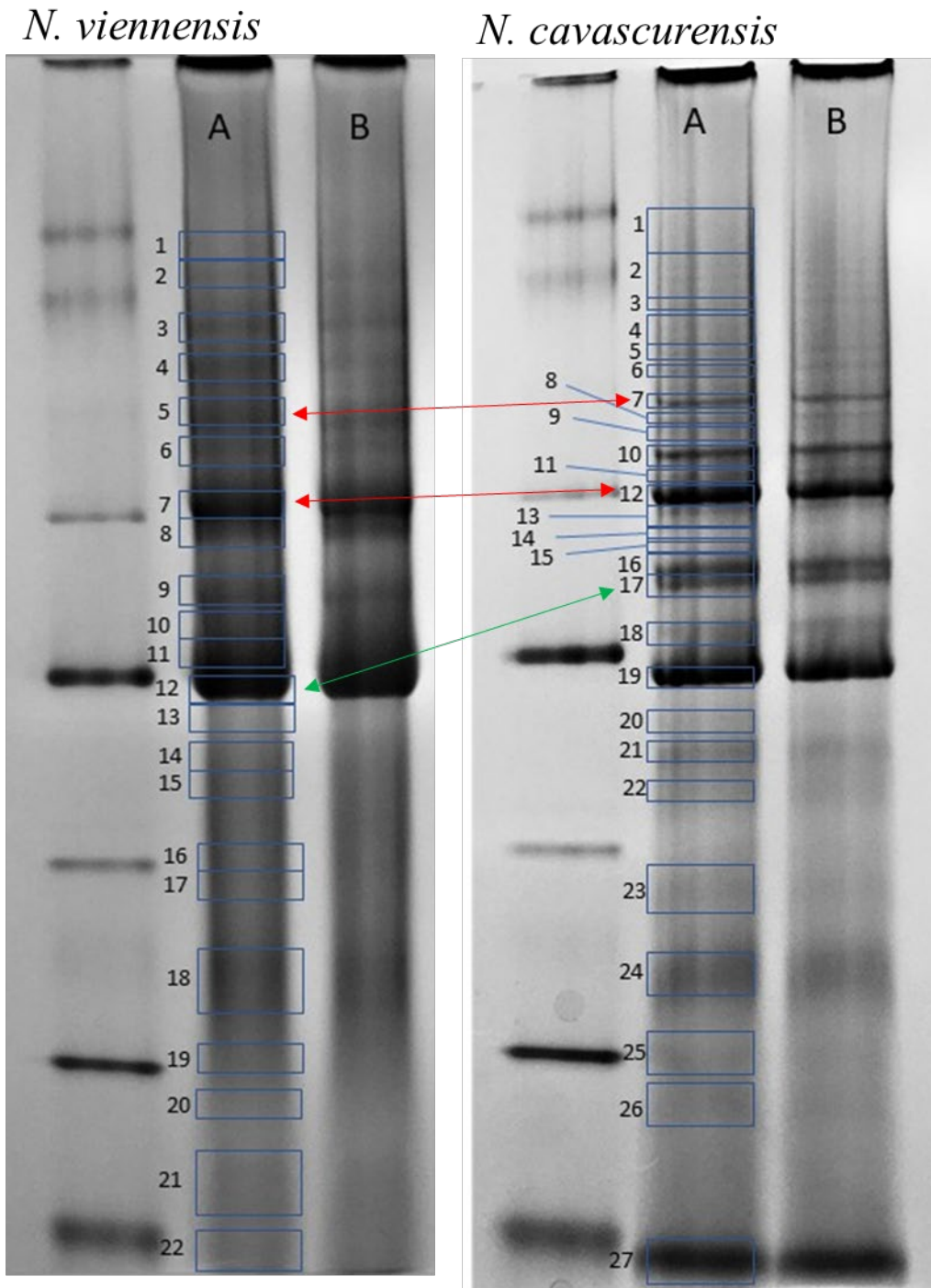
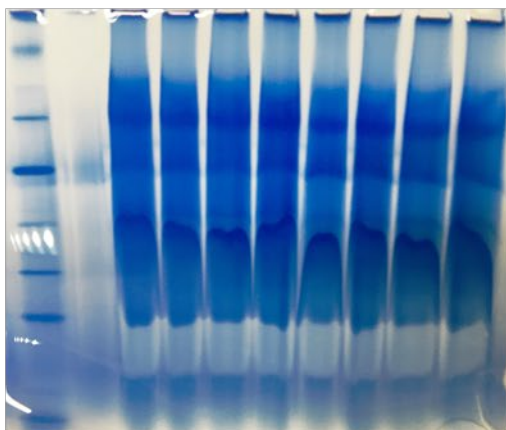


Figure S6: BN-PAGE gels of *N. viennensis* and *N. cavascurensis*. Comparison of protein membrane fractions between two species on BN-PAGE gels. Ladder and gels are the same. Red arrows indicate bands of high AMO content that correspond between the two species. Green arrow points to a third AMO peak in both species that is found at different heights. A and B represent identical gel and samples conditions in each gel.

N. viennensis

Extraction Yield Improvements

Date	Biomass (mg)	Protein (μg)	Approx. Yield (μg/mg)
02.03.2021	1022.9	1356.8	1.33
17.03.2021	728	1092	1.50
24.03.2021	207.3	410	1.98
01.04.2021	531	2042.04	3.85
16.04.2021	454.8	1977.75	4.35

Improved parameters:

- extraction parameters (buffers, lysis, centrifuge, etc.)
- extraction buffers
- DDM concentration
- running gel temperature
- Coomassie concentration
- change of running buffers
- protein loading
- gel staining

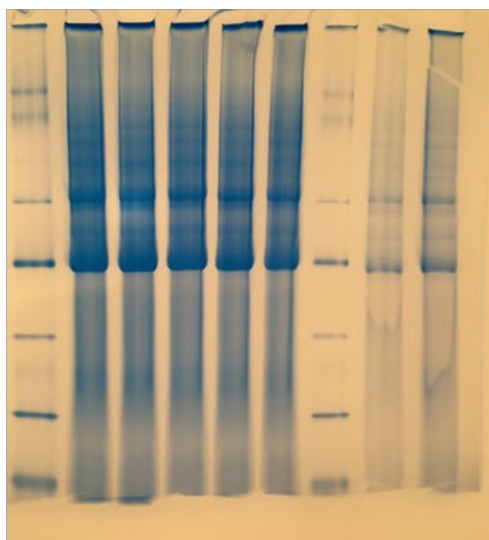
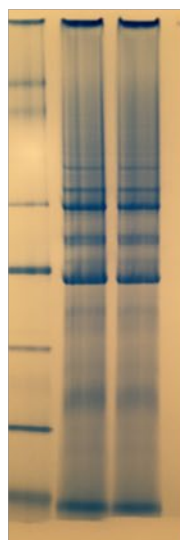
N. viennensis*N. cavascurensis*

Figure S7: Improvement of BN-PAGE gels of membrane fractions from ammonia oxidizing archaea. Top gel represents initial BN-PAGE gel from *N. viennensis* following published protocols. Subsequent modification of the listed parameters to improve the gel as seen in bottom left. Bottom right represents an initial gel of *N. cavascurensis* utilizing the parameters modified through testing in *N. viennensis*.

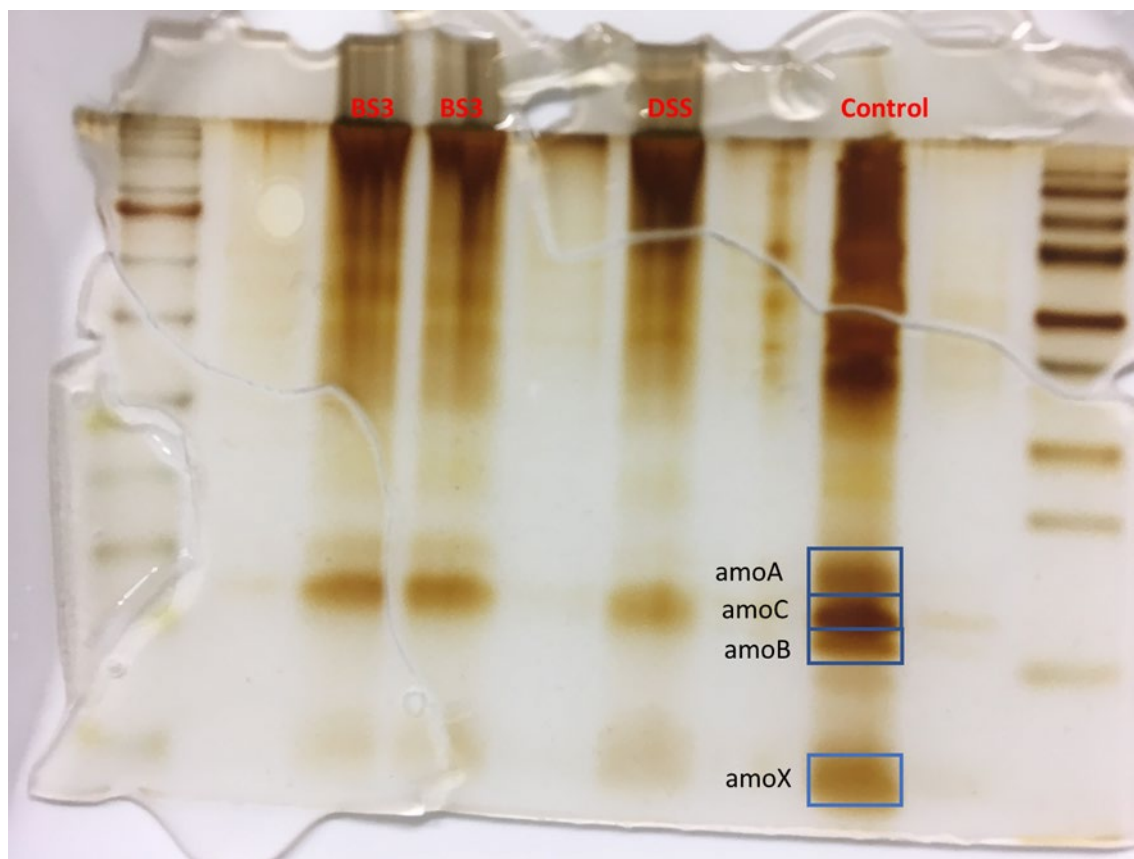


Figure S8: Proof of concept for in-gel cross-linking of BN-PAGE bands from *N. viennensis*. The SDS-Tricine-PAGE gel shows the resulting bands of proteins from BN-PAGE AMO bands treated with cross-linkers (BS3 and DSS) compared to an untreated band (control). Bands for AmoA, AmoB, AmoC, and AmoX (which was found to include AmoY and AmoZ) are indicated by blue boxes. Upon treatment by cross-linkers, the prominent AMO bands clearly become fainter. The cross-linker used in the manuscript, DSSO, is a mass spec. cleavable analog of DSS and should therefore be comparable. DSSO was chosen to allow for better identification due to its cleavability within the mass spectrometer.

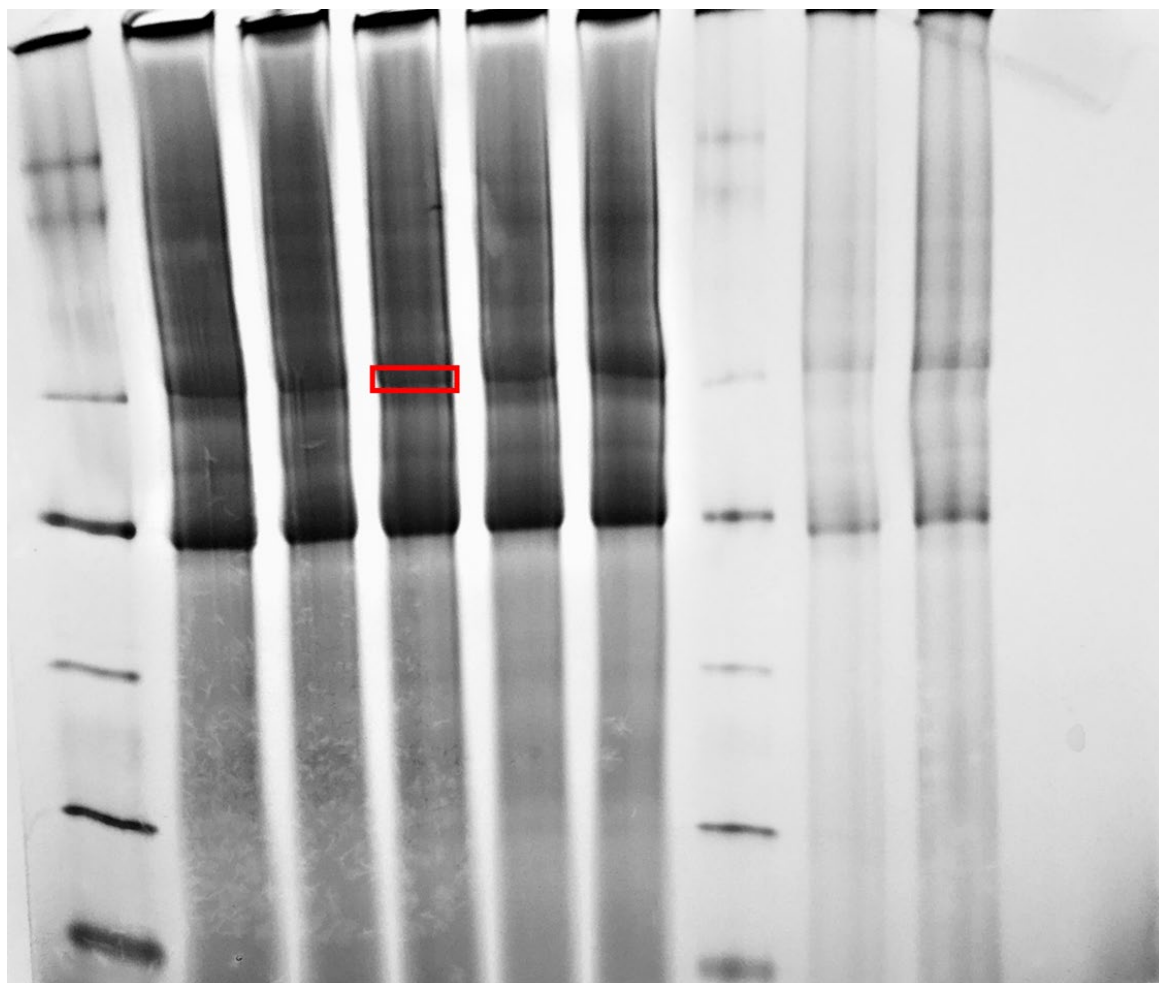


Figure S9: BN-PAGE gel of *N. viennensis* used for in-gel cross linking. The red box indicates the cut-out used for in-gel cross-linking with DSSO and subsequent mass spectrometry analysis. This cut-out corresponds to band 7 within Figure 1.

Chapter III

Genome Wide Transcriptomic Analysis of the Soil Ammonia Oxidizing Archaeon *Nitrososphaera viennensis* upon Exposure to Copper Limitation

Carolina Reyes^{1,2,3,†}, Logan H. Hodgskiss^{2,3,†}, Melina Kerou^{2,3}, Thomas Pribasnik^{2,3}, Sophie S. Abby⁴, Barbara Bayer^{3,5,6}, Stephen M. Kraemer^{1,3}, Christa Schleper^{2,3}

Author affiliations:

¹ Department of Environmental Geosciences, Centre for Microbiology and Environmental Systems Science, University of Vienna, Althanstrasse 14, UZA2, 1090 Vienna, Austria

² Department of Functional and Evolutionary Ecology, Archaea Biology and Ecogenomics Unit, University of Vienna, Althanstrasse 14, UZA1, 1090 Vienna, Austria

³ Environmental Science Research Network (ESRN), Faculty for Geosciences, Geography and Astronomy, University of Vienna, Althanstrasse 14, UZA2, 1090 Vienna, Austria

⁴ University Grenoble Alpes, CNRS, Grenoble INP, TIMC-IMAG, 38000 Grenoble, France

⁵ Department of Limnology and Oceanography, Division of Bio- oceanography, University of Vienna, Althanstrasse 14, UZA1, 1090 Vienna, Austria

⁶ Present address: Department of Ecology, Evolution and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106-9620, USA

[†]CR and LH are co-first authors.

Published in the ISME Journal (2020) 14:2659–2674.



ARTICLE

Genome wide transcriptomic analysis of the soil ammonia oxidizing archaeon *Nitrososphaera viennensis* upon exposure to copper limitation

Carolina Reyes^{1,2,3} · Logan H. Hodgskiss^{2,3} · Melina Kerou^{2,3} · Thomas Pribasnig^{2,3} · Sophie S. Abby⁴ · Barbara Bayer^{3,5,6} · Stephan M. Kraemer^{1,3} · Christa Schleper^{2,3}

Received: 28 February 2020 / Revised: 9 June 2020 / Accepted: 2 July 2020 / Published online: 14 July 2020

© The Author(s) 2020. This article is published with open access

Abstract

Ammonia-oxidizing archaea (AOA) are widespread in nature and are involved in nitrification, an essential process in the global nitrogen cycle. The enzymes for ammonia oxidation and electron transport rely heavily on copper (Cu), which can be limited in nature. In this study the model soil archaeon *Nitrososphaera viennensis* was investigated via transcriptomic analysis to gain insight regarding possible Cu uptake mechanisms and compensation strategies when Cu becomes limiting. Upon Cu limitation, *N. viennensis* exhibited impaired nitrite production and thus growth, which was paralleled by downregulation of ammonia oxidation, electron transport, carbon fixation, nucleotide, and lipid biosynthesis pathway genes. Under Cu-limitation, 1547 out of 3180 detected genes were differentially expressed, with 784 genes upregulated and 763 downregulated. The most highly upregulated genes encoded proteins with a possible role in Cu binding and uptake, such as the Cu chelator and transporter CopC/D, disulfide bond oxidoreductase D (*dsbD*), and multicopper oxidases. While this response differs from the marine strain *Nitrosopumilus maritimus*, conserved sequence motifs in some of the Cu-responsive genes suggest conserved transcriptional regulation in terrestrial AOA. This study provides possible gene regulation and energy conservation mechanisms linked to Cu bioavailability and presents the first model for Cu uptake by a soil AOA.

Introduction

Ammonia-oxidizing archaea (AOA) perform the first and rate-limiting step of nitrification along with ammonia-

oxidizing bacteria (AOB) and comammox [1–3]. AOA are ubiquitous and abundant in most aerobic environments [4, 5]. Cultivated AOA strains belonging to the class *Nitrososphaeria* of the phylum Thaumarchaeota [6] represent some but not all of the currently recognized lineages of AOA based on environmental gene sequences [7].

A common feature of AOA is the central role of copper (Cu) in their metabolism. Copper is used as a cofactor to many key proteins in AOA including the ammonia monooxygenase (AMO), which performs the energy-gaining step

These authors contributed equally: Carolina Reyes, Logan H. Hodgskiss

Supplementary information The online version of this article (<https://doi.org/10.1038/s41396-020-0715-2>) contains supplementary material, which is available to authorized users.

- ✉ Carolina Reyes
creyes6@gmail.com
- ✉ Christa Schleper
christa.schleper@univie.ac.at

- ¹ Department of Environmental Geosciences, Centre for Microbiology and Environmental Systems Science, University of Vienna, Althanstrasse 14, UZA2, 1090 Vienna, Austria
- ² Department of Functional and Evolutionary Ecology, Archaea Biology and Ecogenomics Unit, University of Vienna, Althanstrasse 14, UZA1, 1090 Vienna, Austria

- ³ Environmental Science Research Network (ESRN), Faculty for Geosciences, Geography and Astronomy, University of Vienna, Althanstrasse 14, UZA2, 1090 Vienna, Austria
- ⁴ University Grenoble Alpes, CNRS, Grenoble INP, TIMC-IMAG, 38000 Grenoble, France
- ⁵ Department of Limnology and Oceanography, Division of Bio-oceanography, University of Vienna, Althanstrasse 14, UZA1, 1090 Vienna, Austria
- ⁶ Present address: Department of Ecology, Evolution and Marine Biology, University of California Santa Barbara, Santa Barbara, CA 93106-9620, USA

of ammonia oxidation, plastocyanins, which are encoded in high numbers by AOA and used presumably as electron carriers, and multicopper oxidases, whose role although conserved in AOA remains unclear [8, 9]. While the AMO of AOB also uses Cu as a cofactor, mostly heme-containing proteins such as hydroxylamine oxidoreductase (HAO) and cytochrome C_{M552} (CycB) make up the bacterial ammonia oxidation electron transport chain [10]. AOA, in contrast, lack homologs for both HAO and CycB. The enzyme replacing HAO in AOA has yet to be determined, while Cu containing plastocyanins are hypothesized to act as electron carriers in place of CycB [8]. Therefore, AOA should depend more highly on Cu containing enzymes than their bacterial counterparts, which would necessitate elaborate regulatory mechanisms to respond to rapidly changing local concentrations of this metal. Indeed, the marine AOA *Nitrosopumilus maritimus* SCM1, with Cu adsorbed on its surface, exhibited exothermic enthalpies in contrast to neutral or endothermic enthalpies for AOB, suggesting a higher surface selectivity of AOA for Cu [11].

The abundance of Cu in the Earth's crust is significantly lower than that of other biologically important metals [12]. In terrestrial environments, its bioavailability is influenced primarily by organic matter content and pH [13]. Recently, Gwak et al. [14] showed that mixtures of organic compounds like yeast extract depleted Cu in culture experiments with AOA and AOB. However, only AOA growth was inhibited by reduced Cu bioavailability. These results support the idea that Cu is a more important trace element for AOA than AOB.

While the mechanisms by which different AOA clades take up Cu from their environment are not well understood, genomic analysis of the transporter complement of six different AOA, including the soil archaeon *Nitrososphaera viennensis* (EN76) [15, 16], showed that orthologs of multicopper oxidases and ZIP family permeases are conserved among AOA and could have a role in making Cu more bio-available and in Cu uptake respectively [9]. In addition, all known AOA harbor proteins belonging to the copper resistance CopC and CopD families, typically involved in Cu resistance mechanisms [17–20]. In general, Cop proteins are suggested to import Cu into cells [21, 22], to export Cu out of cells [23], and to bind Cu in the periplasm [24].

The effects of Cu-limitation on cell physiology, metabolism, and uptake have been studied on pure cultures of *N. maritimus* SCM1 [25, 26] and *N. viennensis* (EN76) [13]. These studies support the idea that Cu is essential for AOA and show that as Cu concentration becomes limiting, ammonia oxidation activity and cell growth decrease. While Qin et al. [26] provides a first outline to the Cu-limitation transcriptomic response in marine AOA, the corresponding mechanisms in terrestrial lineages remain

unexplored. Terrestrial AOA generally have larger genomes that have evolved under different selection pressures and exhibit different capabilities than the streamlined marine AOA [9].

The aim of this study was to identify genes potentially involved in Cu-uptake and to identify physiological and regulatory responses in *N. viennensis* under Cu-limiting conditions. As such, a genome-wide transcriptomic analysis of *N. viennensis* in response to Cu-limiting conditions was performed and is presented along with supporting proteomic and quantitative PCR results. While the reaction of *N. viennensis* markedly differs from that of *N. maritimus* with respect to possible Cu-sequestration, conserved regulatory motifs in downregulated genes show conservation among related soil strains.

Methods

Detailed cultivation procedures, measurement of ammonia oxidation, RNA extraction, protein extraction, quantitative reverse transcription PCR (RT-qPCR), sequencing methods, sequencing analysis, proteomic methods and analysis, motif analysis, and phylogenetic analysis are described in the Supplementary material.

Cu-limiting conditions for differential gene expression

The Cu chelator TETA (1,4,8,11-tetraazacyclotetradecane N, N', N'', N''' -tetraacetic acid hydrochloride hydrate) with selective affinity for Cu^{2+} was used to induce Cu limitation as previously described [13]. Five replicate 500 mL cultures were each inoculated either with a Cu-limited (Cu12) or a standard *N. viennensis* culture to establish Cu-limited and Cu-replete cultures respectively. It was estimated using the PhreeqC program [27] that Cu-limited cultures Cu10 and Cu12 contained 6×10^{-15} mol/L and 8×10^{-16} mol/L of bioavailable Cu^{2+} , respectively [13].

RNA and proteome analysis

Five replicate cultures of each growth condition of *N. viennensis* were harvested by filtration followed by extraction of RNA from the filters. Extracted RNA (cDNA) of each sample was sequenced by the Vienna Biocenter Facility [VCBF], Vienna, Austria, using an Illumina HiSeq2500 instrument. Raw read files were processed and then analyzed for differential expression using DESeq2 [28]. The Benjamini Hochberg DESeq2 default step was used to control for false discovery rate (FDR). Genes with an adjusted p value (p -adj) < 0.01 were considered differentially expressed.

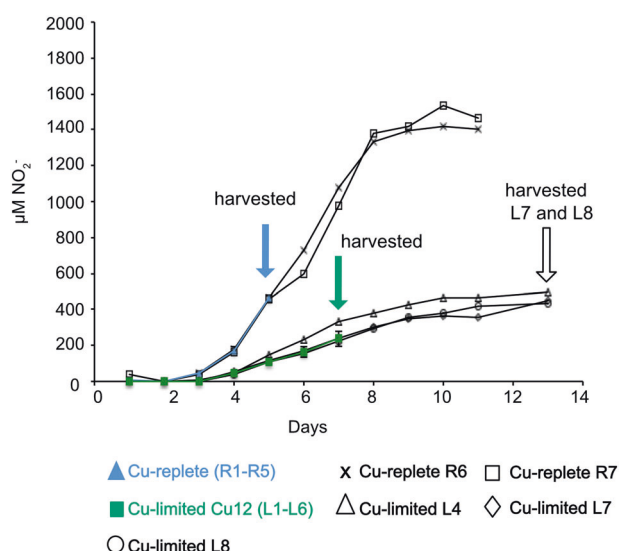


Fig. 1 Nitrite concentrations in batch cultures of *N. viennensis* grown in 500 mL in 1 L polystyrene bottles under Cu-replete and Cu-limited (Cu12) conditions. Cu replete cultures, labeled (R1–R5) refer to the five biological replicates R1, R2, R3, R4, and R5 that were harvested on day 5. Cu limited cultures, labeled (L1–L3, L5–L6) refer to the five biological replicates as L1, L2, L3, L5, and L6 that were harvested on day 7. Solid arrows indicate when cultures were harvested for RNA extraction. Open arrow indicates when cultures were harvested for protein extraction. The data and error bars for Cu-replete and Cu-limited cultures represent the average and standard deviation of biological replicates.

Proteomic and RT-qPCR analyses were performed on two additional Cu-limited cultures (L7 and L8) that were harvested on Day 13 (Fig. 1). While the proteomes from two Cu-limited cultures verified the presence of multicopper oxidases, further critical analysis of the proteomic data was not carried out due to the lack of more biological replicates and the lack of proteome data from the control cultures.

All raw sequence runs from the RNA-Seq analysis were deposited in the DNA Databank of Japan (DDBJ) under the accession number DRA008627. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [29] partner repository with the dataset identifier PXD018461 and 10.6019/PXD018461.

Results and discussion

Growth response under Cu-limiting conditions

Initial growth tests with *N. viennensis* were performed in the presence and absence of TETA in order to determine if this Cu chelator affects growth of the organism. TETA alone did not have a toxic effect on cell growth, as cultures supplemented with TETA and excess Cu produced nitrite

concentrations comparable to the Cu-replete cultures (Fig. S1). When TETA was applied to Cu-limited cultures, a decreased ability to oxidize ammonia compared to Cu-replete cultures was observed, while re-introduction of Cu to Cu-limited (Cu10) cultures restored their ability to produce nitrite comparable to the Cu-replete culture (Fig. S1) [13]. These results demonstrate that TETA can be used to chelate Cu and induce Cu-limitation in *N. viennensis* and that Cu-limited cells retain the ability to oxidize ammonia following Cu addition. Metabolic recovery following Cu starvation has also been observed in studies with *N. maritimus* [25], where cells that were starved of Cu for up to 6 days recovered within 24 h of Cu addition.

For the transcriptome analyses, under Cu-limiting versus Cu-replete conditions, five replicates each were grown in parallel. At the time cultures were harvested for RNA extraction, Cu-replete cultures had produced 436–493 μM nitrite and Cu-limited (Cu12) cultures had produced 202–334 μM nitrite (Fig. 1).

RNA sequencing analysis

RNA sequencing results are reported in Table S1. To analyse variation in gene expression between Cu-replete and Cu-limited conditions and between replicates of each condition a PCA analysis was performed. PCA results showed that the five replicates of each condition grouped together (Fig. S2) thus demonstrating high reproducibility. One principle component (PC1) could explain ~51 % of the variance between the Cu-limited and Cu-replete cultures. Comparison of protein-coding genes between Cu-limited and Cu-replete samples resulted in 1,547 differentially expressed genes out of 3,180 genes identified from reads mapped to the genome (Dataset S1) (Benjamini Hochberg adjusted p value cut-off of 0.01). After careful evaluation of the data a p value cut-off of 0.01 was chosen instead of the more commonly used 0.05 to define differentially expressed genes, which allowed for a better focus of the results of the DESeq2 analysis. Because of the low adjusted p values and therefore high statistical reproducibility, genes meeting this criteria were considered even if they showed small changes (in terms of the $\log_2$ fold change) between the two conditions. This is in line with previous archaeal studies that have identified and/or observed genes of physiological significance with a focus on a p value cut-off irrespective of the low $\log_2$ fold changes [30, 31]. Of the differentially expressed genes, 784 were upregulated (24.7 % of detected genes) and 763 were downregulated (24.0% of detected genes) in Cu-limited cultures. Overall, these results showed that 50% of the differentially expressed genes were upregulated in response to Cu-limitation compared to 30% of genes upregulated in Cu-limited *N. maritimus* cultures [26]. Next, to gain insight into the general biological function of

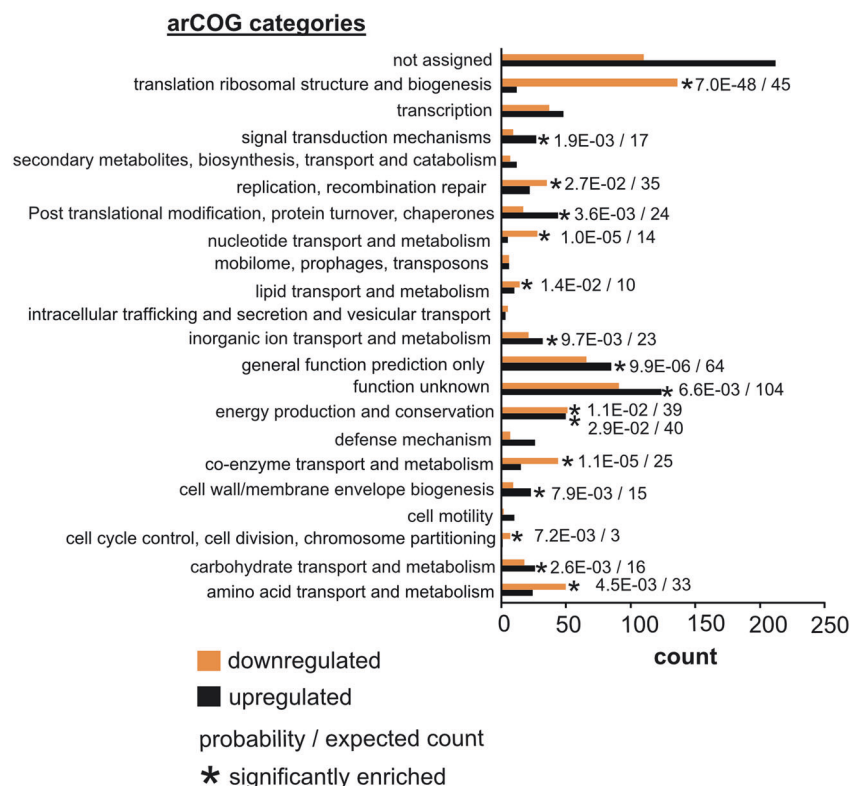


Fig. 2 Bar charts showing up and downregulated genes assigned to functional groups in Cu-limited samples. Genes were assigned to functional categories based on their archaeal clusters of orthologous

group (arCOG) categories. A p value < 0.05 cutoff was used to determine significance (see Supplemental Methods).

differentially expressed genes in Cu-limited samples, a hypergeometric test was used to determine which arCOG categories were significantly enriched (p value < 0.05) in up and downregulated groups. The hypergeometric test revealed enriched arCOG functional categories within the upregulated group including: inorganic ion transport and metabolism; signal transduction mechanisms; and post-translational modification, turnover, and chaperones (Fig. 2). In the downregulated group, enriched categories included: nucleotide transport and metabolism; replication, recombination and repair; and translation/ribosomal structure and biogenesis (Fig. 2). An enrichment of the inorganic ion transport and metabolism category within the upregulated group suggests that the Cu-limited cultures directly responded to Cu limitation on a transcriptional level. Furthermore, the enrichment of translation, replication, and nucleotide metabolism genes within the downregulated group denote a global downregulation of genes associated with growth. These results and those of the growth experiments, indicate that the cells were unable to effectively replicate and oxidize ammonia to nitrite, a proxy for growth [15], due to insufficient levels of Cu [26].

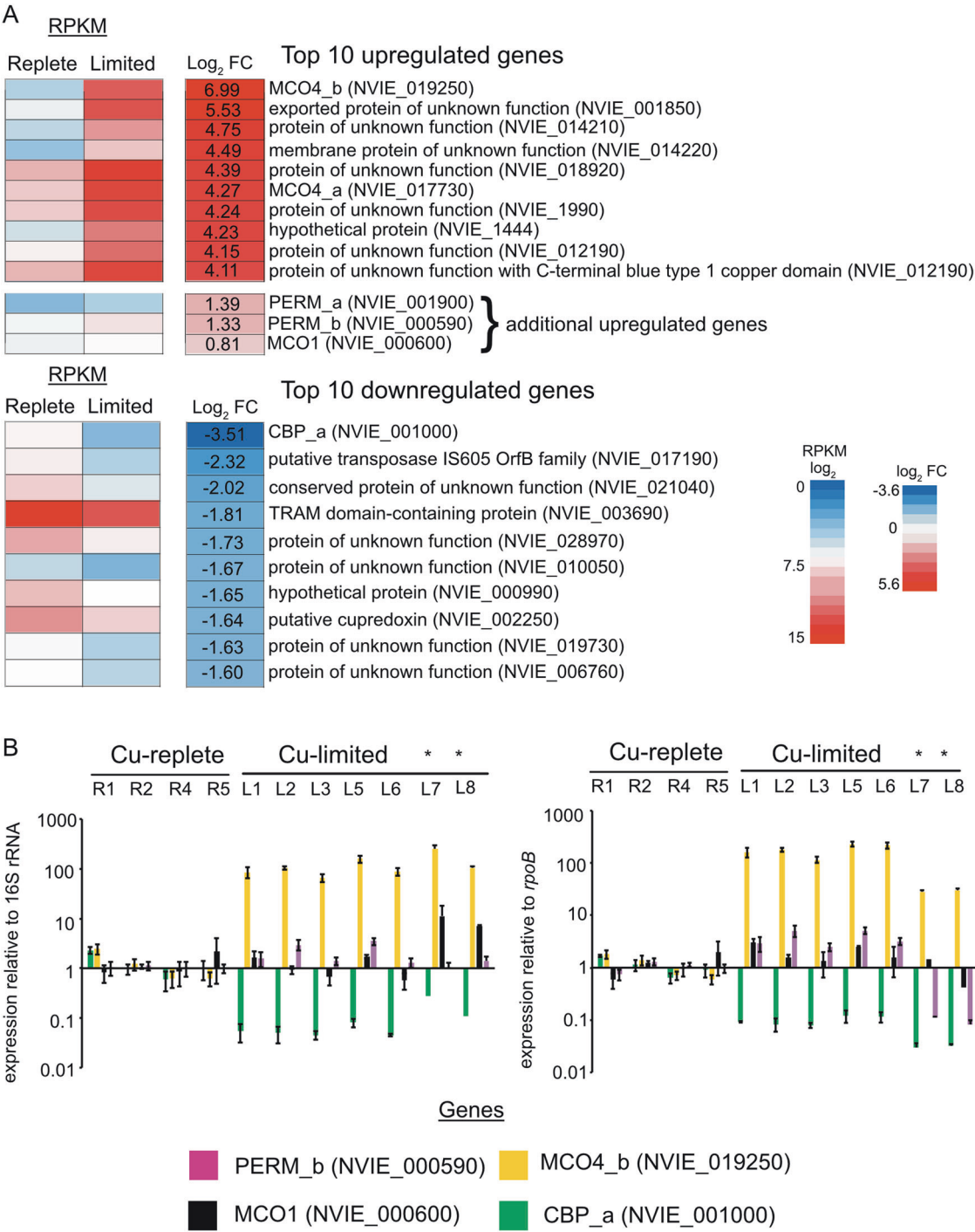
The top upregulated and downregulated genes are shown in Fig. 3a. In Cu-limited cultures, the multicopper oxidase

gene MCO4_b (NVIE_019250), was the most upregulated gene ($\log_2$ fold change (FC) of 6.99) and CBP_a (NVIE_001000), a gene encoding a putative surface-associated Ca^{2+} binding protein, was the most downregulated ($\log_2$ FC -3.51).

Genes encoding MCOs and associated proteins

Cu-limitation induced the expression of several proteins possibly involved in Cu uptake and transport. Previously, the two domain MCOs (2d MCOs) [32, 33] of *N. viennensis* were phylogenetically compared to other archaeal and bacterial 2d MCOs and were divided into four separate lineages (MCO1-4) [9]. MCOs belonging to lineage 1 (MCO1) were only found in thaumarchaeal genomes, whereas MCOs belonging to lineage 4 (MCO4) were only present in the order Nitrososphaerales and the genera *Nitrosopumilus* and *Nitrosarchaeum* [9]. MCOs of the type encoded by MCO1 and MCO4 genes contain three types of Cu centers (T1–T3), while MCO4 has an extra blue copper domain [9].

In this study, two highly upregulated MCO4 genes, MCO4_a (NVIE_017730; $\log_2$ FC 4.27) and MCO4_b (NVI_019250; $\log_2$ FC 6.99), and one moderately



* Duplicate cultures harvested on Day 13 and subjected to proteomic and RT-qPCR analyses.

Fig. 3 **Top upregulated and downregulated genes.** **a** Top 10 up and downregulated genes and additional upregulated genes. A p value < 0.01 cutoff was used to determine significance. **b** qPCR results of select up and downregulated genes relative to the 16S rRNA and *rpoB*

gene in biological replicates of Cu limited and Cu replete cultures. A logarithmic scale was used for the y-axis. Samples L7* and L8* are Cu-limited cultures collected on Day 13 (see Fig. 1) and subjected to proteomic and RT-qPCR analysis.

upregulated MCO1 (NVIE_000600; log₂ FC of 0.81) were identified. In addition, a putative zinc/iron permease gene adjacent to MCO1, PERM_b (NVIE_000590; log₂ FC

1.33), and another zinc/iron permease, PERM_a (NVIE_001900; log₂ FC 1.39), were also differentially expressed. Both MCO4s and the MCO1 were also detected

in the proteome of the two tested Cu-limited cultures (Dataset S1). The two zinc/iron permeases were absent from the proteome, likely attributed to the protein extraction method that was not specifically adapted for membrane proteins. In RT-qPCR, general trends of selected up and downregulated genes were also confirmed across all copper limited samples used for transcriptomics or proteomics: MCO4_b (NVIE_019250) was highly expressed, and PERM_b (NVIE_000590) and MCO1 (NVIE_000600) were moderately expressed, whereas CBP_a (NVIE_001000) was under-expressed relative to the housekeeping genes 16S rRNA and *rpoB* in Cu-limited cultures (Fig. 3b; see Table S2 for primers).

Comparison of MCO1 and MCO4 with other 2d MCOs from bacteria (Fig. S3) suggests that the T1 Cu site in these MCOs are coordinated by two histidines, a cysteine and either a methionine (in MCO4_a; NVIE_017730) or leucine (in MCO1; NVIE_000600 and MCO4_b; NVIE_019250). In bacteria, the T2/T3 Cu sites of MCOs can reduce oxygen to water and oxidize organic (e.g., siderophores, phenolic compounds) or inorganic (e.g., Cu^+ , Fe^{2+}) electron acceptors [17, 32] via the T1 Cu site [34]. MCOs in archaea could use a similar mechanism to oxidize Cu^+ to Cu^{2+} , for example to make Cu available to cells as discussed below.

Interestingly, genes for proteins requiring Cu (like MCOs) were upregulated while the organism was under Cu limitation. One possible explanation for this observation is that Cu limitation does not affect transcription of these MCO genes directly, but rather their stability and proper function following translation. This is true of MCOs with ferroxidase activity in Cu-deficient mammals [35]. Similarly, in expression studies of MCOs in the yeast *Candida albicans*, some MCO transcripts were upregulated under Cu limitation while others were unaffected [36]. This implies the cell recognizes that the proteins are defective and increases transcription to make new proteins while attempting to remedy the situation. In this scenario, improperly folded proteins rather than direct Cu-limitation would be inducing a transcriptional response.

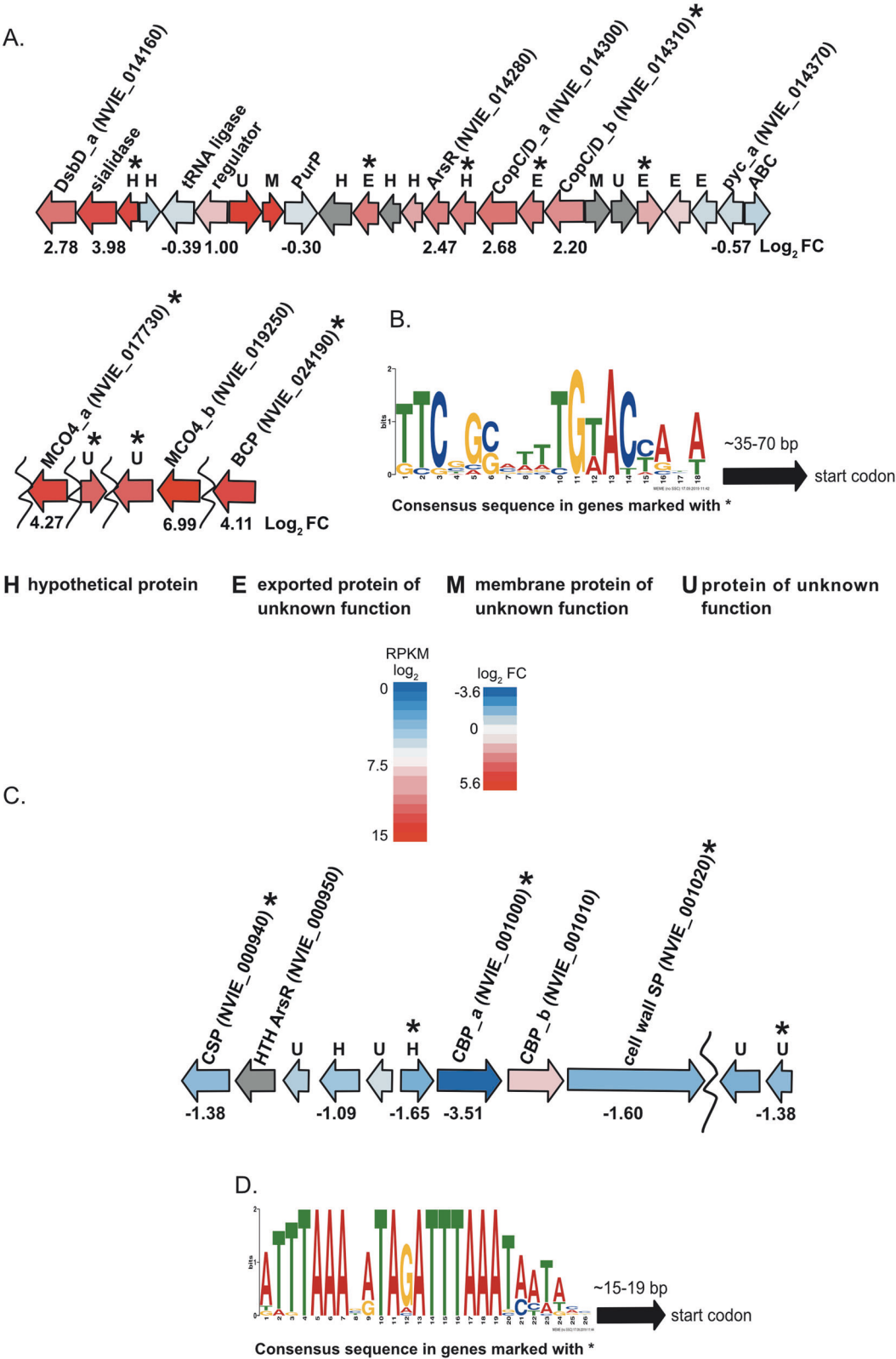
Another possibility is that MCOs make Cu more bio-available to cells. In *Saccharomyces cerevisiae*, the MCO Fet3p oxidizes iron which is then transported into the cell via a permease [37]. Perhaps a similar strategy is employed for Cu transport into AOA cells using MCOs and permeases, whose genes are directly linked in the genome (i.e., NVIE_000600 and NVIE_000590; NVIE_026950 and NVIE_026940). In addition, *Escherichia coli* has an MCO, CueO, that has been shown to oxidize Cu^+ to Cu^{2+} [38]. MCOs in AOA could potentially oxidize Cu^+ to Cu^{2+} which in turn could be more readily transported by the ZIP permeases. However, the hypothesis that these MCOs and permeases are involved in Cu transport has to be experimentally validated.

