Exploring the Role of Copper in Novel Ammonia-Oxidising Archaeal Isolates and using Dynamic Structural Biology Tools to Characterise Copper Monooxygenases

Aidan Meekings

A thesis submitted for the degree of Doctor of Philosophy in Biological Sciences

School of Life Sciences

University of Essex

October 2025

Abstract

The oxidation of ammonium by ammonia oxidising archaea (AOA) is the first step of nitrification and is critical to the global nitrogen cycle. However, little is known about AOA physiology, especially their responses to bioavailable copper and copper enzyme mechanisms. Whilst AOA genome data is sparse, it does suggest that copper proteins could play a central role in nitrogen transformation pathways, and the biology of copper in soil and marine systems is central to this thesis. This thesis successfully enriched two soil AOA (C2, C4-2) relating to *Nitrosocosmicus* using a combination of liquid and solid cultivation methods. CARD-FISH confirmed AOA were predominant in both cultures. Physiological characterisation showed C4-2 was mesophilic, neutrophilic with optimum growth between 1-50 mM NH_4^+, 0-20 mM NO_2^- (Chapter 2) and bioavailable Cu(II) as low as (10^{-16} M) (Chapter 4). Amplicon and metagenomic sequencing (Chapter 3) showed that C2 and C4-2 were related to *N. Franklandus*, and *N. hydrocola* respectively. Metagenomic-assembled genomes (MAGs) identified key proteins involved in ammonia transport, carbon fixation and oxidative stress defence (Chapter 3). Several multicopper oxidases (MCOs) found in marine (*Nitrosopumilus maritimus* SCM1) and terrestrial AOA (*Nitrosocosmicus franklandus* C13), were also identified and selected for recombinant protein over-expression trials, however these proved unsuccessful (Chapter 5). Although not directly linked to the nitrogen cycle, soil dwelling copper containing two-domain lytic polysaccharide monooxygenases (LPMOs) from *Streptomyces lividans* were studied.

The LPMO domain carries out oxidative chemistry creating chain breaks in polysaccharides and the carbohydrate binding domain facilitates the polysaccharide binding. Active site copper coordination geometries were studied using EPR spectroscopy and SEC-SAXS has been used to investigate the spatial arrangement of the two domains in solution. Finally, in Chapter 7 a LPMO domain was studied using serial femtosecond X-ray

crystallography, and experiments to initiate molecular oxygen activation using an oxygen releasing photocage carried out .

Acknowledgments

I am deeply grateful for my mother, her belief in me and financial support has made it possible for me to focus on my studies. I would not be in the fortunate position I am now and this thesis would not have been possible without her. My Mother has been a role model to me ever since a young age, she did not only tell me the value of working hard, she led by example excelling in her own studies and career, inspiring me in the process. She has also taught me how to be resilient during tough times, not only during writing this thesis but every aspect of life. I feel incredibly lucky to have a mother who has instilled such great values and thanks to her I hold the same values today. I am very grateful to my sister, who has also led by example excelling with her studies, I thank her for her faith in me, eagerness to help me and her unwavering encouragement.



I would like to express my sincere gratitude to my supervisors, Professor Jonathan Worrall, Professor Corinne Whitby and Professor Mike Hough for their guidance during my

PhD journey. I would like to give a special thanks to both Jonathan and Corinne for their kindness and calm composure during both setbacks and successes. I am especially grateful to Jonathan for introducing to me the world of metalloproteins and guiding me through beamline trips. Despite my interests leaning towards structural biology at the start of my PhD, the microbiology aspect Corinne provided to the project was equally as cool and interesting for me, for which I am very grateful. A special thanks to Mike who initially learned of my interest of perusing a PhD and encouraged me to apply. I would also like to give a massive thanks to Dr Marcus Edwards and Dr Dimitri Svistunenko for their assistance in SEC-SAXS and EPR techniques. I would like to thank the I24 beamline team, Dr Robin Owen, Dr Danny Axford and Dr Sofia Jaho for helping at beamline trips, as well as the beamline team at I04 and BL21. A massive thank you to Dr Hiroshi Sugimoto and Professor Takehiko Tosha for their assistance at the XFEL beamline at SACLA, Japan. I would also like to give enormous gratitude to my lab friends in 5.08, Dr Marina Lučić and Lewis Williams for being such great people to work alongside. A massive thank you to Marina for teaching me so many techniques in the lab, my research journey would have been very different without her friendship and guidance. A big thanks to Lewis for always being up for a beer and a pub crawl to blow off steam.

Statement of Contributions

We thank collaborators and technical staff for providing valuable assistance during this work as follows: Mr John Green, (University of Essex) who provided technical support for the autoanalyser (NH_4^+ and NO_2^- concentration) and ICP-MS (Cu concentration) measurements. Dr Dimitri Svistunenko (University of Essex) who performed the EPR experiments and analysis of spectral simulations. Dr Amy Thompson (Diamond Light Source) who assisted with the processing of the serial crystallography data and Dr Fabien Cholet (University of Glasgow) who generated the phylogenetic trees presented in this thesis. Dr Marcus Edwards for his assistance with the size-exclusion small-angle X-ray scattering data analysis. Ms Keri-Anne Moss, Dr Negin Behboodi and Dr Philippe Laissue for their technical assistance with the CARD-FISH and microscopy. Finally, Mr Harry Lang and Mr Farid Benyahia for their assistance in maintaining the microbial cultures.

List of Abbreviations

AA - Auxiliary Activity

AMT - Ammonium Transporter Protein

ANI - Average Nucleotide Identity

ANOVA - Analysis of Variance

AOM - Ammonia oxidising microorganisms

AOA - Ammonia oxidising archaea

AOB - Ammonia oxidising bacteria

ATP - Adenosine Triphosphate

BUSCO - Benchmarking Universal Single-Copy Orthologs

Card-FISH - Catalysed Reporter Deposition–Fluorescence In Situ Hybridisation

CAZymes - Carbohydrate Active Enzymes

CBM - Carbohydrate-Binding Module

CBD - Catalytic-Binding Domain

CDS - Coding Sequence

DEAE - Diethylaminoethyl

DFAST - DDBJ Fast Annotation and Submission Tool

DNRA - Dissimilatory Nitrate Reduction to Ammonium

DOPE-FISH - Double-Labelled Oligonucleotide Probes–Fluorescence *In Situ* Hybridisation

DNA - Deoxyribonucleic Acid.

dNTP - Deoxyribonucleotide Triphosphate

$\mathbf{D_{max}}$ - Maximum Dimension

DTT - Dithiothreitol

EDTA - Ethylenediaminetetraacetic Acid

EPR - Electron Paramagnetic Resonance

EPS - Extracellular Polymeric Substances

FWM - Fresh Water Medium

GAJOE - Genetic Algorithm Judging Optimisation of Ensembles

GG - Gellan Gum

GH - Glycoside Hydrolase

GTDB - Genome Taxonomy Database

HAA - Hydrogen Atom Abstraction

HAO - Hydroxylamine Oxidoreductase

HMM - Hidden Markov Model

HPBC - (l-peroxo)(l-hydroxo)bis[bis(bipyridyl)cobalt(III)]

HRP - Horseradish Peroxidase

ICP-MS - Inductively Coupled Plasma - Mass Spectrometry

IEX - Ion-Exchange Chromatography

IMAC - Immobilised Metal Affinity Chromatography

IPTG - Isopropyl β-D-1-Thiogalactopyranoside

LPMO - Lytic Polysaccharide Monooxygenase

MAG - Metagenome-Assembled Genomes

MBP - Maltose-Binding Protein

MCO - Multicopper Oxidase

MES - 2-(N-morpholino)ethanesulfonic Acid

MiGA - Microbial Genome Atlas

MW - Molecular Weight

NirK - Nitrite Reductase

PCR - Polymerase Chain Reaction

PDB - Protein Data Bank

Pfam - Protein Family Databse

pMMO - Particulate Methane Monooxygenase

PTM - Post-Translational Modifications

qPCR - Quantitative Polymerase Chain Reaction

LB - Luria-Bertani Broth

SEC - Size-Exclusion Chromatography

SEC-SAXS - Size-Exclusion Chromatography - Small Angle X-Ray Scattering

SEM - Scanning Electron Microscope

SDS-PAGE - Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

SD - Standard Deviation

SFX - Serial Femtosecond Crystallography

SSX - Serial Synchrotron Crystallography

TBE - Tris-Borate-EDTA Buffer

TETA - Triethylenetetramine

TRIS - Tris(hydroxymethyl)aminomethane

tRNA - Transfer RNA

UV - Ultraviolet Light

Ranch - RANdom Chain

Rg - Radius of Gyration

RMSD - Root Mean Square Deviation

RNA - RiboNucleic Acid

ROS - Reactive Oxygen Species

rRNA - Ribosomal Ribonucleic Acid

RT - Room Temperature

XFEL - X-ray Free-Electron Laser

Table of Contents

Abstract 2

Acknowledgments 3

Statement of Contributions 5

List of Abbreviations 6

Table of Contents 9

Chapter 1 Introduction 16

1.1 Importance of Nitrogen 16

1.2 The Nitrogen Cycle 18

1.3 Introduction to Ammonia Oxidising Archaea (AOA) 20

1.4 Phylogenetics and Ecological Distribution 22

1.5 Metabolism and Biochemistry of AOA 25

1.6 Copper Dependence in Marine and Soil AOA 27

1.7 Multicopper Oxidases and Nitrite Reductases in AOA 28

1.8 Aims and Objectives 35

Chapter 2 Isolation, purification and physiological characterisation of AOA from agricultural soils. 36

2.1 Introduction 36

2.1.1 Aims 38

2.2 Methods 39

2.2.1 Agricultural soil samples 39

2.2.2 Enrichment cultures and isolation 40

2.2.3 CARD-FISH and microscopic analysis of AOA enrichments 44

2.2.3.1 Cell fixation and hybridisation 44

2.2.3.2 Microscopy 45

2.2.4 Physiological Characterisation of enrichment C4-2 45

2.2.5 Statistical analysis 46

2.3 Results 47

2.3.1 Abundance of AOA in agricultural soils 47

2.3.2 Fluorescence Microscopy (CARD-FISH) 49

2.3.3 Physiological Characterisation of enrichment C4-2 52

2.3.3.1 Effect of ammonia concentration on growth 52

2.3.3.2 Growth inhibition of enrichment C4-2 to increasing nitrite concentration 55

2.3.3.3 Effect of temperature on growth of enrichment C4-2 57

2.3.3.4 Effect of pH on growth of enrichment C4-2 58

2.4 Discussion 59

Chapter 3 Amplicon and Genome Sequencing of Putative Ammonia-Oxidizing Archaeal Isolates. 69

3.1 Introduction 69

3.1.1 Aims 70

3.2 Methods 71

3.2.1 Amplicon sequencing and phylogenetic analysis of AOA *amoA* and 16S rRNA genes 71

3.2.2 Whole Genome Sequencing (Novogene) of AOA isolates 71

3.2.3 Metagenome-Assembled Genome Construction of AOA 73

3.2.4 Screening of Functional Genes for Metabolism and Stress Response 74

3.3 Results 75

3.3.1 PCR Amplification of 16S rRNA and AOA *amoA* genes 75

3.3.2 Assembly and Annotation of AOA Metagenome-Assembled Genome 79

3.3.3 Screening of metabolic and adaptive proteins in AOA MAGs 83

3.3.4 Phylogenetic Analysis of AOA Isolates based on 16S rRNA and amoA Sequences obtained *via* Genome Sequencing 86

3.4 Discussion 90

Chapter 4 Ammonia-Oxidizing Archaea Growth Response to Copper 97

4.1 Introduction 97

4.1.1 Aims 99

4.2 Methods 99

4.2.1 Metal Speciation Analysis 99

4.2.2 Effect of copper on AOA growth. 99

4.2.3 qPCR analysis of *amoA* genes 101

4.2.4 Statistical Analysis 101

4.3 Results 102

4.3.1 Copper bioavailability effects on growth 102

4.3.2 Growth of *N. franklandus* (C13) at varying copper concentrations 103

4.3.3 Growth Response of AOA isolate C4-2 at varying copper concentrations 104

4.3.4 Changes in AOA and AOB amoA gene abundance during growth at varying copper concentrations 106

4.4 Discussion 108

Chapter 5 Molecular biology and over-expression of putative Archaeal multicopper oxidases 113

5.1 Introduction 113

5.1.1 Aims 117

5.2 Methods 117

5.2.1 Expression of Recombinant Proteins 117

5.2.2 Purification from pET28a constructs 118

5.2.3 Protein Purification using Ammonium Sulfate "Cut" precipitation 119

5.2.4 Extraction and Refolding of Inclusion Bodies 119

5.2.5 SDS-PAGE and Western Blot 120

5.3 Results 122

5.3.1 Target Protein Characteristics 122

5.3.2 Protein expression and purification of targets 123

5.3.3 Western Blot Analysis 129

5.3.4 Inclusion body extraction of Nmar_1131 131

5.4 Discussion 132

Chapter 6: Spectroscopic and structural characterization of two-domain LPMOs 136

6.1.1 Polysaccharide Degradation for Industrial Biotechnology 136

6.1.2 Lytic Polysaccharide Monooxygenases 138

6.1.3 Carbohydrate-Binding Modules in Polysaccharide Degradation 143

6.1.4 Structural Dynamics of Linkers in Two-domain LPMOs 146

6.1.4.1 Aims 148

6.2 Methods 148

6.2.1 LPMO Cloning and Site-directed Mutagenesis 148

6.2.2 Recombinant Overexpression and Purification 149

6.2.3 Concentration determination and preparation of copper solutions. 150

6.2.4 Fluorescence Spectroscopy 151

6.2.5 Substrate Binding Assay 151

6.2.6 Electron Paramagnetic Resonance Spectroscopy 151

6.2.7 Simulated to experimental spectra fitting methodology 153

6.2.8 Crystallization and X-ray Crystallography of Copper Binding Domain 153

6.2.9 Size-Exclusion Small Angle X-ray Scattering (SEC-SAXS) 156

6.2.10 Processing and Analysis of SEC-SAXS Data 156

6.2.11 Ensemble Optimisation Methods and AlphaFold Model Alignment to Experimental SAXS Envelope 157

6.3 Results 158

6.3.1 Protein purification 158

6.3.2 Tryptophan quenching of LPMOs upon Cu(II) binding 160

6.3.3 Chitin and Cellulose Binding Assays 162

6.3.4 EPR spectroscopy on the two-domain LPMOs 163

6.3.5 X-ray Crystal Structure of the LPMO Domain of *Sli*LPMO10G 169

6.3.6 SEC-SAXS of LPMOs 173

6.3.7 Scattering Profiles of LPMOs 175

6.3.8 Guinier Analysis of SEC-SAXS Data 176

6.3.9 Kratky Profiles Analysis of LPMOs 178

6.3.10 Pair-Distribution Analysis of SEC-SAXS Data 179

6.3.11 *Ab Initio* Reconstruction of SEC-SAXS Envelopes 181

6.3.12 Ensemble Optimization Method Fitting of SEC–SAXS Data 183

6.3.13 Comparison of Experimental SEC-SAXS and EOM-Derived Rg and D_{max} 185

6.3.14 Alignment of LPMO Models to SEC-SAXS Envelopes 185

6.3.15 Multiple Sequence Alignment of LPMO Linkers 187

6.4 Discussion 188

Chapter 7 Serial X-ray crystallography of *Sli*LPMO10E: Towards Time-Resolved Studies with a Molecular Oxygen Releasing Photocage 193

7.1 Introduction 193

7.1.1 Aims 195

7.2 Methods 195

7.2.1 *Sli*LPMO10E Cloning and Site-directed Mutagenesis 195

7.2.2 Recombinant Overexpression and Purification of *Sli*LPMO10E 195

7.2.3 Concentration determination of *Sli*LPMO10E and Copper Solution Preparation ... 196

7.2.4 Crystallization and Cryo-X-ray Crystallography of *Sli*LPMO10E 197

7.2.5 Crystallisation of *Sli*LPMO10E microCrystals 198

7.2.6 Spectroscopic Validation of HPBC on Bovine Myoglobin 199

7.2.7 Sample Preparation and Crystallographic Data Measurements at SACLA XFEL .. 199

7.2.8 Data processing at XFEL 201

7.3 Results 202

7.3.1 Microscopy of *Sli*LPMO10E Micro-Crystals 202

7.3.2 Cryo and Room Temperature X-ray Structures of *Sli*LPMO10E 203

7.3.3 Spectroscopic Validation of HPBC on Bovine Myoglobin 207

7.3.4 Active Site Copper Coordination under Reduced States and HPBC Soaks 208

7.4 Discussion 209

Chapter 8 Concluding Remarks 211

References 216

Appendix A List of Tables 301

Appendix B List of Figures 311

Chapter 1 Introduction

1.1 Importance of Nitrogen

Nitrogen is fundamental for sustaining life, as it is a key component of amino acids, nucleic acids (DNA/RNA), proteins and chlorophyll (Wang, *et al*., 2024; Zayed, *et al*., 2023; Anas, *et al*., 2020). Nitrogen gas (N_2) constitutes ca.78% of the Earth's atmosphere and is the fourth most abundant element in cellular biomass. However, atmospheric N_2 is inert and interchangeability of inert N_2 gas into reactive N_2 is essential for cellular growth and metabolism, and this process is entirely carried out by microbial activity (*Pi et al*., 2022). Bioavailable N_2 enters organism systems through N deposition and/or fixation, performed by microbes and plants (Hoang, *et al*., 2022; Hutchins and Capone, 2022; Ladha, *et al*., 2022; Schlesinger and Bernhardt, 2020). These processes convert inert N_2 gas to reactive forms of bioavailable N species (*e.g*. NH_4^+, NO_2^- or NO_3^-). The deposition process is the direct input of atmospheric ammonia NH_3, nitrite NO_2^-, nitric oxide NO and nitrate NO_3^- into organisms (Tang, *et al*., 2023; Hutchins and Capone, 2022). Because of the reliance of reactive/bioavailable N, N frequently controls primary productivity of ecosystems (Threatt and Rees, 2023). Numerous studies show how crop yields and plant biomass are often constrained by bioavailability of reactive N in the biosphere (Feng, *et al*., 2023; Yokamo, *et al*., 2023; Guo, *et al*., 2022).

Global nitrogen fixation has contributed to the production of ca.240 Tg of reactive nitrogen species (N_r) to marine and terrestrial ecosystems annually, of which, most was as a result of anthropogenic inputs (Galloway and Cowling, 2021). The majority of N_r emissions derives from nitrogen exploitation for intensive farming techniques to increase crop yield as a result of the growing human population. Previous estimations conclude that potentially half of the world's population depend on crops that have been treated using nitrogen containing fertilizer (Beaudor, *et al*., 2023). Additionally, nitrogen can be obtained naturally through the

Earth's atmosphere *via* nitrogen fixing bacteria. Despite this, reliance on natural nitrogen fixation alone would not be sufficient to meet crop demands (Ladha, *et al*., 2022). To overcome the increase in demand, non-natural processes of nitrogen fixation would need to be implemented, such as fixing nitrogen industrially through the Haber-Bosch process (Galloway and Cowling, 2021). Inert atmospheric nitrogen is also fixed unintentionally through daily human activities, including the burning of fossil fuels by combustion engines, and the production of electricity and industrial activities (Erisman. *et al*., 2013). The overall magnitude of the global nitrogen cycle has since doubled over the last century due to both anthropogenic (210 Tg N yr^{-1}) and natural fixed nitrogen sources (203 Tg N yr^{-1}) (Melillo, 2021; Fowler, *et al*., 2013).

Alterations in the supply of nitrogen will significantly influence certain aspects of life on Earth, as nitrogen is a major nutrient with higher availability leading to greater contribution of radiative forces on regional and global climate (Gong, *et al*., 2024). Compounds derived from nitrogen, such as nitrogen oxides (NO_x) act as precursors to the tropospheric ozone and promote the emission of particulate materials into the atmosphere, thus degrading air quality and increasing human mortality (Nguyen, *et al*., 2022; Monks, *et al*., 2015). Anthropogenic influence has undoubtedly helped to enhance crop yields through production of synthetic N fertilisers, thus allowing an unprecedented growth of the world population (Galloway *et al*., 2021). Despite this, Sanjuan *et al*., (2020) revealed that 80% of N fertiliser applied to crops does not reach human mouths. Excess N is lost to the environment, from animal waste, water run-off from field and gas emission (Sanjuan *et al*., 2020). This excess use of N fertiliser is particularly high in developed countries, which has massively perturbed the biogeochemical N cycle. Accumulation of N in the biosphere threatens biodiversity and human health, through eutrophication and increased emissions of potent greenhouse gas N_2O (de Vries, 2021; Sanjuan, *et al*., 2020).

1.2 The Nitrogen Cycle

The nitrogen cycle is one of the most important nutrient cycles on the planet, comprising of four microbiological processes: nitrogen fixation, mineralisation/decomposition, nitrification, and denitrification (Mosley, *et al*., 2022; Stein and Klotz, 2016). Nitrogen fixation is the conversion of atmospheric N_2 into bioavailable N_2, such as NH_4^+ (Stein and Klotz, 2016). Mineralisation is the conversion of organic N_2 to NH_4^+ during decomposition of organic matter (Geisseler, *et al*., 2021). Nitrification is the oxidation of NH_4^+ to NO_2^- and then NO_3^-, followed by denitrification, which is the reduction of NO_3^- to NO and N_2O, ultimately returning to atmospheric N_2 (Prosser and Nicol, 2012).

Firstly, inert atmospheric N_2 is fixed to bioavailable NH_4^+, carried forward by di-nitrogenase reductase (NifH), this process can occur in both anoxic and oxic environments (Rutledge and Tezcan, 2020). The first step in nitrification; $NH_3 \rightarrow NH_2OH$ is catalysed by ammonia monooxygenase (AMO), consisting of canonical subunits amoA/amoB/amoC (Wright and Lehtovirta-Morley, 2023), NH_2OH is further oxidised into NO_2^- by hydroxylamine oxidoreductase (HAO), commonly found in ammonia oxidising bacteria (AOB) (Soler-Jofra, *et al*., 2021; Norton, *et al*., 2002). The NO_2^- is then further oxidised into NO_3^-, carried out by nitrite oxidoreductase (NXR) (Chicano, *et al*., 2021). Alternatively, nitrification can be carried out through the comammox pathway, which involves the complete oxidation of NH_4^+ into NO_3^- by a single organism, such as *Nitrospira* AOB (**Figure 1.1**). These microbes possess all three AMO, HAO and NXR genes to fully carry out nitrification in a single cell (Daims, *et al*., 2015), rather than relying on other NH_4^+ and NO_2^- oxidising microorganisms. Denitrification is an anaerobic pathway, in which NO_3^- is reduced to NO_2^- by nitrate reductase (NarA/NarG), followed by further reduction from NO_2^- to NO by either copper bound nitrite reductase (NirK) or heme containing nitrite reductase (NirS) (Zhang, *et al*., 2023). NO is then reduced to N_2O by cytochrome c-dependent nitric oxide reductase (cnorB) or quinol-dependent nitric oxide

reductase (qnorB) (Gopalasingam, 2019; Heylen, *et al*., 2007). Lastly, N_2O is converted back to atmospheric N_2, by nitrous oxide reductase (NosZ) (**Figure 1.1**) (Hiis, *et al*., 2024). Recently, the dirammox pathway was discovered (Qiu and Liu, 2025; Hu, *et al*., 2024b; Lv, *et al*., 2024), which is an aerobic reaction carried out by heterotrophic bacteria that also possess the ability to convert reactive N species into atmospheric N_2, (*i.e*. converting NH_4^+ back into atmospheric N_2) (Qiu and Liu, 2025). The dirammox pathway is carried out by a di-iron oxygenase, encoded by the *dnfA* gene for the oxidation of NH_2OH back to atmospheric N_2 (**Figure 1.1**) (Xu, *et al*., 2022). Alternatively, NO_3^- can undergo the six-electron reduction to NH_4^+ in the dissimilatory nitrate/nitrite reduction (DNRA) pathway, carried out by NrfA penta-haem cytochrome (Hird, *et al*., 2025). Other nitrite reductases NasAB, NarB can convert NO_3^- to NO_2^-, with NO_2^- being reduced back to NH_4^+ by NirA nitrite reductase (**Figure 1.1**) (Sanz-Luque, *et al*., 2015; Rubio, *et al*., 1996; Lin, *et al*., 1993). The anaerobic oxidation of ammonium (anammox) pathway results in the conversion of NO → N_2H_4 by hydrazine synthase, followed by the conversion of N_2H_4 → N_2 which is driven by hydrazine dehydrogenase (HZO) (Underwood, *et al*., 2022; Harhangi, *et al*., 2012) (**Figure 1.1**).