Genes encoding Cu-containing Cop proteins and their possible regulation

A putative Cu-storage protein, Csps (NVIE_028200), was identified and upregulated in *N. viennensis* under Cu limitation. This Csps contains 18 cysteine residues (15.3% of the amino acid sequence) and is similar (28.44% identity; BLASTP 2.10.0+ analysis [39, 40]) to a Cu-storage protein found in *Methylosinus trichosporium* OB3b, MtCsp1. A study of *M. trichosporium* OB3b showed that under Cu limitation MtCsp1 stores Cu^+ for use by its particulate methane monooxygenase, allowing cells to grow on methane when Cu is limiting [41]. Perhaps a similar strategy is used by *N. viennensis* to gather enough Cu for its AMO complex.

Cop proteins also seem to play a role in Cu limitation, as several of their genes were upregulated including: two *copC/D* genes, *copC/D_a* (NVIE_014300) and *copC/D_b* (NVIE_014310). Both proteins are fusions of CopC and CopD domains, and their genes together with neighboring genes form part of a larger cluster comprising 25 genes of which 15 were upregulated (Fig. 4a). In bacteria, CopC is a hypothetical chaperone located in the periplasm that brings Cu to CopD, an inner membrane protein, perhaps helping to transport Cu into the cytoplasm [21]. Alignment of the *N. viennensis* CopC/D amino acid sequences to CopC representative sequences from the Lawton et al. 2016 study [21] showed they were most similar to CopC from *M. trichosporium* OB3b (*Mst-CopC*, Fig. S4). *Mst-CopC* has one binding site for Cu^{2+} (Fig. S4). CopCs similar to *Mst-CopC*, with one Cu^{2+} binding site, are the most common type of CopC proteins [21]. In these CopC proteins, a methionine rich region that would bind Cu^+ is missing [21]. The missing methionine residues have been noted to occur in species (including *N. viennensis*) that are missing *copA* and *copB*, encoding for a multicopper oxidase and outer membrane protein respectively, from the classical *copABCD* operon [42]. Furthermore, Gram-positive bacteria were observed to not only lack the methionine residues, but to also contain a fusion of CopC/D [42]. This fusion was suggested to be relevant for Gram-positive bacteria due to the absence of a periplasm (defined as a space between two membranes) when compared to Gram-negative bacteria [42]. Therefore, the lack of a true periplasm in Thaumarchaeota might also account for the CopC/D fusion protein that is conserved across the phylum (Fig. S5).

CopC and CopD are primarily involved in Cu resistance, but some exceptions are known. In *Pseudomonas fluorescens* SBW25, an organism also lacking CopA and CopB homologs, overexpression of *copCD*, resulted in hypersensitivity to Cu [24] supporting the idea that CopCD is involved in Cu uptake in some microorganisms. In two marine AOA species, the expression of CopC/D was



increased as a response to oxidative stress, and was suggested to represent a strategy to reduce oxidative damage through copper accumulation [43].

The expression of CopCD under Cu limitation was also proposed to be regulated by CopRS [24], a two-component regulatory system that is known to control *copABCD*

◀ **Fig. 4 Upregulated and downregulated gene clusters.** **a** Upregulated gene clusters along with their gene expression information. Curved line indicates genes are in another gene cluster relative to the other genes. Gene arrow colors represent $\log_2$ fold change of genes in copper limited cultures. Gray arrows are not differentially expressed. **b** A motif consensus sequence found in certain upregulated genes marked with an asterisk (*) in **a**. DsbD, disulfide bond oxidoreductase D family protein; sialidase, putative sialidase neuraminidase family protein; tRNA ligase, putative threonyl/alanyl tRNA ligase; regulator, putative transcriptional regulator; PurP, 4-carboxamide-1-(β)-D-ribofuranosyl 5' monophosphate ligase; ArsR, transcriptional regulator ArsR; pycA, putative plastocyanin Cu containing protein/azurin family; ABC, putative ABC transporter; MCO, multicopper oxidase; BCP, protein of unknown function with C-terminal blue type(1) copper domain. Locus tag information for all genes in this figure can be found in Dataset S1. **c** Gene clusters of top downregulated genes that were found to harbor a motif sequence along with their gene expression information. Curved line indicates genes are in another gene cluster relative to the other genes. Gene arrow colors represent $\log_2$ fold change of genes in copper limited cultures. Gray arrows are not differentially expressed. **d** A motif consensus sequence found in certain upregulated genes marked with an asterisk in **c**. PUA, PUA domain containing protein; Lrp, transcriptional regulator Lrp family; psmB1, psmB1 proteasome subunit beta; CSP, cell surface protein w/ DUF 11 domain; HTH-ArsR, hypothetical protein w/HTH ArsR; CBP, cell surface-associated Ca^{2+} binding domain protein aemolysin type; cell wall SP, putative cell wall surface anchor family protein. Gene locus information for all genes in figure can be found in Dataset S1.

operons [44]. Indeed, an upregulated potential two-component system under Cu limitation was found in *N. viennensis* (NVIE_030040, $\log_2$ FC 1.77; NVIE_030060, $\log_2$ FC 1.51) with NVIE_030040 sharing a 27.97% identity to CopR from *P. fluorescens* SBW25 (PFLU_1576 or cusR, NCBI accession number CAY47824.1) [39, 40]. Potentially, these two proteins could act as CopR and CopS respectively. If so, a regulatory system induced by Cu limitation, similar to that proposed in *P. fluorescens* SBW5, is likely.

Alternatively, it was suggested that in the bacterium *Corynebacterium glutamicum* CopC could participate in the assembly of cytochrome aa_3 oxidase (complex IV) as an intermediate Cu chaperone before Cu is incorporated into the cytochrome oxidase [45]. This could also be true for *N. viennensis* as the cytochrome oxidase in Thaumarchaeota is classified as type aa_3 [46, 47]. Whether or not Cu needs to be transported into the cell before incorporation (as for *cbb_3* oxidase in the bacterium *Rhodobacter capsulatus* [48, 49]) or whether this takes place in the (pseudo-)periplasm (as in the model for cytochrome aa_3 oxidase in *Paracoccus denitrificans* [50]) is unclear.

Genomic organization in *N. viennensis* revealed that the Cop genes are spread out in the genome with CopC/D genes residing in one genomic region, CopA1 (NVIE_008380) in another, and CopA2 (NVIE_012900), CopZ (NVIE_012910), and one of three CopT genes (NVIE_012920) residing in a third genomic region (Fig. S6). In *N. viennensis*, CopA1 and CopA2 are P1B-type

(also known as CTPx-type) ATPase proteins (Fig. S7a, b respectively) and not the CopA-MCO discussed above. This discrepancy is a result of redundant naming in the literature. In *N. viennensis*, CopA1 and CopA2 refer exclusively to P1B-type ATPase proteins. A number of studies have shown that P1-type ATPases can function as exporters and importers of ions [20]. CopZ and CopT are located in the same gene cluster as CopA2 (Fig. S6). CopZ could deliver Cu^+ to CopA1 or CopA2 [51] for export. CopT is similar (38% identity) to a novel archaeal transcriptional regulator (SO2652) that is hypothesized to control the expression of Cop genes in *Sulfolobus solfataricus* P2 [19]. Of these Cop genes in *N. viennensis*, only the CopC/D genes and the CopA2 gene were differentially upregulated under Cu-limitation (Dataset S1). This is unusual for CopA2 as it is assumed to be an exporter, but perhaps it is constitutively expressed to respond to sudden fluctuations in Cu availability. How these Cop genes are transcriptionally regulated remains an open question.

Shared promoter motifs in differentially upregulated genes

To investigate possible regulatory mechanisms of the upregulated genes, upstream regions from translational start sites were analyzed for shared motifs. Initially, no motifs were identified in the upstream regions of the top 25 upregulated genes. The analysis was repeated focusing on 15 upregulated genes known or suspected to be involved in Cu metabolism. These included the upregulated genes with a $\log_2$ FC > 1.5 found in close proximity to the two CopC/D genes (NVIE_014300 and NVIE_014310) in the genome, upregulated genes next to the highly upregulated MCO4_b (NVIE_019250), and genes specifically associated with Cu from the top 25 upregulated genes (see Dataset S1 for a full list of genes).

A motif with a consensus sequence of 5'-TTCSGSMTTTTGWACYANT-3' was found 12 times starting at position -35 to -70 with respect to the potential translational start codon of the respective gene (Fig. 4b). In particular, it was found twice in front of one of the CopC/D genes, *copC/D_b* (NVIE_014310) (Fig. S8a). For some of the analyzed genes, the motif was found in the upstream region on the opposite strand in reference to the gene orientation (Fig. S8a). This was the case for an upregulated exported protein of unknown function near the CopC/D genes (NVIE_014340), a protein of unknown function (NVIE_019240) that directly follows the highly upregulated MCO4_b (NVIE_019250), a highly upregulated gene of unknown function near the *copC/D* region of the genome (NVIE_1444), and MCO4_a (NVIE_017730). Regarding NVIE_1444, the motif was also repeated. A motif recruiting a transcriptional activator on the opposite strand in

reference to the transcribed gene was previously observed upstream of the *malZ* gene in *E. coli*. [52]. Once found, the specific motif sequence was screened against the top 25 upregulated genes. This resulted in the additional discovery of the motif in the second most upregulated gene (NVIE_001850, log₂ FC 5.53), an exported protein of unknown function. The presence of this motif in various upregulated genes could indicate that it is involved in transcription under Cu limitation.

Most of the upregulated genes with the motif, or in a putative operon with the motif, encode hypothetical or unknown proteins. An ArsR family transcriptional regulator (NVIE_014280), a family known to be involved in metal stress [53], was one of the upregulated genes near the CopC/D cluster and is presumably co-transcribed with NVIE_014290 containing the identified motif (Fig. 4a). Another gene, NVIE_014160 (DsbD_a), a disulfide bond oxidoreductase D (DsbD) family protein, was located in a putative operon starting with NVIE_1444. DsbD family proteins shuttle electrons from the cytoplasm to the periplasm to reduce oxidized disulfide bonds and/or to adjust the redox status of prokaryotic cells. Since further counterparts of the redox system are missing, the upregulation of *dsbD* could be a cellular response to rectify damage to disulfide bonds induced by a lack of Cu rather than a mechanism to adjust the cell's redox state. Alternatively, DsbD, like its homolog ScsB, could bind Cu⁺ before passing it to a Cu chaperone in the periplasm [54, 55], thus sequestering Cu rather than removing it (as in *S. enterica* [56]).

Downregulated genes

A common motif was also searched for in the upstream region of the top 25 downregulated genes. A motif with a sequence of 5'-RATTTAAABATAGATTTAAATAATAAS-3' was identified in five of the top 25 downregulated genes (Fig. 4c, d and S8b): NVIE_001000 (the most downregulated gene), NVIE_0009400, NVIE_0009900, NVIE_001020, and NVIE_019740. NVIE_019740 appears to form an operon with NVIE_019730, which is also in the top 25 downregulated genes. Therefore, only the upstream sequence of NVIE_019740 was used for the motif search. NVIE_001000 and NVIE_001020 are suggested to code for proteins involved with the cell surface while the rest are annotated as hypothetical proteins or proteins of unknown function.

The motif was also present in similar genes from *Candidatus Nitrososphaera evergladensis* and *Candidatus Nitrososphaera gargensis*, the two most closely related Thaumarchaeal species to *N. viennensis* (see Dataset S1 for a list of the identified genes and Fig. S6 for alignments of the motif across all three species). The conservation of this

promoter in these other species suggests that a similar transcriptional pattern would be observed in both *Ca. N. evergladensis* and *Ca. N. gargensis* under Cu limitation.

The most strongly downregulated gene in Cu-limited cultures was a putative surface-associated Ca²⁺-binding protein (CBP_a, NVIE_001000; log₂ FC -3.51) (Fig. 3a). Perhaps the cell remodels its surface to accumulate Cu in response to Cu limitation as previously reported [11, 57]. It is also possible that the Cu-limited cells inhibit cell aggregation by downregulating the putative surface-associated Ca²⁺ binding protein to enhance Cu uptake.

Responses in energy metabolism and carbon fixation

In accordance with growth inhibition upon Cu limitation, most genes central for energy metabolism, i.e., the AMO complex, electron transport chain (ETC), and genes involved in carbon fixation, were downregulated. For the ETC, these included all subunits of complex I, most of the genes encoding for the archaeal ATP synthase (complex V), and genes potentially encoding plastocyanins [8]. It is hypothesized that plastocyanins form part of the archaeal HURM (hydroxylamine: ubiquinone redox module) and may shuttle electrons between protein complexes, similar to c-type cytochromes [58, 59]. Because of its ammonia-oxidizing metabolism, *N. viennensis* does not rely on complex I as a source of electrons for the ETC. Therefore, it has been suggested that complex I could be operating in reverse, possibly supplying a source of reducing equivalents to the cell for carbon fixation and biomass accumulation [58, 60, 61]. In this hypothesized model, complex I would consume H⁺ generated by the ETC and deplete the available proton motive force (pmf) (Fig. 5).

In contrast to complex I and complex V, other ETC components that contribute to pmf, (i.e., complex III (cytochrome b/b6 and Rieske [2Fe-2S]) and complex IV (*coxA1* and *coxB*)) were upregulated under Cu limitation (Fig. 5). A possible explanation for the ETC complexes being regulated in different ways is that the cell is attempting to conserve pmf, given that the pmf consuming complexes (complex I and complex V) were downregulated while the pmf contributing complexes (complex III and complex IV) were upregulated. A second possibility is that complex III and complex IV are upregulated to maintain the quinone and plastocyanin electron carriers in an oxidized state. The shift towards oxidized electron carriers could help to ensure that any electrons introduced into the ETC by the AMO complex are channeled directly towards the ETC because of the high supply of available electron acceptors.

Ammonia monooxygenase genes *amoA*, *amoB*, *amoX*, and *amoC1* were downregulated while AMO subunits

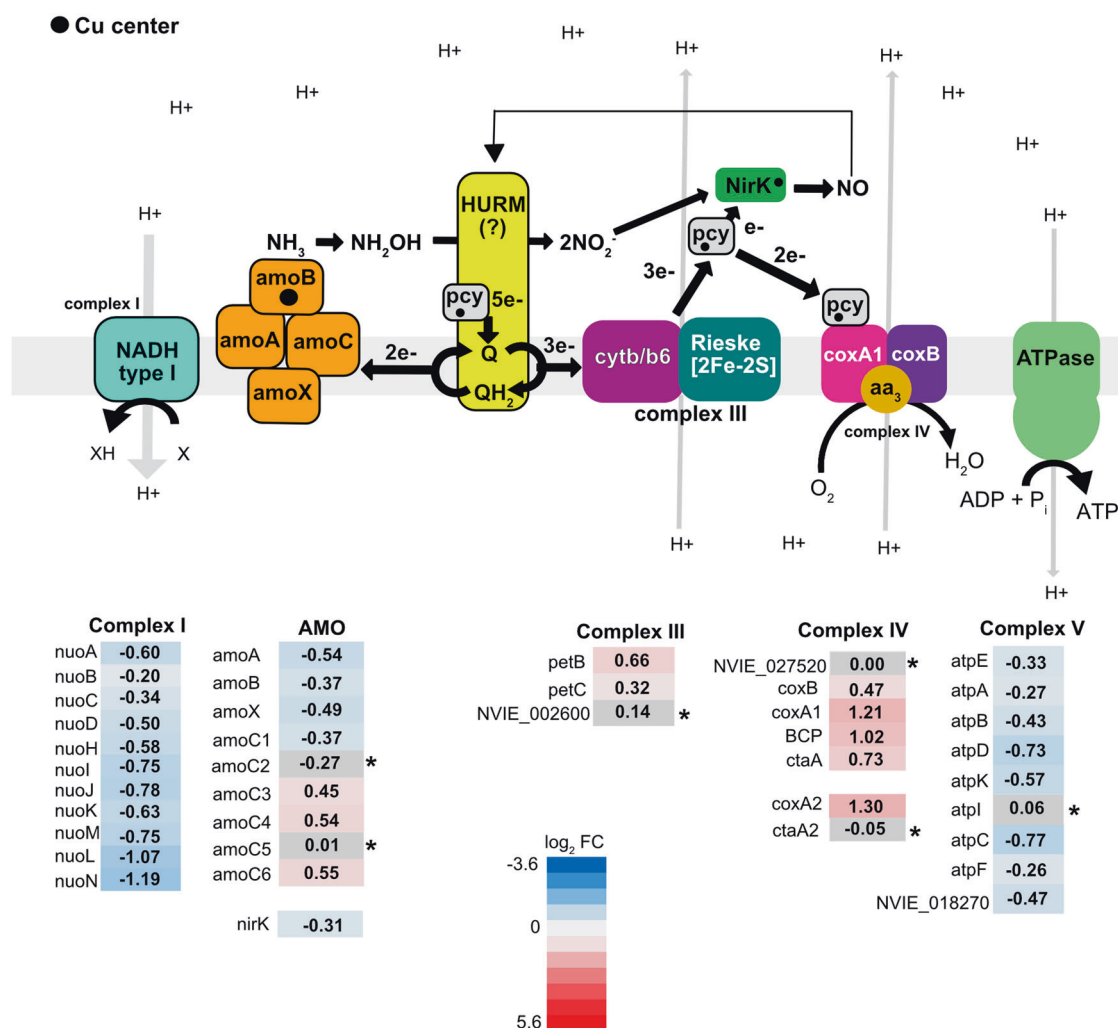


Fig. 5 A schematic showing the expression of Cu and non-Cu containing genes potentially involved in the ammonia oxidation and electron transport pathways. Blue color is downregulated and red is upregulated; ● circle stands for Cu with an unknown redox state. A grey background with an asterisk represents genes that were not differentially expressed. A p value < 0.01 cutoff was used to determine significance. (pcy) NVIE_002600 and NVIE_027550 (putative blue

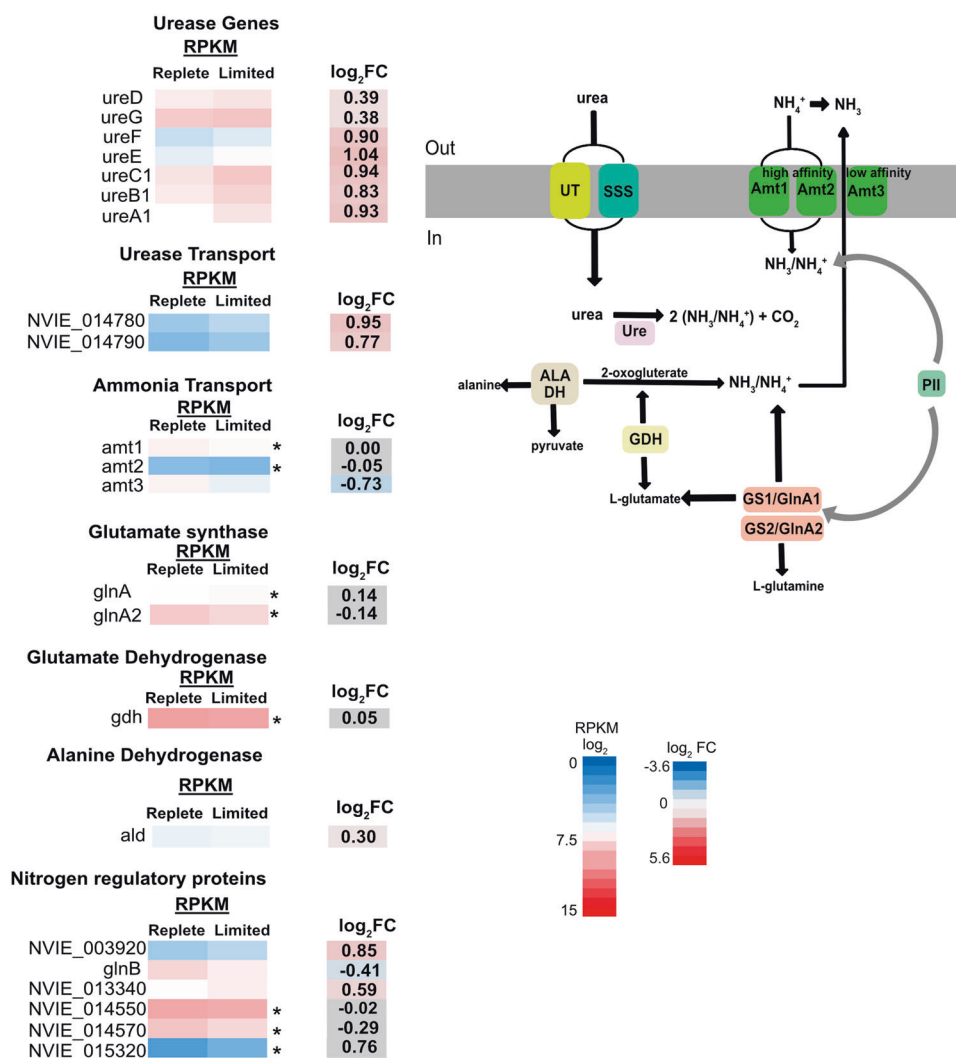
(type 1) copper domain containing proteins), NVIE_027520 cytochrome c oxidase subunit, NVIE_018270 putative ATPase subunit H, HURM; represents the hydroxylamine ubiquinone redox module [58] yet to be identified in AOA. NVIE gene information for all genes in this figure can be found in Dataset S1. Figure adapted from Qin et al. [26], Kerou et al. [9], and Kozłowski et al. [59].

amoC3, *amoC4*, and *amoC6* [9] were upregulated (Fig. 5). There are six copies of *amoC* in the genome [9] and it is unknown if all or only some interact with the AMO complex in *N. viennensis*. However, it was suggested that additional copies of *amoC* could ensure that the AMO enzyme is stable under conditions of stress [9] similar to findings in AOB [26, 62]. Although the cell seems to be reacting to energy limitation, AMO, the only source of energy, is downregulated. This is counter-intuitive but can be explained by limited Cu available to the cell, which is needed for this enzyme complex. Consequently, a transcriptional response that lowers the transcription of AMO is initiated. It should be noted that although transcription of AMO genes is lower under Cu-

limited conditions compared to Cu-replete conditions, they still have some of the highest overall transcription levels in the cell (Dataset S1).

Genes involved in nitrogen uptake also showed distinct expression patterns. *N. viennensis* encodes three ammonia transporters, two urea transporters, and a urease operon (Fig. 6). Two of the ammonia transporters are thought to be high-affinity transporters and the other to be a low-affinity transporter [63]. The low-affinity transporter, *amt3* (NVIE_023840) (Fig. 6) was downregulated under Cu limitation while the presumed high-affinity transporters were unaffected. The differential regulation of ammonia transporters in response to environmental stimuli was also observed in *N. maritimus* [64]. Genes within the urease

Fig. 6 A schematic showing the expression of ammonia transporter, urease, and transcriptional regulator genes. For RPKM $\log_2$ and $\log_2$ FC, blue color is downregulated and red is upregulated. A grey background with an asterix represents genes that were not differentially expressed. A p value < 0.01 cutoff was considered as significantly expressed.



operon and urea transporters were all upregulated. The downregulation of the low-affinity transporter and the upregulation of the urease genes and urea transporters (even in the absence of urea) initially suggest a cellular response to ammonia limitation. This response regarding the low-affinity ammonia transporter was shown in previous studies of *N. viennensis* and other AOA when the cells were starved of ammonia [26, 63, 64]. However, in this study, the cells appear to be starved of energy, not ammonia. This is shown by the clear presence of ammonium in the medium at the time the cultures were harvested (Fig. 1). Therefore, it is likely that these cellular responses are regulated by an arrest of the ETC, rather than the physical presence or absence of ammonia.

As previously mentioned, the expression of the presumed high-affinity ammonia transporters was unaffected. If these importers responded to energy limitation, they would likely be upregulated similar to the urea transporters. This not being the case, it implies that their

expression is regulated by a different system that could potentially sense the physical presence or absence of ammonia despite the energy-limited state of the cell. Moreover, the lack of differential expression in ammonia assimilation enzymes indicates that these proteins are unaffected by the low energy state of the cell and/or support the assumption that ammonia assimilation is not a limiting growth factor for the cell. A complex regulatory system involving nitrogen and energy should not be surprising as these two substrates are fundamentally intertwined in ammonia-oxidizing microorganisms.

Cu-limitation resulted in downregulation of genes encoding enzymes that consume NADPH (e.g., *asd*, *msr*) in the 3-hydroxypropionate/4-hydroxybutyrate (3HB/4HB) carbon fixation cycle (Fig. S9). Due to the low energy status of the cell, NADPH reducing equivalents, normally available to reduce inorganic carbon into organic carbon compounds, is likely scarce. As discussed above, one source of reducing equivalents could be the reverse complex I

reaction (Fig. 5). However, reverse complex I would be unable to directly produce NAD(P)H due to the lack of *nuoE*, *nuoF*, and *nuoG* subunits in *N. viennensis* [9] and other archaea [65].

The non-oxidative pentose phosphate pathway (which produces precursors that can be converted to nucleotides [66]), ribosomal proteins, information processing genes, and most of the lipid biosynthesis pathway were also downregulated (Dataset S1). These observations match the functional enrichment of translation, replication, and nucleotide metabolism genes within the downregulated group (Fig. 2) and the general growth impairment of cells upon Cu limitation (Fig. 1).

Different reactions to Cu limitation in *N. viennensis* and *N. maritimus*

While *N. maritimus* showed similar overall responses to Cu limitation by downregulating general energy and carbon fixation pathways and slowing growth, Cu-acquisition or coping with Cu-limitation clearly differs from that in *N. viennensis* [26] (Fig. 7). Opposite reactions of complex III and complex IV in the ETC in response to the same environmental stimuli indicate a distinct regulation of energetic pathways between *N. viennensis* and *N. maritimus*, although the pathway itself is highly conserved among AOA. This distinction in regulatory patterns is also seen by the differing response in expression of homologous MCO4 [9] genes between the two organisms. Curiously, homologous MCO1 genes [9] show conflicting patterns in *N. maritimus*. In the case of MCO4, this would suggest that MCO4 plays a different physiological role in *N. maritimus*, possibly one that is not Cu related. In addition, each organism contains at least one CopC/D fusion protein and a lone CopD protein. Based on this comparison, *N. viennensis* relies on upregulation of the CopC/D fusion protein while *N. maritimus* rather upregulates *copD*. As CopD is a membrane bound transporter, *N. maritimus* likely utilizes a chaperone protein other than CopC to deliver Cu to CopD for transport into the cell. Alternatively, CopC/D could be linked to the proper assembly and function of complex IV as both are upregulated in *N. viennensis* and both are downregulated in *N. maritimus*. A final key difference is the upregulation of DsbD family proteins (arCOG02398) in *N. viennensis*. Perhaps, the upregulation of MCO4 and DsbD in *N. viennensis*, which are presumed to interact with Cu⁺, and the downregulation or lack of response respectively in *N. maritimus* is a reflection of the different availability of Cu⁺ in the very different environmental niches that these organisms occupy. Although Cu²⁺ is known as the bioavailable form in aerobic soil [67] and oxygenated seawater environments [68], Cu⁺ is more likely to be found in soil environments that occasionally experience anoxic conditions and can possibly be stabilized at least to a certain

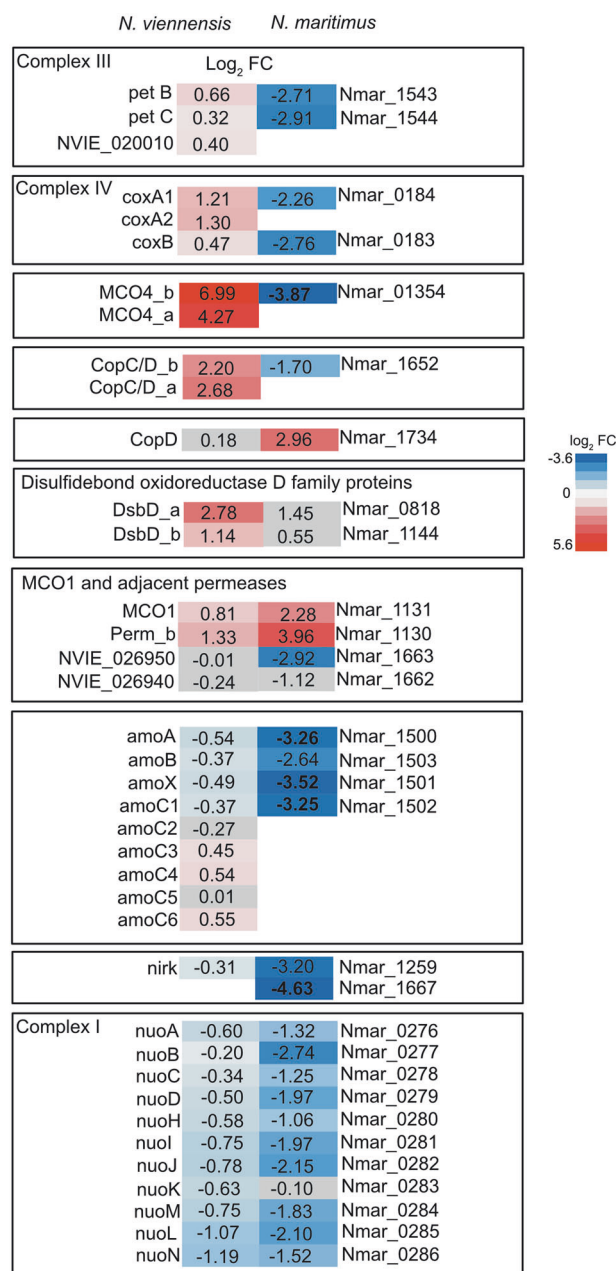


Fig. 7 Comparison of selected genes from *N. viennensis* and *N. maritimus* under Cu limitation. Values represent the log₂ fold change of genes compared to Cu-replete conditions. A grey background represents genes that were not differentially expressed in *N. viennensis* (*p* value < 0.01) or *N. maritimus* (*p* value < 0.05) [26] respectively.

extent by humic acids upon soil re-oxidation [69, 70]. Such environments include wetlands and riparian floodplain soils.

In neutral and alkaline soils, where organic and inorganic soil particles bind Cu²⁺ in soil solution and lower its bioavailability [71–73], these conditions could trigger the expression of CopC/D genes in soil organisms like *N. viennensis*. Another soil environment in which Cu concentrations are low and where expression of Cu-uptake

genes could be triggered are heavily weathered and eroded soils. A study by Jenkins et al. [74] observed low *amoA* gene abundances of AOA in such soils with lowered total Cu concentrations.

Cu-C ratios and the types of ligands found in the soil environment differ from those in the marine environment [69]. In marine surface waters 10% of dissolved Cu has been shown to be Cu^+ and in estuarine environments up to 80% of Cu could be Cu^+ [75, 76]. Copper in these environments can be bound by thiol ligands, reduced sulfur species or colloidal ligands that lower its bioavailability [77–79]. In surface waters of Hood Canal, Washington, bioavailable Cu^{2+} and AOA *amoA* gene abundances were found to be low in surface waters [78]. Thus, *N. maritimus* inhabiting these environments might trigger the expression of genes encoding transporters that differ from those observed in this study to facilitate uptake of ligand-bound Cu.

Conclusions

Results show that Cu-limitation in *N. viennensis* leads to high expression of several genes encoding MCOs and CopC/D proteins, which may be involved in Cu uptake. This study also puts forward the idea that Cu limitation triggers an energy conservation response in *N. viennensis*. This energy conservation response appears to downregulate the expression of genes involved in the consumption of the pmf, ammonia oxidation, and carbon fixation pathways while upregulating genes involved in ammonia acquisition by inducing the urease operon.

In summary, Cu uptake could involve binding and uptake of Cu^{2+} by CopC/D fusion proteins (Fig. 8). CopC could also act as a chaperone helping to assemble cytochrome *aa*₃ oxidase and AMO by transferring Cu^+ to these proteins. MCO1 and MCO4 could bind and oxidize Cu^+ to Cu^{2+} , which could then be taken up either through a

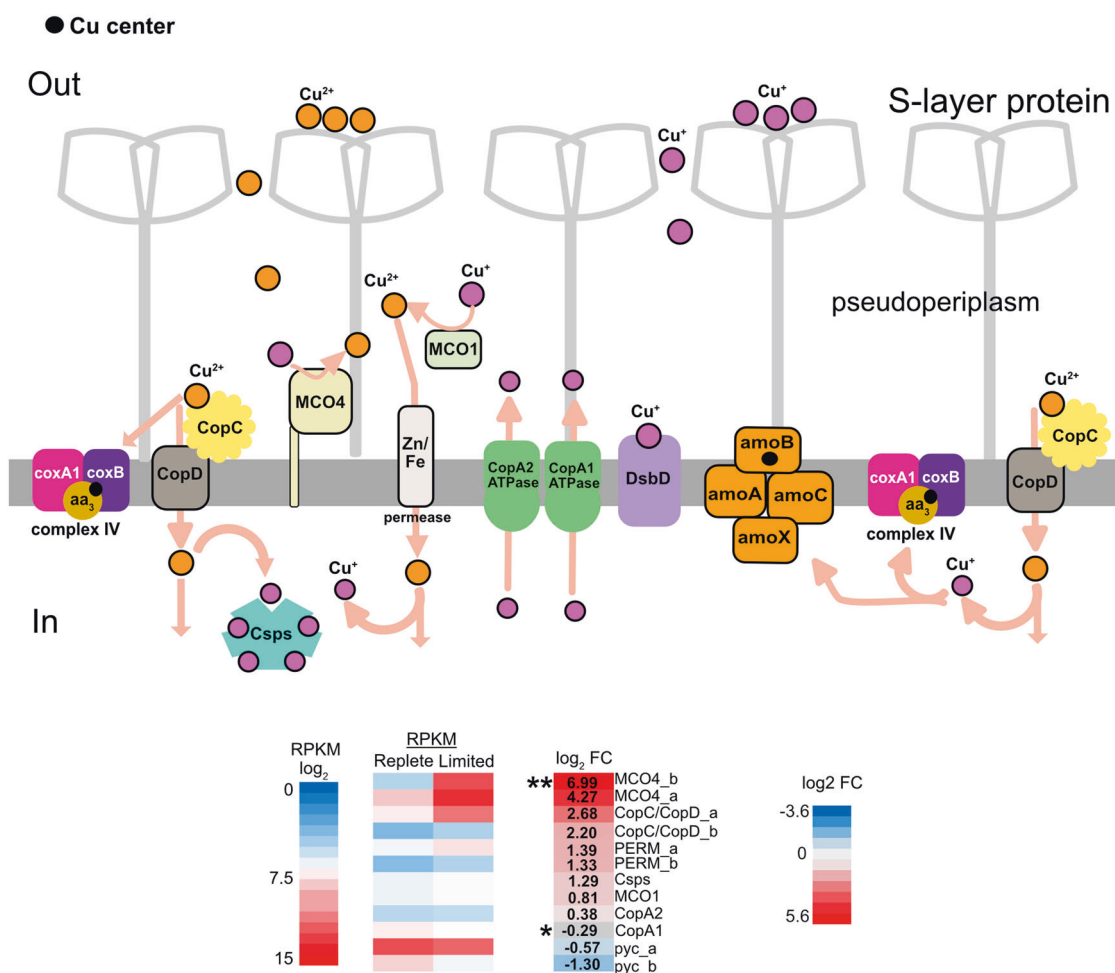


Fig. 8 A schematic showing the expression of MCOs and Cop genes under Cu limiting conditions. ● circle stands for Cu with an unknown redox state, ● Cu^{2+} , and ● Cu^+ . For RPKM log₂ and log₂ FC, blue color is downregulated and red is upregulated. (**) log₂ fold

change value is higher than pictured in the scale. A grey background with an asterisk represents genes that were not differentially expressed. A *p* value < 0.01 was considered as significantly expressed.

permease (encoded adjacent to MCO1) or an unknown divalent transporter. Once acquired, Csps could sequester Cu inside the cell. CopA1/A2 could function as Cu exporters and DsbD could possibly bind Cu⁺ and transfer it to chaperones in the periplasm for assembling Cu containing proteins or to import Cu into the cell.

Overall, this analysis proposes possible regulatory networks in terrestrial AOA, while highlighting distinct differences in adaptation strategies employed by different AOA clades in accordance to their ecophysiology.

Acknowledgements The authors would like to thank Herwig Lenitz, Sonja Tischler and Romana Bittner for technical support, Lokeshwaran Manoharan for bioinformatics input, Isabelle Anna Zink and Gabriella Paredes for their advice regarding qPCR methods, Thomas Rattei and Florian Goldenberg for providing access to the CUBE servers (U. Vienna), Sung-Keun Rhee (Chungbuk National University) for discussions and input on the copper model, and Amy Schmid (Duke University) for suggestions regarding analysis of the data. This study was supported by the Marie Skłodowska Curie Individual Fellowship (IF) CORA Project (grant no. 700788) (C.R.), by the Austrian Science Fund (FWF) DK+ project “Microbial Nitrogen Cycling: From single cells to ecosystems” (W1257) (L.H., B. B. and CR), MK was supported through the European Research Council Advanced Grant ERC-ADG 695192 and BB by the Uni:docs fellowship of the University of Vienna.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Kuypers MMM, Marchant HK, Kartal B. The microbial nitrogen-cycling network. *Nat Rev Microbiol*. 2018;16:263–76.
- Daims H, Lebedeva EV, Pjevac P, Han P, Herbold C, Albertsen M, et al. Complete nitrification by *Nitrospira* bacteria. *Nature*. 2015;528:504–9.
- van Kessel MAHJ, Speth DR, Albertsen M, Nielsen PH, Op den Camp HJM, Kartal B, et al. Complete nitrification by a single microorganism. *Nature*. 2015;528:555–9.
- Schleper C, Nicol GW. Ammonia-oxidising archaea—physiology, ecology and evolution. In: Poole RK, editor. *Advances in microbial physiology*. Academic Press; Amsterdam. 2010. p. 1–41.
- French E, Kozłowski JA, Mukherjee M, Bullerjahn G, Bollmann A. Ecophysiological characterization of ammonia-oxidizing archaea and bacteria from freshwater. *Appl Environ Microbiol*. 2012;78:5773–80.
- Kerou M, Schleper C. *Nitrososphaera*. In: Whitman, WB, Rainey, F, Kämpfer, P, Trujillo, M, Chun, J, DeVos, P, Hedlund, B, Dedysh S, editors. *Bergey's manual of systematics of archaea and bacteria*. John Wiley & Sons, Inc. & Bergey's Manual Trust; 2016.
- Alves RJE, Minh BQ, Urich T, Von Haeseler A, Schleper C. Unifying the global phylogeny and environmental distribution of ammonia-oxidising archaea based on *amoA* genes. *Nat Commun*. 2018;9:1517.
- Walker CB, de la Torre JR, Klotz MG, Urakawa H, Pinel N, Arp DJ, et al. *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc Natl Acad Sci USA*. 2010;107:8818–23.
- Kerou M, Offre P, Valledor L, Abby SS, Melcher M, Nagler M, et al. Proteomics and comparative genomics of *Nitrososphaera viennensis* reveal the core genome and adaptations of archaeal ammonia oxidizers. *Proc Natl Acad Sci*. 2016;113:E7937–E7946.
- Arp DJ, Chain PSG, Klotz MG. The impact of genome analyses on our understanding of ammonia-oxidizing bacteria. *Annu Rev Microbiol*. 2007;61:503–28.
- Gorman-Lewis D, Martens-Habben W, Stahl DA. Cu(II) adsorption onto ammonia-oxidizing bacteria and archaea. *Geochim Cosmochim Acta*. 2019;255:127–43.
- Lide DR. *CRC handbook of chemistry and physics: a ready reference book of chemical and physical data*, 84th ed. CRC Press Taylor and Francis Group. Boca Raton. 2003.
- Reyes C, Hodgskiss LH, Baars O, Kerou M, Bayer B, Schleper C, et al. Copper limiting threshold in the terrestrial ammonia oxidizing archaeon *Nitrososphaera viennensis*. *Res Microbiol*. 2020;171:134–42.
- Gwak JH, Jung MY, Hong H, Kim JG, Quan ZX, Reinfelder JR, et al. Archaeal nitrification is constrained by copper complexation with organic matter in municipal wastewater treatment plants. *ISME J*. 2019;14:335–46.
- Tourna M, Stieglmeier M, Spang A, Könneke M, Schintlmeister A, Urich T, et al. *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc Natl Acad Sci USA*. 2011;108:8420–5.
- Stieglmeier M, Klingl A, Alves RJE, Rittmann SKMR, Melcher M, Leisch N, et al. *Nitrososphaera viennensis* gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *Int J Syst Evol Microbiol*. 2014;64:2738–52.
- Soliz M. Copper and bacteria: evolution, homeostasis and toxicity. Cham: Springer; 2018.
- Villafane AA, Voskoboynik Y, Cuevas M, Ruhl I, Bini E. Response to excess copper in the hyperthermophile *Sulfolobus solfataricus* strain 98/2. *Biochem Biophys Res Commun*. 2009;385:67–71.
- Ettema TJG, Brinkman AB, Lamers PP, Kornet NG, de Vos WM, van der Oost J. Molecular characterization of a conserved archaeal copper resistance (*cop*) gene cluster and its copper-responsive regulator in *Sulfolobus solfataricus* P2. *Microbiology*. 2006;152:1969–79.
- Coombs JM, Barkay T. New findings on evolution of metal homeostasis genes: evidence from comparative genome analysis of bacteria and archaea. *Appl Environ Microbiol*. 2005;71:7083–91.
- Lawton TJ, Kenney GE, Hurley JD, Rosenzweig AC. The CopC family: structural and bioinformatic insights into a diverse group

- of periplasmic copper binding proteins. *Biochemistry*. 2016;55:2278–90.
22. Rensing C, McDevitt SF. The copper metallome in prokaryotic cells. In: *Metallomics and the cell*. Dordrecht: Springer Netherlands; 2013.
 23. Ladomersky E, Petris MJ. Copper tolerance and virulence in bacteria. *Metallomics*. 2015;7:957–64.
 24. Zhang X-X, Rainey PB. Regulation of copper homeostasis in *Pseudomonas fluorescens* SBW25. *Environ Microbiol*. 2008;10:3284–94.
 25. Amin SA, Moffett JW, Martens-Habben W, Jacquot JE, Han Y, Devol A, et al. Copper requirements of the ammonia-oxidizing archaeon *Nitrosopumilus maritimus* SCM1 and implications for nitrification in the marine environment. *Limnol Oceanogr*. 2013;58:2037–45.
 26. Qin W, Amin SA, Lundeen RA, Heal KR, Martens-Habben W, Turkarslan S, et al. Stress response of a marine ammonia-oxidizing archaeon informs physiological status of environmental populations. *ISME J*. 2018;12:508–19.
 27. Parkhurst DL. User's guide to PHREEQC, a computer program for speciation, reaction-path, advective-transport, and inverse geochemical calculations. Water-Resources Investigation Report; 1995.
 28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
 29. Yasset P-R, Attila C, Bai J, Bernal-Linares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Res*. 2019;47:D443–D450.
 30. Esteves AM, Graça G, Peyriga L, Torcato IM, Borges N, Portais J-C, et al. Combined transcriptomics-metabolomics profiling of the heat shock response in the hyperthermophilic archaeon *Pyrrococcus furiosus*. *Extremophiles*. 2019;23:101–18.
 31. Coker JA, DasSarma P, Kumar J, Müller J, DasSarma S. Transcriptional profiling of the model Archaeon *Halobacterium* sp. NRC-1: responses to changes in salinity and temperature. *Saline Syst*. 2007;3:6.
 32. Kosman DJ. Multicopper oxidases: a workshop on copper coordination chemistry, electron transfer, and metallophysiology. *J Biol Inorg Chem*. 2010;15:15–28.
 33. Lawton TJ, Sayavedra-Soto LA, Arp DJ, Rosenzweig AC. Crystal structure of a two-domain multicopper oxidase: Implications for the evolution of multicopper blue proteins. *J Biol Chem*. 2009;284:10174–80.
 34. Quintanar L, Stoj C, Taylor AB, Hart PJ, Kosman DJ, Solomon EI. Shall we dance? How a multicopper oxidase chooses its electron transfer partner. *Acc Chem Res*. 2007;40:445–52.
 35. Prohaska JR. Impact of copper limitation on expression and function of multicopper oxidases (ferroxidases). *Adv Nutr*. 2011;2:89–95.
 36. Cheng X, Xu N, Yu Q, Ding X, Qian K, Zhao Q, et al. Novel insight into the expression and function of the multicopper oxidases in *Candida albicans*. *Microbiology*. 2013;159:1044–55.
 37. Taylor AB, Stoj CS, Ziegler L, Kosman DJ, Hart PJ. The copper-iron connection in biology: structure of the metallo-oxidase Fet3p. *Proc Natl Acad Sci*. 2005;102:15459–64.
 38. Singh SK, Grass G, Rensing C, Montfort WR. Cuprous oxidase activity of CueO from *Escherichia coli*. *Society*. 2004;186:7815–7.
 39. Altschul S, de la Cruz F, Schaffer A, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389–402.
 40. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, Schäffer AA, et al. Protein database searches using compositionally adjusted substitution matrices. *FEBS J*. 2005;272:5101–9.
 41. Vita N, Platsaki S, Baslé A, Allen SJ, Paterson NG, Crombie AT, et al. A four-helix bundle stores copper for methane oxidation. *Nature*. 2015;525:140–3.
 42. Arnesano F, Banci L, Bertini I, Thompson AR. Solution structure of CopC: a cupredoxin-like protein involved in copper homeostasis. *Structure*. 2002;10:1337–47.
 43. Bayer B, Pelikan C, Bittner MJ, Reinthaler T, Könnike M, Herndl GJ, et al. Proteomic response of three marine ammonia-oxidizing archaea to hydrogen peroxide and their metabolic interactions with a heterotrophic alphaproteobacterium. *mSystems*. 2019;4:e00181–19.
 44. Mills SD, Jasalavich CA, Cooksey DA. A two-component regulatory system required for copper-inducible expression of the copper resistance operon of *Pseudomonas syringae*. *J Bacteriol*. 1993;175:1656–64.
 45. Morosov X, Davoudi C-F, Baumgart M, Brocker M, Bott M. The copper-deprivation stimulon of *Corynebacterium glutamicum* comprises proteins for biogenesis of the actinobacterial cytochrome bc₁-aa₃ supercomplex. *J Biol Chem*. 2018;293:15628–40.
 46. Hemp J, Gennis RB. Diversity of the heme-copper superfamily in archaea: insights from genomics and structural modeling. *Bioenergetics*. Berlin, Heidelberg: Springer; 2008. p. 1–31.
 47. Sousa FL, Alves RJ, Ribeiro MA, Pereira-Leal JB, Teixeira M, Pereira MM. The superfamily of heme-copper oxygen reductases: Types and evolutionary considerations. *Biochim Biophys Acta - Bioenerg*. 2012;1817:629–37.
 48. Ekici S, Pawlik G, Lohmeyer E, Koch H-G, Daldal F. Biogenesis of cbb(3)-type cytochrome c oxidase in *Rhodobacter capsulatus*. *Biochim Biophys Acta*. 2012;1817:898–910.
 49. Trasnea P-I, Andrei A, Marckmann D, Utz M, Khalfaoi-Hassani B, Selamoglu N, et al. A copper relay system involving two periplasmic chaperones drives cbb 3 -type cytochrome c oxidase biogenesis in *Rhodobacter capsulatus*. *ACS Chem Biol*. 2018;13:1388–97.
 50. Dash BP, Alles M, Bundschuh FA, Richter OMH, Ludwig B. Protein chaperones mediating copper insertion into the CuA site of the aa₃-type cytochrome c oxidase of *Paracoccus denitrificans*. *Biochim Biophys Acta - Bioenerg*. 2015;1847:202–11.
 51. González-Guerrero M, Argüello JM. Mechanism of Cu⁺-transporting ATPases: soluble Cu⁺ chaperones directly transfer Cu⁺ to transmembrane transport sites. *Proc Natl Acad Sci USA*. 2008;105:5992–7.
 52. Tapio S, Yeh F, Shuman HA, Boos W. The malZ gene of *Escherichia coli*, a member of the maltose regulon, encodes a maltodextrin glucosidase. *J Biol Chem*. 1991;266:19450–8.
 53. Osman D, Cavet JS. Bacterial metal-sensing proteins exemplified by ArsR-SmtB family repressors. *Nat Prod Rep*. 2010;27:668.
 54. Gupta SD, Wu HC, Rick PDA. *Salmonella typhimurium* genetic locus which confers copper tolerance on copper-sensitive mutants of *Escherichia coli*. *J Bacteriol*. 1997;179:4977–84.
 55. Cho S-H, Parsonage D, Thurston C, Dutton RJ, Poole LB, Collet J-F, et al. A new family of membrane electron transporters and its substrates, including a new cell envelope peroxiredoxin, reveal a broadened reductive capacity of the oxidative bacterial cell envelope. *MBio*. 2012;3:e00291–11.
 56. Subedi P, Paxman JJ, Wang G, Ukuwela AA, Xiao Z, Heras B. The Scs disulfide reductase system cooperates with the metallo-chaperone CueP in *Salmonella* copper resistance. *J Biol Chem*. 2019. <https://doi.org/10.1074/jbc.RA119.010164>.
 57. Karlsen OA, Larsen Ø, Jensen HB. The copper responding surface of *Methylococcus capsulatus* Bath. *FEMS Microbiol Lett*. 2011;323:97–104.
 58. Simon J, Klotz MG. Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. *Biochim Biophys Acta - Bioenerg*. 2013;1827:114–35.

59. Kozłowski JA, Stieglmeier M, Schleper C, Klotz MG, Stein LY. Pathways and key intermediates required for obligate aerobic ammonia-dependent chemolithotrophy in bacteria and Thaumarchaeota. *ISME J*. 2016;10:1836–45.
60. Spang A, Poehlein A, Offre P, Zumbrägel S, Haider S, Rychlik N, et al. The genome of the ammonia-oxidizing candidate nitrososphaera gargensis: Insights into metabolic versatility and environmental adaptations. *Environ Microbiol*. 2012;14:3122–45.
61. Chain Patrick, Lamerdin J, Larimer F, Regala W, Lao V, Land M, et al. Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph *Nitrosomonas europaea*. *J Bacteriol*. 2003;185:2759–73.
62. Berube PM, Stahl DA. The divergent amoc3 subunit of ammonia monooxygenase functions as part of a stress response system in *Nitrosomonas europaea*. *J Bacteriol*. 2012;194:3448–56.
63. Offre P, Kerou M, Spang A, Schleper C. Variability of the transporter gene complement in ammonia-oxidizing archaea. *Trends Microbiol*. 2014;22:665–75.
64. Nakagawa T, Stahl DA. Transcriptional response of the archaeal ammonia oxidizer *Nitrosopumilus maritimus* to low and environmentally relevant ammonia concentrations. *Appl Environ Microbiol*. 2013;79:6911–6.
65. Friedrich T, Scheide D. The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Lett*. 2000;479:1–5.
66. Bräsen C, Esser D, Rauch B, Siebers B. Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation. *Microbiol Mol Biol Rev*. 2014;78:89–175.
67. Sauvé S, McBride MB, Norvell WA, Hendershot WH. Copper solubility and speciation of in situ contaminated soils: effects of copper level, pH and organic matter. *Water Air Soil Pollut*. 1997;100:133–49.
68. Gledhill M, Nimmo M, Hill SJ, Brown MT. The toxicity of copper (II) species to marine algae, with particular reference to macroalgae. *J Phycol*. 1997;33:2–11.
69. Maurer F, Christl I, Fulda B, Voegelin A, Kretzschmar R. Copper redox transformation and complexation by reduced and oxidized soil humic acid. 2 Potentiometric titrations dialysis cell. *Exp Environ Sci Technol*. 2013;47:10912–21.
70. Fulda B, Voegelin A, Maurer F, Christl I, Kretzschmar R. Copper redox transformation and complexation by reduced and oxidized soil humic acid. 1. X-ray absorption spectroscopy study. *Environ Sci Technol*. 2013;47:10903–11.
71. Logan EM, Pulford ID, Cook GT, McKenzie AB. Complexation of Cu^{2+} and Pb^{2+} by peat and humic acid. *Eur J Soil Sci*. 1997;48:685–96.
72. Minnich MM, McBride MB. Copper activity in soil solution: I. measurement by ion-selective electrode and donnan dialysis. *Soil Sci Soc Am J*. 1987;51:568.
73. Pendias AK. Elements of group 11: previously group Ib. In: Pendias AK, editor. Trace elements in soils and plants, 4th ed. Boca Raton: CRC Press Taylor and Francis Group; 2010. p. 584.
74. Jenkins SN, Murphy DV, Waite IS, Rushton SP, O'Donnell AG. Ancient landscapes and the relationship with microbial nitrification. *Sci Rep*. 2016;6:30733.
75. Moffett JW, Zika RG. Measurement of copper(I) in surface waters of the subtropical Atlantic and Gulf of Mexico. *Geochim Cosmochim Acta*. 1988;52:1849–57.
76. Buerge-Weirich D, Sulzberger B. Formation of Cu(I) in estuarine and marine waters: application of a new solid-phase extraction method to measure Cu(I). *Environ Sci Technol*. 2004;38:1843–8.
77. Whitby H, Hollibaugh JT, van den Berg CMG. Chemical speciation of copper in a salt marsh estuary and bioavailability to Thaumarchaeota. *Front Mar Sci*. 2017;4:178.
78. Jacquot JE, Horak REA, Amin SA, Devol AH, Ingalls AE, Armbrust EV, et al. Assessment of the potential for copper limitation of ammonia oxidation by Archaea in a dynamic estuary. *Mar Chem*. 2014;162:37–49.
79. Tang D, Warnken KW, Santschi PH. Organic complexation of copper in surface waters of Galveston Bay. *Limnol Oceanogr*. 2001;46:321–30.

Supplementary Material

Genome wide transcriptomic analysis of the soil ammonia oxidizing archaeon

***Nitrososphaera viennensis* upon exposure to copper limitation**

Carolina Reyes¹ (a)(b)(c)(#), Logan H. Hodgskiss¹ (b)(c), Melina Kerou (b)(c), Thomas Pribasnig (b)(c), Sophie S. Abby (d), Barbara Bayer (c)(e)*, Stephan M. Kraemer (a)(c), Christa Schleper (b)(c)(#)

(a) University of Vienna, Centre for Microbiology and Environmental Systems Science, Department of Environmental Geosciences, Althanstrasse 14, UZA2, 1090 Vienna, Austria. (b) University of Vienna, Department of Functional and Evolutionary Ecology, Archaea Biology and Ecogenomics Unit, Althanstrasse 14, UZA1, 1090 Vienna, Austria. (c) University of Vienna, Environmental Science Research Network (ESRN), Faculty for Geosciences, Geography and Astronomy, Althanstrasse 14, UZA2, 1090 Vienna, Austria. (d) University Grenoble Alpes, CNRS, Grenoble INP, TIMC-IMAG, 38000 Grenoble, France. (e) University of Vienna, Department of Limnology and Oceanography, Division of Bio-oceanography, Althanstrasse 14, UZA1, 1090 Vienna, Austria.

*Present address: University of California Santa Barbara, Department of Ecology, Evolution and Marine Biology, 93106-9620 Santa Barbara, USA

#Corresponding Authors: Carolina Reyes (creyes6@gmail.com) and Christa Schleper (christa.schleper@univie.ac.at)

Running Title: Transcriptome of Cu-limited *Nitrososphaera viennensis*

Keywords: AOA, nitrogen, ammonia oxidation, copper, archaea

¹ CR and LH are co-first authors

This article includes:

Supplementary figure legends

Supplementary dataset description

Supplementary methods

Supplementary references

Supplementary figures and legends S1-S9

Supplementary tables S1-S2

Figure Legends Supplementary Material

Figure S1. Nitrite concentrations in batch cultures of *N.viennensis* grown in 30 mL polystyrene bottles under Cu-replete and Cu-limited (Cu10 and Cu12) conditions as a function of the calculated free Cu^{2+} species concentration. The data and error bars represent the average and standard deviation of multiple biological replicates (≥ 4) from several experiments.

Figure S2. Principal component analysis of rlog transformed expression data of Cu-limited (Lim) and Cu-replete (Rep) samples. The x-axis describes variability between Cu-limited and Cu-replete samples. The y-axis describes variability between the replicates of each condition. Ellipses represent concentration ellipses for each group at an ellipse.level=0.95.

Figure S3. Alignment of 2dMCOs showing the predicted Cu binding sites. 2dMCO sequences from the following microorganisms were aligned using the Clustal Omega multiple alignment web interface (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) using default parameters: WP_084790903.1 copper oxidase; NVIE_019250 [*Nitrososphaera viennensis*], WP_075054901.1 copper oxidase; NVIE_017730 [*Nitrososphaera viennensis*], WP_075053492.1 copper oxidase; NVIE_000600 [*Nitrososphaera viennensis*], NP_711736.1 multicopper oxidase-like protein [*Leptospira interrogans* serovar Lai str. 56601], WP_011610637.1 multicopper oxidase [*Trichodesmium erythraeum*], WP_010998083.1 MULTISPECIES: multicopper oxidase [*Nostocaceae*], WP_011871074.1 copper oxidase [*Herminiimonas arsenicoxydans*], ADV66099.1 multicopper oxidase type 3 [*Deinococcus maricopensis* DSM 21211], EIA07124.1 multicopper oxidase [*Flavobacterium frigidum* PS1], ACA85706.1 multicopper oxidase type 3 [*Shewanella woodyi* ATCC 51908], KXK47905.1 multicopper oxidase [*Nitrosomonas europaea*], WP_013032143.1 multicopper oxidase [*Nitrosococcus halophilus*].

Figure S4. Alignment of the *N. viennensis* CopC/D amino acid sequences to CopC representative sequences from Lawton et al. (Biochemistry 55:2278-2290, 2016). Amino acid sequences were aligned using the Clustal Omega multiple alignment web interface (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) using default parameters. Copper binding residues

are boxed in black. Representative sequences are from the following organisms: *Nitrososphaera viennensis* (NVIE_014300, NVIE_014310), (NTE) *Nitrososphaera evergladensis* SR-1 (IMG ID 2586122348), (PG) *Pseudomonas* sp. GM17 (IMG ID 2503652063), (OB3b) *Methylosinus trichosporium* OB3b (IMG ID 2507406428), (PR) *Pseudomonas fluorescens* R124 (IMG ID 2503652063), (MT) *Mycobacterium tusciae* JS617 (IMG ID 2508745664), and (SY) *Sphingobium yanoikuyae* XLDN2-5 (IMG ID 2549030586).

Figure S5. Phylogeny and protein domain architecture of CopD proteins found in AOA (Nitrososphaeria, Thaumarchaeota). This phylogenetic tree was built on the basis of the alignment of the shared protein domain: CopD. The sequences were retrieved by sequence similarity search from a dataset representative of the diversity of the tree of life (see Supplementary Methods). The protein domain architecture is displayed for each sequence, as retrieved using PFAM domains' annotation. For AOA, the name of the orders is indicated along the tree with colored strips, whereas for other lineages, only the two highest taxonomic ranks are indicated (“phylum” and “kingdom” from left to right). Sequences found in *N. viennensis* are shown in red. DUF4149 (“Domain of Unknown Function”) domains that were overlapping CopD domains were ignored for display for two sequences: *Sphaerobacter thermophilus* DSM 20745 (WP_012872233.1) and *Candidatus Nitrosocosmicus exaquare* G61 (WP_148686115.1). “CopC” stands for protein domain PF04234.8 (“Copper resistance protein C”), “CopD” for protein domain PF05425.9 (“Copper resistance protein D”), “SGL” for PF08450.8 (“SMP-30/Gluconolactonase/LRE-like region”), “YtkA” for PF13115.2 (“YtkA-like”) and “FixH” for PF05751.7. Support values correspond to UltraFast bootstraps computed by IQ-Tree.

Figure S6. (A) Cop A1 and A2 (copper-exporting P-type ATPase A proteins), CopZ and CopT predicted gene clusters along with their gene expression information. A p-value cut-off of 0.01 was used to determine significance. Gene arrow colors represent log₂ fold change of genes in copper limited cultures. Gray arrows are not differentially expressed. U = unknown protein, E = exported unknown protein. (B) Alignment of down-regulated genes with a motif sequence (5'-RATTTAAABATAGATTTAAATAATAAS-3') from *N. viennensis* (NVIE) with other Thaumarchaeota genes containing a homologous motif. Ngar = *Nitrososphaera gargensis*, NTE

= *Candidatus Nitrososphaera evergladensis*. Predicted translational start codons are boxed and text is in red. Locus tag information for all genes can be found in Dataset S1.

Figure S7. (A) *N. viennensis* CopA1 ATPase predicted metal binding sites. Amino acid metal binding sites information was gathered from several studies and reviews [1][2][3] [4]. N-MBS = N terminal metal binding site, TM-MBS = transmembrane metal binding site, C-domain = C terminal metal binding site, A-domain = actuator phosphatase P-domain = phosphorylation, N-domain = nucleotide binding. Arrow indicates path of Cu delivery by CopZ to TM-MBS (B) *N. viennensis* CopA2 predicted metal binding sites. Amino acid metal binding sites information was gathered from several studies and reviews cited above. TRASH= motif involved in metal coordination, N-MBS = N terminal metal binding site, TM-MBS = transmembrane metal binding site, C-domain = C terminal metal binding site, A-domain = actuator phosphatase, P-domain = phosphorylation, N-domain = nucleotide binding. Arrow indicates path of Cu delivery by CopZ to TM-MBS.

Figure S8. (A) Alignment of *N. viennensis* up-regulated genes suspected to be involved in Cu metabolism with a motif consensus sequence of 5'-TTCSGSMTTTGWACYANT-3'. Curved line indicates a break in the alignment. * indicates a reverse complement sequence and red letters indicates the motif repeats itself twice upstream or downstream of the motif. Predicted translational start codons are boxed and text is in red. Boxed green nucleotides indicate the location of a second motif and are included with the other motif regions at the bottom of the alignment. MCO, multicopper oxidase; BCP, protein of unknown function with C-terminal blue type (1) copper domain. (B) Alignment of top down-regulated genes from *N. viennensis* (NVIE) with a consensus motif sequence (underlined in gray). Predicted translational start codons are boxed and text is in red. CSP, cell surface associated Ca^{2+} binding domain protein, aemolysin type; CBP_a, putative surface associated Ca^{2+} -binding protein; CWSP, putative cell wall surface anchor family protein; Lrp, transcriptional regulator Lrp family. Locus tag information for all genes can be found in Dataset S1.

Figure S9. A schematic of the carbon fixation and the tricarboxylic acid (TCA) pathways of *N. viennensis* showing the genes expressed in each pathway. FC in blue color is downregulated and red is upregulated. A p-value < 0.01 was considered as significantly expressed. Asterisk (*) symbol and a grey background indicates gene was not significantly expressed. Locus tag information for all genes can be found in Dataset S1.

Dataset in Supplementary Material

Dataset S1: Excel sheets with all up and downregulated genes at adjusted p-values of 0.01, excel sheets with genes used to find the identified motifs, excel sheet with list of homologous downregulated genes with the identified motifs, excel sheets listing gene locus information for genes in figures, excel sheets of ribosomal proteins, the non-oxidative pentose phosphate pathway, and genes encoding proteins containing Cu or Fe cofactors, excel sheet with list of species used for phylogenetic analysis of CopD, and excel sheet summarizing proteomic results.

Supplementary Methods

Media Preparation

Ultrapure water (18.2 MΩ cm, 2 ppb TOC) was used to prepare all media and to rinse plastic bottles. To reduce Cu contamination all media components were prepared in plastic bottles that had been washed with a mild detergent and then acid washed for several weeks, first in 3.8 % (v/v) HCl, and then in 7.4 % (v/v) nitric acid. Bottles were rinsed five times with ultrapure water in between acid washes and following the nitric acid wash. Preparation of all media components and handling of cultures was done in a trace metal free clean room environment with a HEPA filtration system. One liter of FWM medium [1 g NaCl, 0.4 g MgCl₂·6H₂O, 0.1 CaCl₂·2H₂O, 0.2 g KH₂PO₄, 0.5 g KCl], amended with 1 ml trace element solution [HCl 100 mM, H₃BO₃ 0.5 mM, MnCl₂·4H₂O, CoCl₂ · 6H₂O, NiCl₂· 6H₂O, ZnSO₄ ·7H₂O, Na₂MoO₄ ·2H₂O, 0.01 mM CuCl₂ · 2H₂O], 1 ml vitamin solution [0.08 mM biotin; 0.05 mM folic acid, 0.48 mM pyridoxin HCl, 0.14 mM thiamine HCl, 0.13 mM riboflavin, 0.4 mM nicotinic acid, 0.2 mM D-L-panthothenic acid, 0.36 mM p-aminobenzoic acid, 1.4 mM choline

chloride, 7.3×10^{-6} M vitamin B12], 7.5 μ M ferric sodium EDTA, 2 mM NH_4Cl , 2 mM NaHCO_3 (as carbon source) and 1 mM sodium pyruvate (as a free radical scavenger), was used for growth. The medium was buffered with 10 mM HEPES/ 6 mM NaOH to pH 7.5. FWM, trace element mixture, vitamin solution, NH_4Cl , NaHCO_3 , sodium pyruvate and HEPES buffer were filter sterilized with a sterile 0.2 μ m PVDF filter and stored in the attached polystyrene bottles at 4°C.

Culture Conditions

The main culture of *N. viennensis* was initiated from a 0.5 mL 40 % glycerol stock by first spinning down the cells at 16,168 x g for 30 min at 4 °C, removing the supernatant, and re-suspending the cell pellet in 0.5 mL FWM. The resuspended cells were used to inoculate a 500 mL flask containing FWM medium supplemented with kanamycin (100 μ g/mL) to prevent bacterial contamination. *N. viennensis* cultures were incubated at 42 °C, in the dark without shaking. Growth was monitored by following NH_4^+ consumption and NO_2^- production using two modified colorimetric assays [5][6] described below. Once Cu-replete cultures had consumed ~ 1 mM NH_4^+ , all cultures (including Cu-limited cultures) were split into subcultures by transferring cultures to 20 mL of FWM in polystyrene tubes supplemented with kanamycin (100 μ g/mL) and carbenicillin (100 μ g/mL) using a 0.25 % inoculation volume. To ensure a normal growth cycle, subcultures were continuously transferred for several weeks before PCR testing the cultures for bacterial contamination and initiating the Cu-limitation experiments.

Nucleic Acid Extraction and PCR Testing

Nucleic acids were extracted from 1 mL of cells in exponential growth phase. Cells were centrifuged at 16,168 x g (4 °C) for 30 minutes. The supernatant was discarded and cells were resuspended in SDS Extraction Buffer [0.14 M NaCl, 0.02 M Na_2SO_3 , 0.1 M Tris/HCl, 0.5 M EDTA pH 8 and 1 % SDS]. This mixture was transferred to sterile 2 mL microtubes containing ~ 0.5 g of 0.1 mm zircona beads that had been baked at 180 °C for at least 3 hours. Phenol-chloroform-isoamyl alcohol (0.5 mL) [PCI, 25:24:1 (vol:vol:vol); pH 6.7-8] was first added to facilitate DNA recovery. Cells were lysed by bead-beating at 4 m/s for 30 s using a Fast Prep-24 instrument (MP Biomedicals). After bead beating and removal of PCI, chloroform-isoamyl-alcohol (0.5 mL) [CI, 24:1 (vol:vol)] was added to samples. During the PCI and CI extraction

steps, tubes were centrifuged at 16,168 x g for 10 min at 4 °C. Between PCI and CI steps, before centrifugation, tubes were shaken heavily. Finally, nucleic acids were precipitated from the aqueous phase by adding 1 µL of glycogen and 2 volumes of 30 % polyethylene glycol (molecular weight 5000-7000) at 4 °C overnight, followed by centrifuging at 16,168 x g at 4 °C for 30 min. Pellets were washed with 1 mL of 4 °C cold 70 % ethanol, air dried for 3-5 min in a vacuum [Eppendorf Vacufuge Plus, 30 °C/AQ setting] and resuspended in 50 µL of DNase-free water. SDS Extraction Buffer, PEG and ethanol solutions were made using DNase-free water. Total DNA concentration was measured using the NanoDrop Spectrophotometer ND-1000 (PeqLab VWR International GmbH) following the manufacturer's instructions.

DNA verification of the culture samples was assessed by performing a PCR using two different archaea primer sets including: 0.2 mM Arc109f [5'-ACKGCTCAGTAACACGT-3'] [7] and Arc 1492r [5'-GYACCTTGTTACGACTT-3'] [8], CamoA-19f [5'-ATGGTCTGGYTWAGACG-3'] originally published by [9] and modified by [10], and TamoA-692r [5'-GCCATCCATCKRTANGTCCA-3'] [11] and possible bacterial contamination was assessed using 0.2 mM of primers 27f [5'-AGAGTTTGATCCTGGCTCAG-3'] and 1492r [5'-GGTTACCTTGTACGACTT-3'] [12]. 2 µL of DNA sample was used along with the following: 200 µM dNTP mix (Thermo Fisher), 0.002 mg/mL BSA, 0.15 µL, Dream-Taq DNA polymerase (Thermo Fisher), 1X Dream-Taq buffer and DNA free water in a 25 µL reaction volume. The following conditions were used for 16S rRNA primers: 95 °C for 5 min followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s and 72 °C for 2 min, and a final elongation step at 72 °C for 10 min. The following conditions were used for *amoA* primers: 95 °C for 5 min followed by 35 cycles of 94 °C for 30 s, 58 °C for 45 s and 72 °C for 45 s, and a final elongation step at 72 °C for 10 min. PCR reactions were checked on a 1 % low melting agarose gel.

Cu-limited Cultures

The trace element solution used in the Cu limited sub-cultures did not contain Cu. CuCl₂·2H₂O was supplemented to cultures at various free Cu and TETA concentrations as described in [13]. Briefly, the PhreeqC program [14] was used to estimate the Cu²⁺ and TETA concentrations needed to induce Cu limitation using the minteq.v4 database, stability constants from Anderegg et al. [15] Cu-limitation was assessed as a decrease in NH₄⁺ consumption and NO₂⁻ production relative to the Cu-replete culture. Total concentration of Cu in the FWM with

additives was measured using inductively coupled plasma mass spectrometry (ICP-MS; Agilent 7700 instrument)[16].

Ammonium and Nitrite Measurements

Ammonium concentrations were determined based on a modified method [5]. Briefly 200 μL of sample was combined with 400 μL of FWM. 300 μL of Color Reagent [5.18 mM sodium salicylate, 2.15×10^{-5} M sodium nitroprusside, 0.1 M NaOH] was added next followed by 120 μL Oxidation Solution [3.91×10^{-5} M of dichloroisocyanuric acid]. After mixing, samples were stored in the dark at room temperature for 30 min. The absorbance of 200 μL samples were measured in a 96-well plate with a UV spectrophotometer [TECAN Sunrise Spectrophotometer] at 660 nm wavelength. NO_2^- concentrations were determined by using a modified method [6]. Briefly, 20 μL sample was combined with 780 μL FWM and to this 200 μL of sulfanilamide/NED Reagent [0.058 M sulfanilamide, 1.9 mM N-(1-Naphthyl)-ethylendiamin dihydrochloride, 2.22 M ortho-phosphoric acid 85 %] was added. After mixing, the samples were stored in the dark at room temperature for 30 min before measuring the absorbance at 545 nm wavelength.

RNA Extraction

For RNA extraction, 1.25 mL of an exponential phase 20 mL culture was used to inoculate a 1 L polystyrene bottle containing 500 mL of FWM. Cultures were harvested on different days as described below (Fig. 1) and total RNA was extracted from five biological replicates. Growth continued to be monitored in the remaining cultures until stationary phase. A sterile 60 mL syringe that was equipped with a 25 mm Swinnex filter holder with gasket with o-ring (Millipore) was used to withdraw culture. In this way, 250 mL (for RNA-Seq experiments) of cells were collected on a 25 mm 0.2 μm mixed cellulose ester filter membrane (Millipore). The filters were added to water treated with 1 % DEPC overnight and immediately autoclaved. Wet filters were used within 1-2 days. Following the filtration step, filters were inserted into 5 mL lysis tubes and frozen on dry ice. After thawing the filters, the extraction procedure was started by first heating the filter to 65 $^{\circ}\text{C}$ for 10 min. RNA was extracted from cells using the

RNeasy Porewater extraction kit (Qiagen) according to the manufacturer's instructions with a few exceptions. Samples were split into 500 μ L sub-samples prior to addition of PWR 3 and 4 solutions. Following addition of PWR 3 and 4, multiple loads for each sample were added to the spin filter. The on-column DNase I digest was omitted. Nucleic acid was eluted in 100 μ L of RNase free water.

DNase I Digest and Total RNA Preparation for Sequencing

Nucleic acid was treated with RQ1 DNase I (Promega) using the manufacturer's instructions. DNase I and any remaining inhibitors were removed from the RNA with Zymo One Step PCR Inhibitor removal kit following the manufacturer's instructions. The DNA High Sensitivity Bioanalyzer Chip (Agilent) was used, following the manufacturer's instructions, to exclude the presence of genomic DNA. The concentration of RNA extracted was estimated using the RNA Qubit HS (Thermo Fisher) assay and the integrity of the RNA was determined with the Agilent RNA 600 Nano Kit.

Reverse Transcription (RT) PCR and Quantitative PCR (qPCR)

The same samples that were sequenced were used in the RT-qPCR analyses. Prior to the RT step, 0.5-1 μ g of total RNA was treated with Turbo DNase I (Thermo Fisher) following the manufacturer's instructions. After checking for DNA contamination using MCO primers in PCR as described below, total RNA was reverse transcribed using the Protoscript II M-MuLV reverse transcriptase (New England Biosciences) following the manufacturer's instructions. Reactions were cleaned up with the Monarch DNA cleanup kit (New England BioLabs) following the protocol for cDNA. As a normalization step, the concentration of cDNA was estimated using the Qubit ssDNA quantitation kit and all samples were diluted to 5 ng/ μ L prior to the qPCR step.

The Luna® Universal qPCR Master Mix (NEB) and the Mastercycler EP Gradient S (Eppendorf) qPCR thermocycler were used for absolute quantification of gene copy numbers. For each biological replicate sample or standard template, three technical qPCR replicates were tested. The primers listed in Table S2 were used to amplify the permease gene *perm_b*

(NVIE_000590), the multicopper oxidase genes *mco1* (NVIE_000600) and *mco4_b* (NVIE_019250), and a putative surface associated Ca^{2+} binding protein encoding gene *cbp_a* (NVIE_001000). Expression of the 16S rRNA gene and *rpoB* (NVIE_013510) was also assessed for comparative purposes.

Standard templates were prepared by amplifying *mco1* (NVIE_000600), *perm_b*, (NVIE_001000), 16S rRNA and *rpoB* genes from DNA extracts of 5 Cu-replete and 5 Cu-limited cell samples of *N. viennensis*, using the primers described above. The following PCR conditions were used for all primers: 98 °C for 3 min followed by 30 cycles of 98 °C for 10 s, 60 °C for 20 s and 72 °C for 30 s, followed by a final elongation step at 72 °C for 4 min. PCR products were cleaned up with the Monarch DNA cleanup kit (New England BioLabs) following the protocol for PCR products and cloned into the pJET/blunt plasmid using the CloneJet PCR Cloning Kit (Thermo Fisher) according to the manufacturer's instructions. Purified plasmids were linearized using Not I HF (NEB) restriction enzyme. Standard curves were generated by 10X serial dilutions of linearized plasmid stocks with a concentration of 0.2×10^8 copies/ μL (*perm_b* and *mco1* plasmids) and 0.2×10^9 copies/ μL for *mco4_b*, *cbp_a*, 16S rRNA, and *rpoB*. Final standards ranged from 10^1 to 10^8 copies of the targeted gene. Each 20 μL reaction mixture was prepared following the Luna Master Mix instructions. The following qPCR cycling conditions were used for all primers: 95 °C for 2 min followed by 40 cycles of 95 °C for 15 s and 60°C for 1 min. The specificity of each qPCR reaction was evaluated via melting curve. Relative difference in expression of genes *mco1*, *mco4_b*, *perm_b*, and *cbp_a* was first assessed by calculating the ΔCt of each gene against the ΔCt of 16S rRNA or *rpoB* (used as stably expressed reference genes). These values were then normalized to the average ΔCt of 16S rRNA or *rpoB* genes from the Cu-replete cultures [17]. Following normalization, one Cu-replete sample (R3) was left out of further analyses as amplification of genes for this sample was one order of magnitude lower compared to other samples.

Sequencing Analysis

Raw reads were downloaded from the Vienna Biocenter Facility website and checked using md5sum. FastQC [18] version 0.11.5 was used to initially analyze sequences and the samtools package version 1.9 [19] (using htslib 1.9) was used to convert file types. Trimmomatic

[20] version 0.36 trimmed adapter sequences and cropped reads using the options: SLIDINGWINDOW:4:15 LEADING:3 TRAILING:3 MINLEN:38 HEADCROP:13. The stand-alone program PrinSeq Lite [21] version 0.20.4 was used to remove sequences < 37 bp and/or with quality scores < 30. SortMeRNA[22] (version 2.1) separated rRNA sequences from total RNA using a curated reference database only containing *N. viennensis* rRNA sequences. Next, [23] “Bowtie2” [24] version 2.3.3.1 was used to align the remaining RNA sequences to the *N. viennensis* reference genome. The “featureCounts” [25] program, version 1.6.2, from the subread package assigned mapped sequencing reads to genomic features. While using featurecounts, a General Feature Format (GFF) annotation file with annotations from GenBank for *N. viennensis* from the National Center for Biotechnology Information (NCBI) website (http://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/698/785/GCA_000698785.1_ASM69878v1) was used as a reference. Feature type was specified as “gene”, and attribute type was specified as “ID”. Classified ambiguous sequences were omitted from the final output prior to further processing. The count matrix from feature counts was used as input for R version 3.2 [26] with RStudio version 0.99.902 [27]. The R bio-conductor package “DESeq2” version 1.10.1 [28] was used to normalize reads and to perform differential expression analysis. In order to lower the false discovery rate (FDR), the default p-value adjustment in DESeq2 was used (Benjamini-Hochberg method) to determine relevant p-values (p-adj in DESeq2 output). For principal component analysis, count data was normalized using the rlog function in DESeq2. A PCA plot was then produced from the normalized data using the R packages ggplot2 [29], ggrepel [30], and factoextra [31]. R scripts for the differential expression analysis and PCA plot production are found on Github (https://github.com/hodgskiss/Nviennensis_copper_limitation). Up and downregulated genes were assigned to functional categories based on their archaeal clusters of orthologous group (arCOG) categories [32]. A statistical hypergeometric test was performed (described below) to determine which categories were significantly enriched (p-value < 0.05) in the up or downregulated gene groups.

arCOG Functional Group Assignment and Enrichment Analysis

Archaeal Clusters of Orthologous Groups, also known as arCOGs [32] (2014 version), were assigned to genes of *Nitrososphaera viennensis* EN76, *Nitrosopumulis maritimus* SCMI,

Candidatus Nitrososphaera gargensis Ga9.2, *Candidatus Nitrososphaera evergladensis* SR1, *Nitrosotenuis chungbukensis* MY2, *Nitrosoarchaeum limnia* SFB1, and *Candidatus Nitrosopumulis sediminis* AR2 using the COGnitor scripts, with an e-value threshold set to 10^{-10} . Manual inspection of the protein families resulted in grouping into custom functional categories. Functional enrichment analysis was performed using an arCOG enrichment R script that utilizes the `phyper()` R function as previously described [33] (https://github.com/amyschmid/histone_arCOG). In brief, a hypergeometric distribution is used to determine the statistical chance of certain categories (in this case arCOGs) being drawn from a set (in this case the total arCOGs found in the genome).

RPKM Calculation

The resulting data from the featurecounts program was used to calculate the reads per kilo base per million (RPKM) for each gene in each sample. The counts for each sample were totaled and divided by one million. This factor was applied respectively for every gene in each sample (i.e. [number of counts for a gene]/factor). Each gene count was also divided by the length of the respective gene in kilobases. This final value, or RPKM, was log transformed with a base of two. Averages for each gene were taken for the copper replete and copper limited conditions and can be found in Dataset S1 for each gene.

Phylogenetic Analysis of the CopD Family in Thaumarchaeota

The sequences from the three proteins harboring CopD domains in *N. viennensis* (NVIE_013130 – WP_084790666.1, NVIE_014300 – WP_075054630.1 and NVIE_014310 – WP_075054632.1) were used to perform a blast similarity search against a database comprising a selection of genome sequences from 39 AOA, plus 341 genomes from representative clades of bacteria, archaea and eukaryotes (Dataset S1). Sequences corresponding to hits with an e-value lower than 10^{-6} were collected, and these hits were then screened with HMMER v3.1b2 [34] against PFAM v28.0 [35] for CopD domains. Sequences harboring CopD domains were selected, and the coordinates (“envelope”) of the CopD domains were used to extract parts of the sequences corresponding to these domains. CopD domains from selected sequences were thus

aligned using the MAFFT program v7.310 (“linsi”) [36], resulting in a 175-sites alignment. A phylogenetic tree was built from this alignment using IQ-Tree v1.6.11 [37], with the best model selected (“-m TESTNEW”) and with the estimation of the support values (“-bb 1000 -alrt 1000 -safe”). The resulting tree was then annotated in iTOL [38] with the protein domain architecture as obtained with the PFAM screen. The figure (Fig. S5) was finalized using the InkScape program (<http://www.inkscape.org/>).

Motif Analyses

The web version of Multiple Em for Motif Elicitation (MEME) software Suite 5.0.5 was used to identify motifs in upstream regions of selected genes. For each analyzed upstream sequence, the first 100 bases upstream of the annotated translational start codon were considered unless the end of an upstream gene caused this region to be shorter. Upstream bases were extracted using The Artemis Software [39]. For genes in a presumed operon, the upstream start region of the operon was considered instead. Genes were presumed to be in an operon if the distance between the end of an open reading frame and the start of the next, on the same strand, was $\sim < 40$ bp. The option “any number of occurrences” was used to find repeated motifs for the selected up and downregulated gene groups. For the list of homologous genes in Thaumarchaeota, the search was limited to “search given strand only” within the advanced options. Once a motif was found in selected upregulated genes, it was used to screen against the top 25 upregulated genes using the web version of Find Individual Motif Occurrences (FIMO) version 5.1.1 [40].

Genes Neighboring Downregulated Genes and Homolog Search in Closely Related Thaumarchaeal Species

Based on the observation that several downregulated genes containing the motif are located in the same genomic region, the motif search was adapted to contain not only the previously identified genes with the motif, but also NVIE_000950, NVIE_000960, NVIE_000970 and NVIE_000980, as most of these genes were also downregulated in the Cu-limited condition and are located between NVIE_000940 and NVIE_001000. The motif was found in the upstream region of NVIE_000970 and in the upstream region on the opposite strand in reference to the gene orientation for genes NVIE_000950 and NVIE_000980. Interestingly,

this promoter region is largely palindromic resulting in NVIE_000990 and NVIE_000980 sharing the same conserved region. Although not confirmed in this study, it seems possible that bi-directional transcription, defined as transcription in opposite directions [41] from one core promoter, may occur.

While most of these additional genes do not have an annotated function, NVIE_000950 stands out as having a helix-turn-helix (HTH) ArsR and 4-vinyl reductase (V4R) signature. Whereas the V4R signature indicates a potential to bind hydrocarbons [42] the HTH ArsR domain corresponds to a group of transcriptional regulators known to respond to the presence of metals. In particular, they are known as repressors that bind to DNA in the absence of metals. In the presence of metals, these repressors are released from the DNA and transcription can proceed [43]. These types of regulators are also known to be involved in Cu regulation systems [44]. It is possible that such a repressor that binds to DNA regulates the downregulated genes identified here. Under Cu replete conditions, the binding of Cu to the repressor would release it from the DNA and allow transcription to proceed. NVIE_000950, or even the upregulated NVIE_014280, could be fulfilling this regulatory function leading to the downregulation of this gene cluster as both are detected in the proteome of copper limited cultures. However, this hypothesis would need to be verified by molecular methods. It should be noted that although NVIE_000950 was found in the proteome of copper limited cultures, it is not differentially expressed between the two conditions based on the p-value cutoff of 0.01.

To further investigate the occurrence of this sequence, the basic local alignment tool (BLAST) [45][46] was used to search for this promoter sequence in other areas of the *N. viennensis* genome. This resulted in a sixth gene being identified with the motif: NVIE_005310, a transcriptional regulator from the Lrp family. This gene, however, is not differentially expressed between the two conditions based on the p-value cutoff of 0.01.