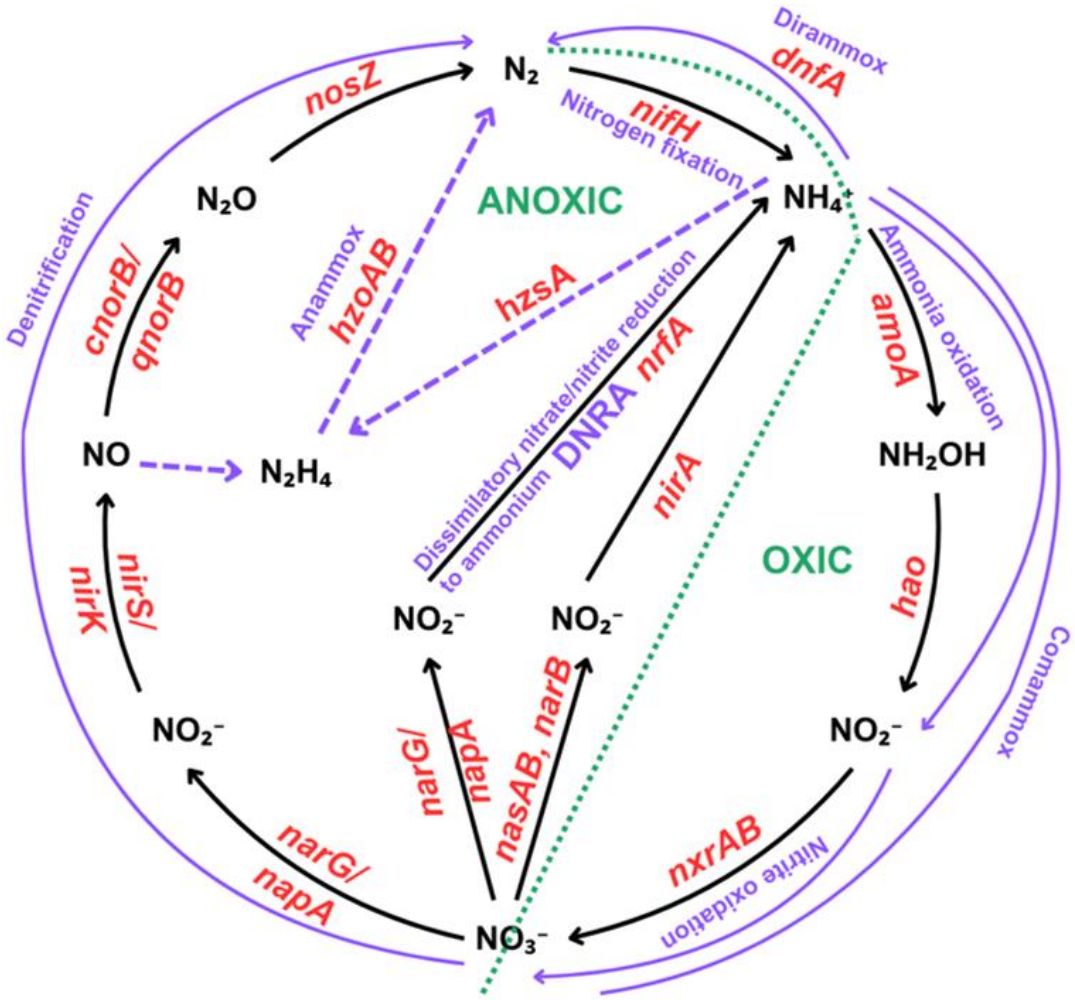


Figure 1.1. Schematic diagram of the global Nitrogen Cycle partitioned by oxic and anoxic domains (green dotted line). Black arrows represent direction of chemical conversion, purple pathways highlight major route groupings and red text indicates commonly used functional genes. (Modified from Underwood, *et al*., 2022).

1.3 Introduction to Ammonia Oxidising Archaea (AOA)

Archaea form one of the three domains of life and are distinguished from eukarya and bacteria by their cellular and molecular features (Woese, *et al*., 1990). At the cellular level, archaeal membranes lack bacterial peptidoglycan and typical consist of a protein rich S-layer (Rodrigues-Oliveira, *et al*., 2017), a feature commonly found in ammonia oxidising archaea (AOA) (Wright and Lehtovirta-Morley, 2023). Interestingly, the machineries archaea use to

express genetic information through DNA replication, transcription and translation, are evolutionarily more closely related to eukaryotes compared to bacteria (van Wolferen *et al*., 2022). However, archaeal lineages share common cellular hallmarks that distinguish them from eukaryotes and bacteria, these features include distinct membranes consisting of lipids containing isoprenoid chains that are ether-linked to a glycerol-1-phosphate backbone, unique motility apparatus called archaellum, and unique types of metabolism, such as methanogenesis (van Wolferen *et al*., 2022).

Originally discovered in extreme environments (high temperatures, salinity, acidity etc.), archaea are now known to be ubiquitous, with prominent members residing in oceanic and terrestrial environments (Moissl-Eichinger, *et al*., 2018), and suggests archaea play central roles in the global carbon and nitrogen cycle (Bayer, *et al*., 2023; Wright and Lehtovirta-Morley, 2023). Furthermore, AOA have emerged as key players in the global nitrogen cycle (Lehtovirta-Morley, 2023). AOA are among the most abundant living organisms on Earth and are responsible for driving the first, rate limiting step of nitrification, through the oxidation of NH_4^+ to NO_2^- (Jung, *et al*., 2022; Huang, *et al*., 2021; Alves, *et al*., 2018). Since their discovery ca. two decades ago with the isolation of the marine *Nitrosopumilus maritimus* (SCM1) (Walker, *et al*., 2010), studies have revealed the ubiquitous nature of AOA, where they have been found in virtually all NH_3 containing environments (Lehtovirta-Morley, 2023; Alves, *et al*., 2018). The high abundance of AOA and their activity, significantly influence NO_2^- bioavailability and N_2O emission, contributing to food security and climate (Wan, *et al*., 2024; Hu, *et al.*, 2022). Therefore, understanding the drivers of NH_3 oxidation, especially in AOA, is of high interest. Despite significant progress in understanding AOA physiology and cultivation, many knowledge gaps remain in understanding AOA, such as their adaptation to environmental changes and cellular metabolism. Successfully cultivating and isolating AOA species from environmental samples would allow researchers to discover novel AOA species and

characterise them to better understand their adaptability and cellular mechanisms (Pereira, *et al*., 2024; Lehtovirta-Morley, *et al*., 2024; Lehtovirta-Morley, 2023; Klein, *et al*., 2022). However, despite AOA being among the most abundant living organisms on Earth, they remain challenging to study as they are difficult to culture and isolate in the laboratory (Klein, *et al*., 2022; Prosser and Nicol, 2008).

1.4 Phylogenetics and Ecological Distribution

Ammonia oxidising microorganisms (AOM) are responsible for the initial and rate-limiting step of nitrification, which involves the sequential oxidation of ammonia (NH_3) to nitrite (NO_2^-) and nitrate (NO_3^-) (Rütting, T. *et al*., 2021; Prosser, *et al*., 2020; Gubry-Rangin, *et al*., 2018; Hink, *et al*., 2018). The oxidation of NH_4^+ is catalysed by the enzyme ammonia monooxygenase (AMO) which is encoded by the *amoA*, *amoB*, and *amoC* genes. Many microbial ecology studies have targeted the *amoA* gene for characterising AOMs in various environments (*e.g.* Ghimire-Kafle, *et al*., 2023; Kikuchi, *et al*., 2023; Baskaran, *et al*., 2020; Wright, *et al*., 2020; Huang, *et al*., 2018; Hink, *et al*., 2017).

Initially, AOA were placed in the archaeal phylum *Thaumarchaeota*, which has since been renamed *Nitrososphaerota*. (Rinke, *et al*., 2021, French, *et al*., 2012). In 2012, phylogenetic analysis of the archaeal *amoA* gene enabled the classification of AOA into five major clusters (Li, *et al*., 2025; Pester, *et al.,* 2012): the *Nitrosopumilus* cluster (formerly known as the marine or 1.1a AOA lineage) (DeLong, *et al*., 1998), the *Nitrososphaera* cluster (previously termed the soil or 1.1b AOA lineage) (DeLong, *et al*., 1998), *Nitrosotalea* (linked to the 1.1a lineage) (Lehtovirta-Morley *et al*. 2011), the *Nitrososphaera* sister cluster and *Nitrosocaldus* (formerly identified as ThAOA or HWCGIII lineage) (Li, *et al*., 2025; de la Torre, *et al*., 2008; Prosser and Nicol, 2008). Recent taxonomic framework of AOA organises AOA into four primary orders: *Candidatus Nitrosomirales*, *Candidatus Nitrosocaldales*,

Nitrosopumilales and *Nitrososphaerales* (Zheng, *et al*., 2024) (**Figure 1.2**). *Nitrososphaerales* are mainly associated with terrestrial systems such as soils and grassland, however, they can also be found in aquifer sediments (Li, *et al*., 2025). *Nitrosopumilales* are predominantly found in aquatic environments, such as marine and freshwater ecosystems (Diamond, *et al*., 2022). Other members such as *Ca*. *Nitrosomirales* demonstrate vast ecological versatility which are distributed across diverse habitats, including terrestrial, marine, geothermal environments, and groundwater environments (Luo, *et al*., 2024). Members of the *Ca*. *Nitrosocaldales* thrive in geothermal environments, exhibiting the ability to live in extreme conditions (Li, *et al*., 2025). The wide taxonomic range and ecological flexibility of AOA emphasise their exceptional adaptability, which enables AOA to thrive in varied habitats and drive the nitrogen cycle at the global scale (Li, *et al*., 2025; 2024).

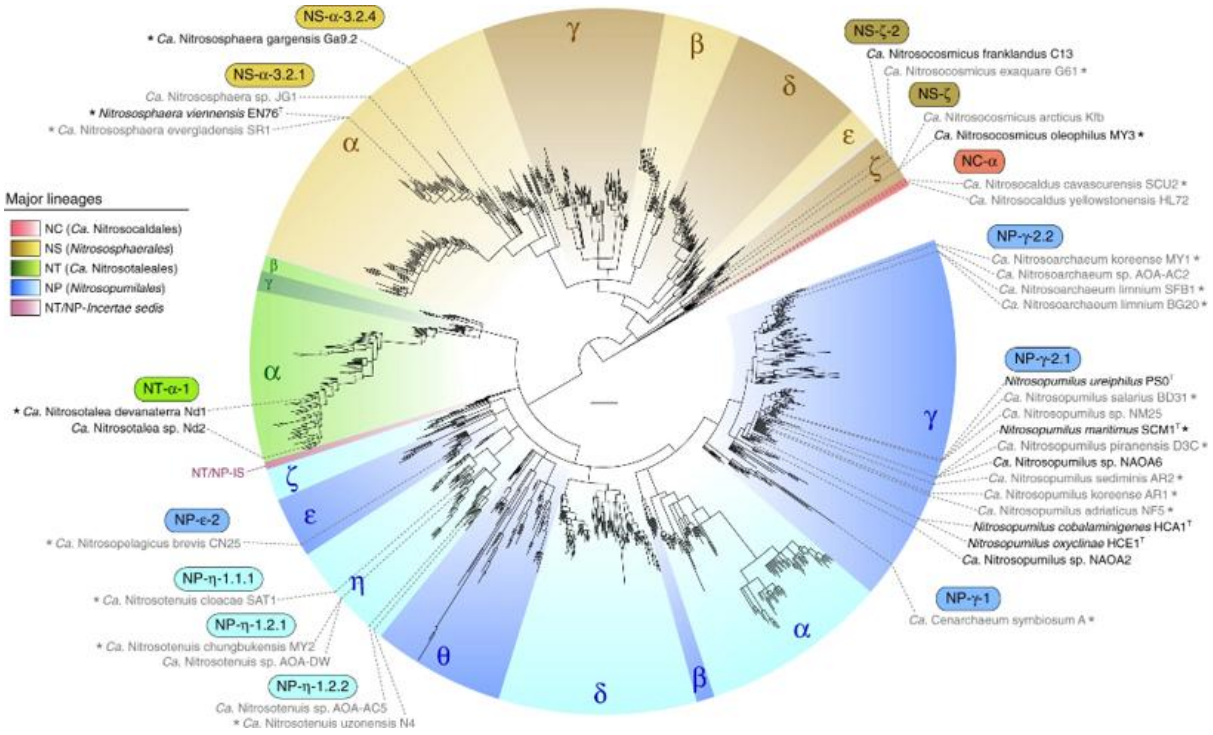


Figure 1.2. Global phylogeny of archaeal *amoA* genes from environmental and pure cultures. Order lineages are indicated with five different colours, with major subclades represented by Greek letters/shades. Black text indicates pure cultures, and grey text represents enrichment cultures. (Figure from Alves, *et al*., 2018).

In contrast to AOA, the AOB consist of fewer lineages: *Nitrosococcaceae* family such as *Nitrosococcus* within the Gammaproteobacteria (Kop, *et al*., 2025), *Nitrosomonadaceae* family such as, *Nitrosomonas* and *Nitrosospira* within the Betaproteobactera (Aigle, *et al*., 2019) and some of the members of the *Nitrospira* comammox bacteria (Ward, *et al*., 2021). Traditionally, AOB were once thought to be the dominant drivers of ammonia oxidation in soils (Prosser, *et al*., 2020; Nicol, *et al*., 2019). However, the discovery and cultivation of AOA, initially in marine environments (Könneke, *et al*., 2005) and later in soils, revealed that AOA also play a significant role in ammonia oxidation (Yamamoto, *et al*., 2010; Prosser, 2020; Banning, *et al*., 2015; Huang, *et al*., 2021). This is especially crucial for soils with low NH_4^+ availability, such

as low or moderately fertilised soils that provide favourable low nutrient conditions for AOA to thrive (Sun, *et al*., 2019).

1.5 Metabolism and Biochemistry of AOA

Most AOA possess at least two Amt transporters (*i.e*. low affinity AMT-1 or high affinity AMT-2) (Rasmussen and Francis, 2024; Lehtovirta-Morley, 2023); Liu, *et al*., 2021). The presence of both AMT-1 or AMT-2 results in greater uptake of NH_4^+ and are found in AOA species that thrive in low NH_4^+ concentrations (μM), such as *Nitrosopumilus maritimus* (SCM1), *Nitrosopumilus koreense* (MY2) *Nitrososphaera viennensis* (EN76) and *Nitrososphaera gargensis* (ga9.2). Contrastingly, AOA from the genera *Nitrosocosmicus*, exhibit growth at higher NH_4^+ concentrations (at the mM level), such as *N. franklandus* (C13), *N. arcticus* (Kfb), *N. oleophilus* (MY3), *N. hydrocola* (G61) and *N. agrestis* (SS) (Han, *et al*., 2024).

All characterised AOA encode a copper-dependent transmembrane enzyme, ammonia monooxygenase (AMO), which belongs to the CuMMO (copper-containing membrane monooxygenase), comprising of ammonia, methane and alkane monooxygenase (Hodgskiss, *et al*., 2023; Lehtovirta-Morley, 2023). Archaeal AMO shows high similarity with bacterial AMO and pMMO (particulate methane monooxygenase), thus it is assumed that archaeal AMO also requires two electrons for the oxidation of NH_3 to NH_2OH, with electrons provided by electron carriers (Lehtovirta-Morley, 2023). Subsequent oxidation steps in the ammonia oxidation pathway release electrons that are recycled back to the AMO, with remaining electrons entering the respiratory pathway in AOA. The reduction potential drives ATP synthesis and assisting in cellular function, such as carbon fixation (Wright and Lehtovirta-Morley, 2023; Könneke, *et al*., 2014).

AMO proteins are difficult to purify, thus predicted structures have been based on homology to characterise AMO proteins, such as pMMO from methanotrophs (Lawton, *et al*., 2014). Structural predictions of AMO reveal a heterotrimeric complex, consisting of three subunits (amoA, amoB and amoC) (Lehtovirta-Morley, 2023). Archaeal AMO shows high divergence compared to AMO that reside in ammonia oxidising bacteria (AOB), this makes subunits such as amoA excellent genetic markers for measuring abundance of AOA versus AOB, in the environment and in liquid culture (Lehtovirta-Morley, 2023). Previous studies also identified additional archaeal AMO subunits, (amoX, amoY and amoZ) (Hodgskiss, *et al*., 2023). The nature and location of the AMO active site remains poorly understood, with subunits amoB and amoC favouring an extracellular active site (**Figure 1.3**) (Lehtovirta-Morley, 2023).

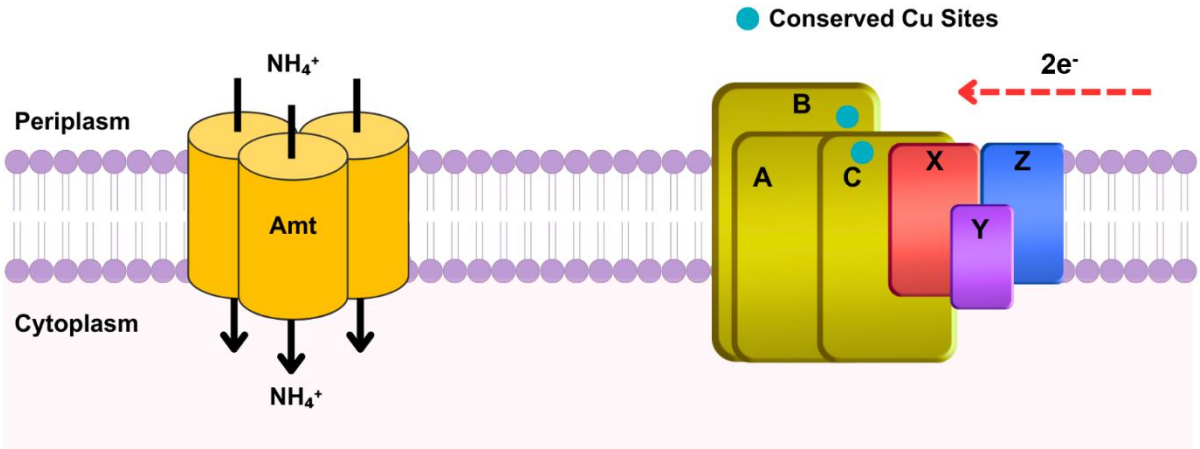


Figure 1.3. Energy dependent Amt transporter for NH_4^+ binding and uptake into cytoplasm. Yellow coloured amoA, amoB and amoC, reveal conserved trimer structure and copper sites, amoX, amoY and amoZ display putative archaeal AMO subunits. Figure adapted from (Wright and Lehtovirta-Morley, 2023).

1.6 Copper Dependence in Marine and Soil AOA

Electron transport in AOA is carried out by a rich network of copper-dependent proteins. Previous studies revealed AOA genomes that encode numerus blue copper proteins (plastocyanin-like electron carriers) and multicopper oxidases (MCOs) (Diamond, *et al*., 2022; Reyes, *et al*., 2020a; 2020b). Interestingly, AOA lack genes that encode canonical heme-rich hydroxylamine oxidoreductase (HAO) (Soler-Jofra, *et al*., 2021; Walker, *et al*., 2010). Despite AOA and AOB sharing a common pathway ($NH_4^+ \rightarrow NH_2OH$), the difference of heme-rich and copper dependent pathways in AOA and AOB make their biochemistry distinctly unique (Lehtovirta-Morley, 2023). To date, no copper-chelating siderophores (chalkophores) have been identified in the genomes of AOA, suggesting a reliance on genes encoding uptake systems (Lehtovirta-Morley, 2023). Reyes, *et al*., (2020a; 2020b) investigated copper limitation as well as transcriptomic analysis of copper protein in marine *N. maritimus* (SCM1) and terrestrial *N. viennensis* (EN76). This study revealed the upregulation of genes involved in copper binding and uptake, as well as slower growth of *N. maritimus* (SCM1) at $10^{-12.7}$ M of bioavailable Cu(II), compared to 10^{-16} M of bioavailable Cu(II) needed to reduce growth of *N. viennensis* (EN76). The ability of the soil dwelling *N. viennensis* (EN76) to tolerate lower levels of bioavailable Cu(II) compared with *N. maritimus* (SCM1) could be attributed to its larger copper proteome, copper transporters and regulatory genes (Reyes, *et al*., 2020b). For example, several copies of copper transporter proteins, such as CopC/D and CopA genes are found in *N. viennensis* (EN76) (Reyes, *et al*., 2020b). Further copper limitation studies, in conjunction with transcriptomic analysis, would provide further insights into how AOA species differ in terms of copper affinity and regulation of their copper proteome under copper stress. Obtaining pure cultures of AOA in liquid media would also be beneficial for studying copper limitation in AOA that reside in soil (Lehtovirta-Morley, *et al*., 2016; Lehtovirta-Morley, *et al*., 2011; Tourna, *et al*., 2011). This is especially important in soils where there is an increase in the

binding of metals to solid, organic, and inorganic matter (clays and metal oxides), as well as dissolved organic molecules (humic/fulvic acids) and dissolved inorganic materials (phosphate, nitrate and sulfate) (**Figure 1.4**), leading to lower levels of bioavailable Cu(II) (Reyes, *et al*., 2020a). This could also explain their high tolerance to such low bioavailable Cu(II) (10^{-16} M) and their higher abundance of copper binding/uptake genes, compared to marine AOA (Reyes, *et al*., 2020b).

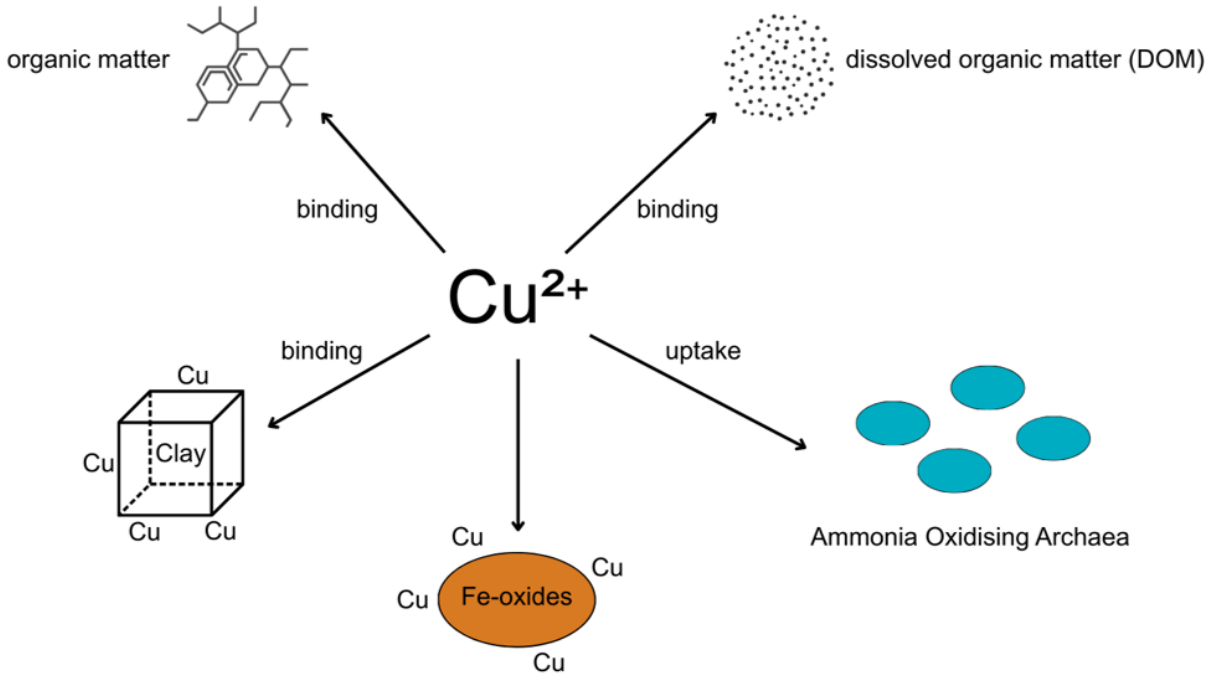


Figure 1.4. Schematic model of Cu(II) speciation in soil environments, showing competing metals, organic and inorganic matter. (Modified from Reyes *et al*., 2020).

1.7 Multicopper Oxidases and Nitrite Reductases in AOA

MCOs are a very diverse group of proteins that play an enzymatic role in coupling the oxidation of a wide range of substrates with the reduction of dioxygen into water. These enzymes can be found in many organisms with the common types being Fet3p, CueO, laccases, ceruloplasmin and ascorbate oxidase (Lawton and Rosenzweig, 2011). The function of MCOs consist of many different processes that are essential in for sustaining life inside its host, such

processes include ascorbate metabolism, phenolic substrate oxidation and metal homeostasis (Lawton and Rosenzweig, 2011). Sequence analysis of MCOs does not reveal their function and despite being well characterised, it is extremely difficult to assign individual MCOs to particular pathways (Takamura *et al*., 2021). On the other hand, it is possible to identify MCO proteins based on metal binding motifs that are highly conserved. Type 1 (T1Cu), type 2 (T2Cu) and type 3 (T3Cu) copper sites are classified according to their coordination geometry, spectroscopic properties and functional roles. T1Cu, often termed blue copper sites, binds a single Cu ion coordinated by two histidine's, one cysteine and another axial ligand such as methionine. Strong absorption at 610 nm can be seen in T1Cu centres from the cysteine to copper charge transfer, causing a blue colour (Jones and Solomon, 2015). On the other hand, T2Cu and T3Cu lack an intense blue absorption, with T2Cu sites containing one Cu ion coordinated by two histidine's, while the T3C site binds two closely spaced Cu ions with each Cu ion being coordinated by three histidine's (Jones and Solomon, 2015; Solomon, *et al*., 2014). Substrates, such as phenols or Cu(I) are oxidised at T1Cu and electrons are transferred to the T3Cu/T2Cu site trinuclear site for the four-electron reduction of molecular O_2 to H_2O ($O_2 + 4\ H^+ + 4\ e^- \rightarrow 2\ H_2O$) (Gräff, *et al*., 2020; Lawton, *et al*., 2009; Sakurai and Kataoka, 2007) *via* Cys-His bridge (Lawton and Rosenzweig, 2011).

AOA encode multiple MCOs that belong to distinct lineages, including MCO1 which are 2-domain (2dMCO) and are found exclusively in AOA genomes (Reyes, *et al*., 2020b). Contrastingly, the MCO4 is also a 2dMCO, but with an extra "dangler" cupredoxin domain (Voland, *et al*., 2025). Alternatively, MCOs can also exist as a monomer containing 3-domains (3dMCO), consisting of mononuclear T1Cu site and trinuclear T3/T2Cu site, with the same role of four-electron reduction of molecular oxygen O_2 to H_2O ($O_2 + 4\ H^+ + 4\ e^- \rightarrow 2\ H_2O$) (Jones and Solomon, 2015; Komori and Higuchi, 2015). T1Cu site contains highly conserved residues, with two His, Met and Cys, while T3/T2Cu site contain 8 highly conserved His

residues, coordinating three copper ions (Orlando, *et al*., 2024; Singha, *et al*., 2023). A recent study has reported the X-ray crystal structure of a 2dMCO from *N. maritimus,* Nmar_1354 of MCO4 lineage, which showed ability for NH_2OH oxidation and bears a unique structural feature termed the "dangler" cupredoxin domain (Voland, *et al*., 2025) (**Figure 1.4**). Importantly, Nmar_1354 demonstrated the formation of nitroxyl (HNO) from NH_2OH oxidation, with HNO chemically dimerising to potent greenhouse gas N_2O and providing the first plausible enzymatic route for N_2O formation in marine AOA (Voland *et al*., 2025). Soil AOA (*N. franklandus* (C13), *N. oleophilus* (MY3), *N. hydrocola* (G61)) all range from 6-8 MCOs (Han *et al*., 2024; Nicol *et al*., 2019; Sauder *et al*., 2017; Jung *et al*., 2016; Kerou *et al*., 2016), compared with the marine AOA *N. maritimus* (SCM1) which contains 6 in total (Walker *et al*., 2010).