With the exception of NVIE_000950, the identified promoter was found at a similar distance from the start codon of its respective gene. Because of the high level of conservation in nucleotide composition of this motif and its distance from the start codon, similar proteins to genes with this motif and/or in this cluster were searched for in other Thaumarchaeota species to determine if this motif sequence was more widespread. Similar proteins were identified by BLAST [45][46] results in the above mentioned Thaumarchaeota species (see “*arCOG Functional Group Assignment and Enrichment Analysis*” section for a list of species) or

identified based on similarity and synteny using the MicroScope web interface Microbial Genome Annotation & Analysis Platform (MAGE) [47]. Similar proteins for NVIE_000960 were not searched for as the motif was not as strong in the promoter region of this gene (Fig. S6B). A complete list of similar proteins can be found in Dataset S1. Upstream regions were taken for each of the identified genes and MEME was used to look for the motif sequence with the search limited to “search given strand only” within the advanced options. An alignment of promoter regions for genes with the identified motif can be seen in Fig. S6B.

Combined Protein and RNA Extraction

Proteins and RNA were extracted using a combined TRIzol (Invitrogen) extraction method from biomass frozen on filters stored at -80 °C from cultures L7 and L8. These methods closely follow those of Ott *et al.* (2019) [48] with the exception of combining RNA and protein extraction, rather than metabolites and proteins, in the same step. Additional slight changes were made in the protocol.

Filters were taken from the freezer and added to 2 mL homogenization tubes (Sarstedt, order number: 72.693.465) containing approximately 0.6 grams of Lysing Matrix B (MP Biomedical). 1 mL of TRIzol reagent was added to each tube and samples were then homogenized using a FastPrep-24 homogenizer (MP Biomedical) set at 4 m/s for 30s. After homogenization, samples incubated at room temperature for 15 min. Following incubation, 200 μ L of chloroform was added and each sample was inverted five times before incubating at room temperature for another three minutes. Samples were then centrifuged at 16,168 x g and 4 °C to separate phases. After centrifugation, the lower phase (chloroform), containing proteins, and the upper phase, containing RNA, were removed into separate low bind 2 mL Protein LoBind tubes (Eppendorf).

To finish the protein extraction, 550 μ L (approximately a 1:1 mix) of filtered water (MilliporeSigma Milli-Q Reference A+ System) was added to each protein sample. Tubes were inverted five times and incubated at room temperature for 3 min. Each sample was then centrifuged at 16,168 x g and 4 °C for 2 min to separate phases. The lower apolar phase containing proteins was transferred to a new 2 mL low bind epi. 1.5 mL of 0.1 M NH_4Cl

dissolved in methanol with 0.5 % β -mercaptoethanol was added to the apolar phase. Samples sat on ice for 3-4 hours before being stored overnight at -20 °C to allow proteins to precipitate. The next day, samples were centrifuged at 16,168 x g for 15 min at 4 °C to allow proteins to pellet. The supernatant was discarded and proteins were resuspended in 1.8 mL of ice cold methanol. The sample was sonicated in a Transsonic 700/H water bath (Elma) until proteins were completely resuspended. The samples were centrifuged at 16,168 x g for 10 min at 4°C and the supernatant was discarded. The washing process with methanol was repeated a second time. After the second methanol washing step, proteins were suspended in 1.8 mL of ice cold acetone. Again, each sample was sonicated in a Transsonic 700/H water bath (Elma) until proteins were completely resuspended. Each sample was centrifuged at 16,168 x g for 15 min at 4°C. The acetone supernatant was discarded and samples were allowed to air dry under a fume hood for 5-10 min (or until just after being fully dried) and then stored at -80 °C.

To finish RNA extraction, 10 μ g of glycogen and 0.5 mL of isopropanol was added to each RNA sample (the previously separated aqueous phase). Each sample was allowed to sit for approximately 60 min at room temperature before being centrifuged at 12,000 x g for 10 min at 4°C. The supernatant was discarded and the RNA was resuspended in 1 ml of 75 % ethanol. Each sample was briefly vortexed and then centrifuged at 7,500 x g for 5 min at 4 °C. The supernatant was discarded and the RNA pellet was allowed to air dry. The dried pellet was resuspended in 25 μ L of diethyl pyrocarbonate (DPEC) treated water and stored at -80 °C for later RT-qPCR analysis.

Protein Digestion

Extracted proteins were resuspended in 500 μ L of extraction buffer (8 M urea in 50 mM HEPES buffer, pH 7.8) and allowed to shake for 30 min at 4 °C at 900 rpm. Protein concentration was then measured using the Bradford assay (Bio-Rad Cat. No. 500-0006) using a pre-made standard curve with bovine serum albumin (BSA) and an absorbance of 545 nm. Based on the results of the Bradford assay, a volume corresponding to 25 μ g of protein was removed to a new low bind epi for each sample respectively. Extraction buffer was used to bring the total volume of each sample to 220 μ L. 1,4-dithiothreitol (DTT) was added to get a concentration of 5 mM in each sample followed by incubation for 45 min at 37 °C and 700 rpm. Iodoacetamide

(IAA) was then added to a concentration of 10 mM followed by incubation in the dark for 60 min at 30 °C and 700 rpm. A final addition of DTT was added to bring the total concentration up to 10 mM followed by incubation in the dark at room temperature for 15 min. To digest proteins, 5 µL of 0.1 µg/µL of mass spec grade rLysC (Promega) was added to each sample. Proteins were digested for 3 hours at 37 °C and 700 rpm. After the rLysC digestion, 660 µL of trypsin buffer (50 mM NH₄HCO₃, 2 mM CaCl₂, 5 mM DTT, 10 % acetonitrile) was added to each sample. This is necessary to dilute the urea concentration to 2 M for the trypsin digestion. 1 µL of trypsin beads from a Poroszyme Immobilized Trypsin Cartridge (Applied Biosystems) were added to each sample. Trypsin digestion was carried out for 16 hours at 30 °C on a rotator. Following digestion, peptides were desalted using OMIX C18 pipette tips (Agilent Technologies). For this, samples were centrifuged briefly to pellet the trypsin beads before desalting. Tips were activated by washing with 100 µL of methanol and then washed twice with 100 µL of 0.1 % formic acid (FA). Peptide solution was then acidified with the addition of FA to a final concentration of 3 %. The solution containing peptides was pipetted through the C18 tips and then washed twice with 100 µL of 0.1 % FA. Peptides were then eluted and saved from the C18 tips by being washed twice with 100 µL of methanol. After desalting, peptides were dried in a ScanSpeed 40 speed vacuum with a ScanVac vacuum control and stored at -20 °C.

Mass Spectrometry of Peptides

To prepare the peptides for mass spectrometry analysis, peptides were resuspended in 250 µL of 2 % acetonitrile (ACN) and 0.1 % formic acid (FA). 5 µL were then injected into an EASY-Spray C18 column (2µm, 100Å 75 µm x 50 µm, Thermo). Peptides were eluted from the column for 150 min using a 90 min linear gradient starting from 96 % solvent A (0.1 % FA) and 4 % solvent B (80 % ACN, 0.1 % FA) to 35 % of solvent B with a flow rate of 0.3 µL/min. The linear gradient was followed by an increase of solvent B to 90 % over 1 min and held at 90 % for 8 min. Solvent B was then adjusted from 90 % to 4 % over 1 min and held at 4 % for the remaining 50 min. Ion source was an EASY-Spray source with a spray voltage of 1.9 kV. Mass spectrometry measurements were taken using an LC-QExactive-Plus (Thermo) with the following settings: 0 to 150 min; MS1: positive polarity, full scan range 380-1800 *m/z*, resolution

70,000, collision-induced dissociation (CID) fragmentation for the 20 most intense ions; MS2: loop count 20, resolution 17,500, scan range 200-2000 *m/z*.

Peptide Analysis

Output files were analyzed using MaxQuant version 1.6.3.3 [49] with a reference proteome for *N. viennensis* downloaded from the Uniprot website (proteome UP000027093). The variable modifications oxidation (M) and acetyl (Protein N-term) were allowed along with the fixed modification carbamidomethyl (C). Digestion mode was set for trypsin and lysC with a maximum missed cleavages of 2. The protein false discovery rate was set at 0.01 and a minimum number of peptides at 1. The mqpar.xml file with relevant parameters is available on Github (https://github.com/hodgskiss/Nviennensis_copper_limitation).

References

1. Raimunda D, González-Guerrero M, Leeb BW, Argüello JM, Argüello JM. The transport mechanism of bacterial Cu⁺-ATPases: distinct efflux rates adapted to different function. *Biomaterials* 2011; **24**: 467–75.
2. Solioz M, Odermatt A. Copper and silver transport by CopB-ATPase in membrane vesicles of *Enterococcus hirae*. *J Biol Chem* 1995; **270**: 9217–21.
3. Orell A, Navarro CA, Arancibia R, Mobarec JC, Jerez CA. Life in blue: Copper resistance mechanisms of bacteria and Archaea used in industrial biomining of minerals. *Biotechnol Adv* . 2010. , **28**: 839–848
4. Zoltner M. Structural and functional characterization of prokaryotic CPx-ATPases: evidence for a novel cobalt P-type ATPase. *PhD Thesis* . 2006. International School Bioscience Ruhr University, Bochum.
5. Kandeler E, Gerber H. Short-term assay of soil urease activity using colorimetric determination of ammonium. *Biol Fertil Soils* 1988; **6**: 68–72.
6. Griess P. Bemerkungen zu der Abhandlung der HH. Weselsky und Benedikt „Ueber

- einige Azoverbindungen". *Berichte der Dtsch Chem Gesellschaft* 1879; **12**: 426–428.
7. Grosskopf R, Janssen PH, Liesack W. Diversity and structure of the methanogenic community in anoxic rice paddy soil microcosms as examined by cultivation and direct 16S rRNA gene sequence retrieval. *Appl Environ Microbiol* 1998; **64**: 960–9.
 8. Delong EF. Archaea in coastal marine environments. *Proc Natl Acad Sci USA* 1992; **89**: 5685–5689.
 9. Tournu M, Stieglmeier M, Spang A, Könneke M, Schintlmeister A, Urich T, et al. *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc Natl Acad Sci U S A* 2011; **108**: 8420–8425.
 10. Pester M, Rattei T, Flechl S, Gröngöft A, Richter A, Overmann J, et al. amoA-based consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of amoA genes from soils of four different geographic regions. *Environ Microbiol* 2012; **14**: 525–539.
 11. Arce MI, von Schiller D, Bengtsson MM, Hinze C, Jung H, Alves RJE, et al. Drying and rainfall shape the structure and functioning of nitrifying microbial communities in riverbed sediments. *Front Microbiol* 2018; **9**: 2794.
 12. Eden P a, Schmidt TM, Blakemore RP, Pace NR. Phylogenetic analysis of *Aquaspirillum magnetotacticum* using polymerase chain reaction-amplified 16S rRNA-specific DNA. *Int J Syst Bacteriol* 1991; **41**: 324–325.
 13. Reyes C, Hodgskiss L, Baars O, Kerou M, Bayer B, Schleper C, et al. Copper limiting threshold in the terrestrial ammonia oxidizing archaeon *Nitrososphaera viennensis*. *Res Microbiol* 2020; 1–9.
 14. Parkhurst DL. User's guide to PHREEQC, a computer program for speciation, reaction-path, advective-transport, and inverse geochemical calculations. *Water-Resources Investig Rep* 1995.
 15. Anderegg, Giorgio, Arnaud-Neu, Francoise, Delgado Rita, Felcman, Judith, Popov K. Critical evaluation of stability constants of metal complexes of complexones for biomedical environmental applications. *Pure Applied Chemistry* . 2005.

16. Technologies A. Agilent ICP-MS (8800 / 7700) Introduction System (ISIS). 2012. Agilent Technologies.
17. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001; **29**: 45e – 45.
18. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010.
19. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.
20. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; **30**: 2114–2120.
21. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 2011; **27**: 863–864.
22. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 2012; **28**: 3211–3217.
23. Team RC. A language an environment for statistical computing. R Foundation for Statistical Computing, Vienna. 2013.
24. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012; **9**: 357–9.
25. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014; **30**: 923–930.
26. R Foundation for Statistical Computing. A language and environment for statistical computing. 2016. Vienna, Austria.
27. Team Rs. RStudio: Integrated Development for R. 2016. Boston, MA.
28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; **15**: 550.
29. Wickham H. ggplot2: Elegant graphics for data analysis. 2016. Springer-Verlag, New York.

30. Slowikowski K. Automatically position non-overlapping text labels with ‘ggplot2’. 2019. , R package version 0.8.1
31. Kassambara, Alboukadel, Mundt F. Multivariate data analyses. 2017. , R package version 1.0.5
32. Makarova K, Wolf Y, Koonin E. Archaeal clusters of orthologous genes (arCOGs): an update and application for analysis of shared features between Thermococcales, Methanococcales, and Methanobacteriales. *Life* 2015; **5**: 818–840.
33. Dulmage KA, Darnell CL, Vreugdenhil A, Schmid AK. Copy number variation is associated with gene expression change in archaea. *Microb genomics* 2018; **4**: e000210.
34. Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol* 2011; **7**: e1002195.
35. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: The protein families database. *Nucleic Acids Res* . 2014. Oxford University Press. , **42**: D222
36. Katoh K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 2002; **30**: 3059–3066.
37. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015; **32**: 268–274.
38. Letunic I, Bork P. Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics* 2007; **23**: 127–128.
39. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice T, Rajandream M, et al. Artemis sequencing visualization and annotation. *Bioinformatics* 2000; **16**: 944–945.
40. Grant CE, Bailey TL, Noble WS. FIMO: Scanning for occurrences of a given motif. *Bioinformatics* 2011; **27**: 1017–1018.
41. Bagchi DN, Iyer VR. The determinants of directionality in transcriptional initiation. *Trends Genet* 2016; **32**: 322–333.
42. Anantharaman V, Koonin E V, Aravind L. Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains. *J Mol Biol* 2001; **307**:

- 1271–1292.
43. Osman D, Cavet JS. Bacterial metal-sensing proteins exemplified by ArsR–SmtB family repressors. *Nat Prod Rep* 2010; **27**: 668.
 44. Rademacher C, Masepohl B. Copper-responsive gene regulation in bacteria. *Microbiol (United Kingdom)* 2012; **158**: 2451–2464.
 45. Zhang Z, Schwartz S, Wagner L, Miller W. A greedy algorithm for aligning DNA sequences. *J Comput Biol* . 2000. , **7**: 203–214
 46. Morgulis A, Coulouris G, Raytselis Y, Madden TL, Agarwala R, Schäffer AA. Database indexing for production MegaBLAST searches. *Bioinformatics* 2008; **24**: 1757–1764.
 47. Vallenet D, Calteau A, Dubois M, Amours P, Bazin A, Beuvin M, et al. MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Res* 2020; **48**: D579–D589.
 48. Ott E, Kawaguchi Y, Özgen N, Yamagishi A, Rabbow E, Rettberg P, et al. Proteomic and metabolomic profiling of *Deinococcus radiodurans* after exposure simulated low Earth orbit vacuum. *Front Microbiol* 2019; **10**: 909.
 49. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008; **26**: 1367–1372.

Supplementary Figures and Tables

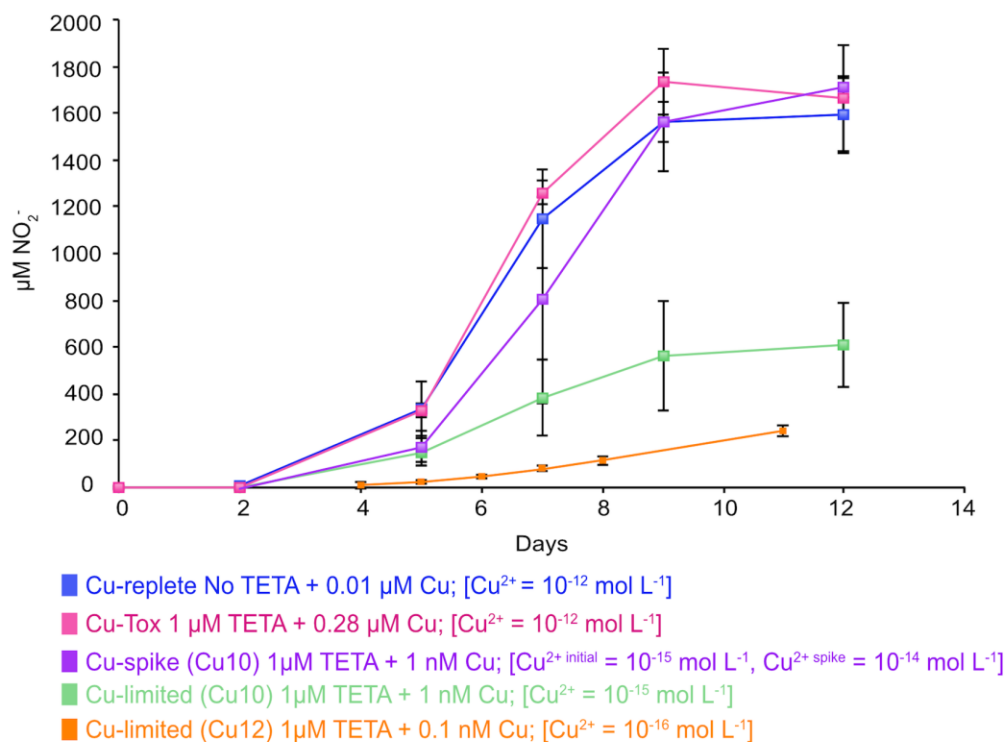


Figure S1. Nitrite concentrations in batch cultures of *N. viennensis* grown in 30 mL polystyrene bottles under Cu-replete and Cu-limited (Cu10 and Cu12) conditions as a function of the calculated free Cu^{2+} species concentration. The data and error bars represent the average and standard deviation of multiple biological replicates (≥ 4) from several experiments.

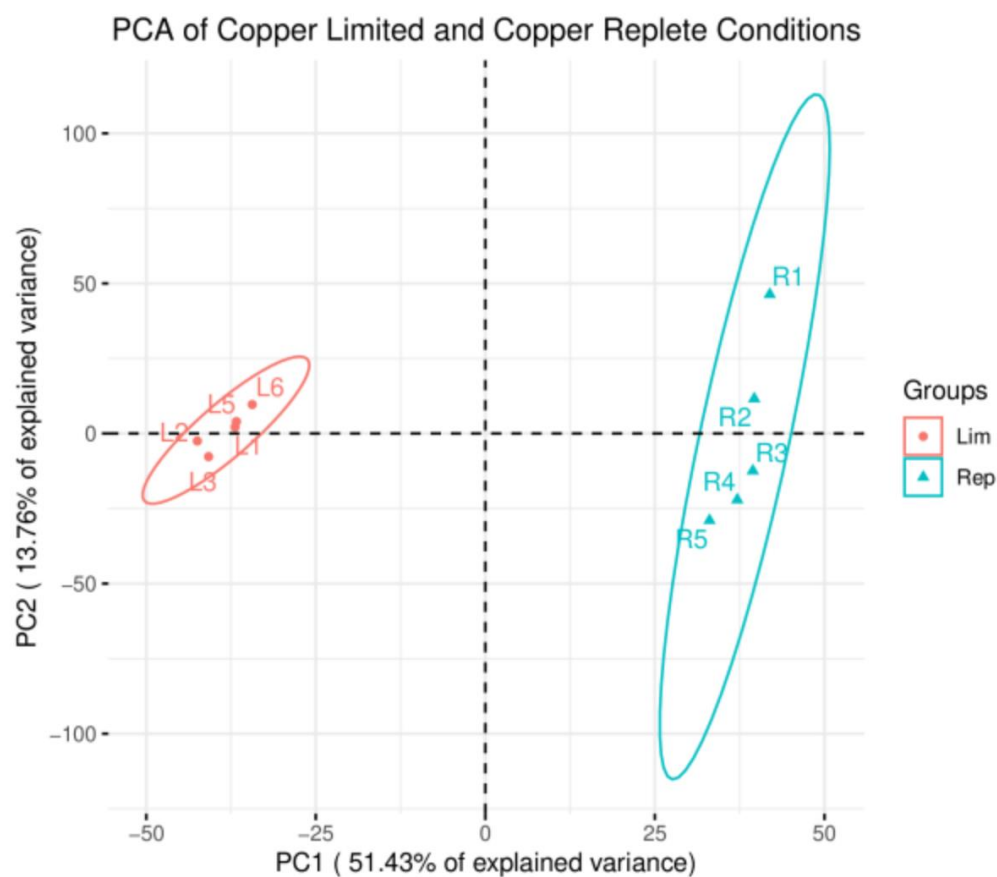


Figure S2. Principal component analysis of rlog transformed expression data of Cu-limited (Lim) and Cu-replete (Rep) samples. The x-axis describes variability between Cu-limited and Cu-replete samples. The y-axis describes variability between the replicates of each condition. Ellipses represent concentration ellipses for each group at an ellipse.level=0.95.

Alignment of 2 domain MCO's

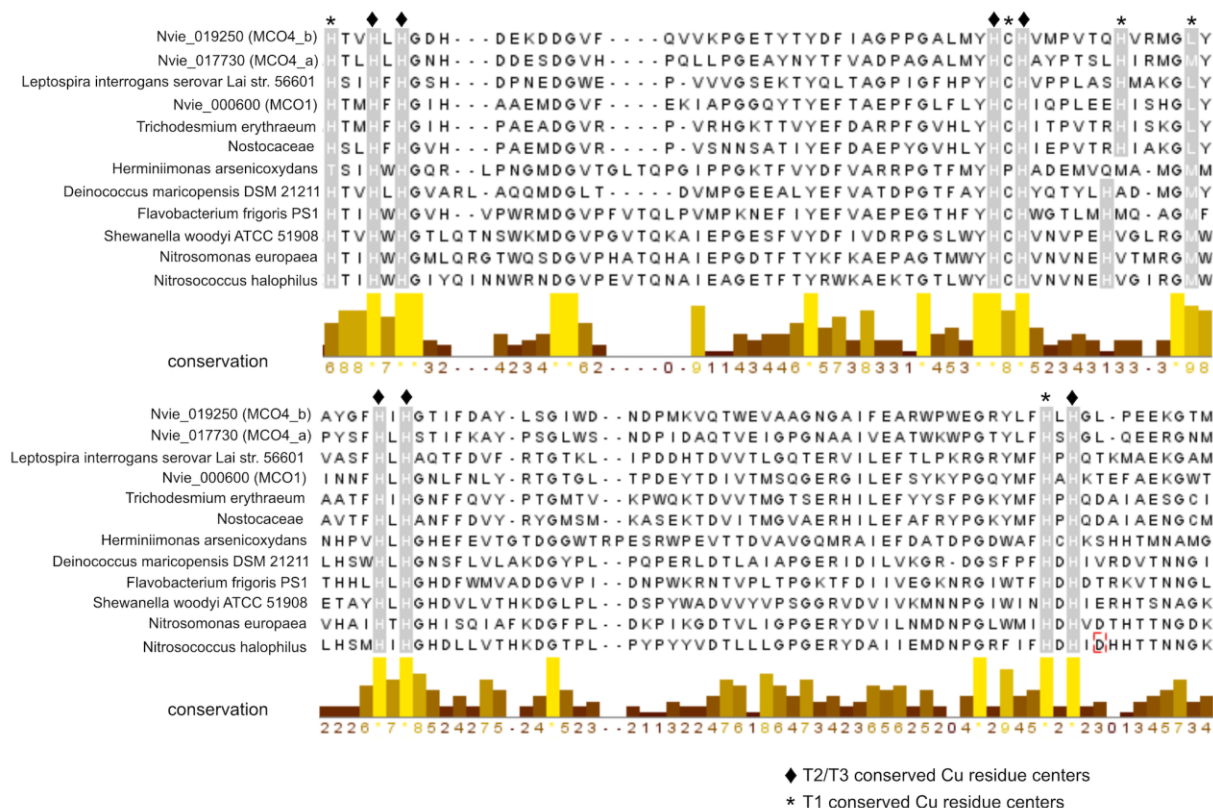


Figure S3. Alignment of 2dMCOs showing the predicted Cu binding sites. 2dMCO sequences from the following microorganisms were aligned using the Clustal Omega multiple alignment web interface (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) using default parameters: WP_084790903.1 copper oxidase; NVIE_019250 [*Nitrososphaera viennensis*], WP_075054901.1 copper oxidase; NVIE_017730 [*Nitrososphaera viennensis*], WP_075053492.1 copper oxidase; NVIE_000600 [*Nitrososphaera viennensis*], NP_711736.1 multicopper oxidase-like protein [*Leptospira interrogans* serovar Lai str. 56601], WP_011610637.1 multicopper oxidase [*Trichodesmium erythraeum*], WP_010998083.1 MULTISPECIES: multicopper oxidase [*Nostocaceae*], WP_011871074.1 copper oxidase [*Herminiimonas arsenicoxydans*], ADV66099.1 multicopper oxidase type 3 [*Deinococcus maricopensis* DSM 21211], EIA07124.1 multicopper oxidase [*Flavobacterium frigidum* PS1], ACA85706.1 multicopper oxidase type 3 [*Shewanella woodyi* ATCC 51908], KXX47905.1 multicopper oxidase [*Nitrosomonas europaea*], WP_013032143.1 multicopper oxidase [*Nitrosococcus halophilus*].

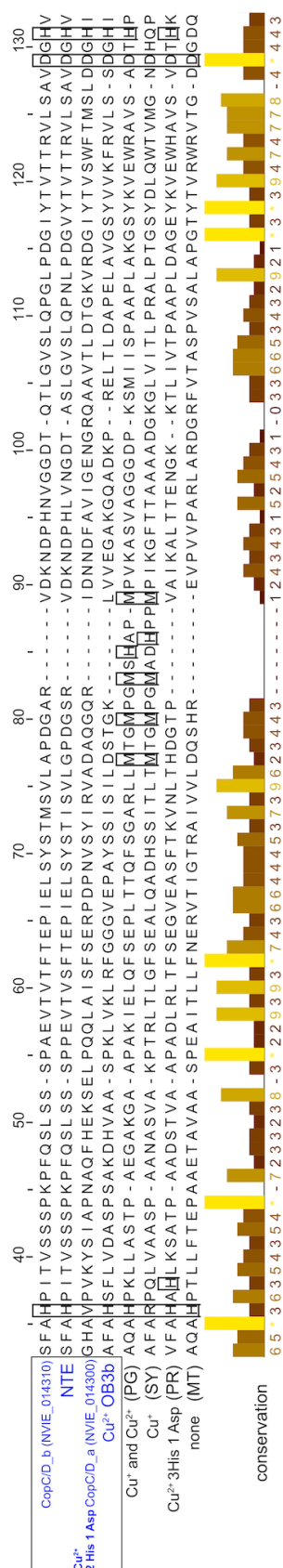


Figure S4. Alignment of the *N. viennensis* CopC/D amino acid sequences to CopC representative sequences from Lawton et al. (Biochemistry 55:2278-2290, 2016). Amino acid sequences were aligned using the Clustal Omega multiple alignment web interface (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) using default parameters. Copper binding residues are boxed in black. Representative sequences are from the following organisms: *Nitrososphaera viennensis* (NVIE_014300, NVIE_014310), (NTE) *Nitrososphaera evergladensis* SR-1 (IMG Id 2586122348), (PG) *Pseudomonas* GM17 (IMG ID 2503652063), (OB3b) *Methylosinus trichosporium* OB3b (IMG ID 2507406428), (PR) *Pseudomonas fluorescens* R124 (IMG ID 2503652063), (MT) *Mycobacterium tusciae* JS617 (IMG ID 2508745664), and (SY) *Sphingobium yanoikuyae* XLDN2-5 (IMG ID 2549030586).

1 subst/site

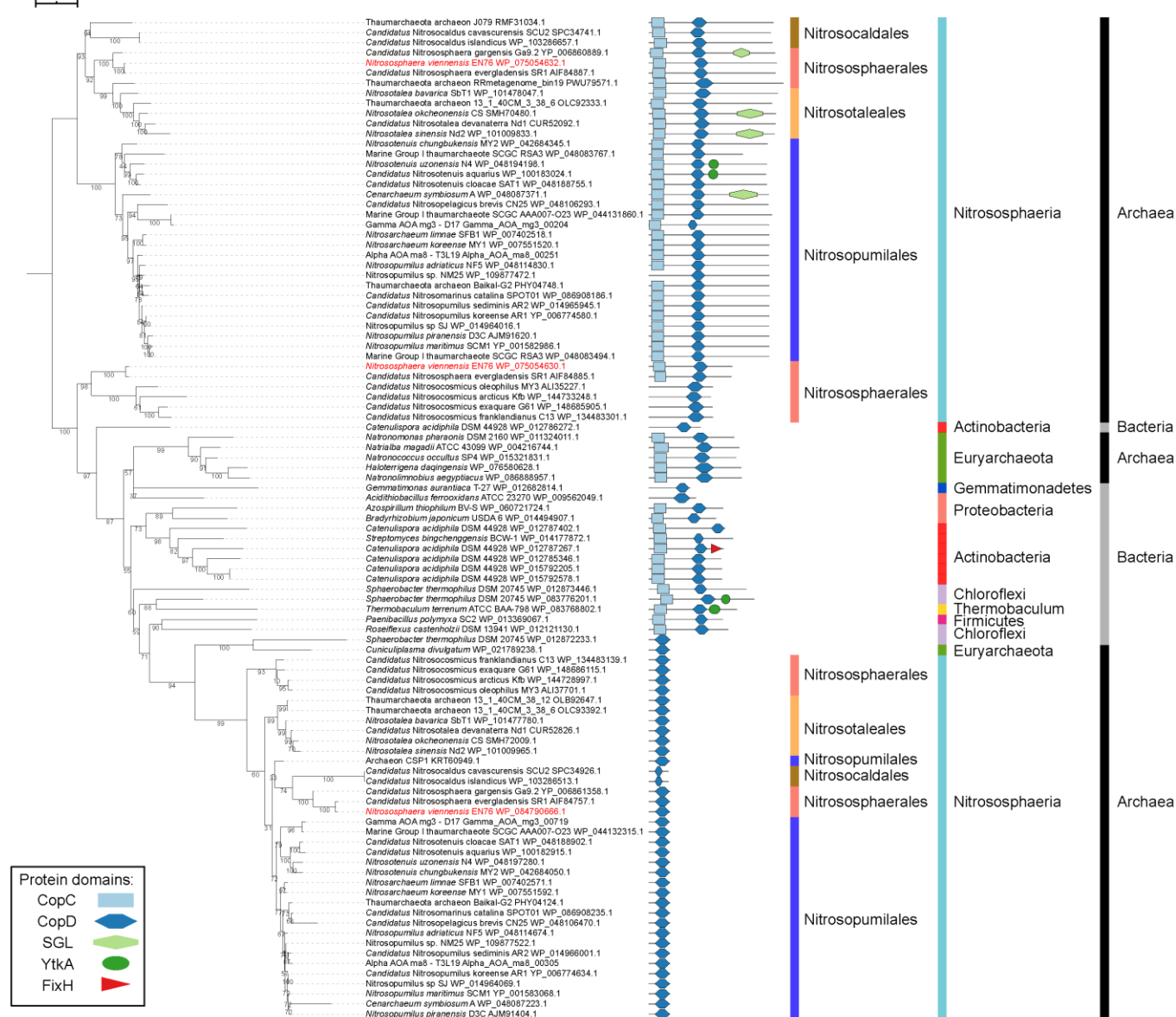
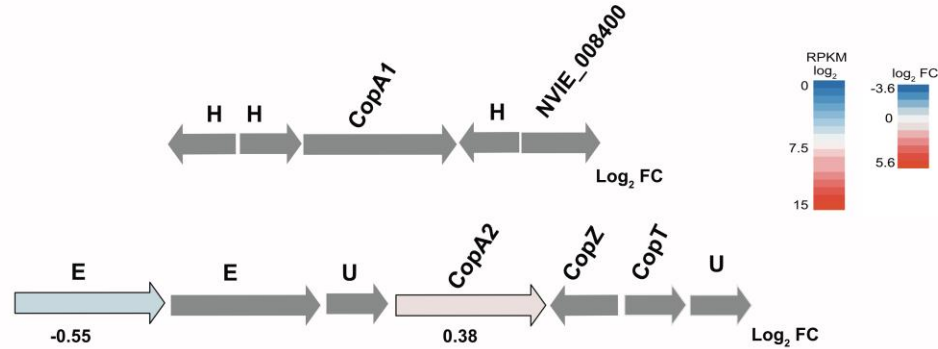


Figure S5. Phylogeny and protein domain architecture of CopD proteins found in AOA (Nitrososphaeria, Thaumarchaeota).

This phylogenetic tree was built on the basis of the alignment of the shared protein domain: CopD. The sequences were retrieved by sequence similarity search from a dataset representative of the diversity of the tree of life (see Supplementary Methods). The protein domain architecture is displayed for each sequence, as retrieved using PFAM domains' annotation. For AOA, the name of the orders is indicated along the tree with colored strips, whereas for other lineages, only the two highest taxonomic ranks are indicated ("phylum" and "kingdom" from left to right). Sequences found in *N. viennensis* are shown in red. DUF4149 ("Domain of Unknown Function") domains that were overlapping CopD domains were ignored for display for two sequences: *Sphaerobacter thermophilus* DSM 20745 (WP_012872233.1) and *Candidatus Nitrososphaera exaquare* G61 (WP_148686115.1). "CopC" stands for protein domain PF04234.8 ("Copper resistance protein C"), "CopD" for protein domain PF05425.9 ("Copper resistance protein D"), "SGL" for PF08450.8 ("SMP-30/Gluconolactonase/LRE-like region"), "YtkA" for PF13115.2 ("YtkA-like") and "FixH" for PF05751.7. Support values correspond to UltraFast bootstraps computed by IQ-Tree.

A.



B.

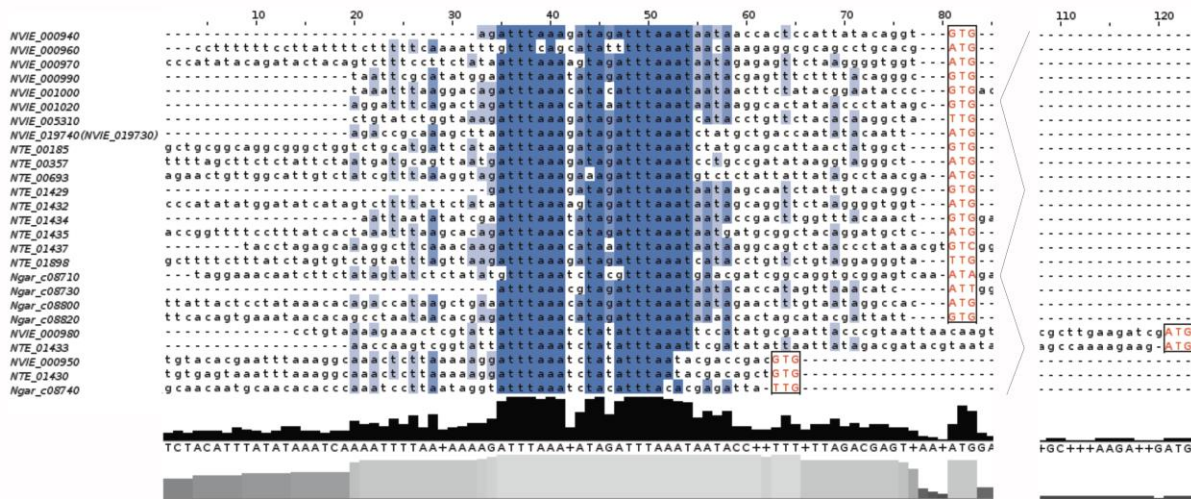


Figure S6. (A) Cop A1 and A2 (copper-exporting P-type ATPase A proteins), CopZ and CopT predicted gene clusters along with their gene expression information. A p-value < 0.01 cutoff was used to determine significance. Gene arrow colors represent $\log_2$ fold change of genes in copper limited cultures. Gray arrows are not differentially expressed. U = unknown protein, E = exported unknown protein. **(B)** Alignment of down-regulated genes with a motif sequence (5'-RATTTAAABATAGATTTAAATAAAS-3') from *N. viennensis* (NVIE) with other Thaumarchaeota genes containing a homologous motif. Ngar = *Nitrososphaera gargensis*, NTE = *Candidatus Nitrososphaera evergladensis*. Predicted translational start codons are boxed and text is in red. Locus tag information for all genes can be found in Dataset S1.

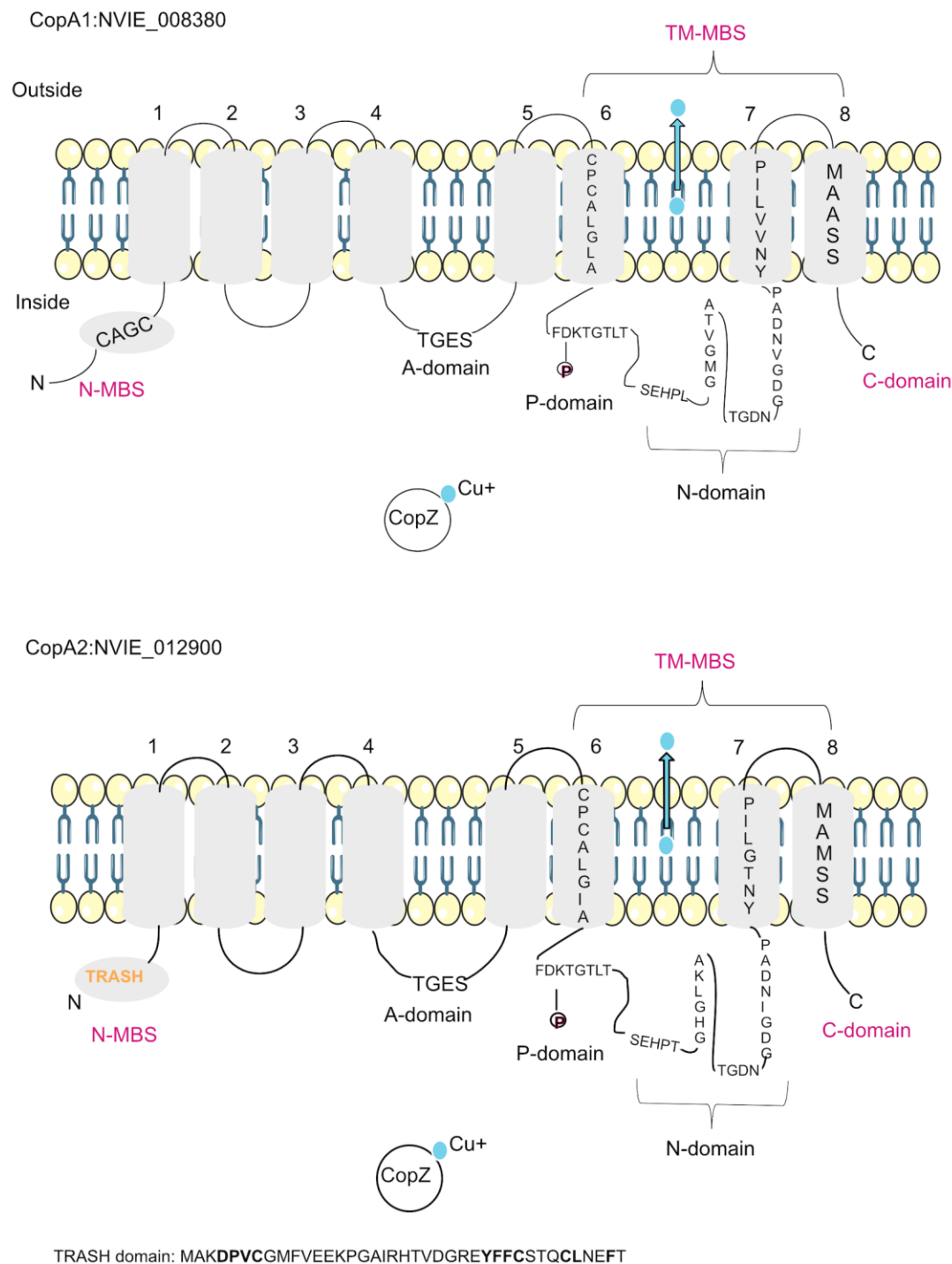
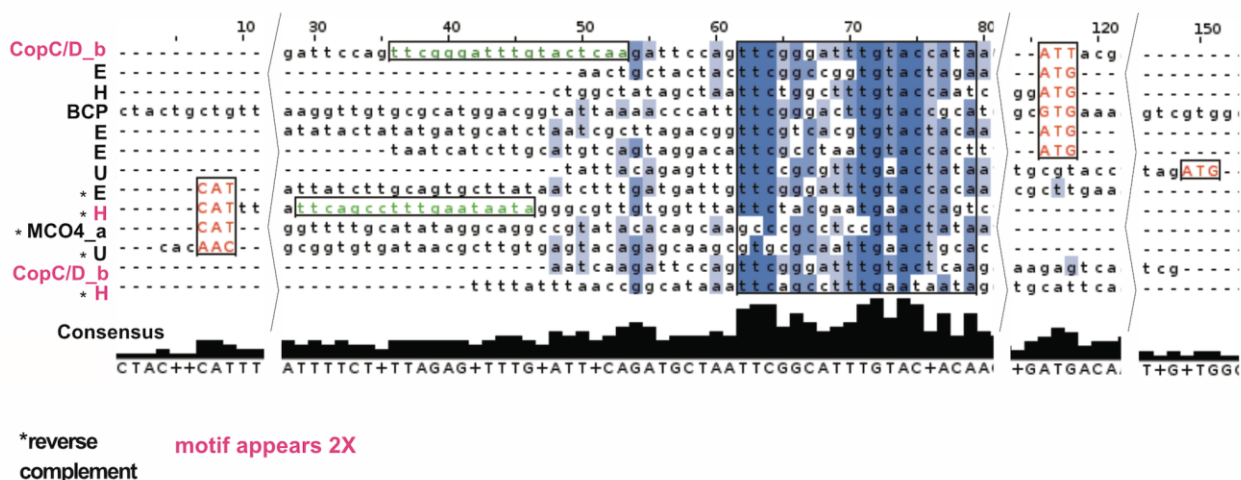


Figure S7. (A) *N. viennensis* CopA1 ATPase predicted metal binding sites. Amino acid metal binding sites information was gathered from several studies and reviews [1][2][3][4]. N-MBS = N terminal metal binding site, TM-MBS = transmembrane metal binding site, C-domain = C terminal metal binding site, A-domain = actuator phosphatase P-domain = phosphorylation, N-domain = nucleotide binding. Arrow indicates path of Cu delivery by CopZ to TM-MBS (B) *N. viennensis* CopA2 predicted metal binding sites. Amino acid metal binding sites information was gathered from several studies and reviews cited above. TRASH= motif involved in metal coordination, N-MBS = N terminal metal binding site, TM-MBS = transmembrane metal binding site, C-domain = C terminal metal binding site, A-domain = actuator phosphatase P-domain = phosphorylation, N-domain = nucleotide binding. Arrow indicates path of Cu delivery by CopZ to TM-MBS.

A.



H hypothetical protein E exported protein of unknown function M membrane protein of unknown function U protein of unknown function

B.

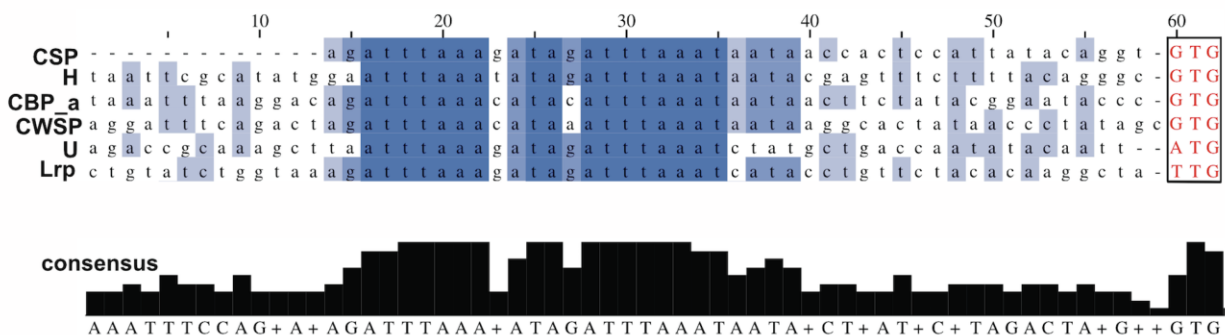


Figure S8. (A) Alignment of *N. viennensis* up-regulated genes suspected to be involved in Cu metabolism with a motif consensus sequence of 5'-TTCSGSMTTTGWACYANT-3'. Curved line indicates a break in the alignment. * indicates a reverse complement sequence and red letters indicates the motif repeats itself twice upstream or downstream of the motif. Predicted translational start codons are boxed and text is in red. Boxed green nucleotides indicate the location of a second motif and are included with the other motif regions at the bottom of the alignment. MCO, multicopper oxidase; BCP, protein of unknown function with C-terminal blue type (1) copper domain. (B) Alignment of top down-regulated genes from *N. viennensis* (NVIE) with a consensus motif sequence (underlined in gray). Predicted translational start codons are boxed and text is in red. CSP, cell surface associated Ca^{2+} binding domain protein, aemolysin type; CBP_a, putative surface associated Ca^{2+} -binding protein; CWSP, putative cell wall surface anchor family protein; Lrp, transcriptional regulator Lrp family. Locus tag information for all genes can be found in Dataset S1.

Table S1. RNA-Seq results summary of samples that were sequenced.

Sample name	•L1Day7	•L2Day7	•L3Day7	•L5Day7	•L6Day7	+R1Day5	+R2Day5	+R3Day5	+R4Day5	+R5Day5
••Trimmed	36,290,278	47,566,588	38,440,347	62,002,562	58,389,026	37,012,985	34,328,560	42,107,829	36,910,739	36,423,799
‡ rRNA sequences	34,671,567	45,422,238	36,818,224	58,629,182	55,664,484	37,753,008	32,419,120	39,749,355	34,829,426	36,423,799
#Non-rRNA sequences	1,618,711	2,144,350	1,622,123	3,373,380	2,724,542	2,259,977	1,909,440	2,358,474	2,081,313	1,672,974
##Sequences assigned to a genomic feature	1,515,278	1,988,513	1,497,003	3,189,422	2,523,110	2,138,075	1,792,743	2,210,469	1,949,417	1,568,005
+Unassigned Sequences	103,433	155,837	125,120	183,958	201,432	121,902	116,697	148,005	131,896	104,969

•Cu-limited cultures

† Cu-replete cultures

•• remaining sequences after adaptor and low quality sequences were trimmed away from raw sequences

‡ number of sequences matching rRNA genes

number of sequences not matching rRNA reads used for genomics feature assignment

number of sequences that could be assigned to genes in the reference genome

+ number of sequences that could not be assigned to genes in the reference genome

Table S2. Primers used in qPCR.

Gene	Primer name	Forward primer 5' to 3'	Reverse primer 5' to 3'	Source or reference
Zn-Fe permease PERM_b NVIE_000590	PERM1	GGCAGCAGGACTACGACGACA	GCCCTCGGTGGTGTGTGTATC	This study
Multicopper oxidase MCO1 NVIE_000600	MCO1	CGCTGAAAAACGCCGTGGTG	CACGAGGTCGCCTTCCGTTG	This study
Putative Ca ²⁺ binding protein	qNV001000	TGTTCCCCACTCGTCTACCACC	ATTGGATCCGACGGCAACAC	This study

CBP_a NVIE_001000				
Multicopper oxidase MCO4_b NVIE_019250	qNV-1019250	GCGCGATATTTGAGGCCAGA	TTGTCGCTCCGGAAGGATCT	This study
16S rRNA	qNV16s	GAGTGGGCACTGAGACAAGGG	CGGCTACGGATGCTTTAGGC	This study
<i>rpoB</i>	qrpoB	CATAATCCCGTACCCCGAGC	CGCCTTCGTGGAAACAACAG	This study

Chapter IV

Carbon Limitation of the Ammonia Oxidizing Archaeon *Nitrososphaera viennensis* Indicates Routes of Cellular Carbon Flow and Cellular Stress Responses

Logan H. Hodgskiss¹, Melina Kerou¹, Barbara Bayer^{2*}, Andreas Maier³, Wolfram
Weckwerth¹, Thomas Naegele⁴, Christa Schleper¹

Author affiliations:

¹Department of Functional and Evolutionary Ecology, University of Vienna, Djerassiplatz 1,
1030 Vienna, Austria

² Division of Bio-Oceanography, Department of Limnology and Bio-Oceanography, Centre
of Functional Ecology, University of Vienna, 1090, Austria

³Department of Geography and Regional Research, Althanstrasse 14 (UZA II), 1090 Vienna,
Austria

⁴Plant Evolutionary Cell Biology, Faculty of Biology, Ludwig-Maximilians-Universität
Munich, Großhaderner Str. 2-4, 82152 Planegg-Martinsried, Germany

*Present address: Department of Microbiology and Ecosystem Science, Centre for
Microbiology and Environmental Systems Science, University of Vienna, Austria

Draft manuscript.

Abstract

In the past decade, ammonia oxidizing archaea (AOA) have proven to be ubiquitous in soil environments. They rely on ammonia oxidation to generate energy and build cell mass from inorganic carbon sources, therefore contributing to both the global nitrogen and carbon cycles. However, little is known about the regulation and interplay of their primary metabolic processes. To this end, a thermodynamic model was developed for *Nitrososphaera viennensis* to estimate the consumption of inorganic carbon in relation to ammonia consumed for energy. Based on the results of this model, cells were grown in carbon-limited and excess conditions. Growth impairment at 0.1 mM supplied bicarbonate supported the calculated threshold of 0.128 mM when grown with 2 mM of ammonia. A combined proteomic and metabolomic approach revealed distinct metabolic responses under different levels of carbon limitation. As expected, flow of carbon through the central metabolisms was clearly regulated, with an induced carbon fixation pathway and parts of the TCA cycle, allowing entrance into gluconeogenesis and amino acid synthesis via malic enzyme and phosphoenolpyruvate carboxykinase. The most limited cultures that displayed strongest growth impairment exhibited a significant accumulation of all amino acids and an induction of proteins involved in nucleotide and amino acid metabolism, while proteins involved in translation were generally not enriched. However, proteins that clustered with the accumulation of amino acids in a combined statistical study, were involved in specific functions of information processing like initiation factor eIF-2a, and one of two methionine sulfoxide reductases, suggesting that accumulated amino acids along with critical enzymes were maintained in carbon-limited cells, while translation was generally slowed down. This study gains a deeper understanding of the metabolism and regulations of autotrophic growth in an ammonia oxidizing archaeon through a combination of modelling and functional omics.

Introduction

Ammonia oxidizing archaea (AOA) have been found in a wide variety of environments often outnumbering their bacterial counterparts, ammonia oxidizing bacteria (AOB)¹⁻⁷. Based on their abundance, their contributions to the global nitrogen cycle could be substantial. Besides interests to unravel the archaeal ammonia oxidation pathway⁸⁻¹³, their carbon

metabolism also draws attention as AOA are autotrophic and therefore represent a direct biological link between the global nitrogen and carbon cycles.

AOA participate in fixation of inorganic carbon by a unique version of the 3-hydroxypropionate/4-hydroxybutyrate cycle. While it appears similar to the 3-hydroxypropionate/4-hydroxybutyrate cycle in hyperthermophilic archaea ¹⁴, it is more efficient due to the use of only one ATP equivalent (rather than two) in selected enzymatic reactions and the use of promiscuous enzymes that can catalyze multiple steps in the pathway ^{14,15}. These benefits, along with an oxygen insensitive 4-hydroxybutyryl dehydratase, make it ideal for efficiently fixing carbon under low ammonia conditions ¹⁵. Most of the biochemical work on the AOA carbon fixation cycle has been carried out with the first isolated AOA, *Nitrosopumilus maritimus* ^{15–17}, which was isolated from a marine aquarium and represents the family *Nitrosopumilaceae* ¹⁸ (GTDB taxonomy ¹⁹, used throughout). Since then, genome comparisons have found this cycle to be conserved in all AOA ^{9,20} including the first AOA isolated from soil, *Nitrososphaera viennensis* ²¹.

N. viennensis in pure culture represents an ideal organism to study the metabolism of AOA in terrestrial environments. Physiological studies have shown that it has a high affinity for ammonia ^{22,23} and can tolerate relatively high levels of the end product of its metabolism, nitrite ^{21,24}. Accordingly, other AOA have been shown to grow on low levels of ²⁵, or even produce ²⁶, the other primary substrate, oxygen. These observations, along with the biochemical efficiency of their carbon fixation pathway, support the notion that AOA can efficiently fix bicarbonate as the sole carbon source. Initial experiments with *N. viennensis*, showed a marked increase in growth and activity with the addition of some organic acids (i.e. pyruvate) to growing cultures ^{21,24}. Although mixotrophy was suggested, this was dismissed with the discovery that pyruvate was acting as a reactive oxygen species (ROS) scavenger ²⁷. However, this abiotic reaction would create carbon dioxide that can be incorporated by *N. viennensis* as supported by previous Nano-SIMS experiments ²¹. In the absence of pyruvate, another ROS scavenger, such as catalase, can be added to the culture. Its dependence on pyruvate or other small organic acids supports the idea that *N. viennensis* is highly sensitive to ROS stress, which is partially attributed to its lack of a catalase ^{9,27}.

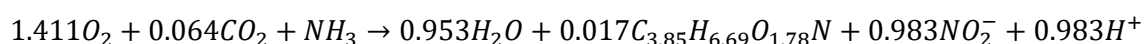
Knowing the primary energetic reactions that drive AOA metabolism (ammonia oxidation and oxygen reduction), an energetic model was calculated based on the Thermodynamic Electron Equivalents Model 2 (TEEM2) ²⁸ method to estimate the amount of

inorganic carbon the organism would need to effectively grow. Once calculated, *N. viennensis* was grown at different carbon concentrations with the addition of either catalase or pyruvate to test the resulting model. Simultaneously, proteins and metabolites were evaluated via a combined extraction protocol and analyzed in a systems biology approach to better understand how *N. viennensis* metabolically behaves under limited carbon supply. The results show a strong induction of proteins involved in carbon fixation, shifts in the metabolome, and highlight unique and distinct roles of methionine sulfoxide reductase (Msr) proteins within the cell.

Results

Growth curves of N. viennensis agree with thermodynamic predictions

A thermodynamic model of the ammonia oxidizing archaeon *Nitrososphaera viennensis* predicted a carbon consumption of 0.064 mmol of inorganic carbon per 1 mmol of ammonia (0.128 mmol carbon/2 mmol ammonia) consumed with an overall growth equation of :



The model takes into account physiological traits of ammonia oxidizing archaea including a monooxygenase reaction, estimated biomass electron equivalents, activation of carbon to acetyl-CoA, and an assumed energy transfer efficiency (see Materials and Methods) ^{28,29}. *N. viennensis* was grown at varying carbon concentrations in a closed system using either catalase or pyruvate to diminish the effect of reactive oxygen species (ROS) while using nitrite production as a proxy for growth.

All cultures incubated with inorganic carbon above the theoretical threshold of 0.128 mM grew without obvious limitations (Figure 1, A,C,D,G). Cultures grown below this threshold were either slightly limited (0.1 mM carbon, Fig. 1 B,F) or extremely limited (0 mM carbon, E). Cultures were harvested in late exponential phase with the exception of condition B (0.1 mM carbon with catalase) and condition E (0 mM with pyruvate), which were stopped earlier before they entered stationary phase due to a lack of carbon. The most limited condition with 0 mM of supplied inorganic carbon presented cultures with an extended lag phase and very slow growth as measured by nitrite production. In these cultures, carbon dioxide was apparently exclusively and slowly supplied from the added pyruvate that was decarboxylated in the presence of hydrogen peroxide. The use of pyruvate or catalase as a ROS scavenger also

produced a notable difference in the other limited cultures. Cultures of condition F with 0.1 mM of carbon that were also treated with pyruvate -albeit slowed in growth- were able to oxidize similar levels of ammonia as the non-limited cultures (i.e. reached similar maximal levels of nitrite, Fig. 1), likely because of the additionally available CO₂ produced from the abiotic reaction of pyruvate with hydrogen peroxide. Contrastingly, cultures with 0.1 mM carbon that were treated with catalase (condition B) were observed to enter stationary phase at an earlier nitrite concentration although there was still ammonia left in the medium (Supplementary Figure S1). This is presumably due to a depletion of carbon.

All cultures were inoculated with comparable amounts of cells but started with very different carbon concentrations. Carbon balances were calculated from measured inorganic carbon in the aqueous phase (beginning and end) and the gas phase (end) of each culture while taking into account potential inorganic carbon released from pyruvate (Materials and Methods). Differences in carbon consumption were observed even among cultures that were not limited but had differing carbon concentrations (Figure 2). Interestingly, based on carbon balances, conditions with an excessive amount of carbon (conditions A and D; 2 mM and 0.75 mM, respectively) consumed a higher amount of carbon/ammonia consumed than cultures that were closer to the theoretical carbon threshold (conditions C and G) or under the threshold (conditions B, F, and E) (Figure 2). Protein and DNA levels normalized to ammonia consumption in cultures of the extremely carbon-limited condition E indicate a lower amount of cells and a decoupling of ammonia oxidation and cell growth in extremely limited cultures (Figure S2).

Proteome analyses reveal essential enzymes in central pathways

Proteomes of the seven different conditions subjected to a principal component analysis (PCA) reflected similar groupings as those observed in the growth curves (Figure 3). 1264 proteins (40.47% of protein coding genes) were identified from the combined data. All conditions above the carbon threshold grouped together on the left side of principle component one (PC1) while the most limited condition (E) was found on the right side of PC1. Conditions B and F, representing limited cultures, were found between these two groupings.

Although the PCA showed overall differences among the proteomes of the conditions, the proteins representing the highest label-free quantification (LFQ) values for each condition were largely the same. Of the top 126 proteins (10%) from each condition, 100 were shared,

likely indicating important functional roles in the cell regardless of carbon limitation. Enrichment analysis according to archaeal clusters of orthologous groups (arCOGS) ³⁰ of these shared proteins (Figure 4) showed a high number of proteins involved in: amino acid transport and metabolism; energy production and conversion; posttranslational modification, protein turnover, and catabolism; and translation, ribosomal structure and biogenesis. Proteins implicated in energy production included AmoB (ammonia monooxygenase enzyme subunit B), NirK (suspected to be involved in ammonia oxidation metabolism ¹⁰), and subunits of ATP synthase (AtpA, AtpC, and AtpE). Numerous proteins involved in carbon metabolism were detected including aspartate semialdehyde dehydrogenase (Asd; carbon fixation cycle), glutamate dehydrogenase (GdhA; ammonium assimilation), malate dehydrogenase (Mdh; TCA cycle) and bifunctional fructose-1,6-bisphosphate/aldolase (Fbp; gluconeogenesis). Although these proteins represent the most highly abundant proteins in all conditions, the majority of them (89%) were also seen to change in relative abundance across conditions (see below) suggesting that they are still under regulatory control with respect to carbon concentration.

Statistical tests (either ANOVA or Kruskal-Wallis) were used to determine which proteins showed a difference among the different conditions. Approximately 80.7% of the total detected proteins showed a change among conditions with 20.3% remaining constant. Notably, among the shared top 100 proteins, 89 showed a statistical difference among conditions; although they were consistently abundant, their relative abundance changed in response to differing carbon concentrations.

Proteins that did not change in response to carbon concentrations were abundant in ribosomal proteins. AmoB, AmoY (a newly identified subunit of the archaeal ammonia monooxygenase (Hodgskiss et al, submitted), and NirK were also included in this category underscoring their importance to cellular function regardless of carbon limitation. Other proteins involved in purine metabolism (PurU, PurA, AdkA, PurD, GuaA), pyrimidine metabolism (ThyX, PyrE, PyrF), the non-oxidative pentose phosphate pathway (RpiA, Tal), and the conserved connection point between C-3 and C-4 metabolism ⁹ in AOA (PckA, ATP-dependent phosphoenolpyruvate carboxykinase) were also found to remain constant. Many of these are involved in carbon metabolism and the production of vital metabolites (i.e. nucleotides). The lack of a response to carbon limitation points towards these proteins serving crucial functions in the metabolism.

Differential protein expression in carbon-limited conditions

A clustering analysis was performed on the proteomic data to identify specific proteins associated with the various conditions. Proteins were divided into 15 clusters after hierarchical clustering (Figure 5). Two clusters with the highest relative abundance for condition E, the most limited condition, were identified (Figure 5). A majority of proteins making up the 3-hydroxypropionate/4-hydroxybutyrate carbon fixation cycle were found in these two clusters. Proteins for nucleotide transport and metabolism were also enriched (Figure 6A). Protein clusters representing high relative abundance in cultures that were not limited (conditions A, C, D, and G) showed enrichment for cell cycle control, transcription, and translation (Figure 6B).

Within proteins of the 3-hydroxypropionate/4-hydroxybutyrate carbon fixation cycle, ANOVA or Kruskal-Wallis tests followed by post-hoc Tukey or Dunn's tests, respectively, showed that the average values for these proteins were significantly different when compared to the standard carbon concentration (2mM; condition A). This is particularly true for the carboxylating half of the cycle. Overall, more proteins appeared to have a stronger response in the carboxylating half of the cycle when compared to the replenishing half of the cycle. A closer look at the primary carboxylating protein, acetyl-CoA carboxylase, revealed a strong reaction in the relative abundance of this protein across all conditions, even if an effect was not observed in the growth curves (Supplementary Figure S3). In the case of the so far unknown proteins for the reduction of acryloyl-CoA and succinic semialdehyde, 10 candidate proteins were proposed based off a genomic analysis⁹. Clustering analysis of the proteomes helped identify potential candidates. It would suggest that Adh4 (NVIE_024420) is the protein fulfilling one or both of these roles in the cycle. Although not found in the clustering analysis, Adh2 (NVIE_016260) is also a possible candidate due to its consistent presence in all conditions. As expected, proteins (PhaE, PhaC) participating in the production of the potential storage compound polyhydroxybutyrate (PHB) were shown to markedly decrease in the proteome under carbon limiting conditions while proteins that would funnel carbon away from PHB and towards the replenishment of acetyl-CoA (PhaA, PhaB) were increased under carbon limitation (not shown).

Metabolomic analysis reveals an accumulation of amino acids under carbon limitation

Selected metabolites were absolutely quantified based on standard curves produced

from known stock concentrations. Unfortunately, several of the organic acids in samples could not be accurately determined due to contamination from the highly purified water system as determined by comparison to background tests. Therefore, these metabolites (citric acid, lactic acid, and oxaloacetate) had to be excluded from the analysis. Pyruvate was also excluded due to its artificial addition to conditions E, F, and G as a ROS scavenger. Metabolites that could be accurately quantified were normalized to the total amount of carbon consumed for each culture and clustered to identify metabolites associated with specific conditions. This normalization would therefore represent metabolites in terms of carbon-moles of metabolite produced per mole of carbon consumed (see Materials and Methods). Surprisingly, almost all amino acids were clustered with the most limiting carbon condition, condition E (Figure 7). Identifiable sugars (glucose, maltose, and trehalose), were strongly associated with conditions that used catalase as a ROS scavenger (A,B,C,D). One sugar, melibiose, a disaccharide, was an exception to this trend and was found predominantly in conditions that were carbon limited. Malic acid, galactose, and fumaric acid had a higher relative abundance within cultures that were not carbon limited.

Systems biology approach with proteomics and metabolomics reveals methionine sulfoxide reductase A as a key protein

Proteomic and metabolomic data were analyzed together in cluster analyses to further investigate the metabolism of *N. viennensis*. To reduce the number of statistical tests, proteins and metabolites identified to be relatively abundant in condition E were analyzed across all conditions to identify significant correlations. This revealed several proteins that consistently correlated with multiple metabolites: 23 correlating with 5 or more. The top three proteins most often correlated with metabolites were methionine sulfoxide reductase A (MsrA), a subunit of succinate dehydrogenase (Complex II, SdhA), and a translation initiation factor (eIf-2a), with 16, 11, and 10 correlations respectively (Table 1). Four of the 23 proteins with predicted functions were involved in translation.

The relationship between MsrA and metabolites was further explored by correlating MsrA with all other proteins to find similar behavioral patterns. MsrA had positive correlations with 59 proteins including other proteins with multiple metabolite correlations. These included eIf-2a, inosine-5' monophosphate dehydrogenase (GuaB), methionine aminopeptidase (Map),

and phosphoribosylformylglycinamidine cyclo-ligase (PurM), all of which also appeared in the combined analysis above (Table 1).

Interestingly MsrB2, another methionine sulfoxide reductase found in the proteome exhibited a very different pattern compared to MsrA (Figure 8). It was not increased under carbon limitation, but was expressed significantly higher in all conditions that were supplemented with pyruvate as a ROS scavenger rather than catalase. Potentially, this could indicate that oxygen stress was higher in those cultures as compared to the catalase containing conditions.

Discussion

Decoupling of ammonia oxidation and carbon fixation in limiting conditions

Growth curves strongly supported the approximate theoretical carbon threshold of 0.128 mM when grown with 2 mM of ammonia as all conditions above this threshold behaved similarly with cultures exhibiting normal growth. As soon as the carbon concentration was lowered to 0.1 mM (conditions F and B), a growth impairment was observed for *N. viennensis*. These tests nicely demonstrate that the addition of pyruvate to the medium alleviates oxidative stress by being decarboxylated in the presence of hydrogen peroxide which leads to an increase in carbon dioxide ²⁷. This produced carbon dioxide is then available as a substrate for autotrophic growth. The benefit of this added carbon dioxide can be seen in the growth of condition F (pyruvate added as ROS scavenger) that is able to oxidize more ammonia than its counterpart, condition B (catalase added) (Supplementary Figure S1). This is also supported by the slow growth of condition E. While no inorganic carbon was initially supplied to these cultures, low amounts became available over time due to the abiotic reaction of hydrogen peroxide and pyruvate. There has been no solid evidence of mixotrophy thus far in AOA. However, previous nano-SIMS results with labeled pyruvate support the conclusion that carbon dioxide from the abiotic reaction with hydrogen peroxide is incorporated into AOA biomass ²¹. The slow growth of condition E also supports a lack of mixotrophy as much faster growth would be expected if the organism was capable of growth on pyruvate directly.

In condition E, there is not enough carbon to produce biomass equivalent with the amount of ammonia that has been oxidized, leading to decoupling of energy generation and biomass production. While the organism is still oxidizing ammonia, there is not enough carbon

available for cell division and growth. This is also reflected in the amount of protein per mmol of ammonia consumed (representing all ammonium/ammonia whether assimilated or oxidized) across cultures, with the most limited (condition E) producing less (Supplementary Figure S2). The fate of the electrons from this ammonia oxidation has yet to be determined, but they could be lost to more ROS or reactive nitrogen species if a biomass sink is not available.

Although growth curves show an expected response to the amount of carbon added, comparison of the theoretical carbon consumption with observed values indicates that the ratio of 0.128 mol inorganic carbon per 2 mol ammonia (0.064 mol inorganic carbon per 1 mol ammonia; also represent as 0.064 mmol C/mmol N, where N represents consumed ammonia/ammonium) is underestimating the amount of carbon that is actually consumed in the conditions that are not under extreme carbon limitation (Figure 2, A,B,C,D,F,G). Observed values range from approximately 0.23 mmol C/mmol N in conditions with excessive amounts of supplied inorganic carbon to values of approximately 0.1 mmol C/mmol N in more balanced conditions (i.e. conditions G and C with 0.25 mM of added carbon). The ratios of 0.1 mmol C/mmol N closely resembles the consumption of dissolved inorganic carbon observed in marine AOA based on isotope studies ^{31,32}. While cells in conditions A and D are fixing twice as much carbon as conditions G and C, protein and DNA content remain relatively similar. This would mean the extra carbon consumed is being directed somewhere else. One possibility is that excess carbon being fixed in conditions A and D is exported from the cell as metabolites. The excretion of metabolites from AOA has been documented in three marine strains ³³. However, DIC calculations accounting for this still revealed a 0.1 mmol C/mmol N ³¹. Alternatively, excess carbon could be incorporated as extracellular polysaccharide (EPS) or PHB. *N. viennensis* has the genomic capability to produce EPS and biofilm ⁹ and has been observed to create biofilms on glass slides ⁹ and in bioreactors when grown as continuous cultures (Melcher et al. unpublished data). The production of polyhydroxyalkanoates (which includes PHB) has also previously been observed in AOA via Raman micro spectroscopy ¹². A high amount of available inorganic carbon in a closed system might have triggered the production of EPS or PHB leading to an increased consumption of carbon.

Comparison of carbon consumption in AOA and AOB

Some variations from the theoretical model are not overly surprising and likely reflect aspects of the metabolism of *N. viennensis* that cannot be accurately accounted for with this

method. The TEEM2 method is based primarily on the inputs and outputs of metabolism^{28,29}. Therefore, this same model would also approximate the carbon consumption within ammonia oxidizing bacteria. As mentioned above, Bayer et al. (2022) used DIC isotope studies to show that marine AOA consume approximately 0.1 mmol C/mmol N while marine AOB consumed approximately 0.04 mmol C/mmol N³¹. The model calculation here falls between these values at 0.064 mmol C/mmol N, in general supporting its validity. Unique metabolic differences between AOA and AOB would likely explain the variation from the prediction. A likely explanation is the differences in efficiency between the carbon fixation cycles of AOA and AOB. With the more efficient thaumarchaeal 3-hydroxypropionate/4-hydroxybutyrate cycle¹⁵, it is plausible that AOA fix more than the predicted amount of carbon while AOB would fix less carbon as they use the less efficient Calvin-Benson-Bassham (CBB) cycle³⁴. Other differences could also stem from the efficiency of the ammonia oxidation pathway, which has not yet been fully resolved in either AOA⁸⁻¹⁰ or AOB^{11,35}.

Proteomic results support succinyl-CoA as connection point to TCA cycle

PCA analyses of the proteome datasets from the different conditions are analogous to observations from the growth curves: conditions above the carbon threshold are overlapping and largely similar, while limited conditions are farther apart indicating a reorganization of the proteins making up the proteome. Cluster analysis identified proteins that were most highly abundant in the most limited condition E (E cluster; 308 proteins) and in conditions A, C, and D (ACD cluster, all of which shared the high relative abundance cluster; 270 proteins). Many of these proteins were also highly abundant in condition G. However, the most highly abundant proteins in condition G were represented by a separate cluster. This separation can likely be attributed to the use of pyruvate instead of catalase as a ROS scavenger.

Comparison of the enrichment analysis between the E clusters and ACD cluster revealed a high enrichment for nucleotide metabolism and lipid metabolism in the E clusters compared to translational, post-translational, transcriptional, and cell cycle enrichment in the ACD cluster. Both groups shared an enrichment for amino acid and coenzyme metabolism. The differences between the clusters point towards different metabolic strategies under carbon replete vs. limited conditions. An abundance of carbon leads to a metabolic strategy of actively producing more proteins, transcripts, and DNA in order for the cells to grow. Under carbon limitation, the proteome appears to shift towards creating the necessary building blocks (i.e.

amino acids and nucleotides) for these processes rather than the processes themselves (translation, replication/transcription). This follows with the hypothesis that less protein and DNA are produced per mmol of ammonia oxidized in carbon limited cultures.

Proteomics reveal regulation patterns in the carbon fixation cycle

Previous work on the 3-hydroxypropionate/4-hydroxybutyrate cycle in *Metallosphaera sedula* showed that the primary connection point of this carbon fixation cycle with the central carbon metabolism occurs at succinyl-CoA rather than acetyl-CoA ³⁶. This has been assumed to be true for AOA as well and is now supported by the proteomic results. Most of the proteins of the proposed 3-hydroxypropionate/4-hydroxybutyrate pathway are found in high relative abundance even for the most limited carbon condition. This would imply that these proteins play a critical role in the metabolism of the organism and are maintained even under extreme carbon limitation. In Figure 9, proteins of this pathway that show a high relative abundance and a statistical difference between the most limited and least limited condition are indicated. Based on these parameters, the carboxylating half of the carbon fixation cycle (i.e. from acetyl-CoA to succinyl-CoA) is more responsive to carbon limitation than the half responsible for replenishment of acetyl-CoA. Redundancy of protein function has also been suggested for this pathway ¹⁵. Two proteins that are likely playing multiple roles in the pathway are found on both the carboxylating and the replenishing side of the cycle (CCH and Adh). With this in mind, it could be postulated that the most responsive part of the carbon fixation pathway is that leading from acetyl-CoA to succinyl-CoA. This is additionally supported by the presence of PhaAB in the selection criteria, which could be directing carbon from stored polyhydroxybutyrate (PHB) that can accumulate within AOA ¹². The flow of carbon from PHB can be assumed based on the drop in relative intensity of the proteins PhaEC that are engaged in the formation of PHB. The increase in PhaAB to replenish the cycle and a high response in proteins from acetyl-CoA to succinyl-CoA indicate a metabolic strategy to input as much carbon as possible to the primary carbon metabolism of the cell.

While some proteins of the replenishing side of the carbon fixation cycle did not fall into the above criteria, it should be noted that these proteins were either maintained at the similar levels across almost all conditions (HBCD and HCD) or responded to carbon limitation at a less drastic limitation (HBCS) which precluded them from the above criteria.

A closer look at the shifts of individual proteins within the cycle also reveals potential regulation points within the metabolism of *N. viennensis* with respect to carbon. The protein responsible for both carboxylation reactions, acetyl-CoA carboxylase, is observed to have strong and significant changes in the relative intensity of its subunits with changes in carbon concentration even before a growth effect is observed (Supplementary Figure S3). This could mean that the expression of acetyl-CoA carboxylase within the cell is highly sensitive to the external carbon concentrations. Acetyl-CoA carboxylase activity is allosterically regulated by coenzyme A³⁷ and has been shown to be inhibited by the PII regulatory protein GlnB but not by the related protein GlnK³⁸. In agreement with this, proteins for the production of vitamin B5 (PanB and NVIE_011900) needed for coenzyme A synthesis are found in cluster E, and GlnB relative abundance sharply drops under carbon limitation while other regulatory PII proteins increase (NVIE_003920, NVIE_014570). *N. viennensis* contains 6 nitrogen regulatory PII proteins⁹, an unusually high number. The high number and implied interaction here is indicative of these proteins coordinating the nitrogen and carbon flux of AOA.

Many of the other proteins either have a more gradual response or do not respond strongly to carbon until after the theoretical threshold is passed, which matches more closely with what is seen in the growth curves (HPSC, CCH, MCM, Mut). Based on these trends the alcohol dehydrogenase that is most likely contributing to the carbon fixation cycle is Adh4 as it is found in the most limited cluster with the other carbon fixation proteins and has a similar pattern. Another possible candidate is Adh2 which is consistently present across all conditions.

Notable exceptions to these trends are succinate-semialdehyde dehydrogenase (GabD), which does not respond to carbon limitation at all until the extreme limitation, and methylmalonyl-CoA epimerase (MCE), which shows a decrease in relative abundance as carbon concentration decreases, an opposite trend compared to the rest of the cycle. In the case of GabD, this could represent a constitutively available protein for basic metabolism that is not strongly affected until major environmental changes persist. As for MCE, an opposite trend could be observed because the expression and/or translation of this protein is controlled based on the presence of substrates (which are likely lacking) rather than on a global response to stress from lack of carbon.

Various nucleotide and amino acid synthesis proteins are also found within the E clusters. Other proteins in the E cluster that have a difference in relative abundance among conditions include enolase (Eno; gluconeogenesis), malic enzyme (YtsJ; possible C-3 to C-4

connection in *Nitrososphaeraceae*), triphosphate isomerase (TpiA; gluconeogenesis), and 3-hydroxyl-3-methylglutaryl-coenzyme A reductase (MvaA; isoprenoid lipid biosynthesis). These could represent points of translational regulation within the cell under carbon limitation.

Metabolomic Response to Carbon Limitation and Reactive Oxygen Species

Amino acids were strongly associated with the most limited carbon condition. This list includes the three aromatic amino acids, tryptophan, phenylalanine, and tyrosine, some of the most expensive amino acids to produce in terms of both carbon and energy. There are two possible explanations for this observation: (1) proteins are broken down leaving pools of amino acids available and (2) amino acids are being synthesized from the little carbon that is available.

It is difficult to distinguish between the two options as proteins for amino acid synthesis and a protein for degradation (proteasome, PsmA1 and PsmB2) are found within the E cluster. Regardless, the higher relative abundance of amino acids is unique to the most limited condition. *N. viennensis* is not known to be capable of degrading amino acids and has no indications of secondary metabolism that would utilize specific amino acids⁹. This data was therefore integrated with the available proteomic data in a systems biology approach to elucidate the behavior of the cells.

Systems Biology Reveals a Unique Function for MsrA and a Cellular Response to Growth Under Carbon Limitation

The protein that strongly correlated with the most metabolites was methionine sulfoxide reductase A (MsrA). As supported by the correlation analysis, the relative abundance profile of MsrA in *N. viennensis* shows an increase in the protein when carbon limiting conditions are encountered (Figure 8A).

MsrA belongs to a group of sulfoxide reductases that include both MsrA and MsrB that are implicated in the reduction of free and protein bound methionine and have been shown to play a crucial role in response to ROS^{39,40}. MsrA and MsrB are specific to S- and R-stereoisomers, respectively, although MsrA has been shown to act more readily on free methionines⁴¹. Methionine represents an important metabolite within the cell as it is often used in the initiation of translation, is a precursor for S-adenosyl-methionine used in methyl transfers

with cobalamin (vitamin B12), and partakes in the cellular use of sulfur^{42,43}. In the archaeon *Haloferax volcanii*, MsrA has also been shown to be involved in ubiquitin-like protein modifications⁴⁴. Ubiquitin-like modification systems have not been studied in ammonia oxidizing archaea although some homologs seem to be present⁴⁵. To explore other possibilities for the function of MsrA a correlation search was performed with other proteins in which elongation factor eIF-2a was identified. This is of particular importance as eIF-2a is responsible for the initiation of translation by providing the methionyl-tRNA to the ribosomal complex^{46,47}, raising an important connection between MsrA and eIF-2a via their interactions with methionine.

In addition to MsrA, *N. viennensis* also contains genes for two MsrB proteins (MsrB1 and MsrB2). MsrB1 is not readily found in the investigated proteomes while MsrB2 is present. The expression pattern of MsrA is distinct from that of MsrB2 in *N. viennensis*. While MsrA increases under carbon limiting conditions, relative abundances of MsrB2 appear to correspond to the choice of ROS scavenger with higher levels of MsrB2 being observed in all conditions with pyruvate regardless of the available levels of carbon (Figure 8). Therefore, while MsrA appears to play a role in cell growth and possibly the translation of proteins, MsrB2 is responding in a redox specific manner within the cell. This response would be counter to that seen in *Haloferax volcanii*, where MsrA was observed to play a greater role under oxidative stress⁴⁴. However, this role was linked to protein sumpylation which may not occur in ammonia oxidizing archaea. This could indicate a separate adaption for Msr proteins for different archaeal clades.

The role of MsrA could be to keep any available methionine in a reduced state and therefore accessible for the start of translation and for use in methyl transfers via vitamin B12. This is supported by the presence of other vitamin B12 synthesis proteins, and methyltransferases found within the E clusters. However, several factors point towards MsrA being specifically relevant for protein translation. These include the correlations of MsrA, eIF-2a, and amino acids as well as the presence of other translation initiation factors in the E clusters (eIF-2B, eIF-2G). Additionally, a putative tRNA methyltransferase complex GCD14 subunit was correlated with 7 metabolites and methionine amino peptidase (Map) was correlated with 5 metabolites. These are predicted to be a subunit of a complex essential in the maturation of tRNA including tRNA-met in *Saccharomyces cerevisiae*⁴⁸ and serve as a post-translational modification to remove the N-terminal methionine for proper protein function

respectively. All together this evidence points towards a coordinated effort of the carbon-starved cell to be primed for translation to occur upon the availability of carbon.

Conclusion

The results of protein abundance under carbon limitation in *N. viennensis* highlight not only unique roles of specific proteins but also possible enzymatic nodes that are important in the carbon metabolism. The results presented support acetyl-CoA carboxylase as a crucial regulatory point in the carbon fixation cycle and succinyl-CoA as the connection point between the TCA cycle and carbon fixation cycle. Carbon is then routed up through gluconeogenesis via both malic enzyme and PEP carboxykinase where carbon can be channeled towards nucleotide and amino acid synthesis. Combined data from proteomics and metabolomics also suggest a role of MsrA in keeping methionine reduced and available for use in translation of proteins in coordination with other translation factors. The analogous protein MsrB2 is shown to react in response to ROS scavenger selection rather than carbon concentration suggesting a more focused role in reduction due to ROS stress. While these routes of carbon flow and possible protein functions are emphasized under carbon limitation, these processes are hypothesized to play the same important roles under normal growing condition of *N. viennensis* but would not have been found without the combined information from both metabolites and proteins under carbon stress.

Data and Code Availability

All data and scripts will be made available upon publication.

Methods

Carbon and Nitrogen Content of Cell

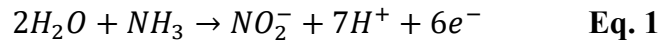
Cells were grown under standard conditions with pyruvate and harvested in late exponential growth phase by centrifugation (16,000 xg at 4°C for 1 h). Cell pellets were washed twice with FWM (no additions of bicarbonate or ammonium), dried for 1h (vacuum concentrator) and stored in a desiccator until further processing. Subsequently, cell pellets were

weighed and packed into tin capsules prior to being analyzed on a CHNS elemental analyzer (Vario MICRO cube, Elementar). The instrument was calibrated with acetanilide following manufacturer protocols.

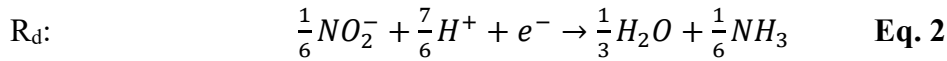
Thermodynamic Model

A thermodynamic model for the metabolism of *N. viennensis* was constructed based off of the Thermodynamic Electron Equivalents Model 2 (TEEM2) method²⁸.

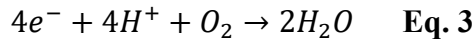
The electron donor reaction was calculated as the oxidation of ammonia to nitrite. Ammonia was chosen for the half reaction instead of ammonium as it is assumed that ammonia, and not ammonium, is the true substrate for the AMO complex⁴⁹ that begins the ammonia oxidation pathway:



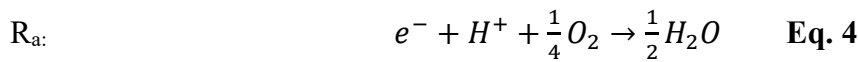
The reaction was then normalized to one electron and arranged in the reducing direction:



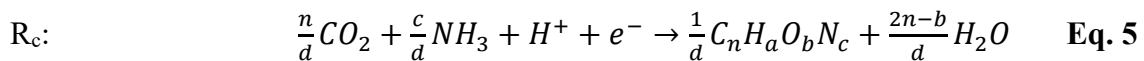
Oxygen reduction was used as the half reaction for the electron acceptor:



This equation was also normalized to one electron:



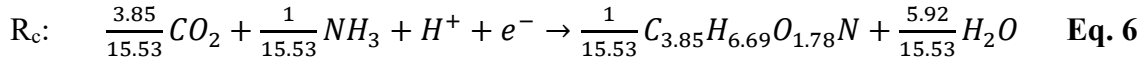
Carbon and nitrogen measurements of washed *N. viennensis* determined the C:N molar ratio to be 3.85:1. Although molar ratios of oxygen and hydrogen were not determined, a biomass composition of $C_{3.85}H_{6.69}O_{1.78}N$ was used. This was based off of values for *Escherichia coli* that had a similar C:N molar ratio of 3.85:1²⁹ (pg. 129). The biomass composition was used with the following formula to estimate a half reaction for biomass formation:



where $d = 4n + a - 2b - 3c$.

This equation is modified from²⁹ (pg. 137). Typically, a generic biomass concentration would include NH_4^+ rather than NH_3 to represent a nitrogen source. The equation was modified

to use NH_3 based on the physiological observation that the vast majority of nitrogen consumed by *N. viennensis* is used for nitrite production and would therefore be utilized in the form of ammonia. The original equation also include HCO_3^- to balance the positive charge of ammonium. Therefore HCO_3^- was also excluded from the equation. In this case all carbon is represent by CO_2 . Utilizing the estimated biomass composition and the modified equation gives the following:



Once the necessary reactions were obtained, $\Delta G'$ values were obtained. Gibbs free energy were calculated for a temperature of 25°C and pH of 7.0. These are not the exact temperature of the growth conditions for *N. viennensis* (42°C and pH ~ 7.2 respectively). However, standard conditions were used as these are the parameters used to calculate the Gibbs free energy for the reduction of oxygen.

$$\text{Oxygen reduction: } \Delta G_a^{0'} = -78.72 \frac{\text{kJ}}{\text{mol e}^-} \quad ^{29}$$

As this is an empirical value based on a hydrogen potential with platinum, it is not readily adjustable. For consistency, the same parameters were used for the calculation of Gibbs free energy for the half reaction of ammonia to nitrite. Gibbs free energy of formation were used for the participating chemical species:

$$\text{H}_2\text{O}_{(l)}: \Delta_f G_{\text{H}_2\text{O}} = -237.13 \text{ kJ/mol} \quad ^{50}$$

$$\text{NH}_3_{(\text{aq})}: \Delta_f G_{\text{NH}_3} = -26.50 \text{ kJ/mol} \quad ^{50}$$

$$\text{NO}_2^-_{(\text{aq})}: \Delta_f G_{\text{NO}_2^-} = -8.9 \frac{\text{kcal}}{\text{mol}} = -37.24 \text{ kJ/mol} \quad (\text{Using } 4.184 \text{ kJ/kcal.}) \quad ^{51}$$

$$\Delta_r G^0 = [\text{products}] - [\text{reactants}] \quad \text{Eq. 7}$$

Using Equation 2:

$$\Delta_r G^0 = \frac{1}{3} \left(-237.13 \frac{\text{kJ}}{\text{mol}} \right) + \frac{1}{6} \left(-26.50 \frac{\text{kJ}}{\text{mol}} \right) - \frac{1}{6} \left(-37.24 \frac{\text{kJ}}{\text{mol}} \right) = -77.25 \frac{\text{kJ}}{\text{mol}}$$

To adjust for pH of 7.0:

$$\Delta G^{0'} = \Delta G^0 + RT \ln \left(\frac{[\text{H}_2\text{O}]^{\frac{1}{3}} [\text{NH}_3]^{\frac{1}{6}}}{[\text{NO}_2^-]^{\frac{1}{6}} [\text{H}^+]^{\frac{1}{6}}} \right), [\text{H}^+] = 10^{-7} \quad \text{Eq. 8}$$

$$\Delta G^{0'} = -77.25 \frac{kJ}{mol} + \left(0.008314 \frac{kJ}{K \cdot mol} \right) (298 K) \ln \left(\frac{[1]^{\frac{1}{3}} [1]^{\frac{1}{6}}}{[1]^{\frac{1}{6}} [10^{-7}]^{\frac{7}{6}}} \right) = -30.66 \frac{kJ}{mol}$$

Ammonia oxidation in reducing direction: $\Delta G_d^{0'} = -30.66 \frac{kJ}{mol e^-}$

For calculation of model:

$$R = f_e^0 R_a + f_s^0 R_c - R_d \quad \text{Eq. 9}$$

$$f_s^0 + f_e^0 = 1 \quad \text{Eq. 10}$$

$$f_s^0 = \frac{1}{1+A} \quad \text{Eq. 11}$$

$$A = -\frac{\Delta G_s}{\varepsilon \Delta G_r} \quad \text{Eq. 13}$$

$$\Delta G_r = \Delta G_a - \Delta G_d - \frac{q}{p} \Delta G_{xy} \quad \text{Eq. 14}$$

$$\Delta G_s = \frac{(\Delta G_{fa} - \Delta G_d)}{\varepsilon^m} + \frac{(\Delta G_{in} - \Delta G_{fa})}{\varepsilon^n} + \frac{\Delta G_{pc}}{\varepsilon} \quad \text{Eq. 15}$$

where ΔG_a is the Gibbs free energy for the electron acceptor half reaction, ΔG_d is the Gibbs free energy for the electron donor half reaction, q represents the number of monooxygenase reactions per substrate, p is the number of electron equivalents per mole of substrate, ΔG_{xy} is the reduction potential for NADH oxidation (representative of the energy input for the monooxygenase reaction; -219.2 kJ/mol²⁸), ΔG_{fa} is the reduction potential for a formaldehyde half reaction, ΔG_{in} is the Gibbs free energy for the reduction potential of acetyl-CoA (representative of carbon activation for autotrophs; 30.9 kJ/mol²⁸) ΔG_{pc} is the Gibbs free energy for intermediate conversion to cells, ε is the energy transfer efficiency, m is 1 if ΔG_{fa} is greater than 0 and equal to n otherwise, and n is 1 if $(\Delta G_{in} - \Delta G_d) > 0$ and m=n²⁸.