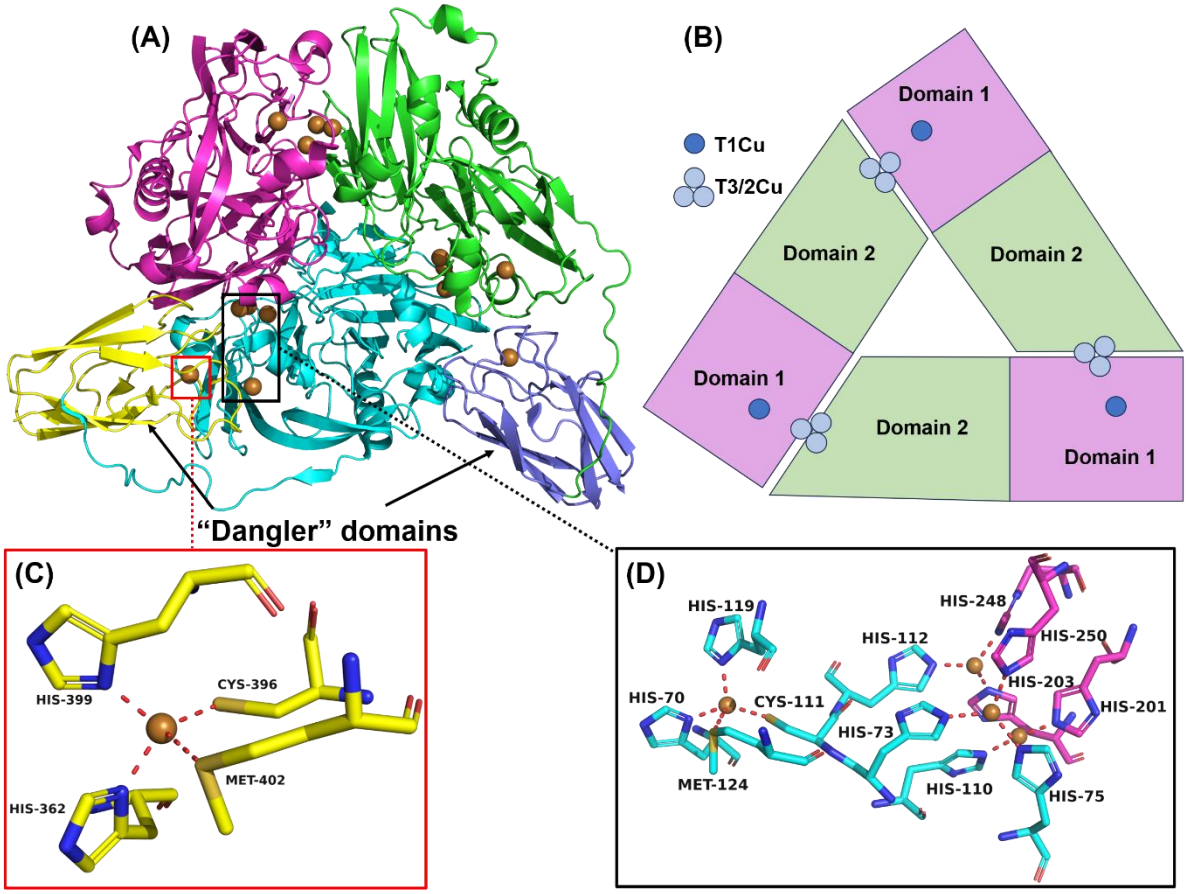


Figure 1.5. Nmar_1354 MCO (PDB 9CCK) (Voland *et al*., 2025). **(A)** Quaternary structure of Nmar_1354 revealing canonical homotrimeric 2dMCO, with additional cupredoxin domains, with each colour representing a monomeric subunit. **(B)** Copper site distribution in 2dMCO (Nmar_1354). **(C)** "Dangler" T1Cu cupredoxin domain, two His, Met and Cys coordinating residues. **(D)** Mononuclear T1Cu site coordinated by two His, a Met and Cys residue, together with a T3/T2Cu trinuclear copper site, containing three copper ions coordinated by eight histidine residues.

Archaeal Copper-containing nitrite reductase (CuNiR) are homologs of MCOs, with a 2dMCO arrangement. Copper containing nitrite reductases NirKs contain two copper sites per subunit, a mononuclear T1Cu electron transfer site and a mononuclear T2Cu copper catalytic site (Horrell, *et al*., 2017; Long, *et al*., 2015). The T1Cu site also shows highly conserved residues consisting of two His, a Met and Cys residue coordinating copper, with T2Cu site possessing four His coordinating residues (Eady and Samar Hasnain, 2022) (**Figure 1.6**). NirK

catalysis is initiated at the T1Cu site, where an electron is accepted by substrate, similarly to 2dMCO (**Figure 1.5**) electrons are relayed (ca.12-13 Å) by the His-Cys-His bridge to the T2Cu site. T2Cu site binds to NO_2^- which is reduced to NO and H_2O by a proton-coupled mechanism ($NO_2^- + e^- + 2\ H^+ \rightarrow NO + H_2O$) (Fukuda, *et al*., 2016; Lawton, *et al*., 2013) (**Figure 1.6**). Solving the structure and functionalities of NirKs derived from AOA will allow for better understanding of CuNirR role and functionality. Previous studies have identified NO as a central intermediate in the AOA NH_4^+ oxidation pathway (Sauder, *et al*., 2016). The same studies revealed that NirK are highly expressed during exponential growth for both marine and soil AOA, with NH_4^+ availability playing an essential role in tightly regulating *NirK* genes (Sauder, *et al*., 2016). Despite NirK showing wide distribution in marine and soil AOA, members of the order *Nitrosocaldales* from geothermal systems, reveal no NirK homologs, with high sensitivity to low concentration of NO-scavenging chemical PTIO (2-phenyl-4,4,5,5,-tetramethylimidazoline-3-oxide-1-oxyl) (Sauder, *et al*., 2016).

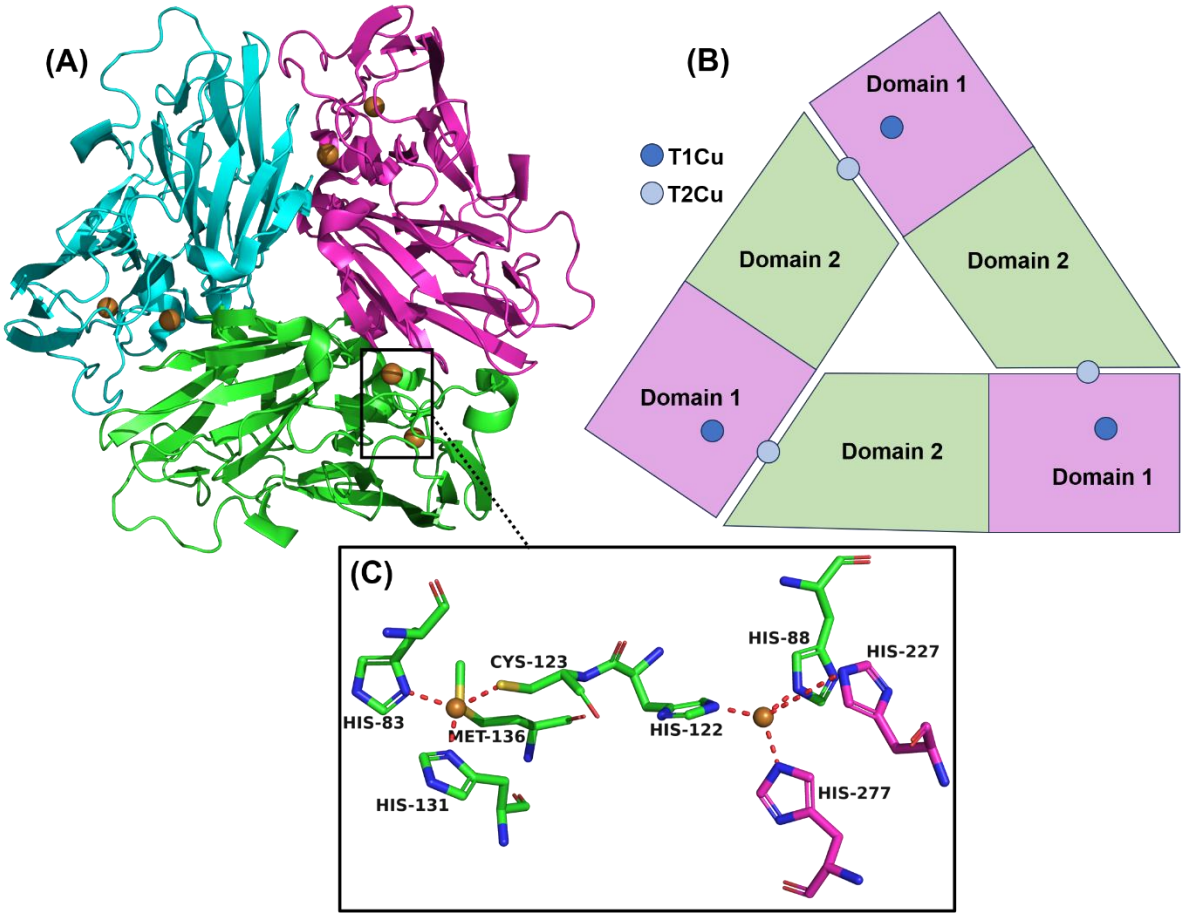


Figure 1.6. 2dMCO NirK from *Nitrosomonas europaea* (PDB 4KNU) (Lawton, *et al*., 2013). **(A)** Quaternary structure of 2 domain (2dMCO) MCO, with different colours representing monomeric subunits. **(B)** Copper site distribution in 2dMCO (4KNU)**. (C)** Mononuclear T1Cu site, coordinated by two His, a Met and Cys residues, together with a T2Cu mononuclear copper site, containing four His resides coordinating the copper ion.

3dMCO MCOs are monomeric and carry out the same catalysis as their 2dMCO counterparts, with substrate oxidation at T1Cu copper site being relayed (ca.12-13 Å) by the His-Cys-His bridge (**Figure 1.7**), followed by the reduction of O_2 to 2 H_2O using four electrons and four protons ($O_2 + 4\ H^+ + 4\ e^- \rightarrow 2\ H_2O$). AOA genomes reveal a high abundance of 3dMCOs (Sakuraba, *et al*., 2011).

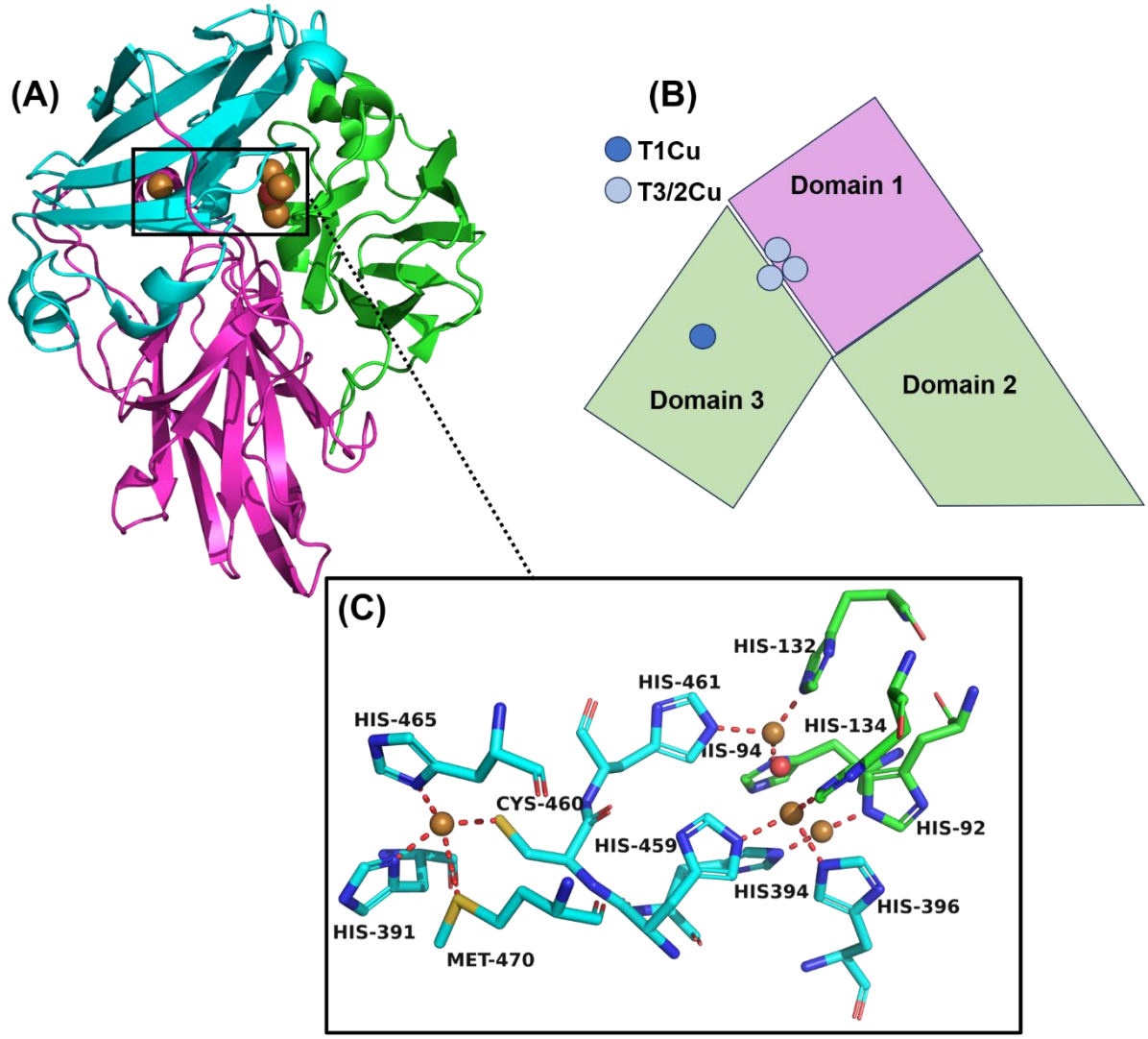


Figure 1.7. 3dMCO from hyperthermophilic archaeon *Pyrobaculum aerophilum* (PDB 3AW5) (Sakuraba, *et al*., 2011). **(A)** Quaternary structure of 3 domain (3dMCO) MCO, with different colours representing monomeric subunits. **(B)** Copper site distribution in 3dMCO (3AW5) **(B)** Mononuclear T1Cu site, coordinated by two His, a Met and Cys residues, together with a T3/T2Cu trinuclear copper site, containing eight His resides coordinating the copper ion.

1.8 Aims and Objectives

A primary aim of this thesis was to isolate potentially novel AOA species from an agricultural farm and to perform physiological experiments with particular focus on the bioavailability of copper, to gain further insights into soil dwelling AOA and the role of copper in AOA. To better understand the role of copper in AOA, this thesis aimed to perform biochemical and structural biology characterisation on putative copper binding MCOs from the marine *N. maritimus* (SCM1) and terrestrial *N. franklandus* (C13). However, difficulties in obtaining sufficient soluble MCOs led to a focus on other copper binding enzymes to identify the role of copper in other microorganisms. To do this, copper binding LPMOs from the soil dwelling bacterium *S. lividan* were investigated using structural biology for characterisation. Three multimodular LPMOs were investigated by X-ray crystallography, size-exclusion chromatography small angle X-ray scattering (SEC-SAXS) and electron paramagnetic resonance (EPR) spectroscopy, in order to provide new insights into the structural biology of the copper binding sites and the dynamic interplay of these highly flexible multimodular proteins. Finally, serial X-ray crystallography was employed to assess whether LPMOs could be used together with a light activated molecular oxygen releasing photocage, to trap intermediate copper-oxygen bound species.

Chapter 2 Isolation, purification and physiological characterisation of AOA from agricultural soils.

2.1 Introduction

Ammonia oxidising archaea (AOA) and bacteria (AOB), play a crucial role in the global nitrogen cycle (Wright and Lehtovirta-Morley, 2023; Scarlett *et al*., 2021; Clark *et al*., 2020; Lehtovirta-Morley, 2018; Yin, Z *et al* 2018; Stein and Klotz, 2016). AOA are ubiquitous in nature, distributed across terrestrial and marine ecosystems (Alves *et al*., 2018; Beeckman *et al*., 2018; Wang *et al*., 2017; de La Torre *et al*., 2008); including hot springs (*Nitrososphaera gargensis* Ga9.2) (Palatinszky *et al*., 2015), coastal waters (*Nitrosopumilus* HCA1 and PS0) (Qin *et al*., 2014), marine surface waters (NAO2 and NAO6) (Elling *et al*., 2015), salt water (*Nitrosopumilus maritimus* SCM1) (Könneke *et al*., 2005), acidic soils (*Nitrosotalea devanaterra* Nd1 and *Nitrosotalea* sp. Nd2 (Lehtovirta-Morley *et al*., 2014; 2011), and soils (*Nitrososphaera viennensis* EN76) (Tourna *et al*., 2011).

Physiological differences between AOA and AOB suggested the possibility of niche differentiation. For example, the dominance of AOA in oligotrophic marine ecosystems and acidic soils (Martens-Habbena *et al.* 2009; Lehtovirta-Morley *et al.* 2011; 2014) and a sensitivity of AOA to nitrite (Prosser and Nicol 2012) and lower (μM) NH_4^+ concentrations, while AOB typically grow at higher (mM) concentrations (Koper *et al.* 2010; Prosser and Nicol 2012). Previous studies investigating inorganic NH_4^+ fertiliser inputs also found a predominance of AOB in soils with higher rates of fertiliser application, whereas AOA were more abundant in soils with lower rates of fertiliser applications (Yin et al 2022; Hink, *et al*., 2018; Lu, *et al*., 2015; Höfferle, *et al*., 2012; Zhang). Interestingly, members of the *Ca*. *Nitrosocosmicus* genus exhibit very high tolerance to elevated concentrations of NH_4^+, compared to other AOA from different lineages. For example, *N. franklandus* (C13) has been observed to grow at NH_4^+ concentrations of up to 100 mM (Lehtovirta-Morley, *et al*., 2016),

and AOA from genus *Ca*. *Nitrosocosmicus* have been shown to contain a single low affinity ammonia transporter gene (AMT-1) (See, Chapter 3, **3.3.3**) (Redondo, *et al*., 2025; Han, *et al*., 2024; Liu, *et al*., 2021). In contrast, AOA that thrive in lower ammonia concentrations (*e.g*. *N. viennensis* EN76 and *N. gargensis* Ga9.2) possess both low affinity AMT-1 and high affinity AMT-2 genes (Redondo, *et al*., 2025).

Ammonia availability is also influenced by pH, which in turn drives AOA and AOB communities (Stempfhuber, *et al*., 2015). For example, AOA play an important role in acidic soils (Lehtovirta-Morley, *et al*., 2011; Hu, *et al*., 2022; Neal, *et al*., 2021), while AOB seem to dominate in neutral pH soils, such as *Nitrosomonas europaea* which grows optimally at pH 7.5 (Jung, *et al*., 2019). Furthermore, an increase in AOA/AOB *amoA* genes at pH levels <4 have been found, and demonstrates how both AOA and AOB can co-exist in very acidic environments (Xi, *et al*., 2017). However, high abundances of AOB *amoA* genes have also been identified in acidic soils (Che, *et al*., 2015) and numerous AOA can grow in at neutral soil pH (*e.g*. *Nitrososphaera viennensis* EN76 and *Nitrosocosmicus franklandus* C13 exhibit an optimal pH of ca.7.5) (Lehtovirta-Morley, *et al*., 2016; Tourna, *et al*., 2011).

Another factor contributing to niche separation of AOA and AOBs in soils is temperature (Mukhtar, *et al*., 2019; Lu, *et al*., 2018; Gubry-Rangin, *et al*., 2017; Ouyang, *et al*., 2017; Taylor, *et al*., 2017). Cultivation of AOA range from mesophiles to hyperthermophiles, with optimal temperatures between 25°C for *Ca*. *N. devanaterra* Nd1 (Lehtovirta-Morley *et al*., 2014) and 72°C for *Ca*. *Nitrosocaldus yellowstonii* (de la Torre *et al*., 2008). Many other AOA, such as *N. viennensis* EN76 (Tourna, *et al*., 2011), *N. franklandus* C13 (Lehtovirta-Morley, *et al*., 2016), and *N. devanaterra* Nd2 (Lehtovirta-Morley, *et al*., 2014) all show optimal growth between 35-40°C. Contrastingly, AOB show lower growth temperature ranges, with optimal temperatures typically between 20-30°C (Hossain, *et al*., 2022), with the exception of some AOB that can grow as low as 4°C (Urakawa, *et al*., 2015).

AOMs have very slow growth rates, making them incredibly challenging to isolate in pure culture, and susceptible to contamination by faster growing heterotrophs (Woo, *et al*., 2022; French, *et al*., 2021; Prosser, *et al*., 2020). Moreover, characterisation of terrestrial AOA is limited compared to AOB, which were first cultured by Frankland and Frankland, (1890). Four AOA isolates previously obtained from soil, of which, three were characterised (*i.e*. *Nitrosoarchaeum koreensis* MY1; *Nitrosotenuis chengbukensis* MY2 (Jung *et al*., 2011; 2014); *N. viennensis* EN76 (Tourna, *et al*., 2011) and *Nitrosocosmicus franklandus* C13), and demonstrated a high NH_4^+ tolerance (Lehtovirta-Morley *et al*., 2016). Recently, a study described the successful recovery of AOA colonies using gellan gum when growing *N. franklandus* C13 and *N. viennensis* EN76 (Klein, *et al*., 2022) and a similar approach was adopted herein.

2.1.1 Aims

This Chapter aims to isolate, purify and characterise novel AOA from agricultural soils using a combination of liquid/solid cultivation and molecular techniques. Such information from these isolates will provide a better understanding on AOA physiology and the molecular mechanisms involved.

2.2 Methods

2.2.1 Agricultural soil samples

Triplicate sandy loam agricultural soils (ca. 5 cm depth; denoted S1-4) were collected using a sterile spatula from an active farm (Poplar Farm Tuddenham St Martins, Suffolk) at the following locations: S1 (52°05'19"N 1°11'55"E), S2 (52°05'09"N 1°11'52"E), S3 (52°04'56"N 1°11'46"E) and S4 (52.0857°N, 1.1974°E) (**Figure 2.1**). Soil NH_4^+/NO_2^- concentrations were as follows: S1 (12.83/14.37 µg g^{-1}), S2 (12.33/14.25 µg g^{-1}), S3 (8.98/66.02 µg g^{-1}) and S4 (6.00/175.80 µg g^{-1}) respectively. Soil pH was between pH 5.2 (S3), 6.8 (S1 and S2) and 6.9 (S4). Soil samples were collected from farmland soils, regularly fertilised using both organic and inorganic fertiliser inputs, with repeated ammonium inputs to increase the likelihood of successful enrichment of AOAs. All samples were stored at 4°C prior to inoculation into liquid media. NH_4^+ and NO_2^- were measured from soils by mixing 6g of dry soil with 18 mL of Milli-Q water, shaken for 15 minutes and samples centrifuged for 30 minutes at 5500 rpm. Supernatant was collected and acidified to pH 5 using 0.1 M phosphoric acid, which was then passed through a Dionex II P cartridge and filtered using a Nalgene syringe. A colorimetric assay was performed using a Continuous Flow AutoAnalyser (SEAL AA3) following the method of Kérouel, *et al*. (1997). The pH of soil samples were measured by resuspending soil in Milli-Q water and pH was recorded using a calibrated pH meter (ThermoFisher).

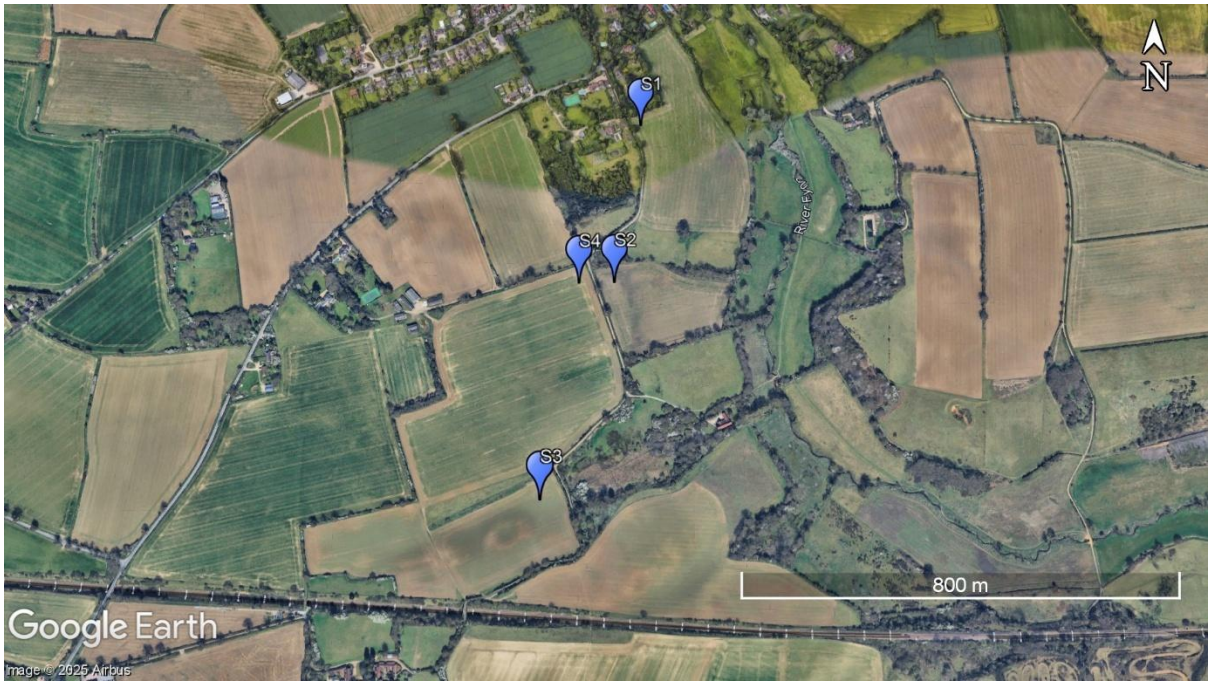


Figure 2.1. Soil sampling locations, Poplar Farm Tuddenham St Martins, UK. (Imagery from Google Earth Pro).

2.2.2 Enrichment cultures and isolation

Triplicate enrichment cultures were set up by inoculating fresh soil (ca. 0.5 g) into sterile 100 mL conical flasks containing 40 mL of sterile freshwater medium (FWM), ammonium chloride (2 mM final concentration), ampicillin (Sigma-Aldrich), streptomycin (Sigma-Aldrich) (both 50 µg mL^{-1} final concentration) and incubated statically at 37°C in the dark as previously described (Reyes, *et al*., 2020a; Lehtovirta-Morley, *et al*., 2016) (**Table A.1**). This enrichment strategy was performed to selectively favour AOA, while supressing fast-growing heterotrophic bacteria through antibiotic supplementations and repeated fresh medium subculturing. A 2mM NH_4Cl concentration was selected to provide sufficient substrate to support ammonia oxidation, whilst limiting excess nitrite accumulation, which can cause a decrease in pH and creates less favourable conditions for AOA. Cultivation conditions of pH 7.5, a temperature of 37°C and copper concentration of $1x10^{-8}$ M were selected as they have been commonly used in previous studies to enrich for neutrophilic soil AOA, such as *N.*

viennensis (EN76) and *N*. *nitrosocosmicus* (C13) (Reyes, *et al*., 2020a; Lehtovirta-Morley, *et al*., 2016; Tourna, *et al*., 2011). Accumulation of NO_2^-/NO_3 was monitored using NO_2^-/NO_3^- test strips (MACHEREY-NAGEL, QUANTOFIX). Once NO_2^- concentration had increased to ca. 40-80 mg L^{-1} (ca. days 20-30), enrichment cultures (2% v/v) were sub-cultured into sterile 100 mL conical flasks containing 40 mL sterile FWM. Subculturing continued in the same manner until later subcultures (*i.e*. subcultures 6-7) were cryopreserved at -80°C by mixing 1:1 vol 40% *(v/v)* sterile glycerol with stationary phase culture. The presence of heterotrophic bacteria was checked by spread-plating 0.8 mL of culture onto Nutrient, Tryptone Soy and LB agar plates (Sigma-Aldrich) at 25%, 10%, 5% and 1.5% *(w/v)* agar and incubated at 37°C in the dark. Antibiotic sensitivity was tested by spread plating 0.5 mL of exponential/stationary phase culture onto solid FWM agar plates containing antibiotic disks (Thermo Scientific) (*i.e*. gentamycin, penicillin, clindamycin, kanamycin, streptomycin and ampicillin at 50 µg mL^{-1} final concentration) and incubated at 37°C in the dark for 30 days. After five successive subcultures, enrichment cultures denoted C2 and C4 (**Figure 2.2**) were grown in solid media (gellan gum (Thermo Scientific)) to obtain single AOA colonies as previously described (Lehtovirta-Morley, *et al*., 2016).

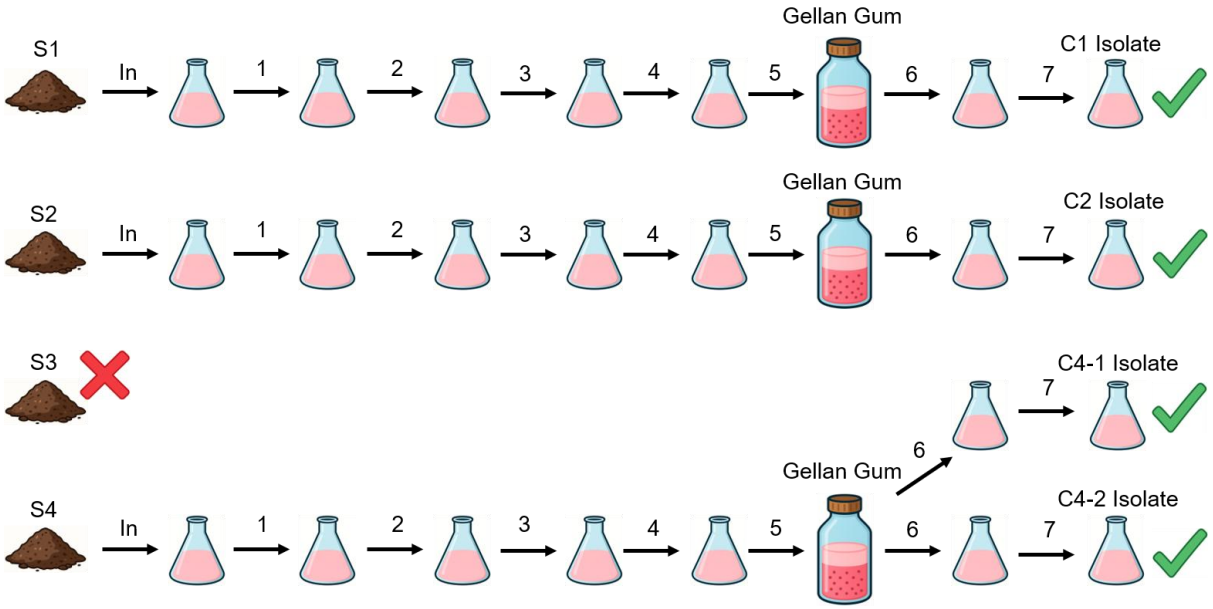


Figure 2.2. Enrichment outcomes of four soil samples (S1-4) into liquid cultures (C1, C2 and C4 isolates). S3 contained no detectable AOA in soil. On the other hand, S1, S2 and S4 grew successfully through multiple subcultures and forming colonies in solid FWM in gellan gum (Thermo Scientific). In = inoculum and numbers 1-7 = subculture number.

Gellan gum (Thermo Scientific) (0.6% w/v) was dissolved in FWM and autoclaved. Once cooled, molten gellan gum was kept in a water bath at 40°C to prevent solidification. Heat-liable antibiotics, streptomycin and ampicillin (both at 50 µg mL^{-1} final concentration, Sigma-Aldrich) were added to 25 mL molten gellan gum in sterile 100 mL Duran bottles and left to solidify (**Figure 2.3**). A further 25 mL of molten gellan gum was overlaid which was then inoculated with 60 µL of exponential or stationary phase culture and left to solidify. A final aqueous layer FWM (25 mL) was added and the cultures were incubated statically in the dark at 37°C for 30 days. NO_2^- accumulation was measured using NO_2^-/NO_3^- testing strips (MACHEREY-NAGEL, QUANTOFIX). Individual separated colonies were obtained (**Figure 2.3**) and approximately 10-20 colonies were picked using a sterile loop and inoculated into a 100 mL sterile conical flask containing 40 mL of sterile FWM and incubated at 37°C for 30-40 days.

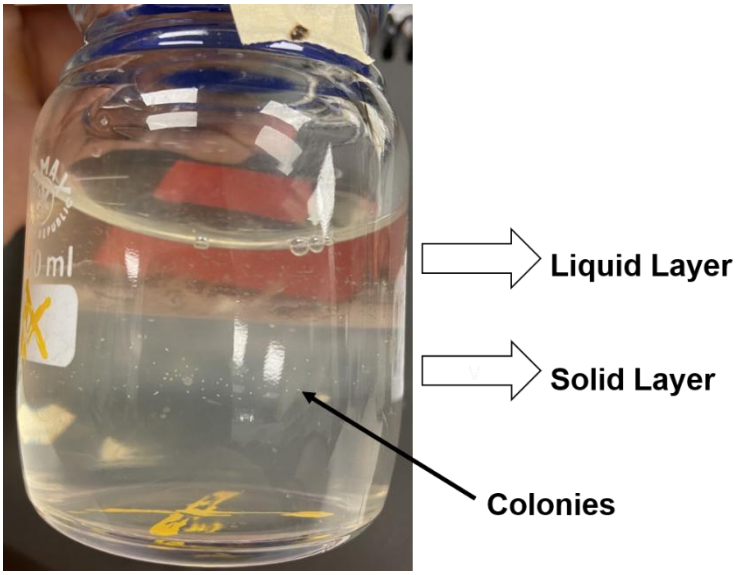


Figure 2.3. Colonies developed in gellan gum from C4 enrichment culture.

2.2.3 DNA extraction and qPCR analysis

Soils, enrichment cultures were harvested in triplicate and DNA extracted using a SoilPro DNA extraction kit (Qiagen) following the manufacturer's instructions. All DNA extracts were stored at -20℃. Abundance of AOA and AOB *amoA* genes were quantified by qPCR using a SensiFAST SYBR No-ROX (Bioline) on a CFX Opus 384 Real-Time PCR System (BioRad) using the primers CrenamoA-23F/CrenamoA-616R for AOA (Tourna, *et al*., 2008) and amoA-1F/amoA-2R for AOB (Rotthauwe, *et al*., 1997). 16S rRNA genes primers were also used for qPCR analysis 21F/1492R for Archaeal 16S rRNA genes (Delong *et al*., 1992; Lane *et al*., 1991). Thermocycling conditions included an initial denaturation at 95℃ for 3 minutes followed by 40 cycles of 95℃ for 5 seconds, 60℃ for 30 seconds and included a combination of annealing and extension time. Melting point analysis was used to confirm product specificity at 52-95℃ with a plate read every 0.5℃. An absolute quantification method against an internal standard curve calibration curve utilising DNA standard to quantify *amoA* abundance. DNA standards of each gene target consisted of 10^1 to 10^8 copies in 10 µL reactions

consisting of 1 µL of DNA template and 200 nM of primers. Efficiency value was 116% slope value was -3.1, Y-intercept was 36.8 and R^2 value was 0.98. Samples duplicate technical replicates were performed. All biological samples were run in triplicate.

2.2.3 CARD-FISH and microscopic analysis of AOA enrichments

2.2.3.1 Cell fixation and hybridisation

To visually check the purity of AOA enrichments (C2 and C4-2), Card-FISH (Catalysed Reporter Deposition Fluorescence *In Situ* Hybridisation) was performed. Stationary phase cells (2 mL) were fixed by centrifugation (20000 x g for 10 minutes (4°C)) and washed three times with 100 µL of PBS. Cells were resuspended in 1:1 100 µL of PBS and 100% *(v/v)* ethanol and stored at 4°C. Fixed cells were pelleted by centrifugation at 20000 xg for 10 minutes (4°C). Cells were then washed three times in 1x PBS, resuspended in lysozyme (10 mg mL^{-1} final concentration) and incubated at room temperature for one hour. Endogenous biotin was removed using "Endogenous Biotin-Blocking Kit" (Invitrogen) following the manufacturer's instructions. Endogenous HRP (Horseradish Peroxidase) was also removed using 2 drops of 3% *(v/v)* hydrogen peroxide solution by either "Invitrogen Alexa Fluor 488 Tyramine Superboost Kit, streptavidin" (Invitrogen) or "Invitrogen Alexa Fluor 555 Tyramine Superboost Kit, streptavidin" (Invitrogen), thoroughly mixed, and incubated at room temperature for 1 hour then washed three times in 1x PBS. Cells were hybridised by centrifugation at 20000 x g for 5 minutes (4°C) and the pellet was resuspended in 75 µL of hybridisation buffer (35mM Tris-HCl pH 7.2 40% *(v/v)* formamide and 0.1 % *(w/v)* SDS) containing the probes CrenamoA-23F for AOA (Tourna, *et al*., 2008) and amoA-1F (Rotthauwe, *et al*., 1997) for AOB (both 50 ng μL^{-1} final concentration). Cells were then incubated at 37°C for 16-20 hours.

Cells were pelleted, washed three times in 100 µL of PBS and pelleted again by centrifugation, 2 drops of streptavidin conjugated HRP added to the cell pellet and incubated at room temperature for 1 hour. Cells were re-washed using 100 µL of PBS. Alexa Fluor 488/555 tyramine stock was prepared by adding 150 µL of DMSO and 10 µL of tyramine added to 75 µL of resuspended cells in PBS for 10 minutes in the dark with a 3x washing steps in 100 µL of PBS after each tyramine addition. Samples were aliquoted into an 8 well chamber slide (ThermoFisher) and counterstained with 75 µL of Hoechst stain (Invitrogen) for visualisation.

2.2.3.2 Microscopy

Cell images were acquired using a Nikon Ti2 Eclipse X-light V3 spinning disc confocal microscope equipped with a Plan Apo λD 60x Oil objective (NA 1.4, DIC-compatible, Nikon) using imaging software NIS-elements and LDF immersion oil (Cargille Laboratories). Imaging was carried out using excitation (Ex) 640 nm, emission (Em) 525 nm for Alexa Fluor 488 (green channel), Ex 640 nm, Em 610 nm for Alexa Fluor 555 (orange channel) and Ex 640 nm, Em 440 nm (blue channel) for Hoechst counterstain. Fluorescence images were captured using a Kinetix, A22M723030 with exposure time of 800 ms (green), 300 ms (orange) and 500 ms (blue).

2.2.4 Physiological Characterisation of enrichment C4-2

Enrichment C4-2 was further characterised in terms of physiology and growth characteristics over a range of temperature, pH, and NH_4^+ concentrations. Cultures were grown in triplicate in 100 mL sterile Durans containing 50 mL FWM as previously described (37°C, static and in the dark, unless otherwise stated). Different NH_4^+ concentrations were tested (0, 1, 2, 5, 10, 20, 50 and 100 mM of NH_4Cl). Inhibition by NO_2^- concentration was also tested using 0, 1, 2, 10, 20 and 50 mM $NaNO_2^-$. Temperature growth ranges were also assessed at

12°C, 20°C, 30°C, 37°C and 40°C. The pH ranges were also measured at 0.5 pH intervals ranging from pH 5.5 to 9.0 and unbuffered. Reduced NH_4Cl concentration was used when investigating the effects of pH to minimise NH_4^+ oxidation and a pH decline and 5 mM MES (2-(N-morpholino) ethanesulfonic acid) was used in pH range 5.5-7 and HEPES (4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid) was used for pH range 7.5-9. Growth was measured by a decrease in NH_4^+ concentration and increase in nitrite NO_2^- using colorimetric determination of NH_4^+ and NO_2^- and measured on a Continuous Flow AutoAnalyser (SEAL AA3), following the methods of Kérouel, *et al*. (1997). Standard curves to measure NH_4^+ and NO_2^- are presented in **Appendix B** (**Figure B.1 and B.2**) and the negative abiotic control (**Figure B.3**).

2.2.5 Statistical analysis

Statistical analysis was performed using a one-way ANOVA followed by Tukey's honestly significant difference HSD post hoc test with the Statistics Kingdom online calculator (https://www.statskingdom.com). A significance threshold of ($\alpha = 0.05$) was applied and Tukey's HSD test was used to identify pairwise group differences. Datasets with multiple groups were displayed using a letter-based grouping system. Groups sharing at least one letter shows no significant difference, whereas groups with different letter are significantly different. Repeated measures ANOVA was carried out on physiological experiments using Statistics Kingdom calculator (https://www.statskingdom.com) to assess changes in NH_4^+ depletion and NO_2^- accumulation over time. The ANOVA output was used to determine whether treatment effects were statistically significant, with a threshold of $\alpha = 0.05$.

2.3 Results

2.3.1 Abundance of AOA in agricultural soils

AOA and AOB *amoA* gene abundance in the unamended agricultural soils (S1-4) and subsequent enrichment cultures (C1, C2 and C4) were quantified (**Figure 2.4**). All soils had greater *amoA* copies per g^{-1} dry weight of AOA compared to AOB, with between ca. 4.5 x 10^7 and 3.5 x 10^7 AOA *amoA* gene copies g^{-1} dry weight for AOA, compared to ca. 1.2 x 10^6 and 9.6 x 10^5 AOB *amoA* gene copies g^{-1} dry weight (with the exception of S3 whereby AOA *amoA* gene copies g^{-1} dry weight was very low, (<300)) **(Figure 2.4 A and B)**. No significant difference was found in AOA *amoA* gene abundance between soil samples (one-way ANOVA/Tukey's HSD, $p > 0.05$), whereas with AOB soil S2 had significantly higher gene copies than S1 ($p < 0.05$), while S4 shows no significant difference between S1 and S2 ($p > 0.05$) (**Figure 2.4 A and B**).

Before isolation on solid media (*i.e*. pre gellan gum (GG) cultivation) AOA abundance was between 2-8 × 10^3 AOA *amoA* gene copies mL^{-1}, which increased in enrichments derived from soil enrichments C2 and C4 gene copies mL^{-1} (3.2 – 17 × 10^3 AOA *amoA* gene copies mL^{-1}). No significant difference was revealed between subcultures 3 and 5 ($p > 0.05$) for AOA *amoA* gene copies mL^{-1}, however a significant increase in AOA *amoA* gene copies per mL^{-1} was seen for cultures C2 and C4, post gellan gum ($p < 0.05$) (**Figure 2.4C**). AOB *amoA* gene copies mL^{-1} revealed no significant difference at subcultures 3, 5 and post gellan gum cultivation (one-way ANOVA/Tukey's HSD, $p > 0.05$) (**Figure 2.4D**), despite C1 showing a 2-fold increase after gellan gum cultivation (ca. 4.7 x 10^3 – 8.9 x 10^3 AOB *amoA* gene copies mL^{-1}). Interestingly, C1 was the only culture to not show an increase in AOA *amoA* gene copies per mL^{-1} after gellan gum cultivation ($p > 0.05$).

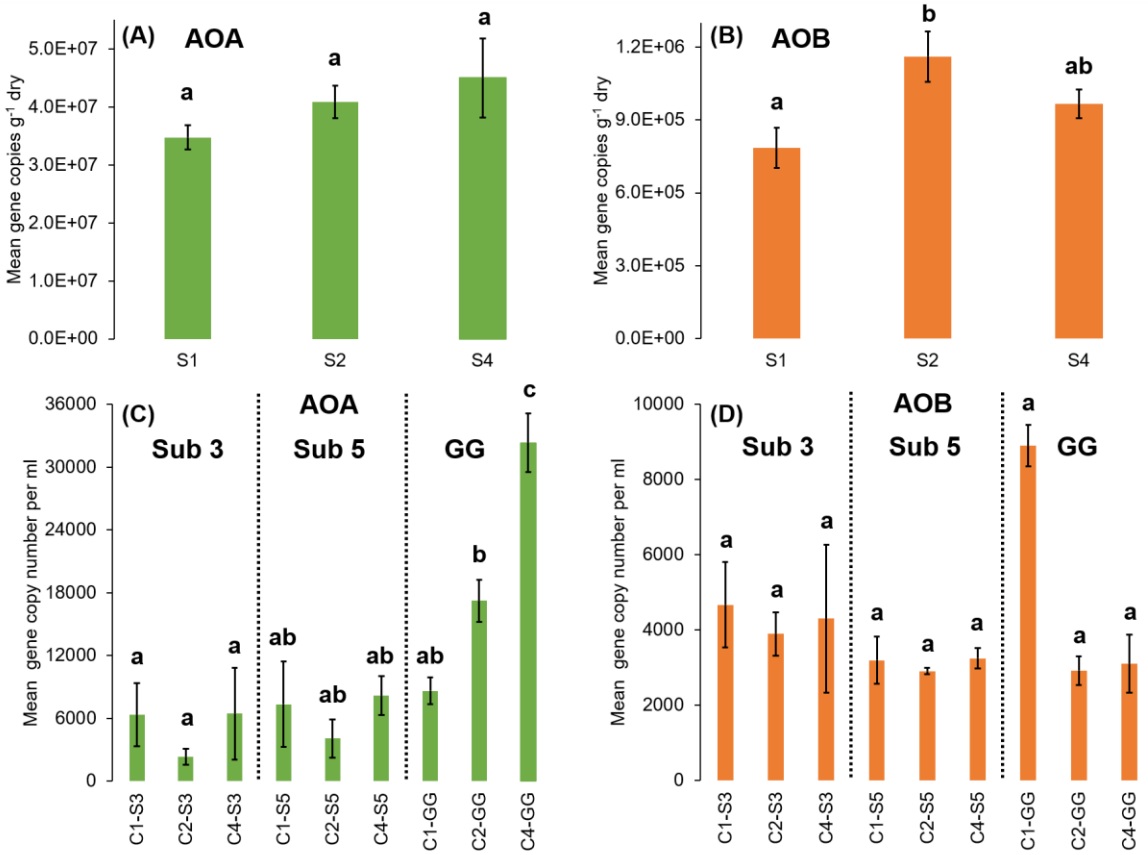


Figure 2.4. Abundance of AOA (**A**) and AOB (**B**) *amoA* gene copies g^{-1} dry weight in soils and AOA (**C**) and AOB (**D**) *amoA* gene copies mL^{-1} in enrichment cultures/ isolates. Green bars denote AOA and orange bars represent AOB. Sub 3 = subculture 3, Sub 5 = subculture 5 and GG = post gellan gum cultivation. Error bars represent mean abundance ± SD (*n*=3). Letters denote significant differences (one-way ANOVA with Tukey's HSD post hoc test (α = 0.05).

Furthermore, enrichment cultures derived from C4 grew consistently with NO_2^- accumulation between 40-80 mg L^{-1} compared to soils from S1 and S2, which sometimes did not show growth in the presence of antibiotics in later subcultures. Subsequent experiments were therefore performed with enrichment C4 due to its highest *amoA* gene copy numbers and consistent growth.

2.3.2 Fluorescence Microscopy (CARD-FISH)

CARD-FISH was performed on enrichments C2 and C4-2 to visualise the spatial organisation of AOA and assess whether AOB were also present and give an insight into other AOB NH_4^+ oxidisers that may have contributed to the results obtained with the physiological experiments (See Section **2.3.3**). Enrichment C2 comprised of highly dense aggregates that were predominantly AOA with some AOB co-occurring. Counter stain Hoechst revealed filamentous structures that co-localised to AOA and AOB. AOA cells had an irregular coccoid morphology and were between 0.8-1.0 μm in size (**Figure 2.5**). Despite AOA *amoA* genes showing a 4-5x greater abundance than AOB *amoA* gene copies per mL^{-1} (**Figure 2.4 C** and **D**), as well as minimal heterotrophic growth observed on Tryptone, Nutrient and LB agar, CARD-FISH revealed a co-occurrence of AOB. (**Figure 2.5B**). However, the higher signal intensity associated with AOA (green signal) and their predominance suggests that the AOA were the likely drivers of NH_4^+ oxidation in the C2 enrichments.

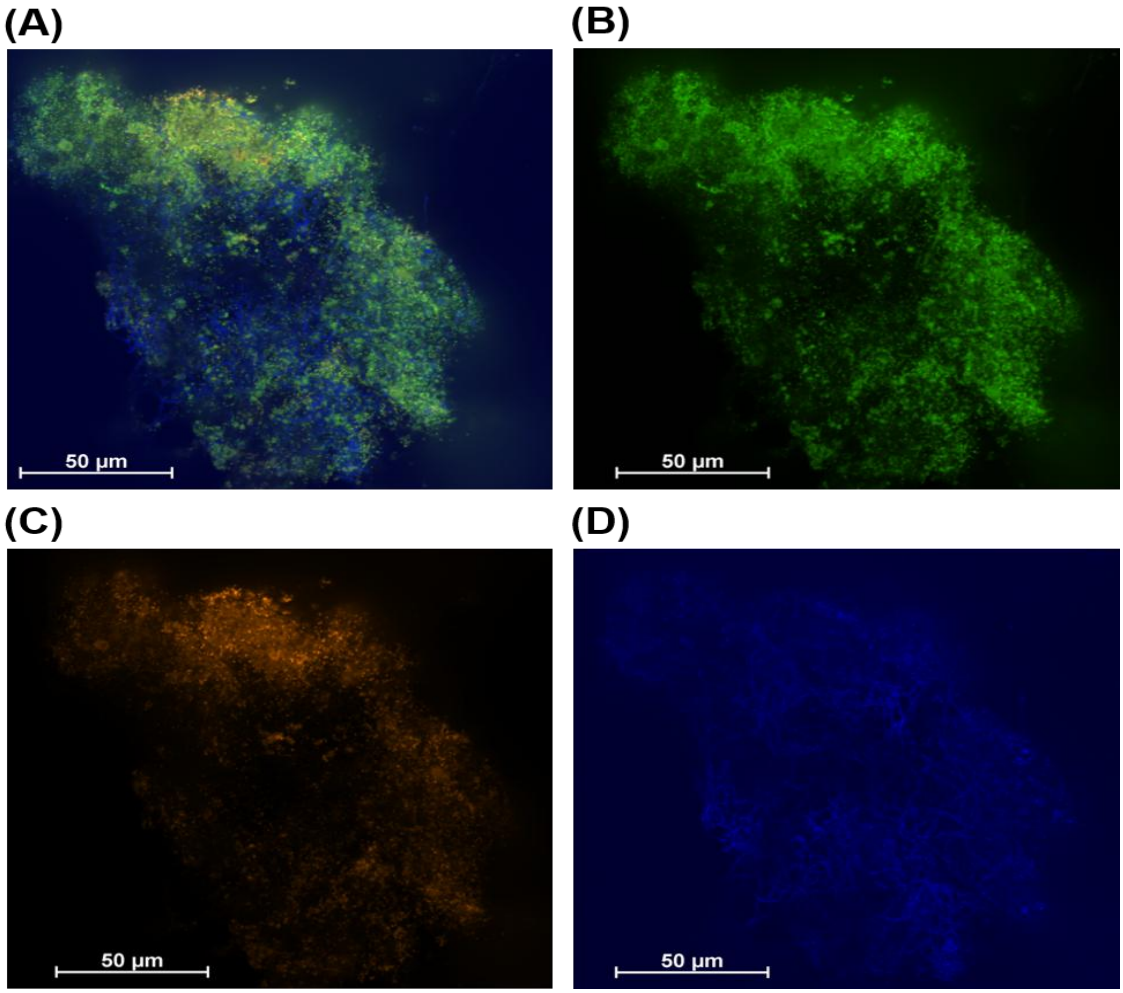


Figure 2.5. CARD-FISH of isolate C2 using a Plan Apo λD 60× oil immersion objective: overlay of all signals **(A)**, AOA (CrenamoA-23F, green) **(B),** AOB (amoA-1F, orange) **(C),** and Hoechst stain (blue) **(D)**. AOA cells were between 0.8-1.0 µm in size.