In the case of *N. viennensis*, $\Delta G_{fa} = 0$ as no formaldehyde is involved and $(\Delta G_{in} - \Delta G_d) > 0$. This causes m=n=1. The Equation 13 for ΔG_s therefore simplifies to:

$$\Delta G_s = \frac{\Delta G_{in} - \Delta G_d + \Delta G_{pc}}{\varepsilon} \quad \text{Eq. 16}$$

The value of p is equal to 6 (6 electrons per ammonia, Eq. 1), q is 1 (one monooxygenase reaction per ammonia), and ΔG_{pc} is the energy to synthesize cells:

$$\Delta G_{pc} = \frac{\left(\text{molecular weight of biomass} \frac{g}{mol}\right) \left(\text{energy per gram of biomass} \frac{kJ}{g}\right)}{\left(\text{electron equivalents per biomass} \frac{eeq}{mol}\right)} \quad \text{Eq. 17}$$

Molecular weight of $C_{3.85}H_{6.69}O_{1.78}N = 3.85*(12)+6.69*(1)+1.78*(16)+1*(14)=95.37 \text{ g/mol}$

Energy per gram of biomass = 3.33 kJ/g ^{29,52}

Electron equivalents per biomass = 15.53 eeq/mol (Equation 6)

Input values for Equation 17:

$$\Delta G_{pc} = \frac{\left(95.37 \frac{g}{mol}\right) \left(3.33 \frac{kJ}{g}\right)}{\left(15.53 \frac{eeq}{mol}\right)} = 20.45 \frac{kJ}{eeq}$$

The energy transfer efficiency, ϵ , is estimated to be 0.57 as observed for other autotrophs ²⁸.

Input values for Equation 16:

$$\Delta G_s = \frac{30.9 \frac{kJ}{mol} - \left(-78.82 \frac{kJ}{mol}\right) + 20.4 \frac{kJ}{mol}}{0.57} = 228.37 \frac{kJ}{mol}$$

Input values for Equation 14:

$$\Delta G_r = -79.82 \frac{kJ}{mol} - \left(-30.66 \frac{kJ}{mol}\right) - \frac{1}{6} \left(-219.2 \frac{kJ}{mol}\right) = -11.63 \frac{kJ}{mol}$$

Input values for Equation 13:

$$A = -\frac{228.37 \frac{kJ}{mol}}{0.57 \left(-11.63 \frac{kJ}{mol}\right)} = 34.44 \frac{kJ}{mol}$$

Input values for Equation 11:

$$f_s^0 = \frac{1}{1 + 34.44} = 0.028$$

Input values for Equation 10:

$$f_e^0 = 1 - f_s^0 = 1 - 0.028 = 0.972$$

Input values and equations (R_a as Eq. 4; R_d as Eq. 2; R_c as Eq. 6) for overall reaction (Equation 9):

$$R = f_e^0 R_a + f_s^0 R_c - R_d$$

$$f_e^0 R_a: 0.972 \left(e^- + H^+ + \frac{1}{4} O_2 \rightarrow \frac{1}{2} H_2O \right); 0.972 e^- + 0.972 H^+ + 0.243 O_2 \rightarrow 0.486 H_2O$$

$$f_s^0 R_c: 0.028 \left(\frac{3.85}{15.53} CO_2 + \frac{1}{15.53} NH_3 + H^+ + e^- \rightarrow \frac{1}{15.53} C_{3.85} H_{6.69} O_{1.78} N + \frac{5.92}{15.53} H_2O \right);$$

$$0.007 CO_2 + 0.002 NH_3 + 0.028 H^+ + 0.028 e^- \rightarrow 0.002 C_{3.85} H_{6.69} O_{1.78} N + 0.011 H_2O$$

$$-R_d: - \left(\frac{1}{6} NO_2^- + \frac{7}{6} H^+ + e^- \rightarrow \frac{1}{3} H_2O + \frac{1}{6} NH_3 \right);$$

$$0.333 H_2O + 0.167 NH_3 \rightarrow 0.167 NO_2^- + 1.167 H^+ + e^-$$

Combine all equations and normalize to 1 mol of ammonia:

$$1.411 O_2 + 0.064 CO_2 + NH_3 \rightarrow 0.953 H_2O + 0.017 C_{3.85} H_{6.69} O_{1.78} N + 0.983 NO_2^- + 0.983 H^+$$

Eq. 18

Equation 18 normalized to 2 mol ammonia:

$$2.821 O_2 + 0.128 CO_2 + 2 NH_3 \rightarrow 1.906 H_2O + 0.033 C_{3.85} H_{6.69} O_{1.78} N + 1.967 NO_2^- + 1.967 H^+$$

Eq. 19

Experimental Setup for Carbon Conditions

With an ammonia concentration of 2 mM ammonia, a theoretical carbon threshold of 0.128 mM carbon was used for the experimental design of growth curves with varying levels of inorganic carbon: 2 mM carbon with catalase (A); 0.75 mM carbon with catalase (D) 0.25

mM carbon with pyruvate (G); 0.25 mM carbon with catalase (C); 0.1 mM carbon with pyruvate (F); 0.1 mM carbon with catalase (B); and 0 mM carbon with pyruvate (E). This spread of carbon concentrations produced a gradient from high to low amounts of available inorganic carbon (Figure 1A). The reaction of pyruvate with ROS would create additional available carbon for the cell. In the case of the 0 mM inorganic carbon concentration, this created a slow release of carbon substrate over time. The pyruvate conditions with 0.25 mM and 0.1 mM of added carbon would therefore have slightly more available carbon than their catalase counterparts. The combination of all conditions gives a gradient of carbon concentrations starting from 2 mM and dropping down to just above 0 mM.

Culture Conditions

Cultures of *N. viennensis* were grown in closed 1 liter pressure safe Schott bottles closed with a grey bromobutyl rubber stopper (GL45 Duran) with 580 mL of fresh water medium (1g/L NaCl, 0.4 g/L MgCl₂·6H₂O, 0.1 g/L CaCl₂·2H₂O, 0.2 g/L KH₂PO₄, 0.5 g/L KCl). The choice of rubber stopper is critical as *N. viennensis* will not grow when certain other rubber stoppers are used. The following was added to FWM: 600 µL trace element solution (100 mM HCl (~12.5M), 0.5 mM H₃BO₃, 0.5 mM MnCl₂·4H₂O, 0.8 mM CoCl₂·H₂O, 0.1 mM NiCl₂·6H₂O, 0.01 mM CuCl₂·H₂O, 0.5 mM ZnSO₄·7H₂O, 0.15 mM Na₂MoO₄·2H₂O; autoclaved, stored in dark), 600 µL vitamin solution (0.02 g/L biotin, 0.02 g/L folic acid, 0.1 g/L pyridoxine HCl, 0.05 g/L thiamine HCl, 0.05 g/L riboflavin, 0.05 g/L nicotinic acid, 0.05 g/L DL pantothenic acid, 0.05 g/L P aminobenzoic acid, 2 g/L choline chloride, 0.01 g/L vitamin B12; adjusted pH to 7 with KOH, filter sterilized, stored in dark), 600 µL 7.5 mM FeNa-EDTA (pH 7), and 6 mL HEPES buffer solution (24 g/L NaOH, 119.2 g/L HEPES; pH ~7.6, filter sterilized).

For all cultures, kanamycin was added to a final concentration of 100 mg/mL and NH₄Cl was added to a final concentration of 2 mM.

A catalase stock solution was made by dissolving catalase from bovine liver (Sigma Aldrich, powder, 2000-5000 units/mg protein) in a potassium phosphate buffer solution (0.6528 g/L KH₂PO₄, 7.83 g/L K₂HPO₄, pH adjusted to 7) at 1 mg/mL. Catalase was filter sterilized to remove chunks before use. This made a stock solution of catalase with 3500U/mL.

For cultures containing catalase, 8.58 mL of stock catalase solution was added to achieve a final concentration of 50U/mL for catalase²⁷. Cultures grown with catalase were observed to have some translucent flocs due to the addition of catalase. These flocs formed upon addition of catalase and were therefore not attributed to be a product of cell growth. Cultures using pyruvate as a ROS scavenger had pyruvate added to a final concentration of 0.3 mM.

Inorganic carbon was added from freshly made solutions of filter sterilized 0.5M sodium bicarbonate. Based on preliminary tests amounts of added bicarbonate are overestimated (i.e. a culture with assumed 2 mM of bicarbonate is actually 1.67 mM bicarbonate). This discrepancy is likely due to the exchange of carbon dioxide between the gas and aqueous phases. The actual concentration of a freshly prepared 0.5 M stock solution was determined to be 0.415 mM. This concentration was used to calculate the amount of stock solution to add to cultures. Therefore, for 600 mL cultures, the following amounts of freshly made 0.5M sodium bicarbonate solution was added: 2 mM carbon culture was supplied with 2.87 mL, a 0.75 mM culture was supplied with 1.08 mL, a 0.25 mM culture was supplied with 0.36 mL, and a 0.1 mM culture was supplied with 0.14 mL.

All 600 mL cultures were inoculated with 1.5 mL (0.25% inoculum) of an actively growing culture in late exponential phase.

After inoculation, the headspace of cultures was flushed with an artificial gas mixture of 80% nitrogen 20% oxygen to remove any residual carbon dioxide in the head space. Immediately after flushing the head space, a water sample was taken from the aqueous phase to determine the total starting carbon concentration (see below). Cultures were then grown in a 42°C incubator on a shaker set at 80 rpm.

Cultures were maintained by taking periodic nitrite measurements. Cultures were harvested when late exponential phase was reached but before entering stationary phase. At the end of a growth curve samples were taken for gas analysis and inorganic carbon analysis of liquid. One mL was collected and spun down for 1 hour at 16,000xg and 4°C. The supernatant was removed and the pellet was stored at -70°C until DNA could be extracted. The rest of the culture was spun down in 250-300 mL volumes using a large Sorvall centrifuge for 45 minutes at 16,100xg at 4°C. Supernatant was poured off and the pellet was resuspended in remaining media and transferred to a new 1.5 mL Eppendorf tube. Cells were then concentrated by centrifuging at 16,100xg for 30 minutes and 4°C. As much supernatant as possible was

removed without disturbing the cell pellet. Eppendorf tubes were weighed before and after cells were added to obtain an approximated wet biomass weight. Pellets were then frozen at -70°C until further metabolite and protein extraction.

Ammonium Measurements

Ammonium was measured colorimetrically. 200 µL of sample was mixed with 400 µL of fresh water medium (FWM) in an Eppendorf tube. 300 µL of color reagent (5.18 mM sodium salicylate, 2.15×10^{-5} M sodium nitroprusside, 0.1 M NaOH; freshly made) was added followed by 120 µL of oxidation solution (3.01×10^{-5} M dichloroisocyanuric acid; freshly prepared). Samples were shaken and incubated in the dark for 30 minutes. After 30 minutes, 200 µL of sample were pipetted into wells in a micro titer plate and absorbance was measured at 660 nm using a Tecan-Sunrise plate reader. Standard curves were made using a stock solution of 1 M NH_4Cl .

Nitrite Measurements

Nitrite was measured colorimetrically. 10 or 20 µL of sample was mixed with 790 or 780 µL fresh water medium (FWM) respectively and 200 µL of a sulfanilamide/NED reagent (150 mL ortho-phosphoric acid, 10 g sulfanilamide, 0.05 g α -naphthylethylenediamine dihydrochloride in 1 L water; stored at 4°C in the dark) in a 1 mL Eppendorf tube. Samples were shaken and stored in the dark for at least 10 minutes. After 10 minutes, 200 µL of each sample was pipetted into wells of a micro titer plate and absorbance was measured at 545 nm using a Tecan-Sunrise plate reader. Standard curves were made from a stock solution of 1 mM NaNO_2 (i.e. 1 mM NO_2^- : 20 µL stock + 780 µL FWM; 0.8 mM NO_2^- : 16 µL stock + 784 µL; 0.6 mM NO_2^- : 12 µL stock + 788 µL; etc.).

IC Analysis

10 mL of sample were taken at the beginning (sterile, with flame and needle) and end of each culture to determine starting and ending inorganic carbon (IC) concentration in the aqueous phase. Samples were filtered to remove cells and stored at -70°C until the inorganic carbon could be measured. IC data were measured using a Shimadzu TOC -LCPH analyzer

equipped with an IC reaction vessel containing a reaction solution of phosphoric acid of about 25% (weight%) and an ASI-L autosampler. The samples were injected into the reaction vessel where all inorganic carbon is converted to carbon dioxide which in the following is volatilized by the sparging process (synthetic air, carbon dioxide free gas) and detected by a NDIR detector. Each sample is measured three times with an injection volume of 100 micro-liter each. The final results are the corresponding mean values of the three injections per sample.

GC Analysis

At the time of harvest, head space pressure was measured and 40-50 mL of head space gas was removed using a 50 mL syringe and transferred to a glass 120 mL serum bottle pre-filled with CH₄ at 1 atm. Mixed gas samples were then analyzed using an Agilent Gas Chromatograph (Agilent 7890A GC, Agilent Technologies, Santa Clara, CA, USA) using a thermal conductivity detector. Gases were separated at 170°C using helium as the carrier gas. The reference flow setting was 10 mL/min. The makeup flow was set to 1 mL/min.

The following gases were used: H₂, CO₂, N₂, CH₄, and H₂N₂CO₂ (mix ratio 7:1:1). All gases were from Air Liquide GmbH, Schwechat, Austria. The standard test gas used for gas chromatography (GC) comprised the following composition: 0.01% volume CH₄ and 0.08% volume CO₂ in N₂ (Messer GmbH, Wien, Austria).

Carbon Balance

The total amount of carbon consumed by each culture can be approximated by the equation:

$$C_{con} = C_{avail} - C_{aq} - C_g \quad \text{Eq. 20}$$

where C_{con} is total carbon consumed, C_{avail} is amount of available inorganic carbon, C_{aq} is the aqueous inorganic carbon concentration at the end of the experiment, and C_g is the gaseous inorganic carbon concentration at the end of the experiment. C_{avail} can be calculated by:

$$C_{avail} = C_{aq,0} + C_{g,0} + C_{pyr} \quad \text{Eq. 21}$$

where $C_{aq,0}$ is the initial aqueous inorganic carbon concentration, $C_{g,0}$ is the initial gaseous inorganic carbon concentration, and C_{pyr} is the amount of inorganic carbon contributed from the interaction of pyruvate and ROS. In all cultures, the $C_{g,0}$ is 0 as the head space of the

cultures is flushed while $C_{aq,0}$ is represented by the initial IC measurement taken. In catalase cultures, C_{pyr} is 0, while in pyruvate cultures, C_{pyr} is estimated by:

$$C_{pyr} = [NO_2^-] * \frac{0.005 \text{ mM } H_2O_2}{0.5 \text{ mM } NO_2^-} * \frac{1 \text{ mM } CO_2}{1 \text{ mM } H_2O_2} \quad \text{Eq. 22}$$

where $[NO_2^-]$ represents the mM of nitrite produced by the culture. The ratio of 0.005 mM H_2O_2 / 0. mM NO_2^- , is based off estimates of hydrogen peroxide production of *N. viennensis* from supplementary material in Kim et al. (2016) ²⁷, and 1 mM CO_2 /1 mM H_2O_2 is based off the assumption of stoichiometric release of carbon dioxide from pyruvate as it interacts with hydrogen peroxide.

DNA Extraction

DNA was extracted from 1 mL of each culture taken at the time of harvest using the NucleoSpin Soil kit from Machery-Nagel and by following instructions for genomic DNA extraction from soil using lysis buffer SL1. Extracted DNA was eluted into 60 μ L of elution buffer SE (5 mM Tris/HCl, pH 8.5). Concentrations were measured using a Qubit DNA Assay with 5 μ L (high sensitivity protocol).

Combined Protein and Metabolite Extraction

Methods of a combined protein and metabolite extraction were followed based on those of Ott et al. (2019) ⁵³. Cell pellets were thawed on ice and resuspended in 500 μ L methanol:chloroform:water (MCW, 2.5:1:0.5). Water for all extractions came from a MilliporeSigma Milli-Q Reference A+ System (MilliQ water). Cells and undissolved pellet were transferred to a Lysing Matrix B tube with Lysing Matrix B (MP Biomedical) filled to 2-3 mm above the cone shape at the bottom of the tube. As an internal control and extraction standard, pentaerythritol (PE) and phenyl- β -glucopyranoside (PGP) (5 μ L of 1 mM solution for each) were spiked into each sample. Cells were lysed with FastPrep-24 homogenizer (MP Biomedical) for 30 seconds at a velocity of 4 m/s. After lysing, samples were cooled for 2 minutes on ice and then spun down for 2 minutes at 16,100 xg and 4°C. Supernatant was collected and saved after centrifugation. Cell lysis and centrifugation with the remaining sand pellet was repeated with 250 μ L of MCW. After centrifugation, the supernatant was added to the previously collected supernatant for each sample. 250 μ L of 80% ethanol was added to the remaining sand pellet. Samples were briefly vortexed and then incubated on a heat block at

80°C and 500 rpm for 30 minutes. After incubation, samples were spun down for 2 minutes at 16,100xg and 4°C. The ethanol supernatant was then collected and combined with the previously collected supernatant. Lysis tubes containing protein, cell debris, and the sand pellet were frozen at -70°C until protein extraction could be performed. 400 µL of water and 100 µL of chloroform were added to each of the collected supernatant tubes. Supernatant samples were then centrifuged for 5 minutes at 16,100xg and 4°C to separate phases. The upper phase (methanol, water, and polar metabolites) was collected into a new tube and dried using a step-wise pressure procedure. Dried samples were stored at -70°C until derivatization.

Metabolite Derivatization and GC-MS Analysis

Dried and frozen samples were allowed to acclimate to room temperature (~20 minutes). Under a fume hood, samples were dissolved in 20 µL of methoximation reagent (40 mg methoxyaminhydrochloride in 1 mL pyridine). Samples were then incubated for 90 minutes at 30°C and 900 rpm. Following incubation, 80 µL of silylation mix (1 mL N-methyl-N-trimethylsilyltrifluoroacetamid spiked with 30 µL of a mix of even-number alkanes (C10-C40)) was added to each sample and samples were incubated for 30 minutes at 37°C and 900 rpm. Following incubation, samples were centrifuged for 2 minutes at 14,000xg. Supernatant was then transferred to gas chromatography microvials and sealed with crimp caps.

Samples were analyzed in three separate runs in split-less injection mode on a Pegasus® BT GC-TOF-MS (LECO Corporation, St. Joseph, MI, USA) with a standard curve of selected metabolites for absolute quantification. Loaded volume for each sample was 1 µL. Metabolites were separated using a Rxi-5 ms column (30 m length, 0.25 mm diameter, 0.25 µm film; Restek, Centre County, PA, USA). The carrier gas as helium with a flow rate of 1 mL/min. Injection was done using a S/SSL injector with an injection temperature of 230°C. The column temperature started with 70°C for 1 minute and then followed a ramp to 340°C with at heating rate of 9°C/min. Temperature was held at 340°C for 15 minutes at the end. For the mass spectrometer, ion source and transfer line temperature was 250°C and acquisition delay was set to 300s. Masses were recorded in the range of 50-600 *m/z* with an acquisition rate of 10 spectra/sec.

Resulting peaks from selected metabolites were manually selected using LECO ChromaTOF software. Selected metabolites were absolutely quantified and normalized based on added extraction standards (see Data Analysis).

Protein Extraction

For the protein extraction, sand pellets containing cellular debris and proteins were thawed on ice. Once thawed, sand pellets were washed with 500 μ L of methanol. Samples were then centrifuged for 5 minutes at 16,100xg and 4°C. Methanol supernatant was then discarded and pellets were allowed to air dry under a fume hood.

To begin the extraction, 1 mL of TRIzol reagent (Thermo Fisher) was added to each pellet and mixed with the sand through gently pipetting. Samples were then incubated at room temperature for 15 minutes. Following incubation, 100 μ L of chloroform was added and tubes were inverted 5 times. Samples then incubated at room temperature for three minutes. Following the incubation with chloroform, samples were centrifuged at 16,100 xg for 2 minutes at 4°C to separate phases. The lower phase, containing the proteins, was removed to a new lo-bind protein Eppendorf tube. 550 μ L of water (or a 1:1 ratio with chloroform phase) was added to each sample. Tubes were inverted 5x and incubated at room temperature for 3 minutes. After incubation, tubes were centrifuged for 2 minutes at 16,100xg and 4°C. The apolar (bottom phase) was transferred to a new 2 mL low-bind protein Eppendorf tube. 1.5mL of ice cold 0.1 M NH_4Cl in methanol with 0.5% β -mercaptoethanol was added to each sample. Samples were then incubated on ice for 1-4 hours before being stored at -20°C overnight to facilitate protein precipitation.

Protein Washing

After overnight incubation at -20°C, samples were centrifuged for 15 minutes at 16,100xg and 4°C. The supernatant was discarded and 1.8 mL of ice cold methanol was added to each sample. Pellets were sonicated in an ice water bath with a Transsonic 700/H water bath (Elma) sonicator for 10-15 minutes or until pellets were completely dissolved. Samples were then centrifuged for 10 minutes at 16,100xg and 4°C. Supernatant was discarded and the washing and sonication steps were repeated with another 1.8 mL of ice cold methanol. After removal of the methanol supernatant from the second washing step, 1.8 mL of ice cold acetone was added to each sample. Pellets were again suspended in an ice water bath via sonication and then centrifuged for 15 minutes at 16,100xg and 4°C. The acetone supernatant was removed and pellets were air dried in a fume hood for 5-10 minutes (while avoiding over-drying). Once dry, proteins were stored at -80°C until further processing.

Protein Digestion

Dried protein pellets were resuspended in 500 μL of extraction buffer (8M urea in 50 mM HEPES, pH 7.8). Samples were then incubated for 30 minutes on a shaker at 4°C and 900 rpm. After incubation, protein concentrations were measured using the Bradford assay (Bio-Rad Cat. No. 500-0006) using a pre-made standard curve with bovine serum albumin (BSA) and an absorbance of 545 nm. The volume to get 25 μg of protein was calculated and taken for each sample and put into new low-bind protein Eppendorf tubes. Extraction buffer was added to each sample to reach a final volume of 220 μL . 5.64 μL of 200 mM dithiothreitol (DTT) was added to each sample and samples were then incubated for 45 minutes at 37°C and 700 rpm. Next, 2.26 of freshly prepared 1 M of iodoacetamide (IAA) was added and samples were incubated for 60 minutes at 30°C and 700 rpm. After the IAA incubation, a final amount of 6.27 μL of 200 mM DTT was added and samples were again incubated in the dark at room temperature for 15 minutes. Next 5 μL of 0.1 $\mu\text{g}/\mu\text{L}$ of mass spec grade rLysC (Promega) was added to each sample and proteins were digested for 3 hours at 37°C and 700 rpm. Following the rLysC digestion, 660 μL of trypsin buffer (2 mM CaCl_2 , 5 mM DTT, 50 mM NH_4HCO_3 , 10% acetonitrile (ACN)) was added to each sample. After the addition of trypsin buffer, 1 μL of trypsin beads from a Poroszyme Immobilized Trypsin Cartridge (Applied Biosystems) were added to each sample. Samples were incubated with trypsin beads for 16 hours at 30°C on a rotator. After digestion, samples were stored at 4°C (briefly, not all can be desalted at once) until they could be desalted. Following digestion, peptides were desalted using OMIX C18 pipette tips (Agilent Technologies). For this, samples were centrifuged briefly to pellet the trypsin beads before desalting. Tips were activated by washing with 100 μL of methanol and then washed twice with 100 μL of 0.1 % formic acid (FA). Peptide solution was then acidified with the addition of FA to a final concentration of 3 %. The solution containing peptides was pipetted through the C18 tips and then washed twice with 100 μL of 0.1 % FA. Peptides were then eluted and saved from the C18 tips by being washed twice with 100 μL of methanol. After desalting, peptides were dried in a ScanSpeed 40 speed vacuum with a ScanVac vacuum control and stored at -20 °C.

LC-MS/MS

To prepare the peptides for mass spectrometry analysis, peptides were resuspended in 250 μL of 2 % acetonitrile (ACN) and 0.1 % formic acid (FA). 5 μL were then injected into an

EASY-Spray C18 column (2 μ m, 100Å 75 μ m x 50 μ m, Thermo). Peptides were eluted from the column for 150 min using a 90 min linear gradient starting from 96 % solvent A (0.1 % FA) and 4 % solvent B (80 % ACN, 0.1 % FA) to 35 % of solvent B with a flow rate of 0.3 μ L/min. The linear gradient was followed by an increase of solvent B to 90 % over 1 min and held at 90 % for 8 min. Solvent B was then adjusted from 90 % to 4 % over 1 min and held at 4 % for the remaining 50 min. Ion source was an EASY-Spray source with a spray voltage of 1.9 kV. Mass spectrometry measurements were taken using an LC-QExactive-Plus (Thermo) with the following settings: 0 to 150 min; MS1: positive polarity, full scan range 380-1800 m/z , resolution 70,000, collision-induced dissociation (CID) fragmentation for the 20 most intense ions; MS2: loop count 20, resolution 17,500, scan range 200-2000 m/z .

Data Analysis

DNA Production, Protein Consumption, and Carbon Consumption

Values for DNA (μ g), protein (μ g), and consumed carbon (mmol) were normalized to the total amount of ammonia/ammonium consumed in each culture (mmol). Normalized values were subjected to a box plot analysis as described below for protein and metabolite values. For carbon consumption, values for two replicates of condition G were manually removed as the carbon balance produced negative values, likely due to sampling errors when collecting the gas phase.

Metabolites

Extracted metabolites were randomly distributed between three batches to be run on the GC-MS. Quality control samples containing known concentrations of metabolites were run for each batch individually. Detected trimethylsilyl (TMS) groups were combined for each metabolite. The internal standard PGP was used to normalize sample as it showed less variation than the PE internal standard. Based on an initial analysis of metabolite areas, some contamination could be seen from blank samples that contained only purified water. Therefore, PGP normalized areas from blank samples were subtracted from samples within their perspective runs. PGP normalized and blank corrected sample values were then used to determine absolute quantification of metabolites.

Standard curves were manually created for each metabolite in each batch. In cases where concentrations on the standard curve became non-linear, only quality curve points that captured the spread of areas from samples were used (i.e. if only low concentrations were detected in samples, high values of standards were disregarded). Linear regression was applied to selected quality control points and slope and intercept values were used to obtain pmol values for metabolites.

Absolute amounts of metabolites for each culture were converted into carbon-moles (C-Moles) and used in conjunction with the amount of carbon consumed for each culture to normalize values to show much consumed carbon was allocated for each metabolite (i.e. C-mol metabolite per C-mol carbon consumed). Carbon normalized values were subjected to a boxplot analysis for each metabolite in each condition. Based on this analysis outliers were removed. Normalized values with outliers removed were then hierarchically clustered using the heatmap.2⁵⁴ function in R version 3.6.3⁵⁵.

Proteins

Protein identification from raw mass spectrometry data for BN-PAGE cut-outs was analyzed using MaxQuant version 1.6.6.0⁵⁶. Max missed cleavages was set to 2 with variable modifications of “Oxidation (M)” and “Acetyl (Protein N-term)”, fixed modifications set to “Carbamidomethyl (C)”, an FDR of 0.01, and the digestive enzyme as trypsin. Minimum number of unique and razor peptides was set to 2 and protein values were LFQ normalized. Reference proteome for *N. viennensis* was downloaded from Uniprot in spring of 2019.

LFQ normalized data was analyzed using a boxplot analysis for each protein in each condition. Identified outliers were removed and averages were calculated for each protein in each group. A PCA plot of the cleaned up data was made using the `fviz_pca_ind` function of the `factoextra` package⁵⁷ in R. Cleaned up and averaged data were hierarchically clustered using the heatmap.2⁵⁴ function in R and split into 15 clusters. For mean comparisons, individual proteins were tested for normality (Shapiro test, `shapiro_test()` function of `rstatix` package⁵⁸) and homogeneity of variance (Levene test, `leveneTest()` function of `car` package⁵⁹). If all conditions passed these tests, one-way ANOVA (`aov()` function, `stats` package⁵⁵) was used to determine if protein averages varied among conditions. If all conditions did not pass these tests, a Kruskal-Wallis test (`Kruskal.test()` function, `stats` package⁵⁵) was used to determine if protein averages varied among conditions. For proteins of interest, Tukey tests

(TukeyHSD() function, stats package ⁵⁵) or Dunn's test (dunnTest() function, FSA package ⁶⁰) was used to determine which conditions were different for proteins with ANOVA and Kruskal-Wallis analysis respectively. In each case, p-values were adjusted using the Benjamini-Hochberg method (p.adjust() function, method="BH", stats package ⁵⁵).

Clusters of interest for the most limited condition (E) and non-limited conditions (ACDG) were identified by choosing the clusters with the highest average values for condition E and ACD respectively.

Proteins were evaluated based on functional categories according to their archaeal clusters of orthologous groups (arCOG) categories as defined in Reyes et al. (2020) ⁶¹.

Systems Biology Analysis

Clusters of high relative abundance for metabolites and proteins in condition E were correlated across all conditions to identify trends in metabolism. The R function cor.test was used with method set to "pearson" and use set to "complete.obs". p-values were adjusted according to the Benjamini-Hochberg method.

A correlation analysis was performed with all proteins for all conditions against MsrA to identify other proteins with similar patterns. The R function cor.test was used with method set to "pearson" and use set to "complete.obs". p-values were adjusted according to the Benjamini-Hochberg method.

Other R packages used in analysis included ggplot2 ⁶², dplyr ⁶³, reshape2 ⁶⁴, tidyverse ⁶⁵, ggVennDiagram ⁶⁶, RColorBrewer ⁶⁷, expss ⁶⁸, magrittr ⁶⁹, plot.matrix ⁷⁰, and viridis ⁷¹.

References

1. Adair, K. L. & Schwartz, E. Evidence that ammonia-oxidizing archaea are more abundant than ammonia-oxidizing bacteria in semiarid soils of northern Arizona, USA. *Microb. Ecol.* **56**, 420–426 (2008).
2. Baker, B. J., Lesniewski, R. A. & Dick, G. J. Genome-enabled transcriptomics reveals archaeal populations that drive nitrification in a deep-sea hydrothermal plume. *ISME J.* **6**, 2269–2279 (2012).

3. Hollibaugh, J. T., Gifford, S., Sharma, S., Bano, N. & Moran, M. A. Metatranscriptomic analysis of ammonia-oxidizing organisms in an estuarine bacterioplankton assemblage. *ISME J.* **5**, 866–878 (2011).
4. Karner, M. B., Delong, E. F. & Karl, D. M. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**, 507–510 (2001).
5. Leininger, S. *et al.* Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**, 806–809 (2006).
6. Nicol, G. W., Leininger, S., Schleper, C. & Prosser, J. I. The influence of soil pH on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea and bacteria. *Environ. Microbiol.* **10**, 2966–2978 (2008).
7. Shi, Y., Tyson, G. W., Eppley, J. M. & Delong, E. F. Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. *ISME J.* **5**, 999–1013 (2011).
8. Walker, C. B. *et al.* Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc. Natl. Acad. Sci.* **107**, (2010).
9. Kerou, M. *et al.* Proteomics and comparative genomics of Nitrososphaera viennensis reveal the core genome and adaptations of archaeal ammonia oxidizers. *Proc. Natl. Acad. Sci.* **113**, E7937–E7946 (2016).
10. 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.* 1–10 (2016) doi:10.1038/ismej.2016.2.
11. Lancaster, K. M., Caranto, J. D., Majer, S. H. & Smith, M. A. Alternative Bioenergy: Updates to and Challenges in Nitrification Metalloenzymology. *Joule* (2018) doi:10.1016/j.joule.2018.01.018.
12. Spang, A. *et al.* The genome of the ammonia-oxidizing candidatus nitrososphaera gargensis: Insights into metabolic versatility and environmental adaptations. *Environ. Microbiol.* **14**, 3122–3145 (2012).
13. Vajjala, N. *et al.* Hydroxylamine as an intermediate in ammonia oxidation by globally

- abundant marine archaea. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1006–1011 (2013).
14. Berg, I. A., Kockelkorn, D., Buckel, W. & Fuchs, G. A 3-Hydroxypropionate 4-Hydroxybutyrate Pathway in Archaea. *Science (80-.).* **218**, (2007).
 15. Könneke, M. *et al.* Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 8239–44 (2014).
 16. Otte, J., Mall, A., Schubert, D. M., Könneke, M. & Berg, I. A. Malonic Semialdehyde Reductase from the Archaeon *Nitrosopumilus maritimus* Is Involved in the Autotrophic 3-Hydroxypropionate/4-Hydroxybutyrate Cycle. (2015) doi:10.1128/AEM.03390-14.
 17. DeMirci, H. *et al.* Structural adaptation of oxygen tolerance in 4-hydroxybutyryl-CoA dehydratase, a key enzyme of archaeal carbon fixation. *bioRxiv* 2020.02.05.935528 (2020) doi:10.1101/2020.02.05.935528.
 18. Könneke, M. *et al.* Isolation of an autotrophic ammonia-oxidizing marine archaeon. (2005) doi:10.1038/nature03911.
 19. Rinke, C. *et al.* A standardized archaeal taxonomy for the Genome Taxonomy Database. *Nat. Microbiol.* **6**, 946–959 (2021).
 20. Abby, S. S., Kerou, M. & Schleper, C. Ancestral Reconstructions Decipher Major Adaptations of Ammonia-Oxidizing Archaea upon Radiation into Moderate. *MBio* **11**, (2020).
 21. Tournai, M. *et al.* *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc. Natl. Acad. Sci. USA* **108**, 8420–8425 (2011).
 22. Dimitri Kits, K. *et al.* Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle. *Nature* **549**, 269–272 (2017).
 23. Jung, M. Y. *et al.* Ammonia-oxidizing archaea possess a wide range of cellular ammonia affinities. *ISME J.* **16**, 272–283 (2022).
 24. Stieglmeier, M. *et al.* *Nitrososphaera viennensis* gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* **64**, 2738–2752 (2014).
 25. Qin, W. *et al.* Influence of oxygen availability on the activities of ammonia-oxidizing
-

- archaea. **9**, 250–256 (2017).
26. Kraft, B. *et al.* Oxygen and nitrogen production by an ammonia-oxidizing archaeon. *Science* (80-.). **375**, 97–100 (2022).
 27. Kim, J.-G. *et al.* Hydrogen peroxide detoxification is a key mechanism for growth of ammonia-oxidizing archaea. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7888–93 (2016).
 28. McCarty, P. L. Thermodynamic Electron Equivalents Model for Bacterial Yield Prediction: Modifications and Comparative Evaluations Perry. *Biotechnol. Bioeng.* **97**, 377–388 (2006).
 29. Rittmann, B. E. & McCarty, P. L. *Environmental Biotechnology: Principles and Applications*. (McGraw Hill Education, 2012).
 30. Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Archaeal Clusters of Orthologous Genes (arCOGs): An Update and Application for Analysis of Shared Features between Thermococcales, Methanococcales, and Methanobacteriales. *Life (Basel, Switzerland)* **5**, 818–40 (2015).
 31. Bayer, B., Mcbeain, K., Carlson, C. A. & Santoro, A. E. Carbon content , carbon fixation yield and dissolved organic carbon release from diverse marine nitrifiers. (2022).
 32. Bayer, B. *et al.* Nitrosopumilus adriaticus sp. nov. and Nitrosopumilus piranensis sp. nov., two ammonia-oxidizing archaea from the Adriatic Sea and members of the class Nitrososphaeria. *Int. J. Syst. Evol. Microbiol.* **69**, 1892–1902 (2019).
 33. Bayer, B. *et al.* Ammonia-oxidizing archaea release a suite of organic compounds potentially fueling prokaryotic heterotrophy in the ocean. *Environ. Microbiol.* **21**, 4062–4075 (2019).
 34. Utåker, J. B., Andersen, K., Aakra, Å., Moen, B. & Nes, I. F. Phylogeny and functional expression of ribulose 1,5-bisphosphate carboxylase/oxygenase from the autotrophic ammonia-oxidizing bacterium Nitrosospira sp. isolate 40KI. *J. Bacteriol.* **184**, 468–478 (2002).
 35. Caranto, J. D. & Lancaster, K. M. Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase. *Proc. Natl. Acad. Sci.* **114**, 8217–8222 (2017).
-

36. Estelmann, S. *et al.* Labeling and enzyme studies of the central carbon metabolism in *Metallosphaera sedula*. *J. Bacteriol.* **193**, 1191–200 (2011).
37. Yeh, L. & Kim, K. Regulation of acetyl-CoA carboxylase : Properties of CoA activation of acetyl-CoA carboxylase. **77**, 3351–3355 (1980).
38. Gerhardt, E. C. M. *et al.* The Bacterial signal transduction protein GlnB regulates the committed step in fatty acid biosynthesis by acting as a dissociable regulatory subunit of acetyl-CoA carboxylase. **95**, 1025–1035 (2015).
39. Ezraty, B. & Aussel, L. Methionine sulfoxide reductases in prokaryotes. **1703**, 221–229 (2005).
40. Maupin-Furlow, J. A. Methionine sulfoxide reductases of archaea. *Antioxidants* vol. 7 (2018).
41. Lin, Z. *et al.* Free methionine- (R) -sulfoxide reductase from *Escherichia coli* reveals a new GAF domain function. **104**, 1–6 (2007).
42. Romine, M. F. *et al.* Elucidation of roles for vitamin B12 in regulation of folate, ubiquinone, and methionine metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1205–E1214 (2017).
43. Vaccaro, J. A. & Naser, S. A. The Role of Methyl Donors of the Methionine Cycle in Gastrointestinal Infection and Inflammation. *Healthc.* **10**, (2022).
44. Fu, X. *et al.* Methionine sulfoxide reductase a (MsrA) and its function in ubiquitin-like protein modification in Archaea. *MBio* **8**, (2017).
45. Makarova, K. S. & Koonin, E. V. Archaeal Ubiquitin-Like Proteins : Functional Versatility and Putative Ancestral Involvement in tRNA Modification Revealed by Comparative Genomic Analysis. **2010**, 9–13 (2010).
46. Kapp, L. D. & Lorsch, J. R. GTP-dependent Recognition of the Methionine Moiety on Initiator tRNA by Translation Factor eIF2. 923–936 (2004)
doi:10.1016/j.jmb.2003.11.025.
47. Schmitt, E., Blanquet, S. & Mechulam, Y. The large subunit of initiation factor aIF2 is a close structural homologue of elongation factors. **21**, 1821–1832 (2002).
48. Anderson, J., Phan, L. & Hinnebusch, A. G. The Gcd10p Tgcd14p complex is the

- essential methyltransferase of *Saccharomyces cerevisiae*. **97**, 5173–5178 (2000).
49. Suzuki, I., Dular, U. & Kwok, S. C. Ammonia or ammonium ion as substrate for oxidation by *Nitrosomonas europaea* cells and extracts. *J. Bacteriol.* **120**, 556–558 (1974).
 50. Tinoco, I. J. *et al. Physical Chemistry Principles and Applications in Biological Sciences*. (Pearson, 2014).
 51. Wagman, D. D. *et al. Selected Values of Chemical Thermodynamic Properties Table for the First Thirty-Four Elements in the Standard Order of Arrangement; Technical Note 270-3*.
<https://babel.hathitrust.org/cgi/pt?id=uiug.30112105089574&view=1up&seq=4&skin=2021&q1=no2> (1968).
 52. McCarty, P. L. Energetics and Bacterial Growth. in *Organic Compounds in Aquatic Environments* (eds. Faust, S. D. & Hunter, J. V.) (Marcel Dekker, 1971).
 53. Ott, E., Fuchs, F. M., Moeller, R., Hemmersbach, R. & Kawaguchi, Y. Molecular response of *Deinococcus radiodurans* to simulated microgravity explored by proteometabolomic approach. 1–13 (2019) doi:10.1038/s41598-019-54742-6.
 54. Warnes, G. R. *et al. gplots: Various R Programming Tools for Plotting Data*. (2020).
 55. R Core Team (2020). R: A language and environment for statistical computing. (2020).
 56. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry–based shotgun proteomics. *Nat. Protoc.* **11**, (2016).
 57. Kassambara, A. & Mundt, F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. (2020).
 58. Kassambara, A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. (2021).
 59. Fox, J. & Weisberg, S. An {R} Companion to Applied Regression, Third Edition. (2019).
 60. Ogle, D. H., Wheeler, P. & Dinno, A. FSA: Fisheries Stock Analysis. (2021).
 61. Reyes, C. *et al.* Genome wide transcriptomic analysis of the soil ammonia oxidizing archaeon *Nitrososphaera viennensis* upon exposure to copper limitation. *ISME J.* **14**,
-

- 2659–2674 (2020).
62. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (2016).
 63. Wickham, H., François, R., Henry, L. & Müller, K. *dplyr: A Grammar of Data Manipulation*. (2021).
 64. Wickham, H. Reshaping Data with the reshape Package. *J. Stat. Software*, 21(12), 1-20 (2007).
 65. Wickham, H. *et al.* Welcome to the tidyverse. *J. Open Source Software*, 4(43), 1686 (2019).
 66. Gao, C.-H. *ggVennDiagram: A ‘ggplot2’ Implement of Venn Diagram*. (2021).
 67. Neuwirth, E. *RColorBrewer: ColorBrewer Palettes*. (2014).
 68. Demin, G. *expss: Tables, Labels and Some Useful Functions from Spreadsheets and ‘SPSS’ Statistics*. (2020).
 69. Bache, S. M. & Wickham, H. *magrittr: A Forward-Pipe Operator for R*. (2020).
 70. Klinke, S. *plot.matrix: Visualizes a Matrix as Heatmap*. (2021).
 71. Garnier, S. *et al.* *Rvision - Colorblind-Friendly Color Maps for R*. (2021).