CARD-FISH of enrichment C4-2 showed cells were distributed across the field of view, rather than forming compact aggregates, with lower density of both AOA and AOB compared to **Figure 2.5** (**Figure 2.6**). However, AOA cells (green signal) were predominant, compared to the AOB (orange signal), suggesting that AOA were likely driving NH_4^+ oxidation in C4-2 enrichments. Counter stain Hoechst revealed a high abundance of dispersed cells (**Figure 2.6**). AOA from enrichment C4-2 had an irregular coccoid to cell morphology and were between 0.8-1.2 µm in size.

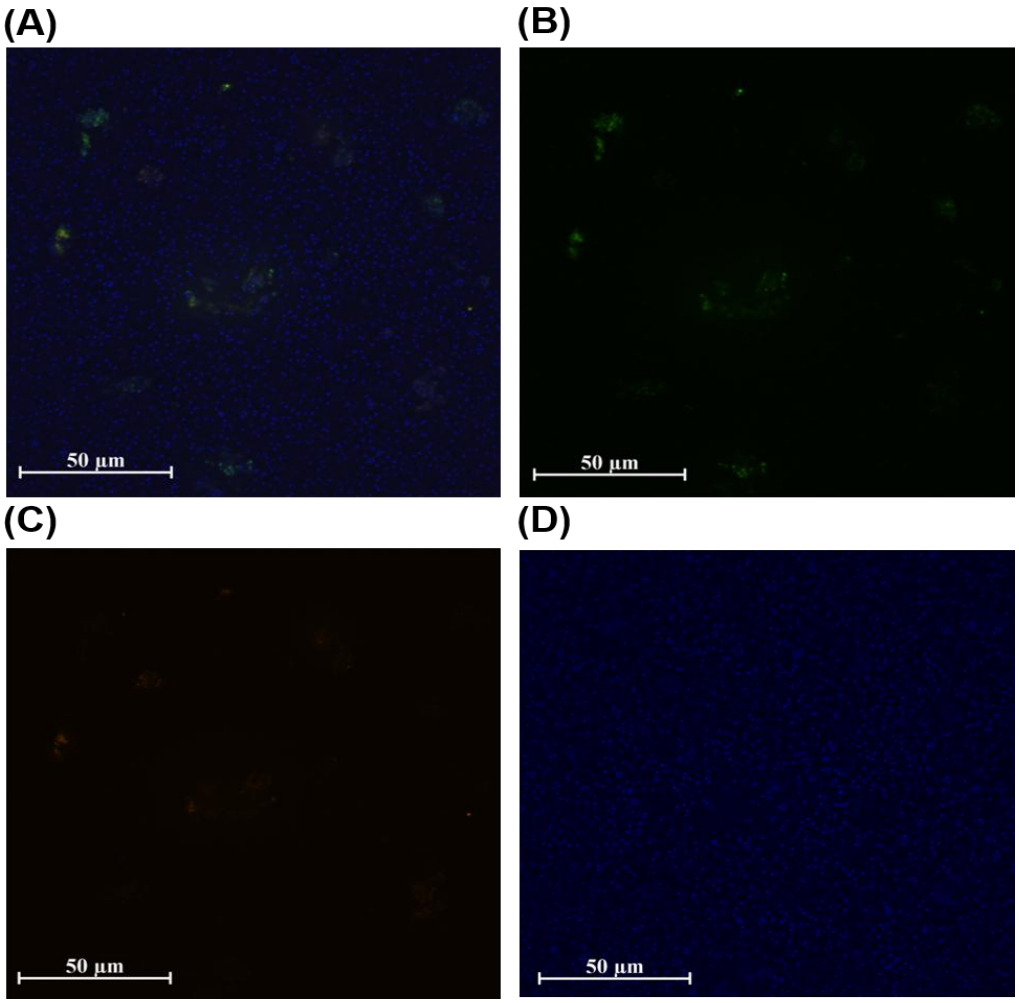


Figure 2.6. CARD-FISH of isolate C4-2 using a Plan Apo λD 60× oil immersion objective **(A)** Overlay of all signals, **(B)** AOA (CrenamoA-23F, green), **(C)** AOB (amoA-1F, orange) and **(D)** Hoechst stain (blue). AOA cells were between 0.8-1.2 µm in size.

2.3.3 Physiological Characterisation of enrichment C4-2

2.3.3.1 Effect of ammonia concentration on growth

AOA growth was assessed by monitoring NH_4^+ depletion and NO_2^- accumulation over time. The effect of NH_4^+ concentration on the growth of enrichment C4-2 was evaluated based on NH_4^+ removal (**Figure 2.7**) and NO_2^- accumulation (**Figure 2.8**). The optimal NH_4^+ concentration, for C4-2 growth was (1-50 mM) ($p < 0.05$), with growth inhibited at 100 mM ($p > 0.05$) (**Figure 2.7** and **2.8**). At concentrations of 1-2 mM NH_4^+ there was a complete removal of NH_4^+, with exponential growth reached by day 14 and stationary phase by day 21 (**Figure 2.7 B - F**). However, when grown in higher NH_4^+ concentrations (*i.e*. 100 mM), little growth was achieved and ammonium oxidation was inhibited, revealing a threshold of NH_4^+ at 100 mM that hinders C4-2 growth (**Figure 2.7 G - H**).

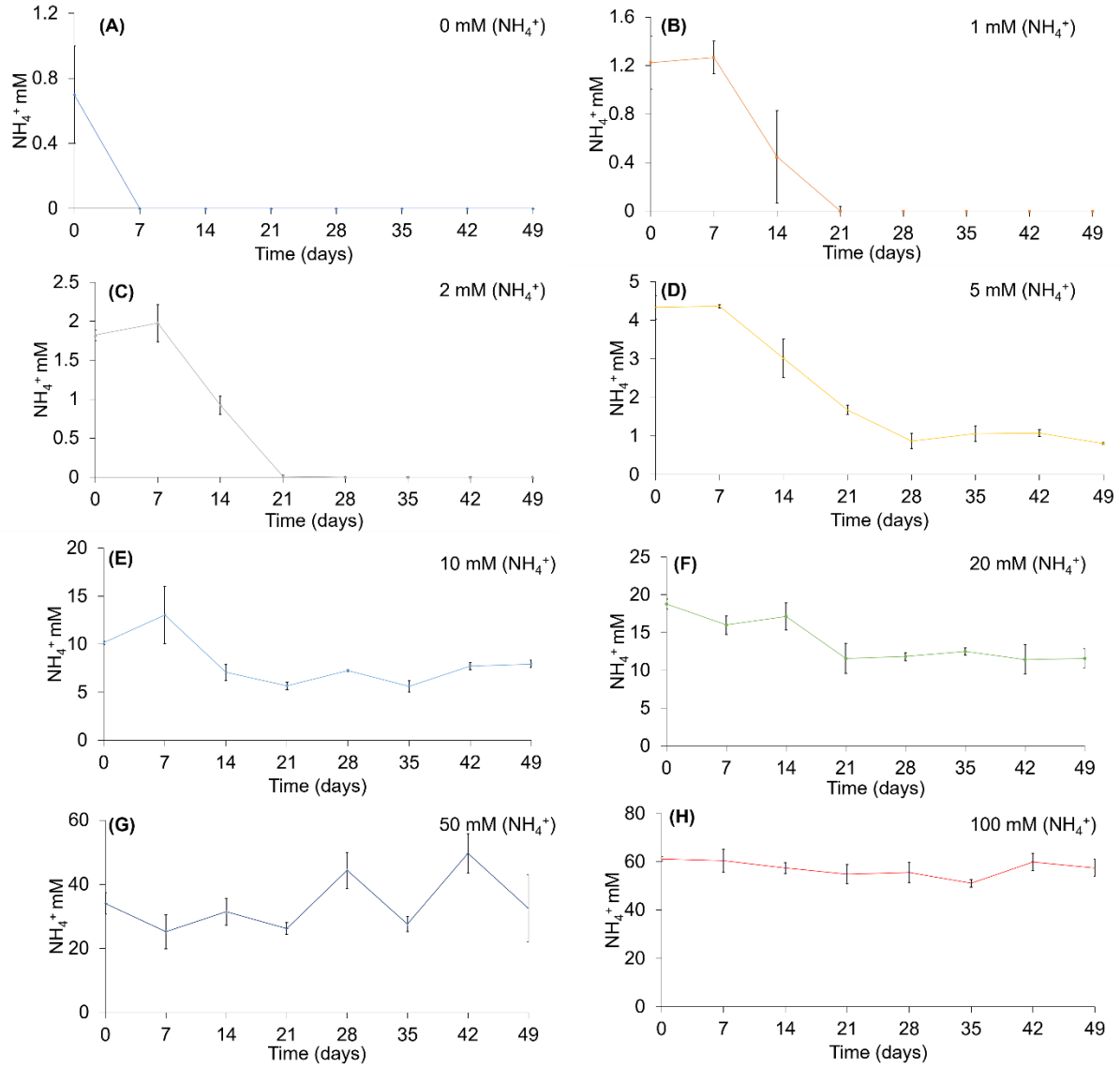


Figure 2.7. Growth of enrichment C4-2 on 0 mM, 1 mM, 2 mM, 5 mM, 10 mM, 20 mM, 50 mM and 100 mM NH_4^+. Error bars represent SD of the mean (*n*=3).

As expected, when no ammonium was added, negligible NO_2^- was detected. However, at lower ammonium concentrations (*i.e.* 1-2 mM NH_4^+), nitrite accumulation was modest, with a maximum of ca. 1.8 mM NO_2^- at 2 mM NH_4^+ addition. NO_2^- accumulation all revealed a significant difference at 1-50 mM NH_4^+ ($p < 0.05$), with an exponential phase at day 14 and stationary phase at day 21-28. The highest NO_2^- accumulation occurred with cultures amended with 5-50 mM NH_4^+, reaching NO_2^- concentrations of 2.35 – 3 mM by day 28. On the other hand, with 100 mM NH_4^+ there was a very small increase in NO_2^- (*i.e.* remained between ca. 60-150 µM nitrite) ($p > 0.05$) (**Figure 2.8**). Cultures amended with 50 mM NH_4^+ also had a longer lag phase and reached exponential phase by day 21 compared to day 14 for 5-20 mM NH_4^+ (**Figure 2.8**).

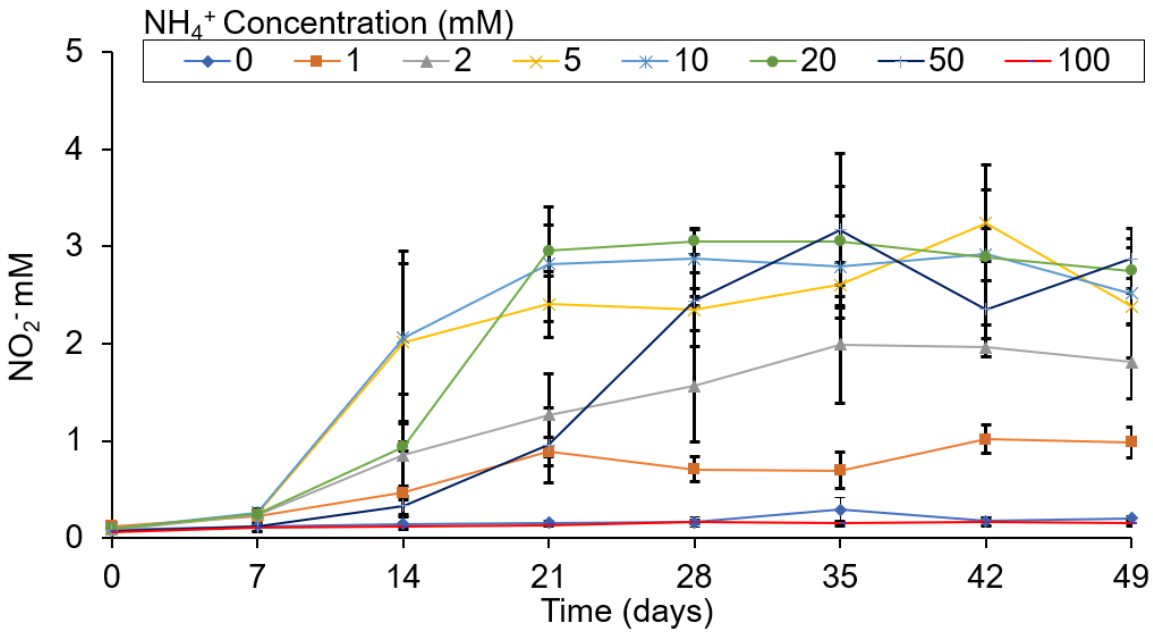


Figure 2.8. NO_2^- accumulation by enrichment C4-2 during growth on varying ammonium concentrations. Error bars represent SD of the mean (n=3).

2.3.3.2 Growth inhibition of enrichment C4-2 to increasing nitrite concentration

To investigate the inhibitory effects of NO_2^- on growth, C4-2 was supplemented with 0, 1, 2, 10, 20, 50 mM NO_2^- and growth was measured through NH_4^+ removal (**Figure 2.9**) and NO_2^- accumulation (**Figure 2.10**). A concentration of 5 mM NH_4^+ was used instead of 2 mM as consistent NH_4^+ oxidation was observed at 5 mM NH_4^+. When grown in 0-20 mM NO_2^-, growth of C4-2 was not inhibited, with NH_4^+ concentration significantly decreasing from ca. 5.2 mM to 2.5 – 1.5 mM between day 7-35 ($p < 0.05$) (**Figure 2.9**). When C4-2 enrichment was grown at higher NO_2^- concentrations (*i.e.* between 10-20 mM) growth was slower and the removal of NH_4^+ over the period of incubation plateaued by day 35 ($p < 0.05$) (**Figure 2.9**). However, when C4-2 was grown on 50 mM nitrite, growth appeared to be completely inhibited with no significant change in NH_4^+ ($p > 0.05$) (**Figure 2.9**).

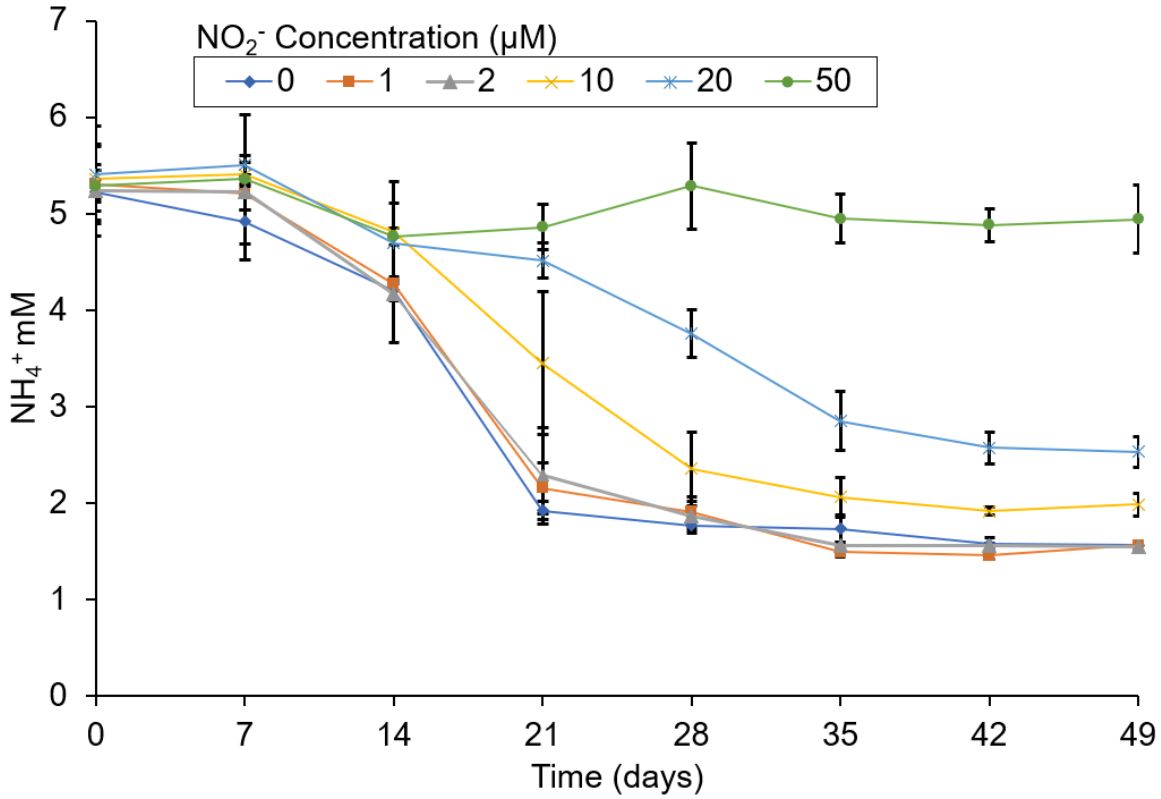


Figure 2.9. Growth inhibition of enrichment C4-2 to increasing NO_2^- concentration. Error bars represent SD of the mean (n=3).

Similar patterns were also observed regarding NO_2^- accumulation, with an exponential increase in NO_2^- production between day 0 and day 21 up to ca. 0 - 4 mM (for 0 mM NO_2^-), ca. 1 - 5 mM (for 1 mM NO_2^-) and ca. 2 - 6 mM (for 2 mM NO_2^-) produced by day 28 ($p < 0.05$) (**Figure 2.10**). Accumulation of NO_2^- (ca. 10-17 mM), (ca. 18–27 mM) in concentration (10-20 mM NO_2^-) ($p < 0.05$), revealed a more gradual increase in NO_2^- compared to (1 – 2 mM NO_2^-) (**Figure 2.10**). When incubated at NO_2^- concentration (50 mM), there was no significant increase in NO_2^- accumulation, with no significant AOA growth ($p > 0.05$) (**Figure 2.10**).

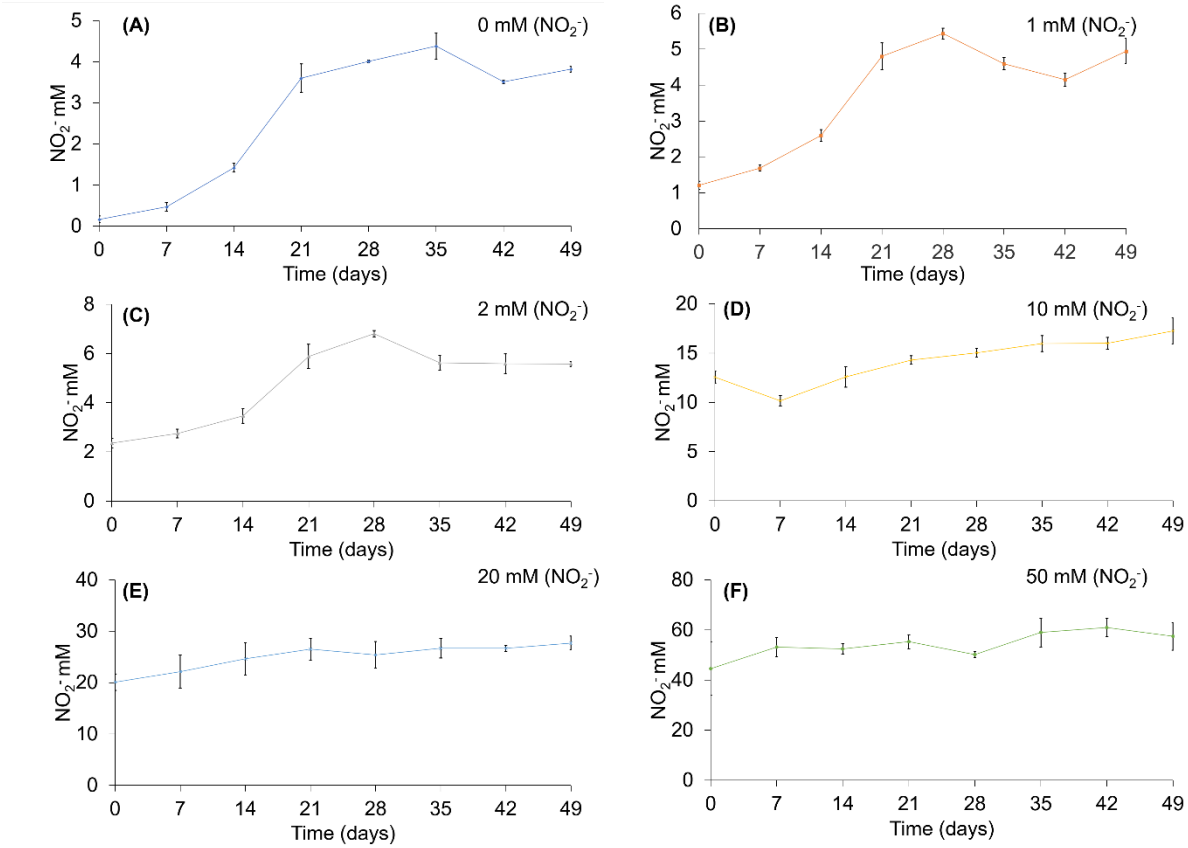


Figure 2.10. Growth inhibition of enrichment C4-2 to increasing nitrite concentration as determined by production of NO_2^-. Error bars represent SD of the mean (n=3). Some SD bars may appear smaller than symbols.

2.3.3.3 Effect of temperature on growth of enrichment C4-2

The temperature growth range of C4-2 enrichment was investigated. The optimum temperature for growth was 37°C with which showed the highest oxidation of NH_4^+ (from ca. 6 to 2 mM) by day 49 (**Figure 2.11**) and a concomitant increase in nitrite production from ca. 0 to >4 mM by day 28 (**Figure 2.12**). At 12°C and 20°C, NH_4^+ depletion showed minimal change in either NH_4^+ (*i.e.* ca. 4 - 6 mM) or NO_2^- (ca. 200-100 µM) levels (*i.e.* ca. 30-100 µM NO_2^- at 12°C, ca. 30-200 µM at 20°C) ($p > 0.05$) during the 49-day incubation (**Figure 2.11**). In contrast, at higher temperatures (*i.e.* 30°C, 37°C and 40°C), NH_4^+ was being rapidly oxidised, decreasing from 5500 to 2500-1500 µM, with a concomitant increase in NO_2^- production from 30 µM to 3.3 - 2.5 mM by day 49 ($p < 0.05$). Exponential growth occurred during days 14-21, with stationary phase reached between days 28-35 (**Figure 2.11 A - B**).

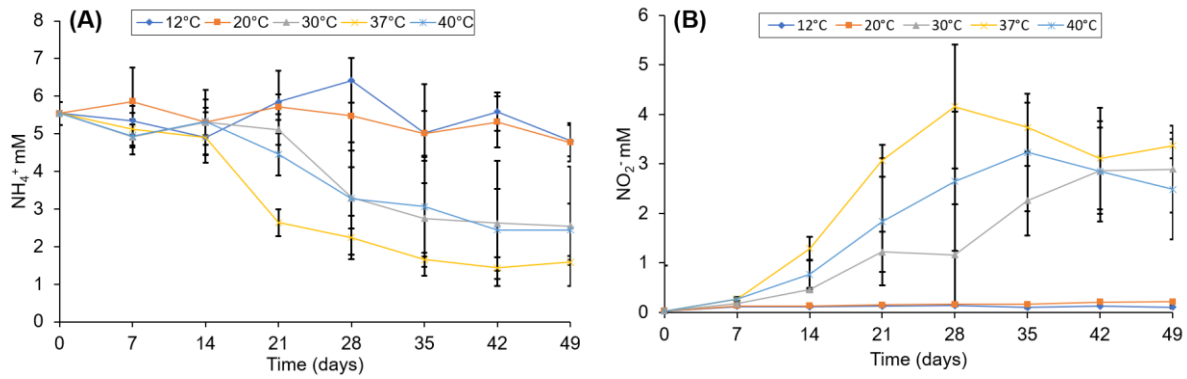


Figure 2.11. Influence of temperature on ammonia oxidation and nitrite production over time. **(A)** representing ammonia concentration and **(B)** showing nitrite production. Cultures were carried out in triplicates and SD was calculated from the mean (*n*=3).

2.3.3.4 Effect of pH on growth of enrichment C4-2

The effect of pH on C4-2 enrichment was carried out at a reduced NH_4^+ (0.15 mM) concentration to minimise pH drift during incubation caused by NH_4^+ oxidation. Clear differences were found when C4-2 was grown at different pH ranges (*i.e.* between pH 5.5-9.0) as well as non-buffered (pH 6.5) (**Figure 2.12,** No B). MES was used in pH range 5.5-7 and HEPES was used for pH range 7.5-9. When grown under acidic pH conditions (*i.e.* pH 5.5) and at a more alkaline pH (*i.e.* pH 8.5-9.0), no growth was observed, with virtually no oxidation of NH_4^+, especially at pH 8.5 whereby NH_4^+ levels were between ca. 0.165-0.12 mM ($p > 0.05$) during the entire incubation period (**Figure 2.12A**). However, optimal pH 6.0-8.0 ranges revealed a rapid decline in NH_4^+ concentration by day 14 and nearly complete NH_4^+ depletion by day 21 ($p < 0.05$). Moreover, with all treatments, an increase in NO_2^- was observed (up to 0.35 mM) over the 49-day incubation period (**Figure 2.12B**). Interestingly, despite nearly no NH_4^+ oxidation occurring at pH 8.5-9.0, an increase from ca. 0 – 0.4 mM can be seen for pH 8.5, together with an increase in NO_2^- (ca. 0-250 µM) can be observed for pH 9.0 ($p < 0.05$) (**Figure 2.12B**). Unbuffered treatment at pH 6.5 showed an increase to pH 6.8 after complete NH_4^+ oxidation.

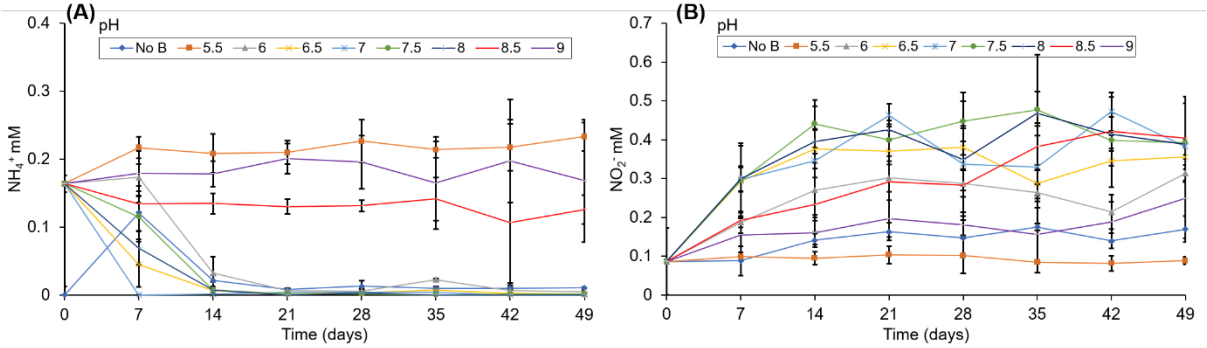


Figure 2.12. Effect of pH on ammonia oxidation **(A)** and nitrite production **(B)** over time. SD was calculated from the mean (n=3). No B: represents non-buffered.

2.4 Discussion

This study successfully enriched and isolated AOA from agricultural soils using a combination of solid and liquid media. QPCR analysis of the unamended soils showed a high abundance of AOA (ca. 4.5 x 10^7 and 3.5 x 10^7 AOA *amoA* gene copies g^{-1} dry weight) which is similar to that found in other soil environments (Sterngren, *et al*., 2015; Tago, *et al*., 2015; Jiang, *et al*., 2014). Moreover, previous studies investigating environmental factors that shape AOA ecological niches, revealed a 1.5 - 230 times greater AOA *amoA* gene copy in topsoils (0-5 cm) compared to AOB *amoA* copies (Erguder, *et al*., 2009). This study shows a ca. 30 – 50-fold great abundance of AOA gene copies g^{-1} dry weight compared to AOB gene copies g^{-1} dry weight (**Figure 2.4 A** and **B**).

The cultures in this study were incubated with 5 mM NH_4^+, which is higher than the NH_4^+ concentrations typically associated with AOA. AOA usually dominate at lower concentration of NH_4^+ in the µM range, compared to AOB which can grow in NH_4^+ concentrations at mM concentrations (Liu, *et al*., 2021). Initial inoculation from the original soils at lower NH_4^+ levels could have potentially lowered AOB co-occurrence. Despite this, AOA display a wide range of kinetic diversity, with the emergence of AOA that can tolerate up to mM concentrations of NH_4^+, such as *Nitrosocosmicus* genera (*N. franklandus* C13 (Lehtovirta-Morley, *et al*., 2016), *N. oleophilus* MYS (Jung, *et al*., 2016), *N. agrestis* SS (Liu, *et al*., 2021) and *Nitrososphaera viennensis* EN76 (Tourna, *et al*., 2011). Therefore, higher NH_4^+ cultures can result in greater AOB competition, ultimately lowering AOA abundance, with greater efforts required to eliminate AOBs, from AOA that contain a high NH_4^+ tolerance. This can be carried forward by the implementation of antibiotic cocktails, continuous subculturing and isolating AOA colonies from solid media (Klein, *et al*.,2022; Tourna et al., 2011).

Interestingly, higher NO_2^- levels were found from soil samples S1-4, especially in samples S3 and S4 (66.02 µg/g and 175.80 µg/g of NO_2^-). This finding was unusual given that NO_2^- is a reactive intermediate and is rapidly oxidised to NO_3^- (Daims, *et al*., 2016). A possible explanation for this could be a low abundance or inactive NO_2^- oxidising bacterial community that were unable to turnover the NO_2^- rapidly, resulting in the accumulation of NO_2^-. Specifically, instances have been found where NO_2^- accumulation occurred in soils when NH_4^+ oxidation was stimulated through increased in NH_4^+ availability from inorganic fertilisers such as urea and anhydrous ammonia (Giguere, *et al*., 2018). Both urea and anhydrous ammonia can negatively impact NO_2^- oxidising bacterial activity by causing spikes in local soil pH, and increasing free ammonia concentration, which is toxic to soil microbes, including NO_2^- oxidising bacteria (Giguere, *et al*., 2018; Ma, *et al*., 2015). Given that the soils in the present study were from active agricultural soils amended with regular ammonia inputs, it is possible that the *in situ* NO_2^- oxidising bacterial communities were negatively impacted in a similar manner.

During the enrichment culturing of C2 and C4, the abundance of AOA decreased from ca. 4.5 x 10^7 - 3.5 x 10^7 AOA *amoA* gene copies g^{-1} dry weight in original soil samples to ca. 3.2 x 10^4 - 1.7 x 10^4 AOA *amoA* gene copies mL^{-1} in liquid enrichment cultures. Previous studies suggest that lower AOA *amoA* genes copies mL^{-1} in liquid culture are due to the loss of soil structure that the AOA reside in (Prosser and Nicol, 2012). The soil structure can create niche environments that favour both AOA and AOB, which are removed when grown in liquid cultures.

Previous studies have shown that using gellan gum, recovered individual colonies of AOA belonging to *Nitrosocosmicus* clade (Lehtovirta-Morley, *et al*., 2016). In our study, similar individual colonies were also obtained with gellan gum. Two out of the three enrichments obtained were selected (denoted C2 and C4-2) and further characterised. Whilst

completely axenic cultures were not obtained with C2 and C4-2 cultures, due to the co-occurrence of AOB, the presence of AOB were extremely low (ca. 2900 – 3100 *amoA* gene copies per mL^{-1}) as determined by qPCR analysis of AOB *amoA* genes and further supported by *amoA* CARD-FISH.

Cell densities (based on *amoA* gene copies) in liquid culture for this study were ca.100 fold lower (10^3 – 10^4 *amoA* gene copies per mL) than other studies culturing terrestrial AOA cultures, where ca. 10^5 – 10^7 *amoA* gene copies per mL have been reported (*N. viennensis* (EN76) (Tourna, *et al*., 2011; *N. franklandus* (C13) (Lehtovirta-Morley, *et al*., 2016; *N. oleophilus* (MY3) (Sauder, *et al*., 2017; *N. agrestis* (SS) (Liu, *et al*., 2021). A possible explanation for this could be the higher NH_4^+concentrations (*i.e*. 5 mM) used in the present study also selectively enriched for AOB (which prefer higher NH_4^+ concentrations), compared with other studies that used lower NH_4^+ concentrations (≤1 mM) (Tourna, *et al*., 2011; Verhamme, *et al*., 2011; Lehtovirta-Morley, *et al*., 2016; Sauder, *et al*., 2017; Liu, *et al*., 2021). Another possible explanation could be an accumulation of H_2O_2 in the media given that AOA lack catalase genes to detoxify H_2O_2, causing oxidative stress (Kim, *et al*., 2016). A previous study investigating H_2O_2 detoxification of AOA *Thaumarchaeota* from marine (DDS1) and terrestrial habits (*N. viennensis*) (EN76) recovered greater AOA biomass when co-cultured with H_2O_2 scavengers, such as α-keto acids (pyruvate, oxaloacetate), as well as co-inoculation with bacteria that express catalase activity (Kim, *et al*., 2016). Despite this, the AOA enrichments obtained in this study did contain genes for oxidative stress defence (Chapter 3, **3.3.3**). Another reason for lower cell densities could be the continuous supplementation of antibiotics ampicillin and streptomycin, potentially eliminating required metabolites from co-cultured bacteria which are required for AOA growth (Lehtovirta-Morley, *et al*., 2016).

Both enrichments C2 and C4-2 had an irregular coccoid cell morphology and ranged between 0.8-1.2 μm in size based on CARD-FISH. Specifically, enrichment C2 (related to *N.*

franklandus, C13 (ANI 99.98%- see Chapter 3, **3.3.2**), was between 0.8-1.0 µm in size and displayed similarities with DOPE-FISH analysis previously reported for *N. franklandus,* C13 (*i.e.* with a size of 1.1 µm) (Lehtovirta-Morley *et al*., 2016). Furthermore, enrichment C4-2 also exhibited an irregular coccoid shape of a similar size to C2 (0.8-1.0 µm) with greater range (0.8-1.2 µm). Both C2 and C4-2 were similar in size to of other *Nitrosocosmicus* AOA namely *N. arcticus* Kfb (ca. 1.0 µm), *N. hydrocola* G61 (ca. 1.3 µm) and *N. oleophilus* MY3 (ca. 1.1 µm) Alves, *et al*., 2019; Sauder et al., 2017; Jung *et al*., 2016). Cells from enrichment C4-2 formed aggregates, and were similar to the aggregate structures found previously using CARD-FISH/DOPE-FISH with *N. hydrocola* G61, *N. arcticus* Kfb, *N. oleophilus* MY3 and *N. franklandus,* C13 (Lehtovirta-Morley, *et al*., 2016; Alves, *et al*., 2019; Sauder et al., 2017; Jung *et al*., 2016).

CARD-FISH was applied in the present study to visually analyse the spatial arrangement and abundance of AOA and AOB cells in enrichment cultures C2 and C4-2. Whilst the technique has several advantages such as the ability to analyse cells spatially and to target slow-growing cells like AOA, there are however several drawbacks to this method. For example, CARD-FISH requires stringent optimisation for probe hybridisation and there is a potential for false-positives from background fluorescence (Kubota, 2013). Although CARD-FISH is more sensitive than FISH, it still needs to meet a per cell rRNA threshold (Hoshino, *et al*., 2008). Hoshino *et al*. (2008) quantified the detection limits of CARD-FISH in terms of 16S rRNA molecule per cell and showed that CARD-FISH was ca. 26-41x more sensitive than conventional FISH (ca. 370 rRNA molecule/cell for fish vs ca. 8.9-14 for CARD-FISH). Whilst detection rates can vary widely across samples using FISH, previous studies using the universal eubacterial probe EUB338 revealed a detection range between 1-100% of total bacteria in aquatic systems and showed that detection is method dependent and often <100% (Bouvier, del

Giorgio, 2003; Daims, *et al*., 1999). As a result of efficiencies being <100%, CARD-FISH may result in the target AOA cells being undetected, generating false negatives.

Another limitation of CARD-FISH is the use of HRP which is a large molecule of approximately 5-6 and a molecular weight of ca. 40 kDa. Consequently, probes labelled with HRP may penetrate fixed cells poorly compared to fluorescently labelled probes (Kubota, 2013). Therefore, it is essential to optimise permeabilization before applying CARD-FISH to identify cells of interest, and optimum permeabilization is very narrow, making it a laborious process (Kubota, 2013). However, optimisation of the permeabilization step helps balance the accessibility to the rRNA molecules, with the potential loss of rRNA molecules or complete cell lysis (Molari and Manini, 2012). In addition, uncultured terrestrial AOA phylogenetic lineages contain cell wall constituents that are broadly unknown, therefore optimised permeabilization treatments remain to be discovered (Molari and Manini, 2012; Gubry-Rangin, *et al*., 2011). In this study lysozyme was used to permeabilise the cells. Alternative enzymatic and chemical treatments include proteinase K, archromopeptidase, recombinant pseudomurein endopeptidase and HCl treatments (Molari and Manini, 2012; Ishii, *et al*., 2004; Teira, *et al*., 2004; Sekar, *et al*., 2003). Proteinase K could be implemented for partial digestion of the protein rich S layer (Ebeling, *et al*., 1974). Triton X-100 and SDS also assist in disrupting the protein rich S layer, while achromopeptidase hydrolyses bonds between *N*-acetylmuramic acid and the D-amino acids which links peptidoglycan strands (Lloyd, *et al*., 2013). Therefore, archaeal cell walls that contain a thick S layer of proteins can potentially be missed using CARD-FISH. Despite this, many AOA that reside in soils do not encode S layer *slp* genes, such as high NH_4^+ tolerant AOA from *Nitrosocosmicus* genera (Han *et al.*, 2024; Jung, *et al*., 2022; Nicol *et al*., 2019). Limitations with probe coverage could also potentially miss lineages of AOA, therefore it is essential for probes to be continually updated so that AOA species of interest are being targeted (Lukumbuzya, *et al*., 2020). To gain a better insight on the

morphology of C2 and C4-2 AOA enrichments, a scanning electron microscope (SEM) could also be used which would provide more detail and greater accuracy for cellular features and size dimensions (Liu, *et al*., 2021; Sauder *et al*., 2017; Lehtovirta-Morley, *et al*., 2016).

In the present study, minimal heterotrophic growth was observed when enrichments were grown on Tryptone, Nutrient and LB agar. However, screening for contaminants using agar plates could miss viable but non culturable heterotrophic nitrifiers which could outcompete AOA and contribute to the nitrification occurring. Other studies however showed that growth of *N. viennensis* (EN76) was significantly higher with the addition of cocultured bacteria (Tourna, *et al*., 2011). In contrast, during the isolation of *N.franklandus* (C13), the addition of numerous antibiotics (ampicillin, carbenicillin, gentamycin, clindamycin and rifampicin) led to a halt in NO_2^- accumulation, presumably because of removal of metabolites produced by co-cultured heterotrophs (Lehtovirta-Morley, *et al*., 2016). Inhibition of AOA growth from heterotrophic elimination using antibiotic cocktails and serial dilution has also been shown with other AOA such as the cultivation of *N. koreensis* (MY1) from soils (Jung, *et al*., 2011).

In the present study qPCR of the AOA and AOB *amoA* genes was also performed (**Figure 2.4**) and corroborated the CARD-FISH findings. Whilst PCR and amplicon sequencing of bacterial 16S rRNA genes was attempted in the present study using the primers Bakt 341F/Bakt 805R (Herlemann *et al*., 2011), PCR products of the correct size were not obtained (data not shown). Archaeal 16S rRNA amplicon sequencing was also performed using the PCR primers 344F/915R (Raskin et al 1994; Stahl et al 1991) and whilst a strong band of the correct size was obtained, the resulting sequence data came back with mixed results (data not shown), suggesting the presence of non-specific sequences and further research into this is required.

The AOA enrichment (C4-2) was physiologically characterised in terms of its growth responses to a range of temperature, pH, NH_4^+ and NO_2^- concentrations. Enrichment C4-2

exhibited a mesophilic, neutrophilic and autotrophic nature. Optimum temperature and pH for growth from enrichment C4-2 was between 30-40°C, and pH 6-8, which is similar to other soil AOA, notably *N. franklandus* C13 (30°C-45°C, pH 6.0-8.5) (Lehtovirta-Morley, *et al*., 2016), *N. viennensis* EN76 (28°C-45°C, pH 6.0-8.5) (Tourna, *et al*., 2011), *Nitrososphaera*" sp. strain JG1 (25°C-45°C, pH 6.0-8.0) (Kim, *et al*., 2012) and *N. chungbukensis* MY2 (25°C–40°C, pH 6.5–8.0) (Jung, *et al*., 2014).

Enrichment C4-2 could tolerate growth at much higher NH_4^+ concentrations, with growth occurring in 50 mM NH_4^+, consistent with other AOA belonging to the *Nitrosocosmicus* genus (Han, *et al*., 2024). However, at much higher NH_4^+ concentrations (*i.e*. 100 mM) growth of enrichment C4-2 was inhibited. Growth of enrichment C4-2 was similar to *N. franklandus* and *N. oleophilus* MY3 which also showed growth at 50 mM NH_4^+ (Lehtovirta-Morley, *et al*., 2016; Jung *et al.,* 2016). In the 50 mM and 100 mM NH_4^+ treatments (**Figure 2.7**), NH_4^+ is very variable during exponential growth which may be because of signal saturation due to the high concentrations of NH_4^+, and could be alleviated by diluting.

Similarities can also be seen for NO_2^- inhibition experiments, with some inhibition of growth at 20 mM NO_2^- and complete inhibition of growth at 50 mM NO_2^-. Lehtovirta-Morley, *et al*., (2016) also showed inhibition of growth with *N. franklandus* C13 at ≥ 20 mM of NO_2^- which is closely related to enrichment C4-2 (See Chapter 3, **3.3.2**). Genomic analysis investigating adaptive traits in *Nitrosocosmicus* species revealed the presence of one single *amt-1* low affinity NH_4^+ transporter gene in *Nitrosocomicus* AOA genomes (*N. oleophilus* MY3, *N. franklandus,* C13 *N. hydrocola* G61 and *N. arcticus* Kfb) (Han, *et al*., 2024; Nicol, *et al*., 2019; Jung, *et al*., 2016; 2014). This is in agreement with the high NH_4^+ tolerance observed with enrichment C4-2. Contrastingly, *N. gargensis* Ga9.2 and the marine *N. maritimus* SCM1 possess both low affinity NH_4^+ transporter genes (*amt-1* and *amt-2*), while having a much lower

NH_4^+ tolerance of 3-2 mM NH_4^+, ca. 25-fold lower than C4-2 enrichment (Wright and Lehtovirta-Morley, 2023).

Interestingly, NH_4^+ depletion and NO_2^- accumulation did not show 1:1 stoichiometry during the physiological experiments, which may be because of the non-axenic nature of C4-2 enrichment cultures meaning the observed NH_4^+ oxidation may not have been attributed solely to AOA. Although metagenomic binning did not recover high confidence MAGs assigned to a NO_2^- oxidising bacterial lineage (**Table A.2/A.3**), there may have been a low abundance of NO_2^- oxidisers. To further clarify this, mining for *nxrA* and *nxrB* genes for nitrite oxidoreductase (Pester, *et al*., 2014) can be carried forward using high-performance computing to identify NO_2^- oxidising bacteria.

In this study axenic AOA isolates were not achieved. To obtain completely or near complete axenic cultures, further techniques such as extinction dilution could be performed, whereby the AOA enrichment is diluted into fresh media for single AOA cell growth (Ordonez, *et al*., 2025). A recent study plating on solid/semi-solid media under controlled conditions was successfully carried out on the soil dwelling *N viennensis* (Klein, *et al*., 2022). Yielding AOA colonies in low nutrient agar and vapor fed ammonia over long incubation periods is also another approach (Abiola, *et al*., 2024). In this study, antibiotics were used in an attempt to remove heterotrophic organisms and further research could implement a greater mixture of antibiotics or use antibiotics at higher concentrations to further eliminate heterotrophic contaminants as was done previously to obtain *N. franklandus,* C13, which used a higher concentration of ampicillin (100 µg mL^{-1}) and kanamycin (500 µg mL^{-1}) for the recovery of seven AOA subcultures (Lehtovirta-Morley, *et al*., 2016). The antibiotic concentrations used in the study by Lehtovirta-Morley, *et al*., (2016) were significantly higher than the 50 µg mL^{-1} of streptomycin and ampicillin used for this study. However, careful consideration should be taken when implementing large antibiotic cocktails at high concentrations. For example, previous

experiments to enrich *N. koreensis* MY1 using antibiotics: streptomycin (100 µg mL^{-1}), kanamycin (50 µg mL^{-1}), ampicillin (50 µg mL^{-1}), penicillin-G (50 µg mL^{-1}), gentamicin (20 µg mL^{-1}), mitomycin C (20 µg mL^{-1}), and lincomycin (50 µg mL^{-1}), led to the inhibition of both AOA and AOB (Jung, *et al*., 2011) and hence concentrations of 50 µg mL^{-1}were used herein.

During AOA enrichment pH 7.5 media was used to isolate individual AOA species, but it is not known how representative the enriched AOA were, as the original soil microbial community was not characterised. Future studies could include the use of media covering a range of pH to enrich other species of soil AOA. A previous study demonstrated significant growth of AOA compared to AOBs in soils at pH 4.5 and 6.5 (Gubry-Rangin, *et al*., 2010), and thus the use of acidic media on soils herein would likely aid in reducing AOB co-cultures to favour axenic AOA cultures. Despite this, acid tolerant AOB from soils have been found including *Nitrosoglobus terrae* (TAO100) that can survive in acidic conditions until pH 2 by forming biofilm-like cell aggregates (Hayatsu, *et al*., 2017). The use of acidic media would also select strongly against *Nitrosocosmicus* and may preferentially enrich acidophilic AOA (*e.g*. *Nitrosotalea* genus, including *N*. *devanaterra* (Nd1) (Nd2) which demonstrated growth at pH 4 – 6.1) (Lehtovirta-Morley, *et al*., 2024).

Unfortunately, due to time constraints a detailed physiological characterisation of the growth range of enrichments C2 and C4-1 could not be performed, and further research is required to investigate these AOA enrichments further. For example, to investigate O_2 affinity (Straka, *et al.,* 2019), and using inhibitors of AOA such as Allylthiourea (ATU) and 2-phenyl-4,4,5,5-tetramethylimidazoline-1-oxyl 3-oxide (PTIO) to determine the selective contributions of AOA and AOB to nitrification (Srithep, *et al*., 2018; Martens-Habbena, *et al*., 2015; Shen, *et al*., 2013).. ATU functions as a copper chelating inhibitor and can inhibit AOA growth at high concentrations (Srithep, *et al*., 2018). PTIO is a scavenger for NO, NO is an intermediate

for NH_4^+ oxidation by AOA (Shen, *et al*., 2013). Therefore, PTIO can be investigated to identify PTIO concentrations that inhibit enrichment C4-2 growth. Previous studies investigating the contribution towards nitrification of AOA and AOB carried out transcriptomic RT-qPCR to quantify archaeal/bacterial *amoA* genes using acetylene to inhibit AOA growth (Gubry-Rangin, *et al*., 2010). Alternatively, 1-octyne could be used to exclusively inhibit AOB growth (Lin, *et al*., 2021). Numerous AOA are also capable of using urea as N source such as *N. viennensis* (EN76) (Tourna, *et al*., 2011) and *N. franklandus* (C13) (Lehtovirta-Morley, *et al*., 2016). Incubating AOA enrichments from this study in different urea concentrations would also provide insights on whether these enrichments utilise urea as an N source and their optimal urea concentrations for growth. However, further work, including *amoA* and bacteria/ archaeal 16S rRNA amplicon sequencing, together with whole genome sequencing would provide a better understanding of their taxonomy and physiological characteristics.

Conclusions

In conclusion, this study utilised a combination of culturing (*e.g*. antibiotic selection, gellan gum), microscopy and molecular techniques to successfully cultivate, enrich and characterise three AOA enrichments denoted C2, C4-1 and C4-2. CARD-FISH showed that enrichments C2 and C4-2 were irregularly shaped cocci and were between 0.8-1.2 µm in size. Physiological characterisation showed that the AOA enrichment C4-2 was mesophilic with growth at temperatures between 30-40°C with an optimum pH range between 6 and 8. Enrichment C4-2 also demonstrated a tolerance to higher concentrations of NH_4^+, with growth occurring at up to 50 mM of NH_4^+ with similar physiological characteristics to members of the *Nitrosocosmicus* clade. Enrichment C4-2 also exhibited NO_2^- inhibition up to 20 mM, which was similar to that found in *N. franklandus* (C13) (Lehtovirta-Morley, *et al*., 2016). Amplicon and genome sequencing of AOA enrichments C2, C4-1 and C4-2 is presented in Chapter 3,

3.3.1 and reveals new insights into the genes responsible for carbon fixation pathway 3-hydroxypropionate-4-hydroxybutyrate (3HP-4HB) cycle, oxidative stress defence, urea metabolism, ammonia oxidation and ammonia transport.

Chapter 3 Amplicon and Genome Sequencing of Putative Ammonia-Oxidizing Archaeal Isolates.

3.1 Introduction

Ammonia oxidising archaea (AOA) are key players in soil nitrogen cycling, where they are responsible for the oxidation of NH_4^+ to NO_2^-, (*i.e.* the first and rate-limiting step in nitrification) (Zhong *et al*., 2023; Huang *et al*., 2021; Ouyang, *et al*., 2018). AOA in soils are phylogenetically diverse with many different clades such as *Nitrosocosmicus* (Lehtovirta-Morley *et al*., 2016), *Nitrososphaera* (Bei, *et al*., 2024) and *Nitrosomirales* (Zheng, *et al*., 2024). Within the *Nitrosocosmicus* clade, *Nitrosocosmicus franklandus* (C13) exhibits a high tolerance to environments with elevated levels of NH_4^+ and has not only been isolated, but has been fully genome sequenced (Nicol *et al*., 2019; Lehtovirta-Morley *et al*., 2016). Agricultural soils which have high fertiliser inputs, may explain why members of the *Nitrosocosmicus* clade can be cultivated under higher NH_4^+ and NO_2^- conditions (Han et al 2024).

Isolating AOA species in liquid culture from soil is considerably more challenging than isolating AOA from marine environments, or indeed isolating AOB from soils (Lehtovirta-Morley *et al*., 2016; Bollmann *et al*., 2011). This is because soils are heterogeneous with highly diverse microbial communities that co-exist with the AOA, including the faster-growing bacteria (AOB), heterotrophic nitrifiers and fungi, which can outcompete or co-occur with the AOA in liquid culture (Wang *et al*., 2019b; Zhalnina *et al*., 2012; Bollmann *et al*., 2011; Morimoto *et al*., 2011). To bypass these issues, techniques can be implemented, including the use of selective media to favour AOA, addition of antibiotics, and growing AOA colonies

within solid medium (Klein *et al*., 2022; Lehtovirta-Morley *et al*., 2016). In addition to the challenges in isolating AOA from soil environments, obtaining a full AOA genome is also difficult, largely due to the co-occurrence of AOB (Qiang *et al*., 2024). Genome-resolved metagenomics can be employed to recover high quality metagenome-assembled genomes (MAGs) (Sarkar *et al*., 2025; Lee *et al*., 2023; Vuong *et al*., 2022) from short read (2x150 bp paired-ends) (Illumina NovaSeq 6000) and whole genome sequencing (Novogene). Whole genome sequencing alongside PCR amplification of highly conserved AOA genes, such as *amoA* and 16S rRNA, can be implemented to study the phylogenetic relationship of potentially novel AOA compared to known species (Moin *et al*., 2009; Purkhold *et al*., 2000). Genome sequenced cultures that display high microbial diversity can also still provide useful metagenomic data in the form of high-quality MAGs which can then be screened for key metabolic proteins and specific taxa.

3.1.1 Aims

The aim of this Chapter is to characterise the genome (or produce high-quality MAGs) of three AOA isolates (C4-2, C4-1 and C2), obtained from agricultural soils to expand the genomic information of AOA for comparison with previously characterised AOA, such as *N. franklandus*. Bioinformatic analysis of genome sequence data, together with PCR amplicon sequencing of *amoA* and 16S rRNA genes, was performed to study the phylogenetic and functional relationships and compare to other known AOA.