Tables

Table 1: Proteins with the highest number of metabolite correlations from the E clusters.

# of Correlations	Gene	Protein
16	msrA	peptide methionine sulfoxide reductase
11	sdhA	succinate dehydrogenase flavoprotein subunit/fumarate reductase
10	elf-2a	translation initiation factor 2, subunit alpha
8	NVIE_006230	protein of unknown function
8	NVIE_015810	protein of unknown function DUF178
8	NVIE_016160	peptide/opine/nickel ABC uptake transporter, substrate-binding protein
8	NVIE_028890	protein of unknown function
8	sdhB	FeS-center protein of succinate dehydrogenase/fumarate reductase
7	NVIE_025600	putative tRNA methyltransferase complex GCD14 subunit
7	NVIE_027310	hypothetical protein
6	guaB	inosine-5'-monophosphate dehydrogenase
6	NVIE_001400	hypothetical protein
6	NVIE_003120	3-hydroxypropionyl-CoA synthetase
6	NVIE_027790	putative protein of unknown function DUF55
5	map	methionine aminopeptidase
5	NVIE_007180	Beta-lactamase domain-containing protein
5	NVIE_012010	exported protein of unknown function
5	NVIE_014480	putative glyoxalase/bleomycin resistance domain protein
5	NVIE_026610	conserved exported protein of unknown function
5	NVIE_030270	conserved exported protein of unknown function
5	NVIE_2290	putative NADPH-dependent FMN reductase
5	purM	phosphoribosylformylglycinamide cyclo-ligase

Associated functions:

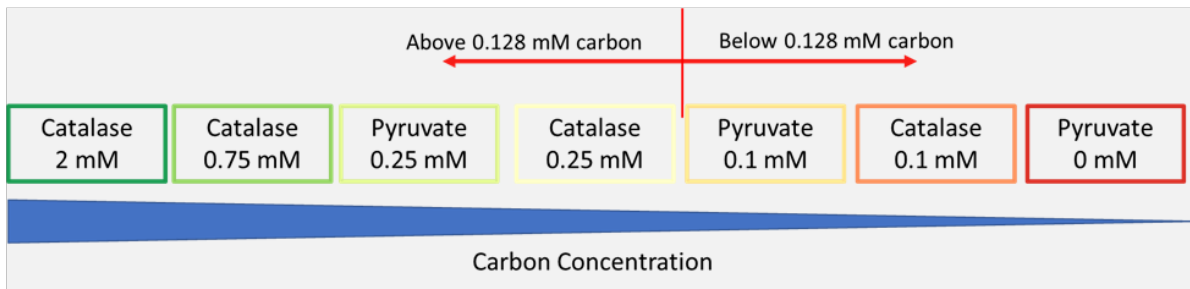
	Translation
	Nucleotide Metabolism
	Energy/Carbon Metabolism

Table 2: Protein identifiers for the carbon fixation cycle in *N. viennensis*.

Identifier	Locus Tag	Gene	Protein Product
AccA	NVIE_013720	accA	biotin carboxyl carrier protein of putative acetyl-CoA/ propionyl-CoA carboxylase
AccB	NVIE_013700	accB	acetyl-CoA/ propionyl-CoA carboxylase, carboxyltransferase subunit
AccC	NVIE_013710	accC	acetyl-CoA/propionyl-CoA carboxylase, biotin carboxylation subunit
Asd	NVIE_024340	asd	aspartate-semialdehyde dehydrogenase
MSR	NVIE_027580	NVIE_027580	malonic semialdehyde reductase
HPCS	NVIE_003120	NVIE_003120	3-hydroxypropionyl-CoA synthetase
CCH	NVIE_008340	NVIE_008340	3-hydroxybutyryl-CoA dehydratase
Adh4	NVIE_024420	adh4	putative NAD-dependent alcohol dehydrogenase
MCE	NVIE_028610	NVIE_028610	methylmalonyl-CoA epimerase
Mut	NVIE_028620	mut	methylmalonyl-CoA mutase, large subunit
MCM	NVIE_028640	NVIE_028640	methylmalonyl-CoA mutase, small subunit, C-terminus
GabD	NVIE_016240	gabD	succinate-semialdehyde dehydrogenase
HBCS	NVIE_024650	NVIE_024650	acetyl-CoA synthetase (ADP-forming)
HBCD	NVIE_017160	NVIE_017160	4-hydroxybutyryl-CoA dehydratase
PhaA	NVIE_020330	NVIE_020330	acetoacetyl-CoA thiolase subunit A
PhaB	NVIE_020320	NVIE_020320	putative beta-ketothiolase beta subunit
PhaE	NVIE_003020	phaE	putative poly(3-hydroxyalkanoate) synthase component
PhaC	NVIE_003010	phaC	poly(R)-hydroxyalkanoic acid synthase
HBD	NVIE_010920	NVIE_010921	hydroxyacyl-CoA dehydrogenase

Figures

A



B

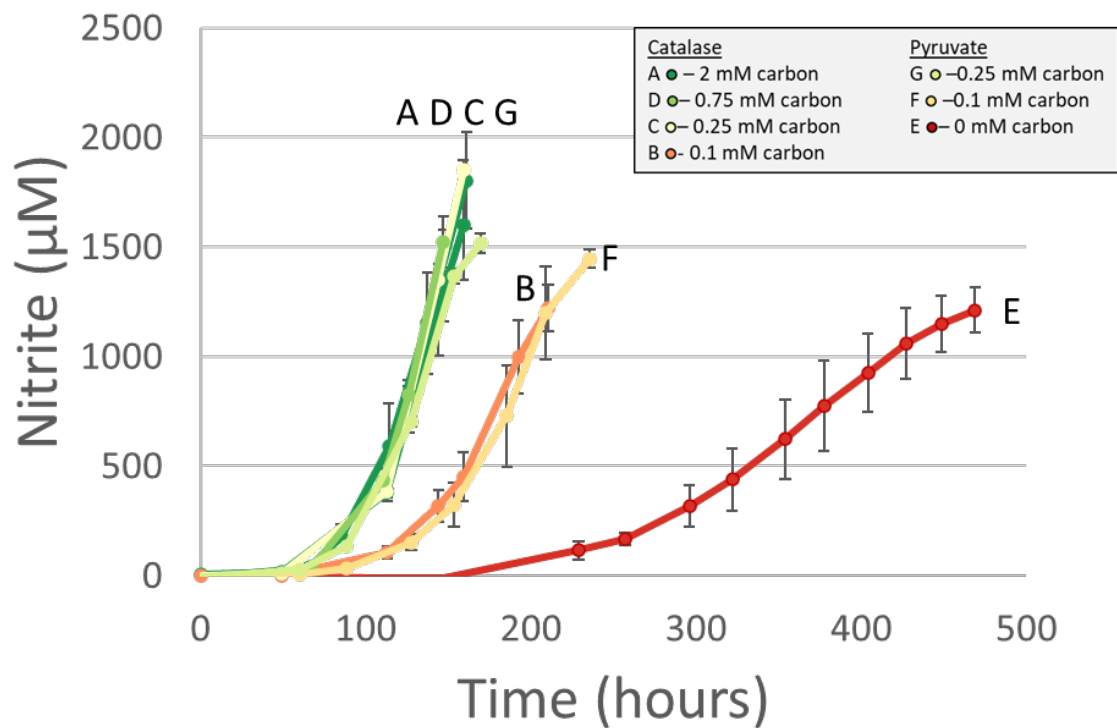


Figure 1: Growth curves of *N. viennensis* with varying amounts of inorganic carbon.

A.) Experimental design of different inorganic carbon concentrations. **B.)** Growth curves of *N. viennensis* exhibiting responses to varying levels of inorganic carbon. Conditions A, D, C, and G are all overlapping and exhibit normal growth. Conditions B and F are slightly slower. Condition E has an extended lag phase and slow growth.

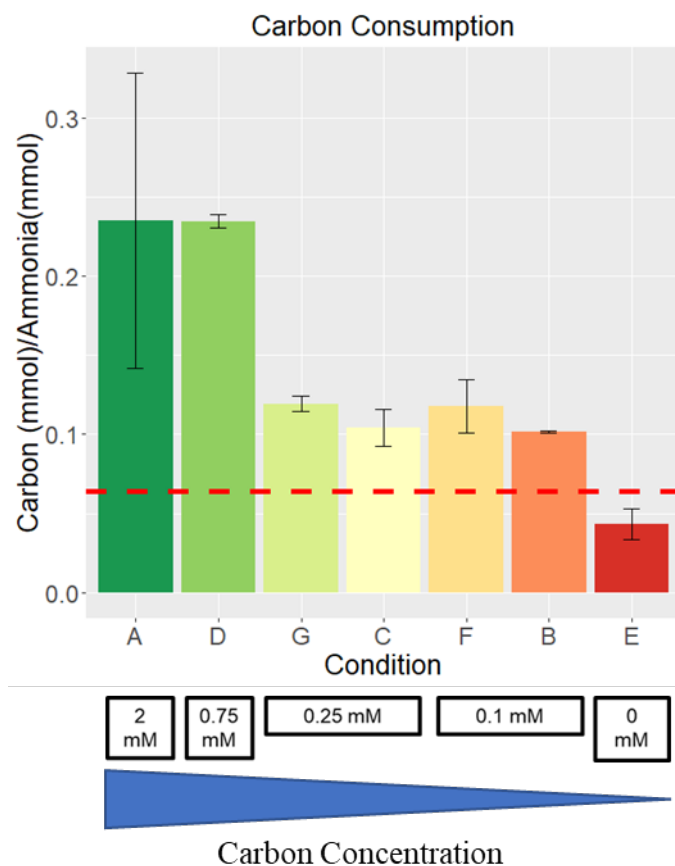


Figure 2: Carbon consumption of *N. viennensis*. Average inorganic carbon consumption per ammonia/ammonium consumed of *N. viennensis* supplied with different inorganic carbon concentrations. Error bars represent standard deviations of cultures. Black boxes indicate the approximate amount of starting inorganic carbon given to each culture. The red dotted line shows the theoretical amount of carbon consumption for 1 mmol of ammonia consumed based on Equation 18.

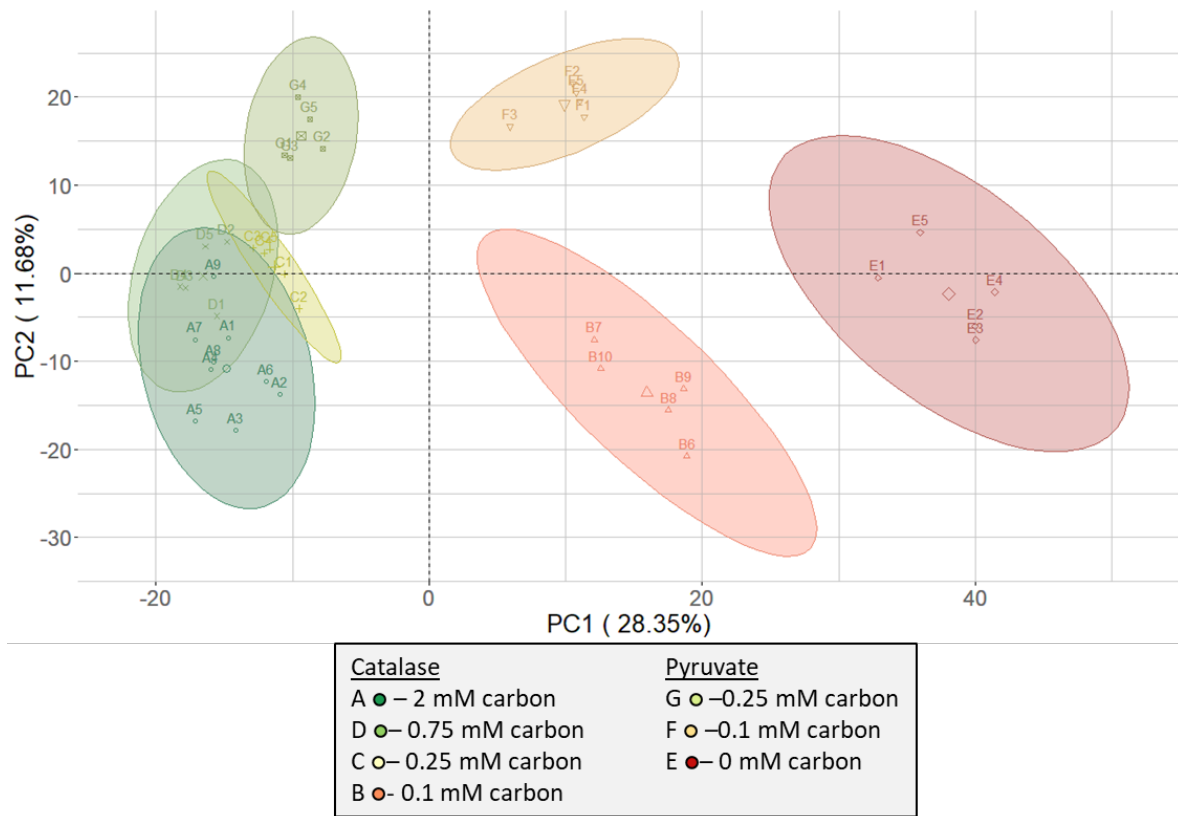


Figure 3: Principal component analysis of *N. viennensis* proteomes. Variance within the proteomes is best described by principal component 1 (PC1) with non-limited conditions A, D, C, and G on the far left of PC1 and the extremely limited culture on the far right. Limited cultures B and F fall between these two groupings on PC1. Ellipses represent confidence ellipses at a level of 0.95. Centers of ellipses are marked by unlabeled points within each ellipse respectfully.

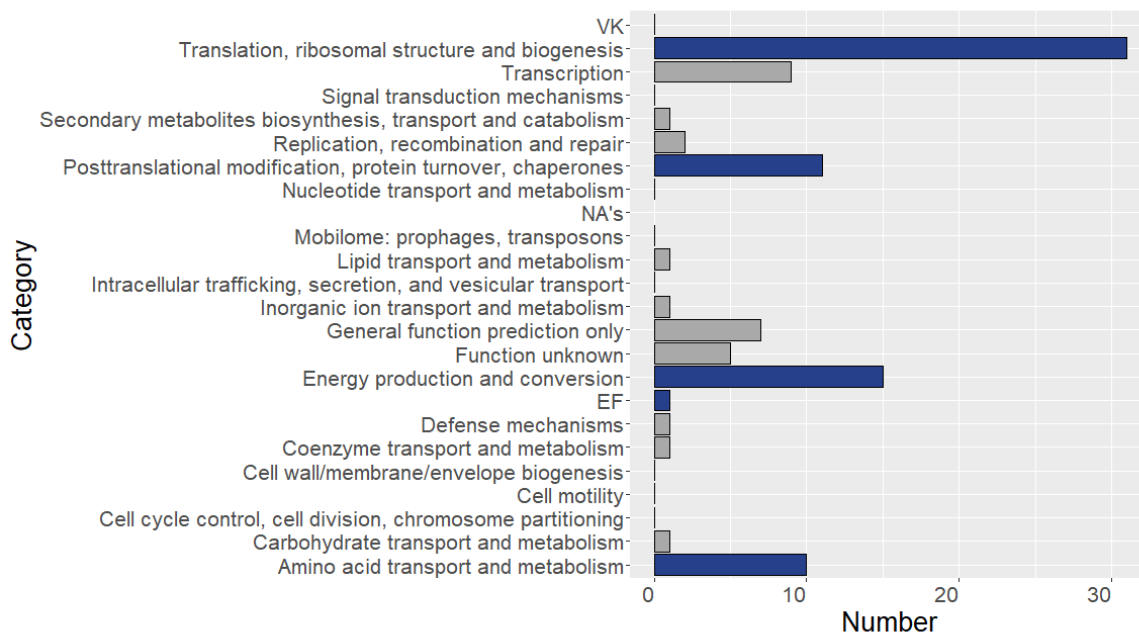


Figure 4 arCOG enrichment of the top 100 proteins found in all proteomes. Proteins were divided into categories based on archaeal clusters of orthologous genes (arCOGs). Bars represent the number of each category found. Bars colored in blue represent a category that is enriched (i.e. more than would be expected) based on a hypergeometric test. Categories VK and EF represent a combination of two categories: VK as defense mechanisms (V) and transcription (K) and EF as amino transport and metabolism (E) and nucleotide transport and metabolism (F).

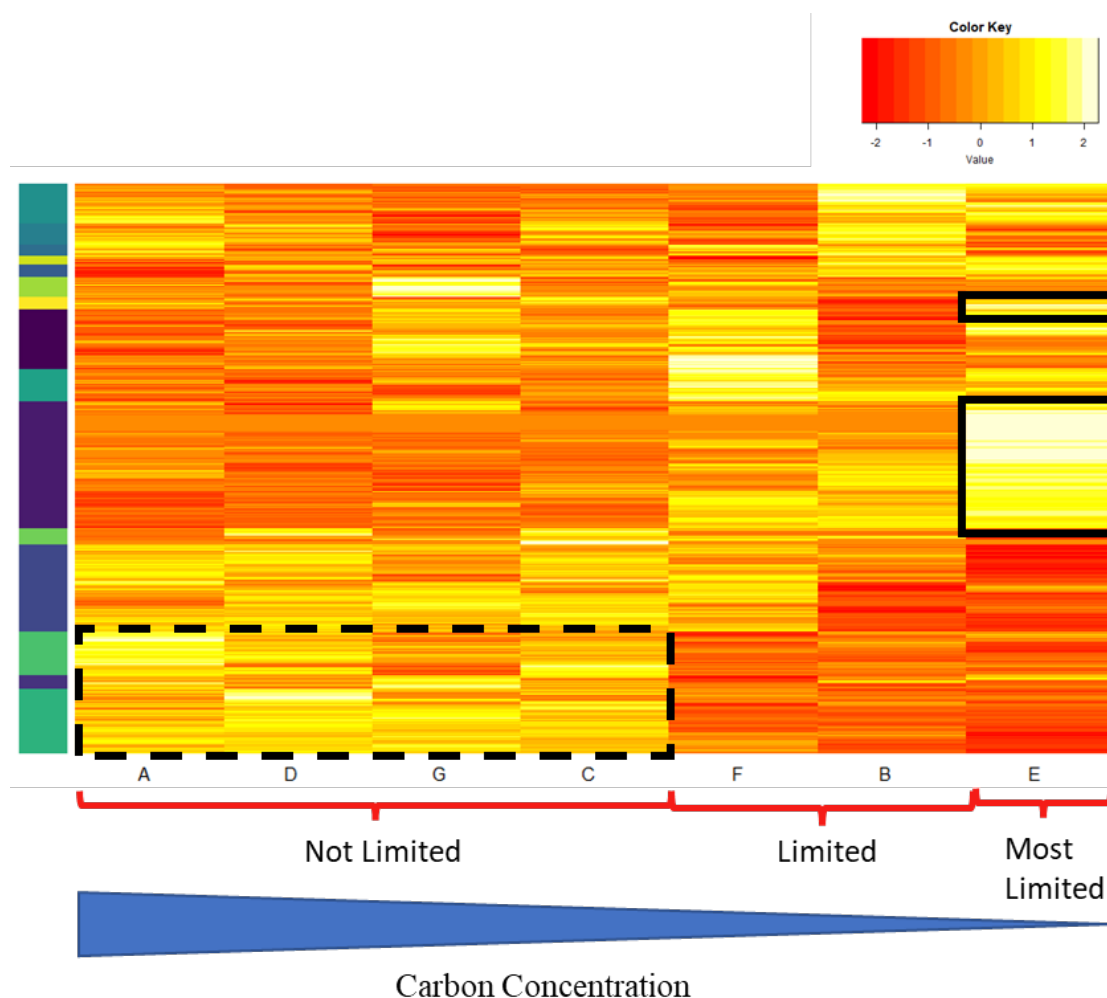


Figure 5: Heatmap of hierarchically clustered proteins. Proteins were clustered after being centered and scaled. Clustered proteins were divided into 15 clusters that are represented by the different colors on the left side of the heatmap. Clusters that have a high relative abundance in the most limited condition E are highlighted by black boxes with solid lines. Clusters that have a high relative abundance for conditions A, D, and C are highlighted by a block box with dashed lines. These clusters were used for subsequent analysis.

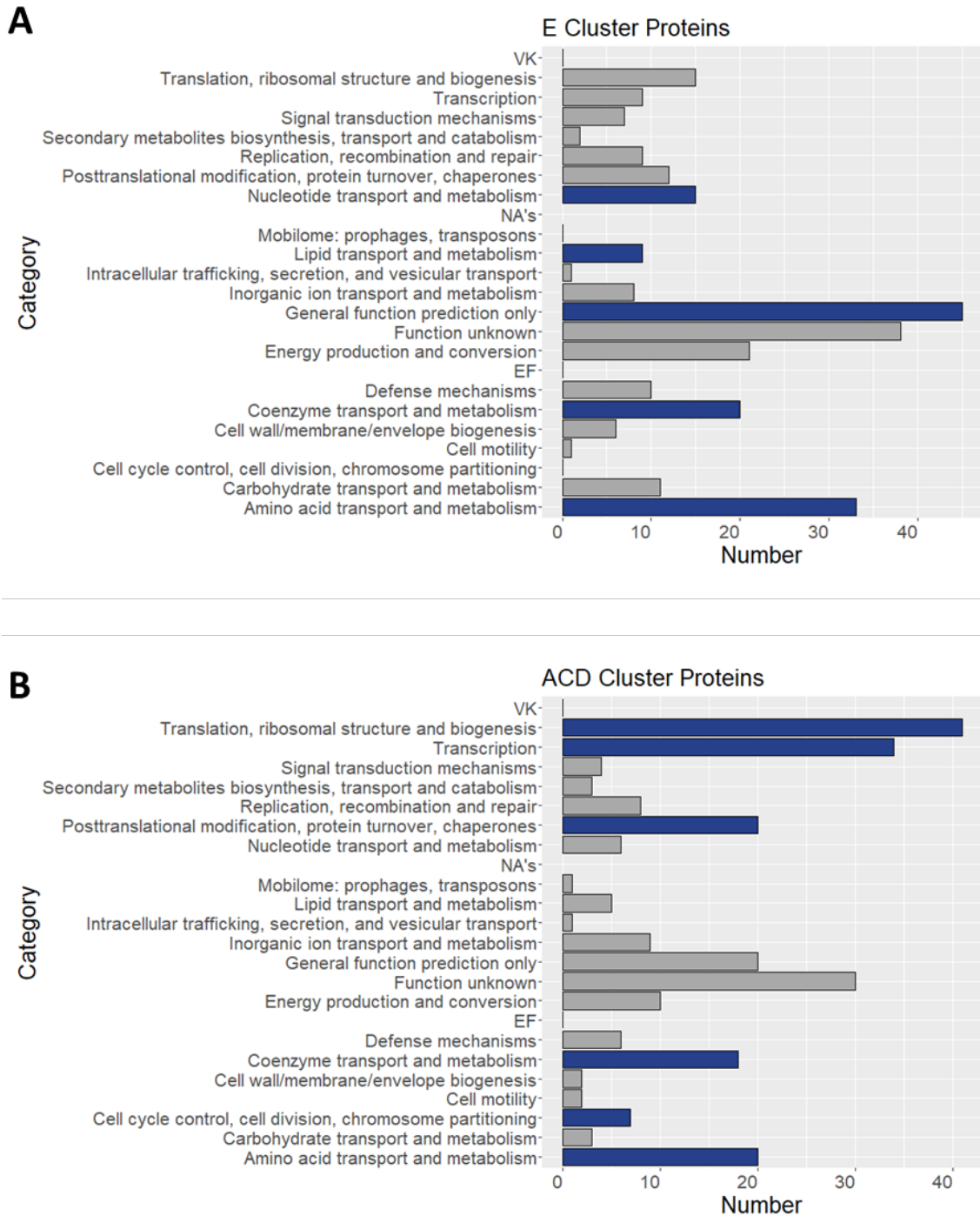


Figure 6: arCOG enrichment of selected protein clusters. A.) Proteins found in clusters representing the most limited condition E. B.) Proteins found in clusters representing the unlimited cultures A, C, and D. Bars represent the number of each category found. Bars colored in blue represent a category that is enriched (i.e. more than would be expected) based on a hypergeometric test. Categories VK and EF represent a combination of two categories: VK as defense mechanisms (V) and transcription (K) and EF as amino transport and metabolism (E) and nucleotide transport and metabolism (F).

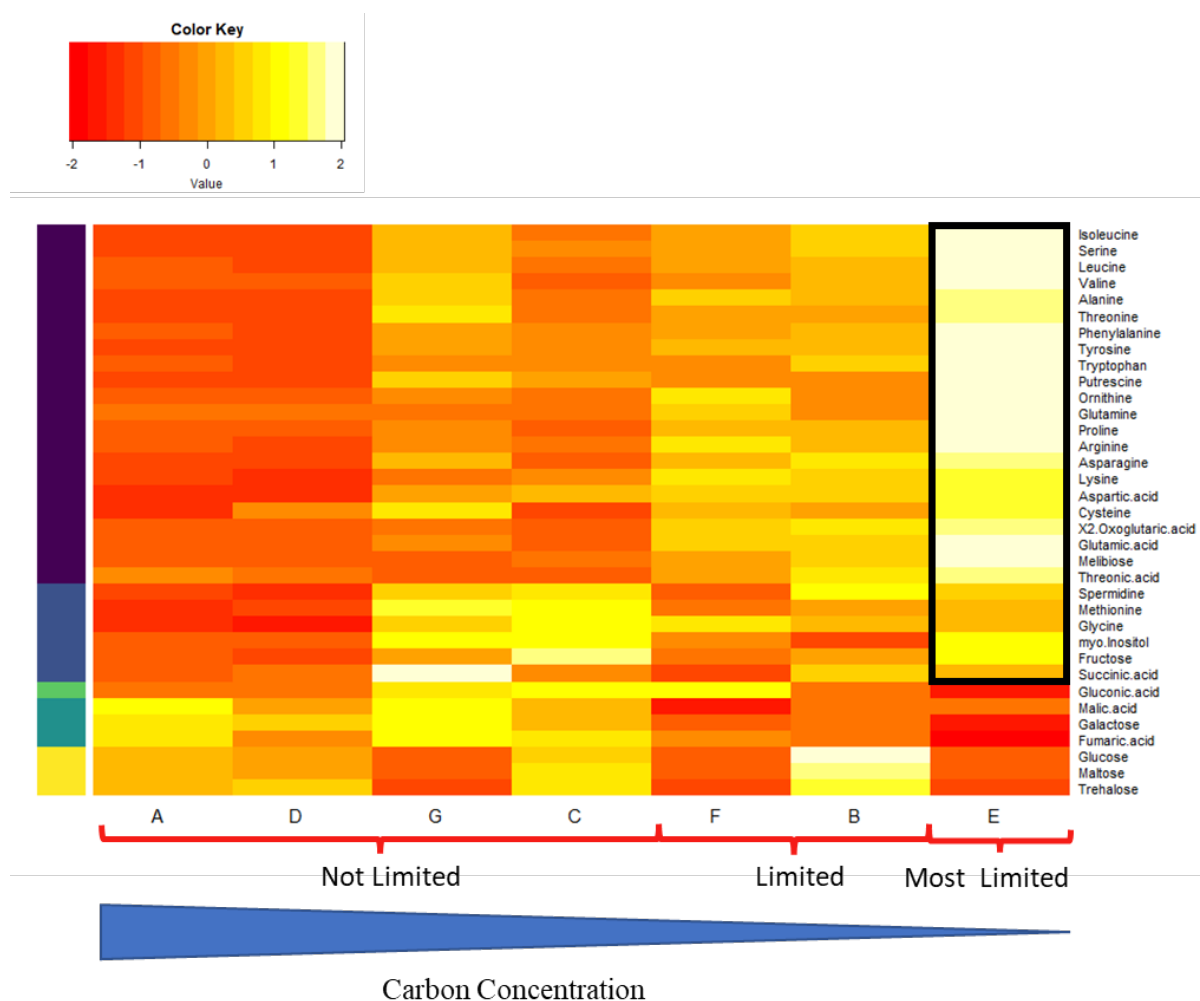


Figure 7: Heatmap of hierarchically clustered metabolites. Metabolites were clustered after being centered and scaled. Clustered metabolites were divided into 5 clusters that are represented by the different colors on the left side of the heatmap. Clusters that have a high relative abundance in the most limited condition E are highlighted by black boxes with solid lines.

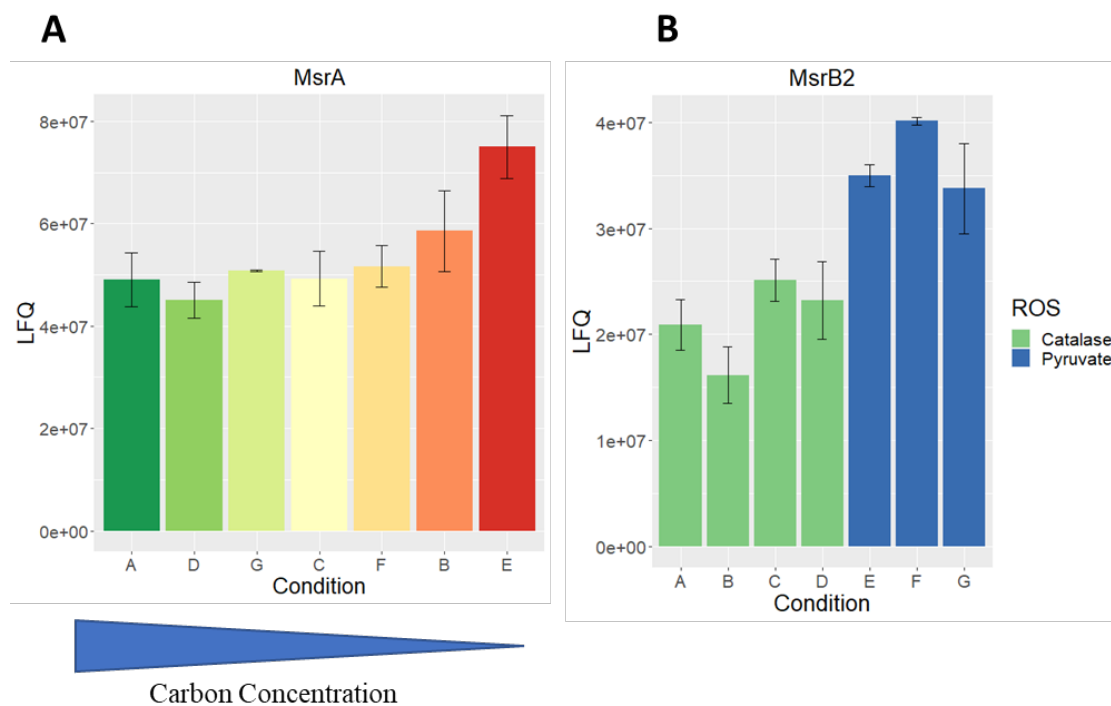


Figure 8: Relative intensities of MsrA (A) and MsrB2 across conditions. A.) Relative intensities of MsrA arranged in order of decreasing carbon concentration (left to right). B.) Relative intensities of MsrB2 arranged by ROS scavenger. Error bars represent standard deviations of conditions.

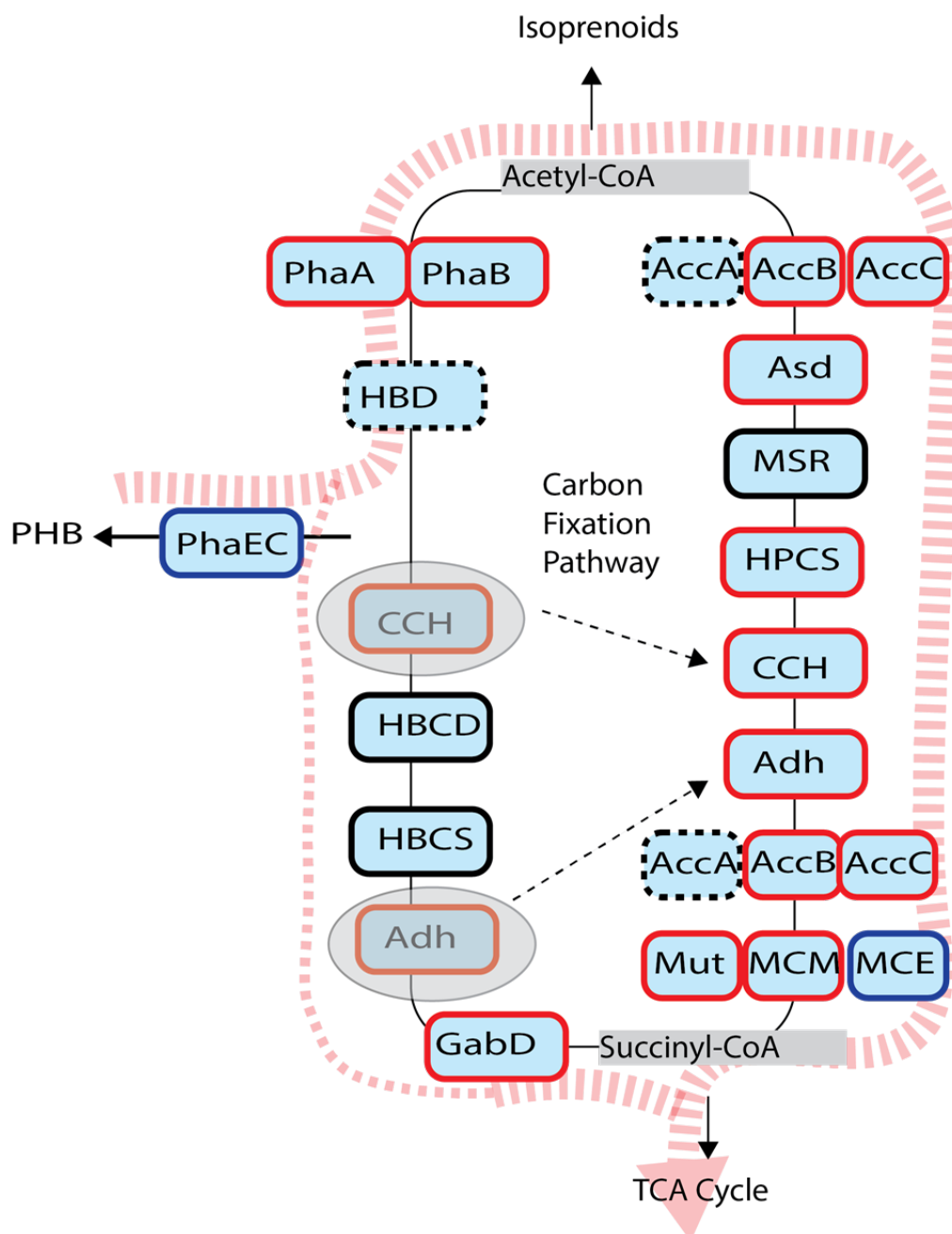


Figure 9: Response of carbon fixation cycle of *N. viennensis* under extreme carbon limitation. Proteins identifiers listed correspond to Table 2. Proteins outlined in red were found in the E cluster and showed a statistically significant difference from the most unlimited condition A. Proteins in blue were found in the ACD cluster and showed a downward trend as carbon became limited. Proteins in dashed black lines represent proteins that were either earlier limitation point (AccA and HBCS) and were not found in the E cluster. Proteins in solid black lines did not show a statistically significant difference between conditions ACD and conditions E. Grey ovals indicate proteins with multiple functions. Dotted arrows indicate where in the cycle these functions are thought to be repeated. The dashed red lines indicate the observed responsiveness of the pathway within the cell, with thicker lines representing a stronger response.

Supplementary Figures

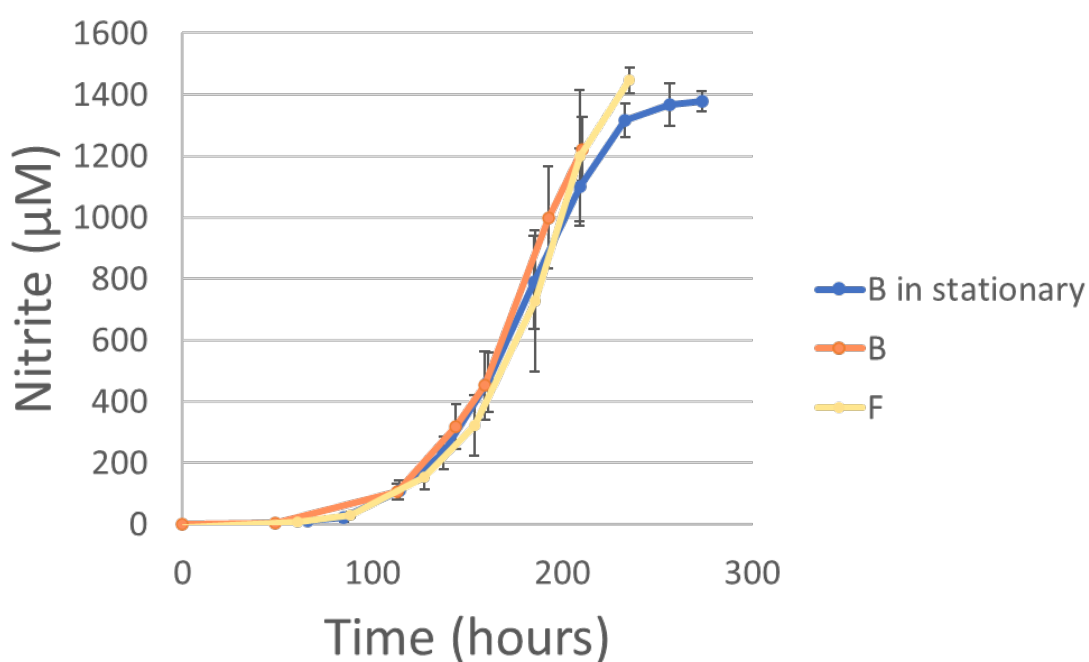


Figure S1: Growth curves of conditions B and F. Growth curves showing condition F (0.1 mM carbon with pyruvate) and condition B (0.1 mM carbon with catalase). Blue line represents condition B when allowed to grow above ~1200-1300 mM nitrite. The cultures begin to enter stationary phase. Condition F is able to oxidize more ammonia without entering stationary phase.

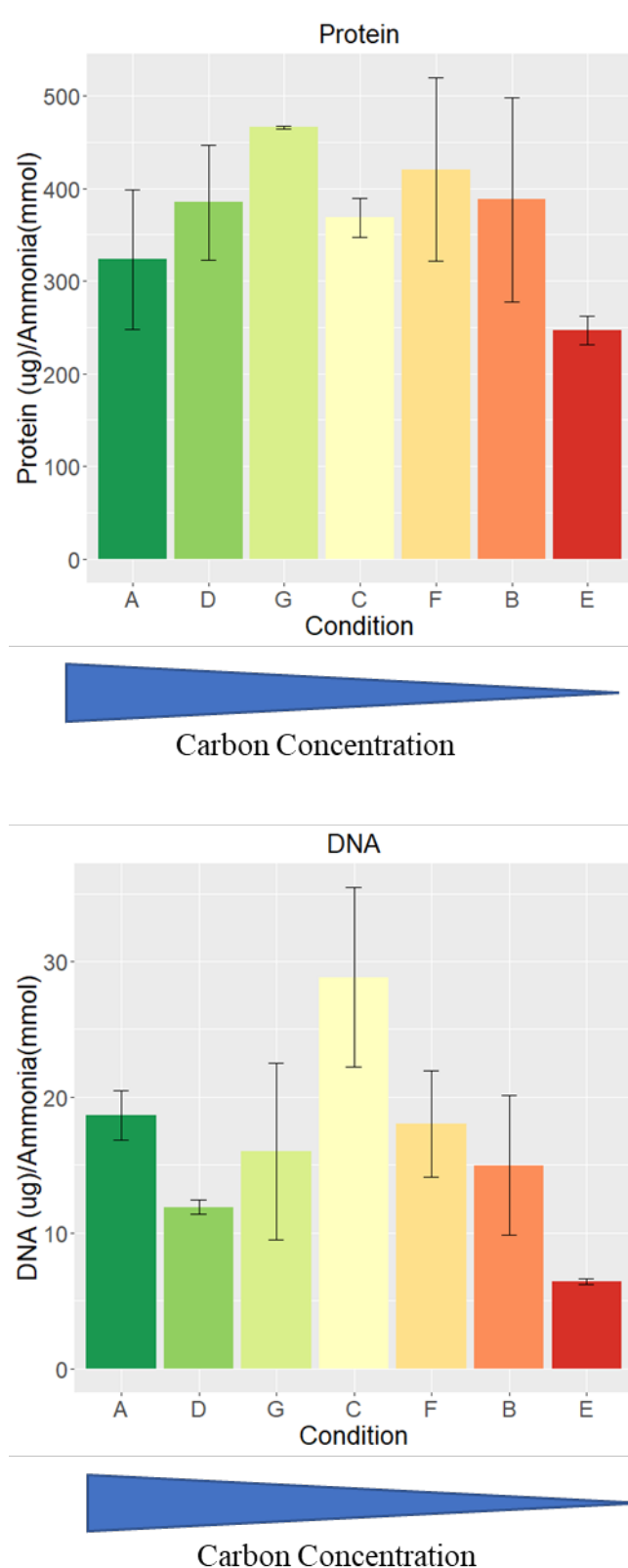


Figure S2: Proteins and DNA content of *N. viennensis*. Average protein and DNA content normalized to amount of consumed ammonium/ammonia in *N. viennensis* with different inorganic carbon concentrations. Error bars represent standard deviations of cultures.

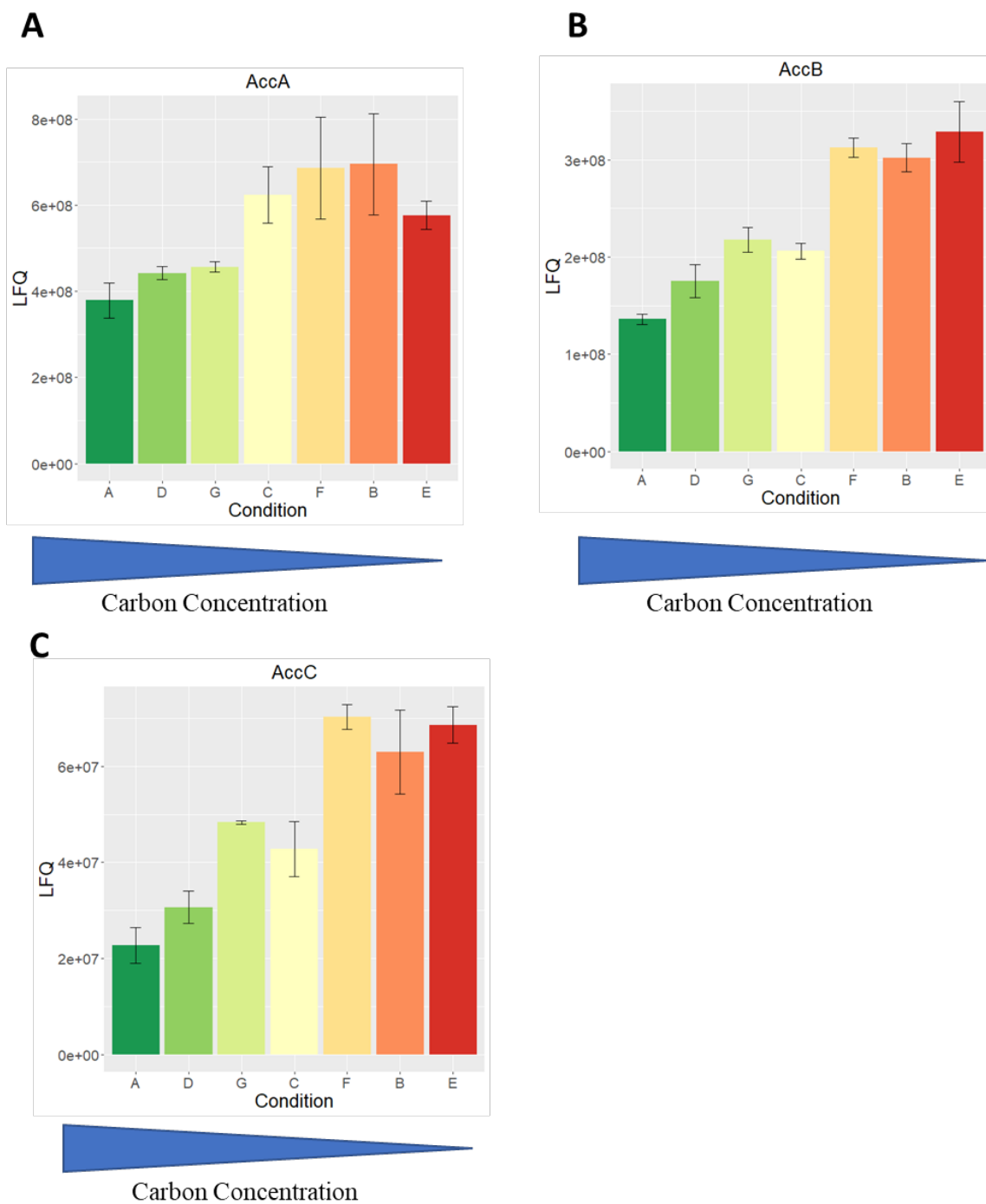


Figure S3: Relative intensities of subunits of acetyl-CoA carboxylase across conditions. A.) Carboxyl transferase subunit (AccA). B.) Biotin carboxyl carrier subunit (AccB). C.) Biotin carboxylase subunit (AccC). Error bars represent standard deviations of conditions.

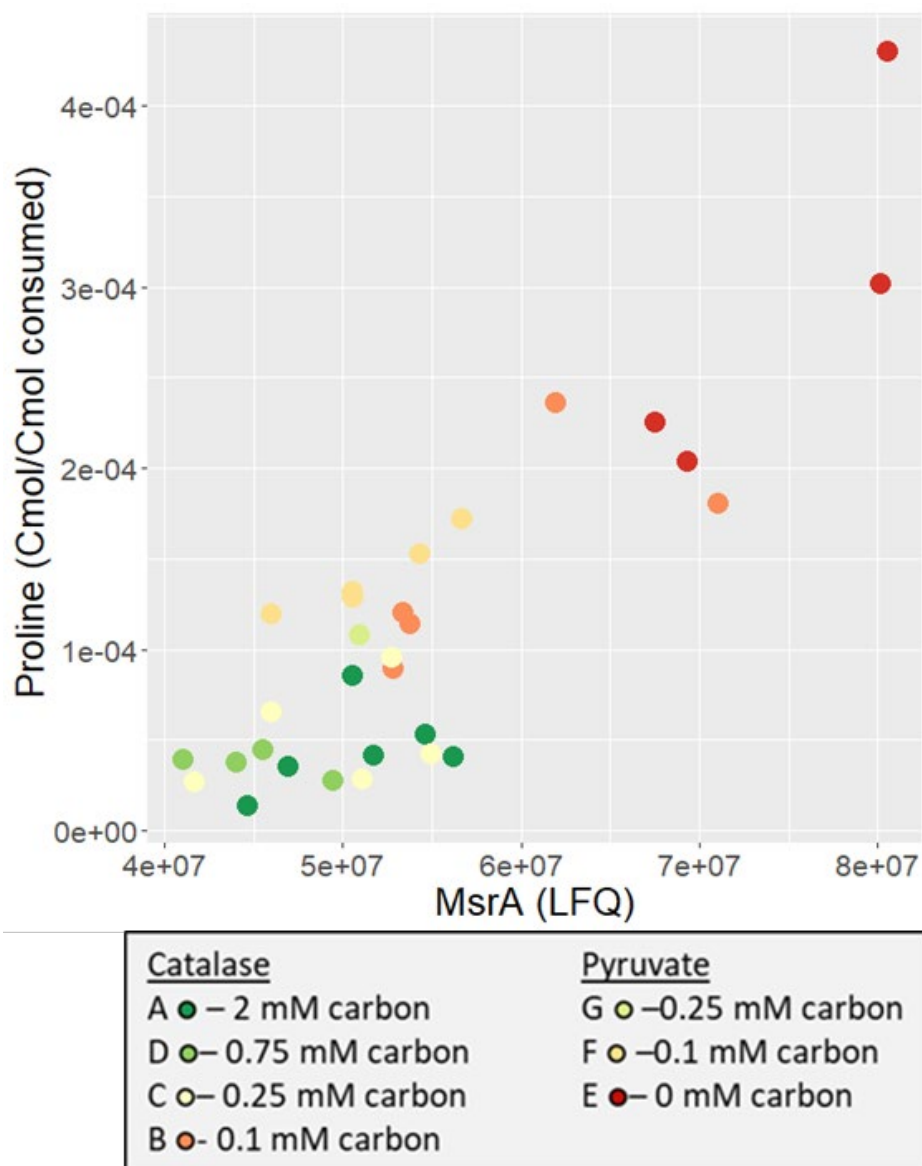


Figure S4: Correlation of MsrA with proline. An example of a protein-metabolite correlation. MsrA LFQ values steadily increase along with normalized proline concentrations as conditions move from unlimited (A) to extremely limited (E).

Chapter V

Discussion

This thesis took steps towards answering open questions in the central metabolisms of ammonia oxidizing archaea (AOA). Chapter II took a direct approach to characterize the archaeal AMO complex. The proteomic and modeling approach not only verified the participation of the canonical subunits (AmoABC) but also confirmed (AmoX) and found two new subunits (AmoYZ) that are conserved across the phylum and together likely replace the missing pieces when compared to the canonical subunits from ammonia oxidizing bacteria (AOB) and methanotrophs. Chapter III revealed several regulatory responses of an AOA from soil, *Nitrososphaera viennensis*, in response to copper limitation and postulated how AOA, which heavily rely on copper, might be able to obtain it from the environment. This work also supplied the first transcriptome of a soil AOA and gave indications of unique responses within the phylum based on different environmental conditions (i.e. marine vs. soil environments). Finally, Chapter IV explored system-wide metabolic reactions to a limiting source of carbon for AOA. This revealed specific responses at the protein level in the carbon fixation cycle, confirmed the ability of AOA to survive on carbon extracted from abiotic reactive oxygen species (ROS) reactions, and, in combination with metabolomics, identified a unique and unexpected role of methionine sulfoxide reductase A (MsrA) within *N. viennensis*. Although this has all pushed the knowledge of AOA further, there are still many open questions and anomalies within the metabolism of AOA that have yet to be addressed.

Unknowns in the electron transport chain in AOA

An aspect of AOA physiology that is not heavily discussed are the components of the classical electron transport chain (ETC). The complexes of the ETC are often overlooked as most of the focus is on unraveling the key steps of ammonia oxidation, in both archaea and bacteria, yet the connections between ammonia oxidization and the ETC are still important to consider. Although the exact steps are not clearly elucidated, it is known that in both archaea and bacteria hydroxylamine is produced from ammonia via the ammonia monooxygenase (AMO) complex^{1,2}. In AOA this appears to take place with three additional subunits (AmoXYZ) that together with AmoABC provide the structural framework for the complex to function (Chapter II). Electrons are then extracted from hydroxylamine via hydroxylamine oxidoreductase (HAO)³ and an unknown protein^{4,5} in bacteria and one or more unknown proteins in archaea^{6,7}. These processes are thought to take place outside of the membrane in the periplasmic or pseudo-periplasmic space^{8,9}. Electrons from these processes need to

somehow make it back into the quinone pool to be used in the generation of a proton motive force (pmf). In bacteria, it has been shown that electrons are passed to cytochrome *c554*^{10,11}, but how they are then passed back into the membrane is ambiguous. A transfer via cytochrome *c552*^{8,11,12} has been proposed but not proven. The situation is even murkier in AOA which do not possess *c*-cytochromes^{6,7} that are available as free electron carriers. Regardless, once in the quinone pool, electrons would be passed to Complex III which can then pump protons while passing the electrons onto another carrier to be taken to Complex IV. Again, in bacteria this role is fulfilled by cytochrome *c* and is unknown for AOA. Specific biochemistry techniques would be needed to solve these questions, but clues can be gained from this thesis work.

It has been proposed that copper containing proteins, such as plastocyanins or blue copper proteins are fulfilling this role of electron carriers^{6,7}. Previous genomic studies of AOA have identified a wealth of copper containing and plastocyanin proteins within the genomes, which also holds true for *N. viennensis*⁷. *N. viennensis* has also been shown to have a low copper limiting threshold, indicating that it is well adapted to extract copper it needs from the soil environment for these processes¹³. Chapter III was able to elucidate several responses of *N. viennensis* to limiting copper conditions giving insights on how the acquisition of copper could be regulated to supply not only these copper containing proteins, but also fundamental complexes in the metabolism (AMO, Complex IV, possibly Complex III) with copper. In terms of putative electron carriers, transcriptional patterns found in the copper replete and copper limited conditions would now allow for a search of candidates involved in electron transfer¹⁴.

Any electrons passed to the quinone pool could then be utilized by Complex III to contribute to the pmf before being further transported to the terminal oxidase, Complex IV. Like AMO, Complex IV utilizes an oxygen in its reaction¹⁵, but in this case the result is reduction of oxygen to water which also contributes to the pmf by translocation of protons. AOA have a type *aa3* cytochrome oxidase as Complex IV^{16,17}. In terms of energetics, this is very beneficial for AOA as type *aa3* oxidases are the most efficient in terms of protons pumped per electron^{17,18}. The pmf created by Complex III and Complex IV can finally be utilized by ATP synthase to produce ATP that can be used for important cellular processes. Mechanistically this process is likely very similar in AOB and AOA. However, the space in which a pmf is created has some important differences. AOB, as gram negative bacteria, contain a classical periplasmic space defined by both an inner and outer membrane where protons are concentrated. In AOA, the outer membrane is replaced by the proteinaceous S-layer and is therefore termed a pseudo-periplasm^{19,20}. Although the S-layer is thought to be

heavily glycosylated²¹, it is also made up of pores depending on the arrangement of the S-layer proteins. The presence of these pores could theoretically create a diffusion effect on the pseudo-periplasm. The S-layer also becomes problematic when focusing on the membrane fraction of archaeal cells. As a protein, it is highly contaminating when using proteomics to identify membrane protein bands in BN-PAGE gels. Regardless, the S-layer has been shown to be critical in other archaea^{22,23}.

Reactive oxygen species in AOA

The generation of pmf and chemiosmosis is a universal quality of biology. In aerobic organisms, like nitrifiers, this is accompanied by the generation of reactive oxygen species (ROS). The generation of ROS is possible at any point where electrons are being transferred and oxygen is capable of interfering and stealing electrons away to produce superoxide^{24,25}. ROS is known to be a problem for AOA²⁶, even in the environment²⁷. A common feature to deal with superoxide is the protein superoxide dismutase which is able to convert superoxide to hydrogen peroxide and oxygen. Hydrogen peroxide is also highly damaging and reactive and most organisms deal with this via the enzyme catalase which will neutralize hydrogen peroxide to oxygen and water. *N. viennensis* uniquely does not have a catalase, making it highly susceptible to ROS production⁷. To effectively grow *N. viennensis* in pure culture either catalase or α -keto organic acids need to be added to the medium^{20,28}. The addition of α -keto organic acids, such as pyruvate, initially was interpreted as a sign of possible mixotrophy. This was later disproved by showing that the abiotic reaction of hydrogen peroxide with pyruvate led to a detoxification of ROS and the production of carbon dioxide and water²⁶. The products of this abiotic reaction were used in Chapter IV to create a slow release of carbon dioxide to *N. viennensis*. While other AOA have been found to contain a catalase²⁹, the addition of pyruvate and other organic acids also help to stimulate growth in those lacking a catalase^{30,31}. Interestingly, there appears to be a distinct proteomic response between cultures grown with pyruvate vs. cultures grown with catalase. This is evident based on the results obtained for MsrB2 in Chapter IV that have a similar response between conditions regardless of extreme carbon limitation. The experimental design of Chapter IV will allow future investigation of the response of *N. viennensis* either between all catalase and all pyruvate conditions irrespective of carbon concentration or between conditions with the same carbon concentration (i.e 0.25 mM carbon (unlimited), catalase vs. pyruvate; 0.1 mM carbon (limited),

catalase vs. pyruvate). Understanding the response of AOA to ROS stress and how to overcome it may also be critical with the prediction of a disulfide bond within the soluble portion of the newly discovered AmoZ subunit (Chapter II). While it could simply serve as a stabilizing bond, it could also serve as a redox sensor for purposes that are yet unknown. The relationship between oxygen, and therefore ROS, and AOA also appears to have a long evolutionary history that can be more thoroughly explored from an experimental point of view in future experiments ^{29,32,33}.

As pointed out above, ROS can come from any electron transfer reactions where oxygen can be present. However, two enzyme complexes stand out as being possible primary sources: Complex III and Complex I ^{25,34–36}. Complex III is already thought to be an important part of the electron transport chain as it connects the quinone pool with the terminal oxidase. In AOA, Complex III accepts electrons from menaquinone rather than ubiquinone ³⁷. Menaquinone has a lower redox potential than ubiquinone and is thought to have been the primary quinone before the great oxidation event and is therefore more associated with low oxygen environments ^{38–40}. This would fit into the context of the last common ancestor of AOA being a hot organism as the amount of soluble oxygen is lower at higher temperatures ²⁹. The use of menaquinone with a low redox potential would also lead to the assumption of a lower redox potential for the Rieske group that the electrons are transferred to in Complex III. However, the *N. viennensis* amino acid sequence encodes for two residues, a serine and tyrosine, whose hydrogen bonding would imply a Rieske center with a high redox potential within Complex III ^{41,42} (Figure 1). Recent work on Complex III in menaquinol reducing organisms has hypothesized that the difference in reduction potentials of the Rieske iron-sulfur center and hemes groups of Complex III is symmetrical about the reduced quinone ⁴³. If true, the large difference between menaquinol and a high redox potential Rieske center would create an equally large difference for the corresponding hemes in Complex III. This difference would be large enough to put these hemes in a range of reduction potentials that would make the creation of superoxide radicals thermodynamically favorable ^{24,43}, therefore increasing ROS production substantially. While this would not be a problem in an anerobic environment, the effects are quite problematic with oxygen present ^{24,43} and could help explain why AOA have difficulty dealing with ROS in pure cultures. It should be pointed out that this hypothesis is specific to *bc₁* complexes and may not hold true for *b₆f* or other complexes. Complexes of the *bc₁* variety are known to oxidize ubiquinone or menaquinone while reducing cytochrome *c*. Alternatively, *b₆f* are often found in photosynthesizing organisms and oxidize plastoquinol

while reducing plastocyanin ⁴⁴. This presents an interesting scenario for AOA that are proposed to oxidize a menaquinol while reducing a plastocyanin. In some cases, such as *N. viennensis*, Complex III has been automatically annotated as a *b₆f* complex. However, it is known that menaquinol is the primary quinol in AOA ³⁷ and the closely related Crenarchaeota are categorized as cytochrome *b* rather than *b₆* ⁴⁵. For these reasons, the complex is more likely a cytochrome *b* complex. As AOA do not have cytochrome *c*, a different heme or metal group is likely compensating; possibly a copper based center as suggested for the marine AOA *N. maritimus* ⁶. A more thorough sequence analysis and direct experimentation on Complex III would be needed to verify the reduction potentials and types of Complex III metal centers in AOA and confirm the proposed excessive ROS production theory.

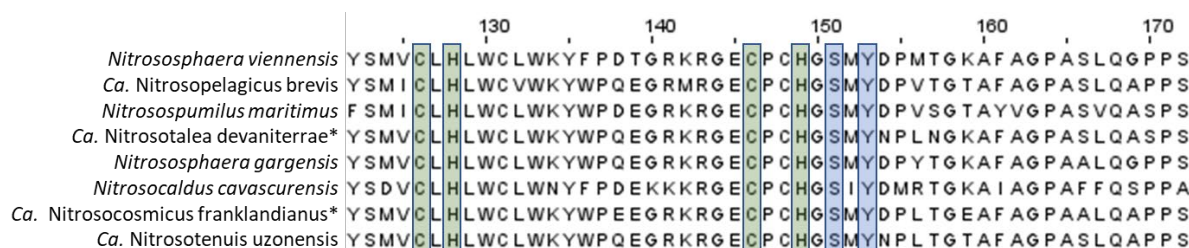


Figure 1: PetC (Rieske subunit of Complex III) Alignment for AOA.

The aligned portion of the Rieske center subunit of Complex III is shown with ligands binding the iron sulfur center highlighted in green and ligands putatively raising the reduction potential of the metal center highlighted in blue. Names marked with a * are named according to corrections suggest in Oren (2017) ⁴⁶.

The second likely candidate for large ROS production is Complex I. Complex I is a fascinating complex in itself, but in the context of AOA physiology it is even more enigmatic. Complex I has been proposed to be the electron input module for AOA via reverse electron flow ^{6,47}. Evidence of reverse electron flow was found in a lithotrophic bacterium *Thiobacillus ferrooxidans* (now known as *Acidithiobacillus ferrooxidans*) ⁴⁸, which also needs reducing equivalents to fix carbon, and has also been proposed in archaea ⁴⁹. Although not experimentally proven, a similar system is assumed to happen in ammonia oxidizers (bacterial and archaeal). Other candidates for electron input have been proposed in AOA based on genomics ^{6,7,47}. One possibility that could be involved in *N. viennensis* include a gene cluster of an NADH:oxidoreductase/hydrogenase gene cluster (NVIE_007050-NVIE_007100). However, this cluster is not transcribed as highly as the canonical Complex I and is not found

in the analyzed BN-PAGE bands of *N. viennensis* (Chapter II). Reverse electron flow has also been extensively studied in mitochondria^{36,50} as well as in *Escherichia coli*^{51,52}. These studies have identified both the FMN site and Q binding site, which interact with NADH and quinone respectively, as likely sites of ROS formation. While archaea are missing the subunits that would contain an FMN site (NuoEFG)⁵³, it is possible that their absence leaves other iron-sulfur sites exposed that could interact with oxygen. AOA would need a large amount of reducing equivalents in order to reduce carbon and produce more biomass. Therefore, a large reverse flow of electrons through Complex I could lead to a large production of ROS.

Reducing equivalents in AOA

If Complex I does prove to be the electron input module as mentioned above, the absence of NuoEFG presents a new challenge. Unlike bacteria, archaea are not capable of transferring electrons from quinols to NAD(P)H via Complex I due to the missing subunits NuoEFG that contain the FMN that would facilitate this transaction. Therefore, electrons would have to be transferred to an intermediate before being transferred to other reducing equivalents (i.e. NAD(P)H). Regardless, NADPH must be produced in considerable amounts to keep the 3-hydroxypropionate/4-hydroxybutyrate carbon fixation cycle running as shown with activity assays of enzymes in the carbon fixation cycle^{54,55}. NADPH is also a necessary co-factor for a variety of anabolic processes required for cell growth. In a typical heterotroph, this NADPH production is closely linked to other cellular processes that would be needed for growth such as the oxidative pentose phosphate pathway or glycolysis via the Embden-Meyerhof-Parnas pathway⁵⁶. These typical routes are not applicable for AOA as they encode for a reductive pentose phosphate pathway and do not have the capability of glycolysis⁷. Other possibilities, such as malic enzyme, could be a possible source of NADPH in *N. viennensis*, but this is not conserved across AOA and it is not currently known if it reacts with NADH or NADPH. One possibility for NADPH production could be from F420-NADPH oxidoreductases. F420 is highly abundant in the cell and can easily be seen through its autofluorescence⁴⁷. How electrons would get to F420 is unclear. Although Complex I has been proposed, it is unclear how one electron iron-sulfur centers would be able to reduce an obligate 2-electron carrier, unless NuoEFG is replaced by an unknown subunit that contains a co-factor that can facilitate this exchange. If F420 is the initially produced reducing

equivalent, electrons would still need to be passed to NADPH. The high abundance of F420 in the cell might also indicate a role in redox signaling and homeostasis.

Another intriguing source of NADPH could be the TCA cycle itself. The TCA cycle is primarily known for producing NADH to supply the electron transport chain with electrons via Complex I. However, in the bacterial lithoautotroph *Hydrogenovibrio crunogenus*, enzymes of the TCA cycle were shown to produce NADPH rather than NADH⁵⁷ as NADPH is the needed co-factor for its growth on a lithotrophic metabolism. This scenario fits nicely with the needs of a lithoautotroph that does not utilize an rTCA cycle for carbon fixation and could be the case for AOA. While this might help balance NADP⁺/NADPH ratios in AOA, all electrons coming from TCA cycle intermediates would still originate from reducing reactions found in the carbon fixation cycle.

A final point on reducing equivalents is their ability to regulate cellular processes. Heterotrophic metabolism is highly regulated by the ratios of NAD⁺/NADH and NADP⁺/NADPH. The addition of the phosphate in NADP⁺ allows for a clear separation between the two pools with catabolic processes being regulated by NAD⁺/NADH concentrations and anabolic processes being regulated by NADP⁺/NADPH concentrations. This is already seen in AOA as enzymes of the carbon fixation cycle are significantly more active with NADPH over NADH^{54,55}. However, this efficient control of metabolism via reducing equivalent concentrations is not nearly as helpful in lithotrophic organisms, including AOA, as NADH is in no way contributing to the electron transport chain. This would imply that the concentration and ratios of other important energetic intermediates play a more important role. In iron-oxidizers, this has been suggested to be fulfilled by ADP/ATP ratios that would affect the flow of electrons through either Complex IV for pmf production or through Complex I for reducing equivalent production⁴⁸. In addition to ATP, a likely candidate specific to AOA could be F420, particularly if F420 is the primary input module for electrons to enter the cell. With the current gaps in information for AOA metabolism it is difficult to conceive how the entire metabolism is regulated and at what primary points regulation may occur, but it does offer a wide range of opportunities for future research.

Possibilities for regulatory points in central pathways are highlighted in Chapter 4 with the strong response of particular proteins to carbon limitation. While extreme carbon limitation allowed for these responses to be identified, it is plausible that under normal growing conditions these proteins are still playing a pivotal role in the flux of carbon.

Findings on translation and cell division through omics approaches

Unexpectedly, the use of proteomics and systems biology highlighted proteins that could be involved in the general cell cycle of *N. viennensis* and other AOA. Very little is known about the AOA cell cycle, and aside from one study⁵⁸ what is postulated is typically derived from comparative genomics and biochemical investigations in other archaeal species. However, it is intriguing that AOA contain multiple possibilities for cell division machinery, i.e. FtsZ proteins as well as homologs for the ESCRT-III system⁵⁹. Enrichment analysis of archaeal clusters of orthologous groups (arCOGs) lead to an inspection of cell cycle proteins in carbon replete cultures in Chapter IV. This led to the identification of cell division proteins (CdvA and CdvC) that had a strikingly similar carbon response pattern to that of a proteosome within *N. viennensis*: all almost completely disappeared from the proteome upon carbon limitation, suggesting an active role of these proteins in cell division and normal cell processes. This is congruent with studies in the marine AOA *Nitrosopumilus maritimus* that identified the use of Cdv proteins for cell division rather than FtsZ⁵⁸. This likely indicates that AOA could utilize the ESCRT-III system in a similar way to the closely related *Sulfolobus acidocaldarius* that divides via the coordinated mechanism of protein degradation and Cdv proteins⁶⁰. However, in the study from *S. acidocaldarius*, this was specific to CdvB rather than CdvA or CdvC. A closer look at the Cdv proteins in AOA, along with a critical examination of any CdvB homologs (if present) would be needed.

Various correlation analyses also pointed towards physiological information regarding translation within AOA. Archaea are known to have translation initiation proteins similar to those of eukaryotes⁶¹, which are much better characterized. The patterns found in Chapter IV suggest a coordinated effort among translation initiation factors, MsrA, and RNA and protein maturation proteins with the availability of reduced methionine playing a potential crucial role. These potential dependencies could serve to direct future information processing studies in AOA and archaea in general.

Unknowns in ammonia oxidation of AOA

While there are many intriguing aspects of AOA physiology, it is undeniable that the biggest unknown is the subsequent steps of the energetic pathway following ammonia

oxidation to hydroxylamine. This has been an open question in the AOA field ever since the first genome was successfully annotated. Over the last several years it has become clear that this question cannot be addressed by comparative genomics alone. The final elucidation of these missing step(s) will have to come from biochemical and genetic techniques. Both are largely problematic with the current state of the field due to the difficulty of obtaining large amounts of biomass and the lack of a genetic system, respectively. The missing proteins and steps likely also contain crucial metal centers and may have important interactions with membrane proteins that could be adapted for the unique lipid membrane of AOA, making heterologous expression also a challenging path to take. Isotopic tracer studies are also difficult to use owing to the highly reactive nature of many of the nitrogen intermediates. Nevertheless, proteomic and transcriptomic data paired with what is already known can at least give a direction of where to focus.

The prevalence of Nirk in both proteomics and transcriptomics points toward this enzyme being somehow involved in the unresolved archaeal pathway. This also appears to corroborate observations that NO is consumed during the archaeal ammonia oxidation pathway and is a critical component based on inhibition of the pathway with the NO scavengers^{62–64}. These scavengers have proved ineffective in bacteria leading to the conclusion that NO is not a critical component in bacteria. However, interpreting NO scavenger data can be tricky. The unique membrane situation in AOA and AOB (S-layer pseudo-periplasm and outer membrane periplasm respectively) make results from AOA and AOB difficult to compare. NO scavenger use in bacteria might provide limited information if the scavenger cannot cross the outer membrane of the gram negative AOB. Additionally, new evidence of the bacterial ammonia oxidation pathway would also suggest that NO is in fact an important intermediate on the way to producing nitrite^{4,5}. Regardless of its role in bacteria, it seems likely that NO is somehow involved in archaea. However, recent genomic evidence from the genus *Nitrosocaldus* has confirmed the absence of an identifiable Nirk^{65–67} and no nitrite reductase activity could be detected in extracts from cells of a pure culture of *N. cavascurensis* (Melcher et al. in preparation). Either this genus has an unknown protein taking its place, or Nirk might not be as critical to the metabolism as is currently thought. Perhaps it plays a more protective role to prevent nitrite toxicity as seen in the AOB *Nitrosomonas europaea*^{68–70}, rather than a metabolic role. Without more information, integrating the role of Nirk in AOA metabolism proves to be difficult.

In summary, there are many questions left to answer. However, this thesis was able to produce transcriptomic, proteomic, and metabolomic data that begin to illustrate important players within the metabolism of *N. viennensis*. Additionally, many of the methods used and developed for *N. viennensis* in this thesis can serve as a basis for the study of other ammonia oxidizing archaea, particularly with regard to bioenergetic investigations of the membrane as evidenced by the successful application of the BN-PAGE methodology to not only *N. viennensis*, but also the distantly related thermophilic AOA, *Nitrosocaldus cavascurensis*. The central themes of nitrogen, copper, and carbon as outlined in Chapters II, III, and IV respectively not only expanded the knowledge for these specific areas within *N. viennensis* but were also utilized in a systematic fashion to begin exploring ways the various fundamental pathways of AOA are connected. In the future, with the development of biomass accumulating and genetic techniques for AOA, many of these open questions will hopefully be readily answered when supplemented with data that can be produced from the currently available techniques.

References

1. Vajrala, N. *et al.* Hydroxylamine as an intermediate in ammonia oxidation by globally abundant marine archaea. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1006–1011 (2013).
2. Dua, R. D., Bhandari, B. & Nicholas, D. J. D. Stable isotope studies on the oxidation of ammonia to hydroxylamine by nitrosomonas europaea. *FEBS Lett.* **106**, 401–404 (1979).
3. Andersson, K. K. & Hooper, A. B. O₂ and H₂O are each the source of one O in NO₂ produced from NH₃ by Nitrosomonas: 15N-NMR evidence. *FEBS Lett.* **164**, 236–240 (1983).
4. Caranto, J. D. & Lancaster, K. M. Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase. *Proc. Natl. Acad. Sci.* **114**, 8217–8222 (2017).
5. Lancaster, K. M., Caranto, J. D., Majer, S. H. & Smith, M. A. Alternative Bioenergy: Updates to and Challenges in Nitrification Metalloenzymology. *Joule* (2018) doi:10.1016/j.joule.2018.01.018.

6. Walker, C. B. *et al.* Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc. Natl. Acad. Sci.* **107**, (2010).
7. Kerou, M. *et al.* Proteomics and comparative genomics of Nitrososphaera viennensis reveal the core genome and adaptations of archaeal ammonia oxidizers. *Proc. Natl. Acad. Sci.* **113**, E7937–E7946 (2016).
8. DiSpirito, A. A., Taaffe, L. R. & Hooper, A. B. Localization and concentration of hydroxylamine oxidoreductase and cytochromes c-552, c-554, cm-553, cm-552 and a in Nitrosomonas europaea. *BBA - Bioenerg.* **806**, 320–330 (1985).
9. Sakoula, D. *et al.* Universal activity-based labeling method for ammonia- and alkane-oxidizing bacteria. *ISME J.* 1–14 (2021) doi:10.1038/s41396-021-01144-0.
10. Arciero, D. M., Hooper, A. B. & Balny, C. Spectroscopic and Rapid Kinetic Studies of Reduction of Cytochrome c554 by Hydroxylamine Oxidoreductase from Nitrosomonas europaea. *Biochemistry* **30**, 11466–11472 (1991).
11. Yamanaka, T. & Shinra, M. Cytochrome c-552 and Cytochrome c-554 Derived from Nitrosomonas europaea. *J. Biochem.* **75**, 1265–1273 (1974).
12. Klotz, M. G. & Stein, L. Y. Nitrifier genomics and evolution of the nitrogen cycle. (2007) doi:10.1111/j.1574-6968.2007.00970.x.
13. Reyes, C. *et al.* Copper limiting threshold in the terrestrial ammonia oxidizing archaeon Nitrososphaera viennensis. *Res. Microbiol.* **171**, 134–142 (2020).
14. Reyes, C. *et al.* Genome wide transcriptomic analysis of the soil ammonia oxidizing archaeon Nitrososphaera viennensis upon exposure to copper limitation. *ISME J.* **14**, 2659–2674 (2020).
15. Yoshikawa, S. & Shimada, A. Reaction mechanism of cytochrome c oxidase. *Chem. Rev.* **115**, 1936–1989 (2015).
16. Hemp, J. & Gennis, R. B. Diversity of the Heme-Copper superfamily in archaea: Insights from genomics and structural modeling. *Results Probl. Cell Differ.* **45**, 1–31 (2008).
17. Sousa, F. L. *et al.* The superfamily of heme–copper oxygen reductases: Types and

- evolutionary considerations. *Biochim. Biophys. Acta - Bioenerg.* **1817**, 629–637 (2012).
18. Wikström, M., Sharma, V., Kaila, V. R. I., Hosler, J. P. & Hummer, G. New perspectives on proton pumping in cellular respiration. *Chem. Rev.* **115**, 2196–2221 (2015).
 19. Albers, S. V. & Meyer, B. H. The archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414–426 (2011).
 20. Stieglmeier, M. *et al.* Nitrososphaera viennensis gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* **64**, 2738–2752 (2014).
 21. Albers, S.-V. & Jarrell, K. F. The archaeellum: how Archaea swim. *Front. Microbiol.* **6**, 23 (2015).
 22. Zink, I. A. *et al.* CRISPR-mediated gene silencing reveals involvement of the archaeal S-layer in cell division and virus infection. *Nat. Commun.* **10**, 1–14 (2019).
 23. Zhang, C. *et al.* Cell structure changes in the hyperthermophilic crenarchaeon *Sulfolobus islandicus* lacking the S-layer. *MBio* **10**, (2019).
 24. Wood, P. M. The two redox potentials for oxygen reduction to superoxide. *Trends Biochem. Sci.* **12**, 250–251 (1987).
 25. Murphy, M. P. How mitochondria produce reactive oxygen species. *Biochem. J.* **417**, 1–13 (2009).
 26. Kim, J.-G. *et al.* Hydrogen peroxide detoxification is a key mechanism for growth of ammonia-oxidizing archaea. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7888–93 (2016).
 27. Tolar, B. B. *et al.* Ammonia oxidation in the ocean can be inhibited by nanomolar concentrations of hydrogen peroxide. *Front. Mar. Sci.* **3**, 1–16 (2016).
 28. Tournu, M. *et al.* Nitrososphaera viennensis, an ammonia oxidizing archaeon from soil. *Proc. Natl. Acad. Sci. USA* **108**, 8420–8425 (2011).
 29. Abby, S. S., Kerou, M. & Schleper, C. Ancestral Reconstructions Decipher Major Adaptations of Ammonia-Oxidizing Archaea upon Radiation into Moderate. *MBio* **11**, (2020).

30. Bayer, B. *et al.* Nitrosopumilus adriaticus sp. nov. and Nitrosopumilus piranensis sp. nov., two ammonia-oxidizing archaea from the Adriatic Sea and members of the class Nitrososphaeria. *Int. J. Syst. Evol. Microbiol.* **69**, 1892–1902 (2019).
31. Qin, W. *et al.* Nitrosopumilus maritimus gen. nov., sp. nov., Nitrosopumilus cobalaminigenes sp. nov., Nitrosopumilus oxyclinae sp. nov., and Nitrosopumilus ureiphilus sp. nov., four marine ammonia-oxidizing archaea of the phylum Thaumarchaeota. *Int. J. Syst. Evol. Microbiol.* (2017) doi:10.1099/ijsem.0.002416.
32. Ren, M. *et al.* Phylogenomics suggests oxygen availability as a driving force in Thaumarchaeota evolution. *ISME J.* (2019) doi:10.1038/s41396-019-0418-8.
33. Sheridan, P. O. *et al.* Gene duplication drives genome expansion in a major lineage of Thaumarchaeota. *Nat. Commun.* **11**, 1–12 (2020).
34. Adam-Vizi, V. & Chinopoulos, C. Bioenergetics and the formation of mitochondrial reactive oxygen species. *Trends Pharmacol. Sci.* **27**, 639–645 (2006).
35. Chen, Q., Vazquez, E. J., Moghaddas, S., Hoppel, C. L. & Lesnefsky, E. J. Production of Reactive Oxygen Species by Mitochondria CENTRAL ROLE OF COMPLEX III*. *Publ. JBC Pap. Press* (2003) doi:10.1074/jbc.M304854200.
36. Onukwufor, J. O., Berry, B. J. & Wojtovich, A. P. Physiologic Implications of Reactive Oxygen Species Production by Mitochondrial Complex I Reverse Electron Transport. 20–24 (2019).
37. Elling, F. J. *et al.* Respiratory quinones in Archaea: Phylogenetic distribution and application as biomarkers in the marine environment. *Environ. Microbiol.* **18**, 692–707 (2016).
38. Zhi, X. Y. *et al.* The futasine pathway played an important role in menaquinone biosynthesis during early prokaryote evolution. *Genome Biol. Evol.* **6**, 149–160 (2014).
39. Schoepp-Cothenet, B. *et al.* Menaquinone as pool quinone in a purple bacterium. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 8549–8554 (2009).
40. Schoepp-Cothenet, B. *et al.* On the universal core of bioenergetics. *Biochim. Biophys. Acta - Bioenerg.* **1827**, 79–93 (2013).
41. Baymann, F., Schoepp-Cothenet, B., Lebrun, E., Van Lis, R. & Nitschke, W.

- Phylogeny of Rieske/cytb complexes with a special focus on the haloarchaeal enzymes. *Genome Biol. Evol.* **4**, 720–729 (2012).
42. Denke, E. *et al.* Alteration of the midpoint potential and catalytic activity of the Rieske iron-sulfur protein by changes of amino acids forming hydrogen bonds to the iron-sulfur cluster. *J. Biol. Chem.* **273**, 9085–9093 (1998).
43. Bergdoll, L., Ten Brink, F., Nitschke, W., Picot, D. & Baymann, F. From low- to high-potential bioenergetic chains: Thermodynamic constraints of Q-cycle function. *Biochim. Biophys. Acta - Bioenerg.* **1857**, 1569–1579 (2016).
44. Sarewicz, M. *et al.* Catalytic reactions and energy conservation in the cytochrome bc1 and b6f complexes of energy-transducing membranes. *Chem. Rev.* **121**, 2020–2108 (2021).
45. Ten Brink, F., Schoepp-Cothenet, B., Van Lis, R., Nitschke, W. & Baymann, F. Multiple Rieske/cytb complexes in a single organism. *Biochim. Biophys. Acta - Bioenerg.* **1827**, 1392–1406 (2013).
46. Oren, A. A plea for linguistic accuracy – also for Candidatus taxa. *Int. J. Syst. Evol. Microbiol.* **67**, 1085–1094 (2017).
47. Spang, A. *et al.* The genome of the ammonia-oxidizing candidatus nitrososphaera gargensis: Insights into metabolic versatility and environmental adaptations. *Environ. Microbiol.* **14**, 3122–3145 (2012).
48. Elbehti, A., Brasseur, G. & Lemesle-Meunier, D. First evidence for existence of an uphill electron transfer through the bc1 and NADH-Q oxidoreductase complexes of the acidophilic obligate chemolithotrophic ferrous ion-oxidizing bacterium *Thiobacillus ferrooxidans*. *J. Bacteriol.* **182**, 3602–3606 (2000).
49. Castelle, C. J. *et al.* The aerobic respiratory chain of the acidophilic archaeon *Ferroplasma acidiphilum*: A membrane-bound complex oxidizing ferrous iron. *Biochim. Biophys. Acta - Bioenerg.* **1847**, 717–728 (2015).
50. King, M. S., Sharpley, M. S. & Hirst, J. Reduction of hydrophilic ubiquinones by the flavin in mitochondrial NADH:ubiquinone oxidoreductase (complex I) and production of reactive oxygen species. *Biochemistry* **48**, 2053–2062 (2009).
51. Schulte, M. *et al.* A mechanism to prevent production of reactive oxygen species by
-

- Escherichia coli respiratory complex I. *Nat. Commun.* (2019) doi:10.1038/s41467-019-10429-0.
52. Esterházy, D., King, M. S., Yakovlev, G. & Hirst, J. Production of Reactive Oxygen Species by Complex I (NADH:Ubiquinone Oxidoreductase) from Escherichia coli and Comparison to the Enzyme from Mitochondria. *Biochemistry* 3964–3971 (2008) doi:10.1021/bi702243b.
53. Friedrich, T. & Scheide, D. The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *Fed. Eur. Biochem. Soc.* (2000).
54. Könneke, M. *et al.* Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 8239–44 (2014).
55. Otte, J., Mall, A., Schubert, D. M., Könneke, M. & Berg, I. A. Malonic Semialdehyde Reductase from the Archaeon Nitrosopumilus maritimus Is Involved in the Autotrophic 3-Hydroxypropionate/4-Hydroxybutyrate Cycle. (2015) doi:10.1128/AEM.03390-14.
56. Spaans, S. K., Weusthuis, R. A., van der Oost, J. & Kengen, S. W. M. NADPH-generating systems in bacteria and archaea. *Front. Microbiol.* **6**, 742 (2015).
57. Quasem, I. *et al.* Peculiar citric acid cycle of hydrothermal vent chemolithoautotroph Hydrogenovibrio crunogenus, and insights into carbon metabolism by obligate autotrophs. *FEMS Microbiol. Lett.* **364**, 1279–82 (2017).
58. Pelve, E. A. *et al.* Cdv-based cell division and cell cycle organization in the thaumarchaeon Nitrosopumilus maritimus. *Mol. Microbiol.* **82**, 555–566 (2011).
59. Lindås, A. C., Karlsson, E. A., Lindgren, M. T., Ettema, T. J. G. & Bernander, R. A unique cell division machinery in the Archaea. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18942–18946 (2008).
60. Risa, G. T. *et al.* The proteasome controls ESCRT-III – mediated cell division in an archaeon. **2532**, (2020).
61. Schmitt, E. *et al.* Recent Advances in Archaeal Translation Initiation. *Front. Microbiol.* **11**, (2020).
-

62. Kozłowski, 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.* 1–10 (2016) doi:10.1038/ismej.2016.2.
63. Sauder, L. A., Ross, A. A. & Neufeld, J. D. Nitric oxide scavengers differentially inhibit ammonia oxidation in ammonia-oxidizing archaea and bacteria. *FEMS Microbiol. Lett.* **363**, fnw052 (2016).
64. Martens-Habben, W. *et al.* The production of nitric oxide by marine ammonia-oxidizing archaea and inhibition of archaeal ammonia oxidation by a nitric oxide scavenger. *Environ. Microbiol.* **17**, 2261–2274 (2015).
65. Abby, S. S. *et al.* Candidatus Nitrosocaldus cavascurensis, an ammonia oxidizing, extremely thermophilic archaeon with a highly mobile genome. *Front. Microbiol.* **9**, 1–19 (2018).
66. Luo, Z.-H. *et al.* Genomic Insights of “Candidatus Nitrosocaldaceae” Based on Nine New Metagenome-Assembled Genomes, Including “Candidatus Nitrosothermus” Gen Nov. and Two New Species of “Candidatus Nitrosocaldus”. *Front. Microbiol.* **11**, (2021).
67. Daebeler, A. *et al.* Cultivation and genomic analysis of ‘Candidatus Nitrosocaldus islandicus,’ an obligately thermophilic, ammonia-oxidizing thaumarchaeon from a hot spring biofilm in Graendalur valley, Iceland. *Front. Microbiol.* **9**, 1–16 (2018).
68. Cho, C. M. H. *et al.* Transcriptome of a Nitrosomonas europaea mutant with a disrupted nitrite reductase gene (nirK). *Appl. Environ. Microbiol.* **72**, 4450–4454 (2006).
69. Beaumont, H. J. E. *et al.* Nitrite reductase of Nitrosomonas europaea is not essential for production of gaseous nitrogen oxides and confers tolerance to nitrite. *J. Bacteriol.* **184**, 2557–2560 (2002).
70. Beaumont, H. J. E., Lens, S. I., Westerhoff, H. V. & Van Spanning, R. J. M. Novel nirK cluster genes in Nitrosomonas europaea are required for NirK-dependent tolerance to nitrite. *J. Bacteriol.* **187**, 6849–6851 (2005).