3.2 Methods

3.2.1 Amplicon sequencing and phylogenetic analysis of AOA *amoA* and 16S rRNA genes

Cells from AOA isolates (C4-2, C4-1 and C2) were pelleted by centrifugation at 20,000 x g for 5 minutes (4℃). DNA was extracted from cell pellets using a SoilProDNA extraction kit (Qiagen), following the manufacturer's instructions, and stored at -20℃. PCR amplification was performed using the primers 21F/1492R (Delong *et al*., 1992; Lane *et al*., 1991) for Archaeal 16S rRNA gene, CrenamoA-23F/CrenamoA-616R (Tourna, *et al*., 2008) for AOA *amoA* gene, and amoA-1F/amoA-2R (Rotthauwe, *et* al., 1997) for AOB *amoA* gene. PCR thermocycling was performed using a 96-well thermocycler (Verti^{TM} Applied Biosystems) as follows: initial denaturation of 95℃ for 3 minutes followed by 35 cycles of 95℃ for 5 seconds, annealing temperature of 56℃ for 30 seconds, elongation temperature of 72℃ for 1 minute and a final elongation step for 5 minutes. DNA was visualised on a 1% *(w/v)* TBE agarose gel and purified using a GenElute^{TM} PCR Clean-Up Kit (Merck). Purity of PCR products was re-checked on a 1% *(w/v)* TBE agarose gel and sequenced (Eurofins). Taxonomic affiliation of archaeal cultures was analysed using Blastn to identify closest matches. Dr Fabien Cholet (University of Glasgow) assisted with the phylogenetic analysis of this study, by providing aligned *amoA* nucleotide sequences from known AOA genera that were retrieved from the database as constructed in Cholet *et al*., (2024) and translated to amino-acid using a custom R scripr (ORF.finder function; available at https://github.com/Fchlt/ARC).

3.2.2 Whole Genome Sequencing (Novogene) of AOA isolates

Genomic DNA was extracted from AOA isolates (C4-2, C4-1 and C2) (as described in Section 3.2.1) and sequenced using the Illumina NovaSeq 6000 platform, generating paired-end short reads (2x150 bp) (Novogene, Cambridge, UK). Library preparation was carried out

using the standard Illumina-compatible kit where raw reads were subjected to quality control, removal of adaptors and low-quality base trimming. After clean reads were assembled into contigs using a proprietary *de novo* assembler from Novogene, draft assemblies were made for each C4-2, C4-1 and C2 samples. As AOA were not completely pure with other co-cultured microbes still present in culture, further bioinformatic processing was carried out to identify key genes/proteins and assemble a high-quality metagenome-assembled genome (MAG) from potentially novel AOA species.

All computational analysis was performed using the University of Essex high performance computing cluster (https://hpc.essex.ac.uk/Grants.hpc). Raw sequence reads generated by Novogene were quality-filtered and adaptor-trimmed using Fastp v0.26.0 (Chen, 2023; Chen *et al*., 2018) with default parameters. High-quality paired-end reads were then suitable for downstream applications. Quality filtered reads were assembled *de novo* into contigs using MEGAHIT v1.2.9 (Li *et al*., 2015), with the optimised –meta preset for complex metagenomic datasets. Nucleotide contigs were translated into amino acids using gene prediction tool Prodigal v2.6.3 (Hyatt *et al*., 2010) in metagenomic mode (-p meta) for archaeal *amoA* identification. Amino acid sequences were screened for archaeal *amoA* using HMMER v3.3.2 (Finn *et al*., 2011) searching against Pfam HMM (hidden Markov model) profile (PF12942) without applying additional threshold cutoffs. High-confidence hits meeting full-sequence E-value $\leq 1\times10^{-5}$, bit score ≥ 30 hits were filtered. In parallel, archaeal 16S rRNA genes were screened using nhmmer, within the HMMER v3.3.2 suite (Finn *et al*., 2011) with Rfam covariance model RF01959. High scoring sequence were filtered the same as previously described, sequence E-value $\leq 1\times10^{-5}$, bit score ≥ 30. AOA *amoA* and 16S rRNA gene sequences were recovered from the genome sequence datasets and underwent Blastn analysis against core nucleotide database (core_nt), with default parameters.

Nucleotide sequences were translated to amino acid using custom R script (ORF.finder function; available at https://github.com/Fchlt/ARC). Additionally, for 16S rRNA, nucleotide sequences of known AOA genera were aligned with 16S rRNA sequences from MiMt database (Cabezas *et al*., 2024). Blastn search was carried out using 16S rRNA nucleotide sequences, acquired from the genome sequence datasets, from isolates C4-2, C4-1 and C2. Full-length *amoA* and 16S rRNA sequences were aligned with reference sequences using Mafft (Katoh *et al*., 2009) and phylogenetic trees was inferred by maximum likelihood using Fasttree (Price *et al*., 2010) courtesy of Dr Fabien Cholet (University of Glasgow) using the reference dataset compiled by Cholet *et al*. (2024).

3.2.3 Metagenome-Assembled Genome Construction of AOA

High quality metagenome-assembled genomes (MAGs) were constructed from raw sequence reads generated by Novogene. A multi-step computational pipeline was implemented using University of Essex high performance computing cluster, as described previously. Paired end reads were again quality filtered using Fastp v0.26.0 (Chen, 2023; Chen *et al*., 2018) and reads were assembled separately with MEGAHIT v1.2.9 (Li *et al*., 2015) as previously described for downstream applications. Coverage profiles were generated by mapping back quality-controlled reads to assembled contigs using Bowtie2 v2.4.5 tool (Langmead *et al*., 2012), (Langmead *et al*., 2019). Samtools v1.16.1 (Danecek *et al*., 2021) was used to convert and sort alignment files and coverage depth per contig was calculated using the jgi_summarize_bam_contig_depths script from MetaBAT2 v2.18 (Kang *et al*., 2019). Generated depth file is based on differential coverage and contigs were filtered using a minimum contig length of 1,000 bp and clustered into genomic bins using binning tool MetaBAT2 v2.18 (Kang *et al*., 2019). Generated bins for C2, C4-1 and C4-2 were manually screened for archaeal *amoA*, of which C2 and C4-2 possessed archaeal *amoA* gene.

MAG completeness and contamination were assessed using both MiGA (Microbial Genome Atlas) (Rodriguez-R *et al*., 2018) and DFAST (Tanizawa, *et al*., 2016), and BUSCO v5.4.4 (Manni, *et al*., 2021), together with the archaeal lineage dataset (archaea_ob10) to estimate genome completeness. Taxonomic classification of MAGs was performed by carrying GTDB taxonomy assignment using DFAST to confirm the MAGs are affiliated with known AOA genera. Further functional gene annotation was performed using Prokka v1.14.6 (Seemann, 2014), to identify coding sequences (CDS), tRNAs and rRNAs. A circular genome map was generated for isolates C4-2 and C2 MAGs using GenBank-formatted outputs produced from Prokka v1.14.6 (Seemann, 2014), using Proksee (Grant *et al*., 2023). Due to poor sequence coverage in C4-1, which was significantly lower than in isolates C2 and C4-2, a circular genome map for isolate C4-1 was not performed.

3.2.4 Screening of Functional Genes for Metabolism and Stress Response

Targeted protein screening was conducted to identify functional proteins for ammonia oxidation, transport, urea metabolism and copper binding/homeostasis. The nucleotide sequences of C4-2 and C2 MAGs were translated into amino acids using Prodigal v2.6.3. Functional annotation was carried out using Prokka v1.14.6 (Seemann, 2014); however, a significant number of target proteins were annotated as hypothetical. To specifically identify target proteins, HMMER v3.3.2 (Finn *et al*., 2011) was employed to identify target amino acid sequences against Pfam profiles. Custom Pfam HMM profiles were obtained through the InterPro website (Blum, *et al*., 2025), HMM profiles were associated with key AOA metabolic functions as listed in **Table 3.3**. HMMER searches were executed using the hmmsearch utility with all hits being filtered to retain only those with e-values = 1e-5 and high alignment scores. While **Table 3.3** represents successful hits, a broader range of Pfam HMM profiles were tested that carried out the same metabolic functions. This Whole Genome Shotgun project has been

deposited at DDBJ/ENA/GenBank under accessions JBPUGK000000000 for the *Candidatus Nitrosocosmicus* sp. C4-2 MAG and JBPUGL000000000 for the *Candidatus Nitrosocosmicus* sp. C2 MAG, under BioProject PRJNA1285782.

3.3 Results

3.3.1 PCR Amplification of 16S rRNA and AOA *amoA* genes

PCR amplification of Archaeal 16S rRNA genes was performed but despite several attempts, bands of the correct size could not be detected using the primers 21F/ 1492R (Delong, *et al*., 1992; Lane, *et al*., 1991) (data not shown). However, when PCR amplification of the AOA *amoA* gene was performed, bands of size between 750-1000 bp were obtained rather than the expected size of ca. 624 bp (Tourna, *et al*., 2008) (**Figure 3.1**). With isolate C4-1, the *amoA* PCR product obtained was slightly bigger compared to isolates C2 and C4-2 (**Figure 3.1**).

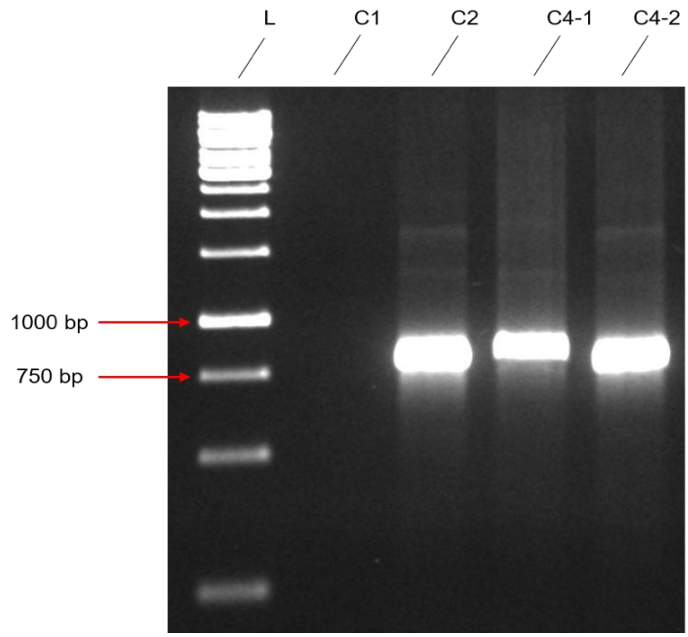


Figure 3.1. PCR amplification of archaeal *amoA* genes from soil isolates, visualised on a 1% *(w/v)* TBE agarose gel. GeneRuler 1 kb DNA ladder (ThermoFisher) **(L)** with AOA isolates denoted as **C1**, **C2**, **C4-1**, and **C4-2**.

The three AOA *amoA* gene PCR products obtained from isolates C2, C4-1 and C4-2 were sequenced (Eurofins) and a Blastn analysis performed (**Table 3.1**). Amplicons ranged from 617-667 bp in length (**Table 3.1**). Each isolate had a relatively high % similarity to either uncultured AOA *amoA* gene sequences or known AOA species, ranging from 92-93% identity (**Table 3.1**). Isolate C2 shows had 93% identity to an uncultured ammonia oxidising archaeon clone J21107021_1-23_M13F (KJ496830.1), and *Nitrosocosmicus franklandus* (NFRAN1) (LR216287.1). Both isolates C4-1 and C4-2, had 92% identity to an uncultured archaeon clone SLFA044 (KF667015.1) and *Nitrosocosmicus hydrocola* strain (ZSH-B). C4-1 and C4-2 had 92% identity for uncultured archaeon clone SLFA044 (KF667015.1) and *Nitrosocosmicus hydrocola* strain (ZSH-B). E-values of 0.0 for all hits indicate highly significant matches. Blastn results suggest that these archaeal isolates are closely affiliated with *Nitrosocosmicus* within the Thaumarchaeota lineage.

PCR amplified *amoA* sequences from isolates C2, C4-1 and C4-2 were aligned with reference *amoA* sequences. All isolates C2, C4-1 and C4-2 were found to cluster within *Nitrosocosmicus* clade (**Figure 3.2**). However, isolates C2 and C4-1 grouped most closely with *N. franklandus* and *N. oleophilus* with a high bootstrap support of 99-100% (**Figure 3.2**). Interestingly, C4-2 revealed a longer branch from C4-1 both of which diverged from isolate C2 within the sub-clade (**Figure 3.2**).

Table 3.1. Blastn analysis of AOA isolates C2, C4-1 and C4-2 based on the PCR amplification and sequencing of the *amoA* gene. Amplicon size represents size of amplicon acquired from PCR, displayed in **Figure 3.1**.

Isolate	Amplicon Size (bp)	Name	% Identify	E-Value	Accession No.	Reference
C2	617	Uncultured ammonia oxidising archaeon (Freshwater Sediment)	93	0.0	KJ496830.1	(Bollmann, *et al.*, 2014)
C2	617	Candidatus *Nitrososcosmicus Franklandus* isolate NFRAN1	93	0.0	LR216287.1	(Nicol, *et al.*, 2019)
C4-1	622	Uncultured ammonia oxidising archaeon (Sewage Sludge)	92	0.0	KF667015.1	(Yan, *et al.*, 2015)
C4-1	622	Candidatus *Nitrososcosmicus hydrocola* strain ZSH-B	92	0.0	MW009676.1	(Sauder, *et al.*, 2017)
C4-2	667	Uncultured ammonia oxidising archaeon (Sewage Sludge)	92	0.0	KF667015.1	(Yan, *et al.*, 2015).
C4-2	667	Candidatus *Nitrososcosmicus hydrocola* strain ZSH-B	92	0.0	MW009676.1	(Sauder, *et al.*, 2017).

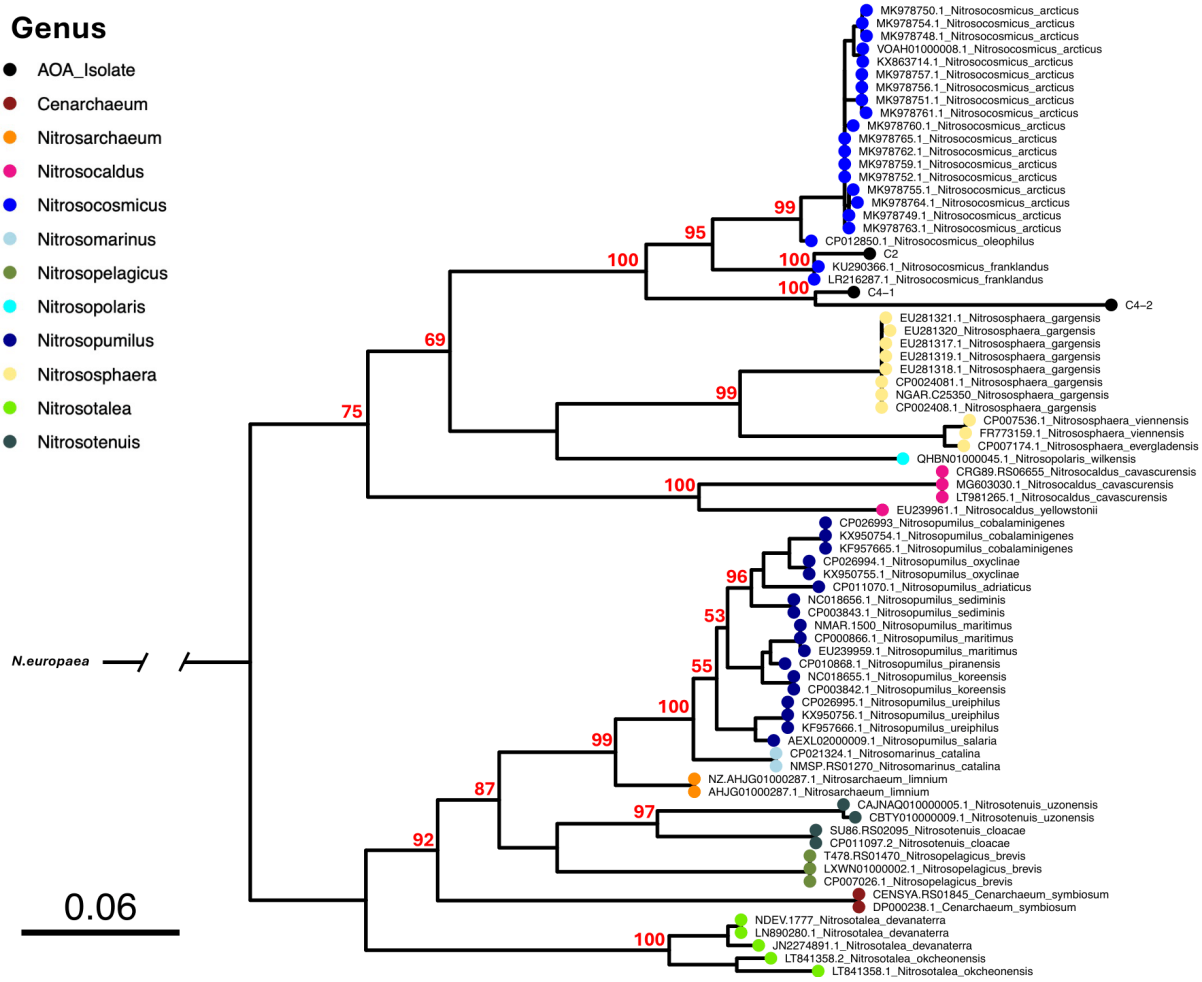


Figure 3.2. Maximum likelihood phylogenetic tree based on the full-length *amoA* PCR amplicon sequence, annotation shows isolates (C4-2, C4-1 and C2) within the *Nitrosocosmicus* clade. Sequences colour coded by genus and bootstrap values of ≥70% shown at branch nodes based on 1000 resampling's. Scale bar represents 0.06 substitutions per base. Figure generated courtesy of Dr Fabien Cholet, (University of Glasgow).

3.3.2 Assembly and Annotation of AOA Metagenome-Assembled Genome

Whole genome sequencing was carried out and high-quality MAGs were obtained following quality filtering, assembling, coverage statistics and genome binning, AOA genomes were isolated by manually searching for archaeal *amoA* genes (See Section 3.2.3). Whilst isolates C4-2 and C2 generated hits for *amoA* genes, isolate C4-1 did not provide sufficient coverage to identify any AOA bins. Isolated MAGs (from Enrichments C2 and C4-2) exhibited high completeness with low contamination and were taxonomically assigned as *Nitrosocosmicus* species. Enrichment C4-1 was excluded from further analysis due to inability to locate any archaeal *amoA* genes generated bins, as coverage was significantly lower than in isolates C2 and C4-2. **Table 3.2** shows very similar genome characteristics for both Enrichments C2 and C4-2, with C4-2 being ca. 400 kbp larger than C2 (**Table 3.2** and **Figure 3.3**). Notably, C2 shows a very high ANI % to *Nitrosocosmicus franklandus* (99.98%), whereas C4-2 contains a very high ANI% to an uncharacterised *Nitrosocosmicus* species, indicative of an uncharacterised AOA species.

Stand-alone Blast searches were carried out to identify other amoA subunits, such as amoB, C, X, Y and Z, within C2 and C4-2 MAGs. Identity matches of 100% were revealed in C2 for amoB, amoC and amoX, with no results found for subunits amoY and Z (**Table A.2**). C4-2 matches revealed 100 % identity for amoB and C, amoC match corresponds to a hypothetical protein and has been treated as a putative amoC associated hit (**Table A.2**. The amoX match shows a 94.6% identity to *N*, *franklandus* (C13), suggesting C2 being related to *N*. *franklandus* (C13) species, whereas C4-2 enrichment revealed a potentially novel AOA, similar to *N*. *hydrocola* (G61).

Bins recovered from C2 and C4-2 MAGs were assessed for taxonomic classification and closest reference genome assignment using GTDB-Tk (Chaumeil *et al*., 2022). In conjunction, CheckM2 (Chklovski, *et al*., 2023) tool was used for Bin quality check to

determine bin completeness and contamination. MAG quality statstics, including ANI %, alignment fraction, genome size (Mb), GC %, total contigs and N50 values are provided in appendix (**Table A.3** and **A.4**). A total of 13 bins were recovered from C2 MAGs and 8 bins were recovered from C4-2 MAGs. Both archaeal (*Nitrosocosmicus*) and bacterial lineages were represented for C2/C4-2 bins. In C2, bin9 was assigned to *N. franklandus* (C13), with an ANI % of 99.98% and an alignment fraction of 0.9758. MAGs were very complete with a 99.35% completeness with very little contamination of 0.03%, indicative of high-quality AOA MAGs. Several high completeness bacterial MAGs were also recovered from C2, affiliated with *Sphingomicrobium*, *Phenylobacterium*, *Nocardioides* and *Mesorhizobium*. However, several C2 MAGs did not contain an assigned closest reference genome using GTDB-Tk (Chaumeil *et al*., 2022) (**Table A.3**). C4-2 enrichment revealed a single archaeal lineage with genus *Nitrosocosmicus* of unknown species, with a closest reference genome (GCF_008389435.1), ANI % of 97.06% and an alignment fraction of 0.8038. CheckM2 (Chaumeil *et al*., 2022) estimated a completeness of 98.99% with a low contamination of 0.33%, indicative of a high-quality archaeal genome. Similar to C2 MAGs, C4-2 also revealed many bacterial lineages with high completeness, such as *Planctomycetota* and *Chloroflexota*. Contrastinly, C4-2 MAGs showed numerous low completeness MAGs which can be treated as partial genomes (**Table A.4)**.

Table 3.2. Assembly and annotation for isolate C4-2 and C2 AOA MAGs

Assembly Parameters	C4-2	C2
Total genome size (bp)	3,192,647	2,784,558
Number of contigs	106	117
Maximum contig length (bp)	112,918	102,237
Minimum contig length (bp)	2,603	2,624
Average contig length (bp)	30,119.3	23,799.6
N50 (bp)	49,422	37,242
GC %	33.34	33,99
BUSCO Completeness %	90.7	90.7
MiGA Completeness %	95.7	95.7
MiGA Contamination %	0.0	0.0
Gene Details		
Number of protein coding genes	3,458	2,986
Average gene length (bp)	652.92	674.796
Gene density per Mb	1,083.29	1,072.34
Protein Features		
Total number of predicted proteins/genes	3,458	2,986
Total number of annotated proteins/genes (not hypothetical)	898	831
tRNAs	21	22
rRNAs	1	0
GTDB Parameter		
Closest reference genome	GCF_008389435.1	GCF_900696045.1
GTDB species	***Nitrosocosmicus* sp008389435**	***Nitrosocosmicus Franklandus***
ANI to reference %	**97.06**	**99.98**
GTDB Taxonomy	Archaea; Thermoproteota; *Nitrosocosmicus*	Archaea; Thermoproteota; *Nitrosocosmicus*

Genomic maps give insight into genome architecture and feature distribution, protein-coding sequences (CDSs), tRNAs, rRNAs, together with GC content and GC skew across assembled genomes. Here, Genbank files generated with Prokka v1.14.6 (Seemann, 2014) gene annotation tool on C4-2 and C2 were used to produce circular genomic maps (**Figure 3.3)** using Proksee (Grant *et al*., 2023). The assembled MAGs of C4-2 revealed a size of ca. 3.19 Mbp, compared to C2 which was ca. 2.78 Mbp. **Figure 3.3** shows distinct regions of skew variation

and local GC content, with both showing a large and even distribution of CDSs around the genomic map.

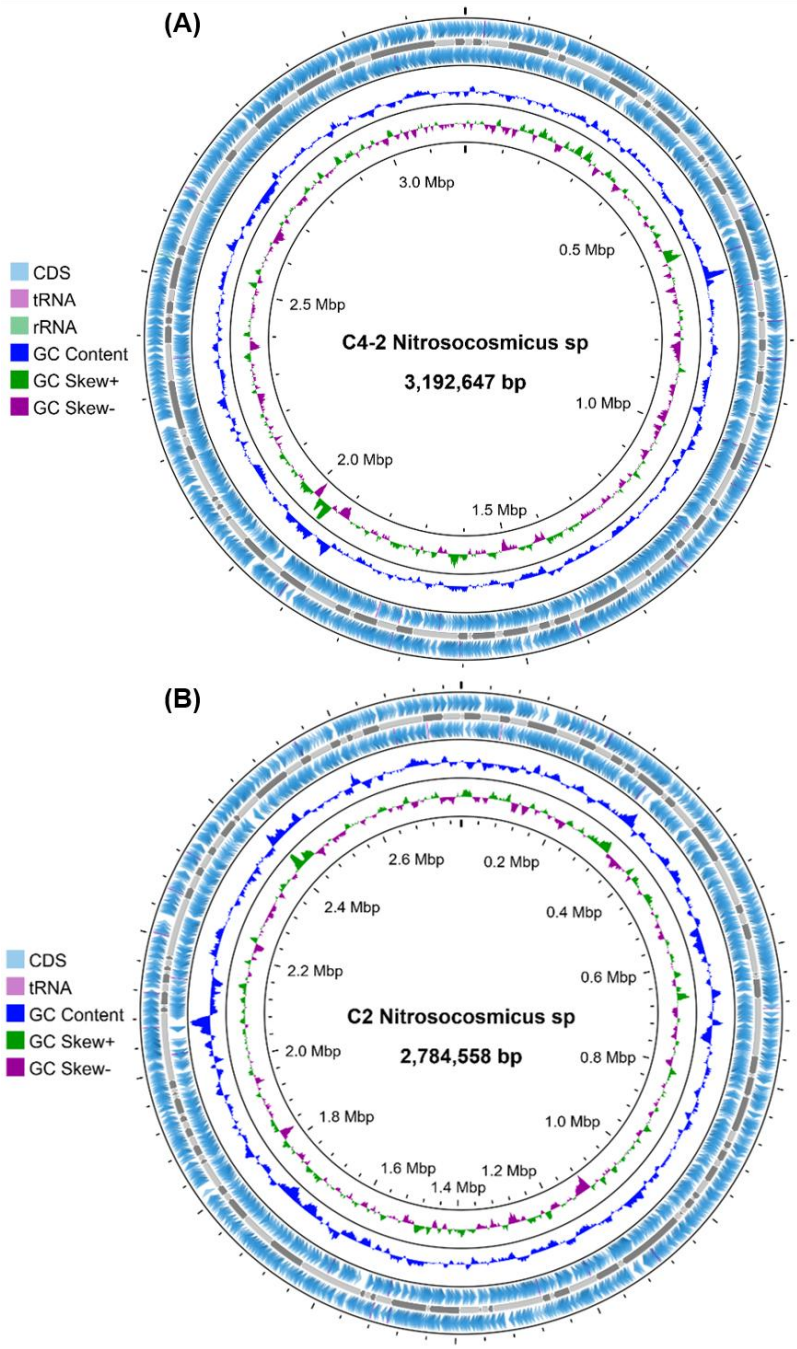


Figure 3.3. Circular genomic map of isolates C4-2 **(A)** and C2 **(B)**. The outer rings indicate the positions of predicted protein-coding genes (CDS, light blue), tRNAs (pink), rRNAs (green) (only found in C4-2 **(A)**. The inner rings represent GC content (dark blue) and GC skew+ (green) and GC skew- (purple). Genome size ruler is located in the centre.

3.3.3 Screening of metabolic and adaptive proteins in AOA MAGs

Functional annotations of *Nitrosocosmicus* MAGS were performed using HMMER v3.3.2 (Finn *et al*., 2011) with curated Pfam HMM profiles from the InterPro (Blum, *et al*., 2025) database to identify essential metabolic and stress response capabilities. A wide range of Pfam HMM profiles were used to screen for ammonia oxidation/transport, carbon fixation, urea metabolism and copper binding. High confidence hits were selected corresponding to their Pfam HMM profile and quantified (**Table 3.3**).

Analysis revealed the presence of ammonia monooxygenase subunit A protein (amoA). Interestingly, the amino acid sequences of the amoA subunit were identical for all three isolates, despite differences in the 16S rRNA gene sequences (**Table 3.3**). Other ammonia monooxygenase (AMO) subunits were screened (amoB, amoC, amoX, amoY and amoZ). However, HMMER v3.3.2 (Finn *et al*., 2011) was unable to obtain any hits for these from the Pfam HMM profiles. The gene coding for the low affinity ammonia transport protein Amt-1 was however identified, with both C4-2 and C2, revealing one Amt-1 copy (**Table 3.3**). No genes were identified for canonical S-layer proteins (*slp*) or cell-anchoring gene (*slaB*), proposed to aid in capturing NH_4^+ from the environment (Han *et al*., 2024; Wright and Lehtovirta-Morley, 2023). However, proteins related to urea metabolism were also recovered in both isolates (Urease alpha subunit), suggesting the potential ability of *Nitrosocosmicus* species to use urea as an alternative nitrogen source, converting the urea to NH_4^+ for chemolithoautotrophic metabolism. Key enzymes involved in the carbon fixation pathway, *i.e.* 3-hydroxypropionate-4-hydroxybutyrate (3HP-4HB) cycle which is widely used by AOA (Hu, *et al*., 2024a; Shi, *et al*., 2023; Destan, *et al*., 2021) were also identified (**Table 3.3**). MAGs from both isolates contained Enoyl-CoA Hydratase and Methylmalonyl-CoA Mutase for the operation of the 3HP-4HB cycle, ultimately reflecting the capacity for autotrophic growth *via*

archaeal-specific carbon fixation. Isolate C4-2 showed double the abundance of these proteins (2 copies) compared to C2 (1 copy) (**Table 3.3**).

A high abundance of proteins involved in oxidation stress defence were also revealed. In particular Iron Superoxide Dismutase was found with 1 copy in both isolates. In addition, Thioredoxins were also found in greater abundance with 3 copies identified in both isolates, whilst AhpC-type peroxiredoxins were found with 6 and 8 copies in C4-2 and C2, respectively (**Table 3.3**). Iron superoxide dismutase is responsible for converting reactive superoxide radicals ($O_2^{\bullet-}$) to $H_2O_2 + O_2$, preventing oxidative damage (Sheng, *et al*., 2014). Thioredoxins maintain the reduced states of protein cysteine, repairs oxidative damage and acts as an electron donor to peroxide-removal enzymes (Hasan, *et al*., 2022). AhpC-type peroxiredoxins reduces H_2O_2 and organic hydroperoxides to H_2O and alcohols, supporting H_2O_2 detoxification after iron superoxide dismutase produces H_2O_2 (Morris, *et al*., 2022; Zhang, *et al*., 2019).

A high abundance of proteins involved in copper binding and copper homeostasis were identified in both C4-2 and C2. Specifically, Multicopper oxidases (MCOs) for copper detoxification and electron transfer were found, with isolate C4-2 having a greater abundance of MCOs with ten identified compared to seven found in isolate C2 (**Table 3.3**). Heavy metal-associated (HMA) domain proteins for intracellular metal chaperoning to transfer copper ions to enzyme targets or exporting systems were identified, with one copy found in C2 and C4-2 isolates (**Table 3.3**). Plastocyanin/azurin-like cupredoxins were found in a high abundance in both isolates, with C4-2 (29 in total) containing five more copies than C2 (24 in total). Additionally, several copper resistance proteins associated with copper homeostasis were identified in both isolates, including CopC, CopD, the and P-type ATPase CopA; however, isolates C4-2 containing an additional copy of CopC and CopD compared with isolate C2. Cytochrome c oxidase is another key copper binding protein for the electron transport chain in *Nitrosocosmicus*, with one copy found in both C2 and C4-2 isolates. However, for both isolates

C4-2 and C2, no genes were discovered associated with chalkophore biosynthesis, for copper limited conditions (Reyes, *et al*., 2020b).

Table 3.3. Key functional proteins from C4-2 and C2 MAGs

Functional Category	Protein	Pfam ID	C4-2 Copies	C2 Copies
Ammonia Oxidation	amoA	PF12942	1	1
Ammonium Transport	Amt	PF00909	1	1
Carbon Fixation	Enoyl-CoA hydratase	PF00378	2	1
	Methylmalonyl-CoA mutase	PF01642	2	1
Urea Metabolism	Urease alpha subunit	PF00449	1	1
ROS Defence	Iron Superoxide Dismutase	PF02777	1	1
	Thioredoxins	PF00085	3	3
	AhpC-type peroxiredoxins	PF00578	6	8
Copper Binding	Multicopper Oxidases	PF07731, PF07732	10	7
	Cupredoxins	PF00127	29	24
	CopC	PF04234	2	1
	CopD	PF05425	3	2
	Heavy Metal-Associated domain Protein	PF00403	1	1
	Cytochrome C Oxidase	PF00115	1	1
	P-type ATPase (CopA)	PF00122	1	1

3.3.4 Phylogenetic Analysis of AOA Isolates based on 16S rRNA and amoA Sequences obtained *via* Genome Sequencing

Archaeal *amoA* gene sequences were identified from the genome sequences of isolates C4-2, C4-1 and C2 using HMMER v3.3.2 (Finn *et al*., 2011) tool and Pfam HMM profile (PF12942). Phylogenetic analysis of the *amoA* gene sequences revealed a monophyletic group within the *Nitrosocosmicus*, which clustered with *N. oleophilus*, *N. franklandus* and *N. arcticus* (but with low bootstrap support: 73%) (**Figure 3.4**). Notably, the marine AOA such as *Nitrosopumilus*, *Nitrosopelagicus*, *Nitrosarchaeum* and *Cenarchaeum* form phylogenetically distinct clades, which clustered away from other terrestrial AOA genera (*i.e*. *Nitrosospharea*, *Nitrosocaldus*, *Nitrosocosmicu*s and isolates C4-2, C4-1 and C2 (**Figure 3.4**). This was also found with the analysis of the 16S rRNA gene for these taxa (**Figure 3.5**).

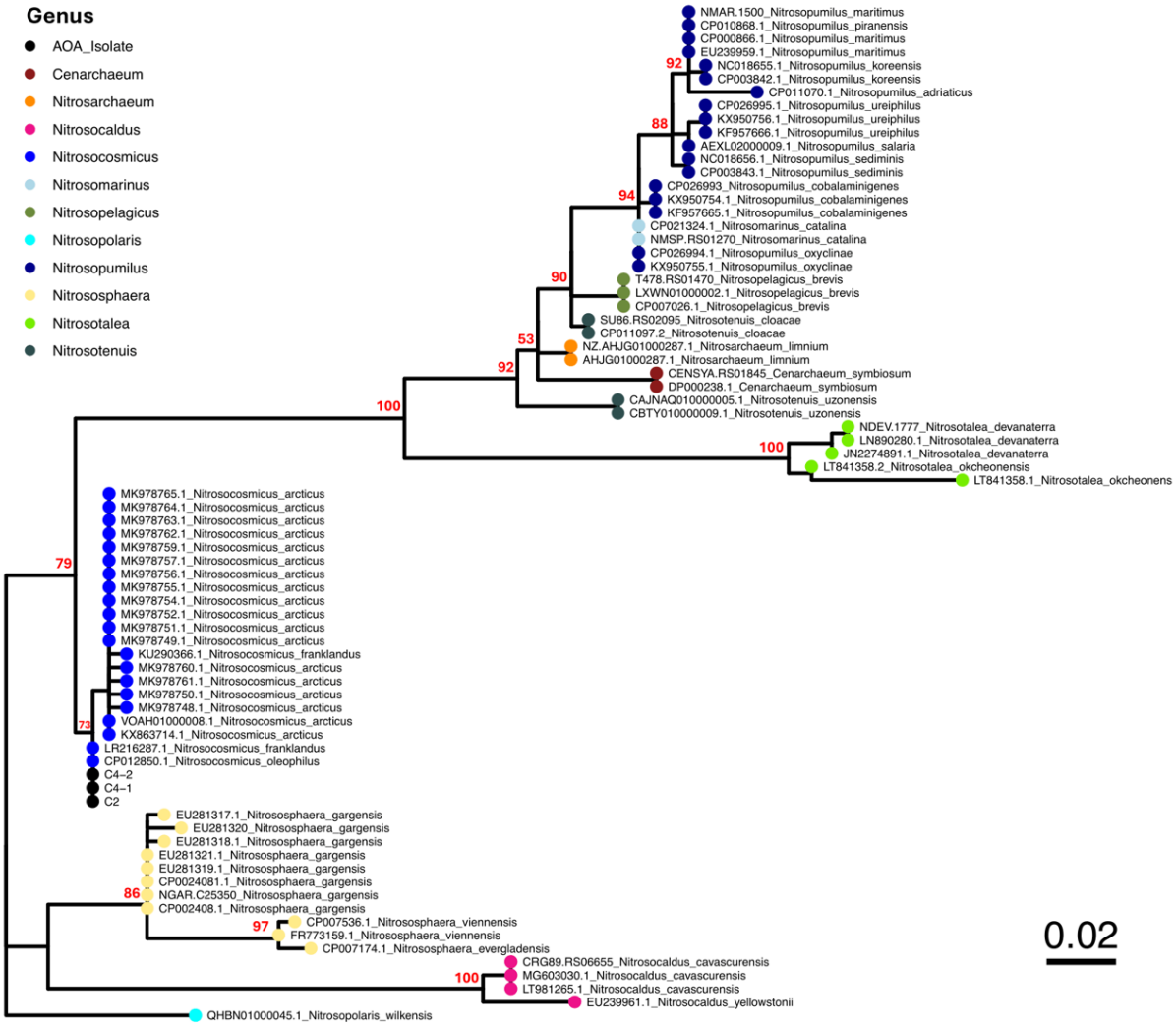


Figure 3.4. Maximum likelihood phylogenetic tree based on the full-length amoA amino acid sequences recovered from genome sequence annotation from isolates C4-2, C4-1 and C2 within the *Nitrosocosmicus* clade. Sequences colour coded by genus and bootstrap values of ≥70% shown at branch nodes based on 1000 resampling's. Scale bar represents 0.02 substitutions per base. Figure generated courtesy of Dr Fabien Cholet (University of Glasgow).

Sequences for the 16S rRNA gene were also recovered from genome sequence annotation from isolates C4-2, C4-1 and C2 and assembled into contigs using nhmmer with Rfam HMM profile (RF01959) and compared against the 16S rRNA gene sequences from known AOA genera from the MiMt database (Cabezas *et al*., 2024). Phylogenetic analysis based on the recovered 16S rRNA gene sequences showed that all three isolates formed a monophyletic clade within the *Nitrosocosmicus*, with strong bootstrap support (100%) and therefore supported the topology obtained with the *amoA* phylogeny. Specifically, isolates C4-2 and C4-1 clustered very closely with *N. hydrocola*, whereas C2 was more closely related to *N. franklandus* (**Figure 3.5**). Therefore, phylogenetic analysis of the *amoA* and 16S rRNA gene showed good congruence and supported the taxonomic identity of these isolates within the *Nitrosocosmicus* clade.

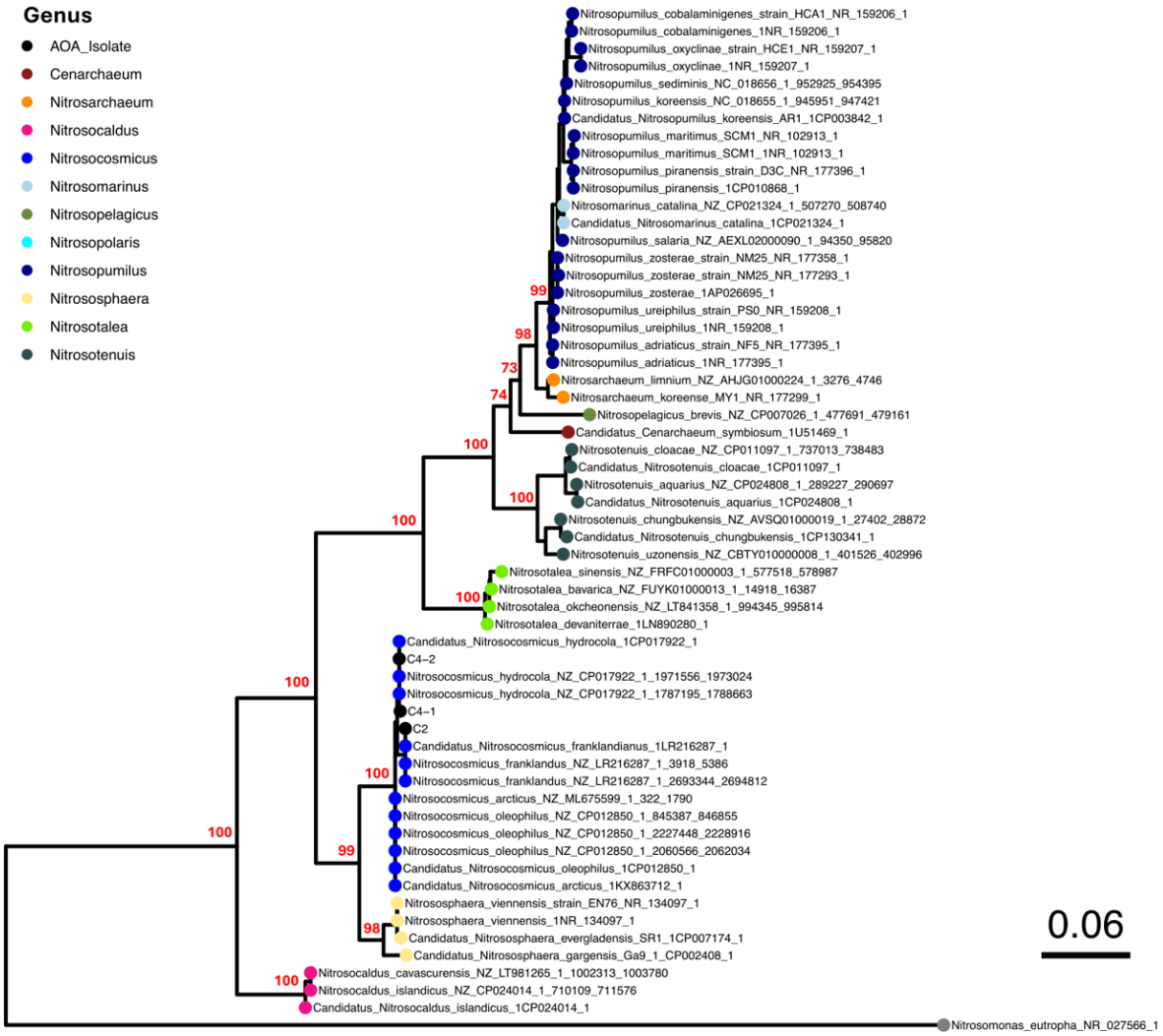


Figure 3.5. Maximum likelihood phylogenetic tree based on the full-length 16S rRNA gene sequences recovered from genome sequence annotation from isolates C4-2, C4-1 and C2 within the *Nitrosocosmicus* clade. Sequences colour coded by genus and bootstrap values of ≥70% shown at branch nodes. Scale bar represents 0.06 substitutions per base. Figure generated courtesy of Fabien Cholet (University of Glasgow).

3.4 Discussion

This study successfully obtained three AOA enrichments (denoted C4-1, C4-2 and C2) from agricultural soils that formed a monophyletic group within the *Nitrosocosmicus* clade (based on almost complete genome sequencing (isolates C4-2, C2), and PCR amplification of the *amoA* genes (from all three isolates). Specifically, C4-2 and C4-1 were similar to each other, and may be potentially novel AOA, related to *Nitrosocosmicus,* particularly *N. hydrocola*, whilst isolate C2 was almost identical to *N. franklandus.* Whilst *N. franklandus* and *N. oleophilus* were also previously cultivated from agricultural soils (Lehtovirta-Morley, *et al*., 2016; Jung, *et al*., 2016), *N. hydrocola* was previously cultivated from wastewater (Sauder, *et al*., 2017).

The assembled MAG sizes were also consistent with the genomes from *Nitrosocosmicus*, whereby, C2 was ca. 2.78 Mbp in size, which is consistent with the genome size of *N. franklandus* (ca. 2.84 Mbp) (Nicol, *et al*., 2019). Additionally, the assembled MAGs of C4-2 was larger (ca. 3.19 Mbp), which is more similar in size to *N. hydrocola* (ca. 3 Mbp) (Sauder, *et al*., 2017) but slightly smaller than *N. oleophilus* (ca. 3.43 Mbp) (Jung, *et al*., 2016). Furthermore, the GC content of C4-2 (33.34%) and C2 (33.99%) were similar to that found in *N. franklandus* (34.07%), *N. oleophilus* (ca. 34%) and *N. hydrocola* (ca. 34%) (Nicol, *et al*., 2019; Jung, *et al*., 2016; (Sauder, *et al*., 2017). Contrastingly, the terrestrial *Nitrososphaera viennensis* which contains a much smaller genome of ca. 2.65 Mbp (You, *et al*., 2011) compared with isolates C2 and C4-2, which may be due to differences between genera, including different ecological adaptations such as *N. viennensis* typically live in acidic soils (Kerou, *et al*., 2016). In the present study, the pH of the soils these isolates were derived from were near neutral at pH 6.8 (for isolate C2) and pH 6.9 (for isolate C4-2).

Despite repeated binning attempts on the genome sequence data from isolate C4-1, MAGs were not successfully obtained from this environmental isolate, which may be due to

inconsistent coverage preventing binning of AOA MAG. However, in this study, PCR amplification and sequencing of the *amoA* genes were successfully amplified from all three isolates including C4-1, which is not surprising given that within the *Nitrosocosmicus* the *amoA* genes are highly conserved (Wright, *et al*., 2023). Blastn analysis of the amplified *amoA* genes revealed that both C4-1 and C4-2 had 92% sequence identity to *N. hydrocola* which was further corroborated by the phylogenetic analysis of the full length 16S rRNA gene sequence from whole genome sequence data which grouped C4-1 and C4-2 with *N. hydrocola.* Furthermore, isolate C4-2 had an ANI% of 97.06% for *Nitrosocosmicus sp008389435*. Interestingly, the amoA amino acid sequences recovered from the genome sequence annotation revealed that C4-2, C4-1 clustered with *N. oleophilus.* Together, this suggests that whilst C4-1 and C4-2 may be very similar to each other, both may be previously unidentified, potentially novel AOA, related to *Nitrosocosmicus,* and more closely related to *N. hydrocola*.

In this study, identical amoA subunit sequences were recovered for all enrichment cultures (C2, C4-1 and C4-2) (**Figure 3.4**) but had different 16S rRNA gene sequences (**Figure 3.5**). A possible explanation for this could be due to the non-axenic nature of the cultures, which likely contain multiple closely related AOA from *Nitrosocosmicus* genus (Rodriguez, *et al*., 2021). It is possible that the identical amoA sequences found in all three MAGs could be due to a dominant lineage with identical amoA sequences, in this case *N. oleophilus* (MY3) and *N. franklandus* (C13). Whereas differing 16S rRNA genes may reflect co-enrichment of additional taxa such as *Nitrosocosmicus* AOA (*N. franklandus* (C13)) for C2 and (*N. hydrocola* (G61)) for C4-1/C4-2. Alternatively, it is known that *amoA* and 16S rRNA genes are not always phylogenetically congruent, as has been observed in *N. franklandus* (C13), where *amoA* and 16S rRNA sequences displayed different evolutionary lineage (Lehtovirta-Morley, 2016). 16S rRNA gene markers in *Nitrosocosmicus* may limit the resolving ability for species level differentiation, which has been seen with *N. arcticus* (Kfb) and *N. oleophilus* (MY3) which

exhibit 100% 16S rRNA sequence identity (Alves *et al*., 2019). Furthermore, *Nitrosocosmicus* AOA have been found to possess multiple copies of rRNA operons, but only one copy of the *amoA* gene (Han, *et al*., 2024). Because 16S rRNA genes are highly conserved, repetitive and with multiple copies, they can be mis-binned during MAG assembly (Song, *et al*., 2022). On the other hand, the *amoA* genes which contain single copies, are easier to bin and link to a dominant genome (Alves, *et al*., 2018).

In contrast to isolates C4-2 and C4-1, Blastn analysis of the *amoA* gene sequence from isolate C2 had a high sequence identity to *N. franklandus*, and together with GTDB taxonomy assignment revealed an ANI% of 99.98% to *N. franklandus*, which suggests that isolate C2 is very likely related to *N franklandus*. However, when the MAGs for isolate C2 (and C4-2) were screened for the ammonia oxidation protein subunits amoA, amoB, amoC, amoX, amoY and amoZ (Hodgskiss, *et al*., 2023), only amoA was identified using HMMER search against Pfam HMM profiles, despite amoX also being identified in *N. franklandus* (Nicol, *et al*., 2019).

Previous studies that analysed the genomes of *Nitrosocosmicus* species found a high resemblance of genomic make-up to their AOB counterparts, which would explain the ability of *Nitrosocosmicus* to be tolerant under high levels of ammonia and nitrite (Han, *et al*., 2024). The same study also discovered *Nitrosocosmicus* species such as *N. oleophilus* and *N. franklandus* possess only one low-affinity Amt transporter proteins, making them less efficient at taking up ammonia, thus making them more favourable towards high ammonia concentrations (Han, *et al*., 2024). In the present study, only one copy of a low affinity ammonia transporter (Amt1) protein was identified in both C4-2 and C2 isolates which is consistent with other *Nitrosocosmicus* genomes (Han, *et al*., 2024; Liu, *et al*., 2019; Nicol, *et al*., 2019) and suggests that isolates C4-2 and C2 may also favour environments with higher ammonia concentrations. A lack of S-layer genes (*slp*) and cell-anchoring genes (*slab*) was also identified from isolate C2 and C4-2 MAGs. These genes have been proposed to function as an NH_4^+

concentrating mechanism (Han *et al*., 2024; Wright, and Lehtovirta-Morley, 2023). To date, all *Nitrosocosmicus* AOA lack any identifiable S-layer proteins, despite being found in marine *N. maritimus* (SCM1) and soil dwelling *N. viennensis* (EN76). Isolates C4-2 and C2 also had a urease alpha subunit which is indicative of their ability to use urea as an alternative nitrogen source to ammonia in soils (Yang, *et al*., 2023), which may explain the ability of these isolates to grow in soils with lower ammonia concentrations. Urea metabolism is another common trait found among other soil dwelling AOA including *N. franklandus* as a potential adaptation in agricultural soils where nitrogen forms frequently fluctuate (Han, *et al*., 2024; Lehtovirta-Morley, *et al*., 2016).

The presence of proteins responsible for ROS detoxification and protection from oxidative stress (*e.g*. Iron Superoxide Dismutase, Thioredoxins and AhpC-type peroxiredoxins) were also found, suggesting that isolates C4-2 and C2 may have the ability to survive in the presence of high nitrite accumulation (Han, *et al*., 2024). This may also explain the observations found herein, that isolates C4-2 and C2 were not inhibited by high levels of nitrite and still showed visible growth for many months in the presence of high nitrite concentrations (See, Chapter 2). The discovery of these oxidative response proteins shows how isolates C4-2 and C2 may be capable of survival without the need for H_2O_2 scavengers such as catalase or co-culturing with microbes that detoxify ROS by producing catalases and peroxidases (Kim, *et al*., 2016). As expected, given that AOA are autotrophic, Enoyl-CoA hydratase and Methylmalonyl-CoA mutase which are involved in the 3-hydroxypropionate-4-hydroxybutyrate carbon fixation pathway (Bei, *et al*., 2024; Alves, *et al*., 2019) were identified, with 1 copy in isolate C2 and 2 copies in isolate C4-2, suggesting that C4-2 and C2 can both use bicarbonate as a sole carbon source like *N. franklandus* (C13) (Lehtovirta-Morley, *et al*., 2016).

It is considered that AOA appear to be highly dependent on copper (as opposed to iron in AOBs) for the maturation of a myriad of copper proteins and enzymes. For example, the respiratory chain and key enzymes in AOB use proteins which require heme (Arp, *et al*., 2002), whilst the respiratory chain in AOA is predicted to be entirely copper based, containing numerous small blue copper proteins and multicopper oxidases which may play a role in electron transfer and nitrogen metabolism (Reyes, *et al*., 2020b). There are also no known examples of chalkophores (copper-binding molecules) in AOA, and differences in affinity for copper uptake has been found between different AOA (Amin et al 2013; Reyes et al 2020a). This study also did not recover any copper-binding chalkophores, thus, confirming the reliance of the vast network of copper-handling proteins (MCOs, HMA domain proteins, cupredoxins and CopC/CopD/CopA) to maintain copper homeostasis (Reyes, *et al*., 2020b). Isolates C4-2 and C2 however contained many copper handling proteins for copper binding/copper homeostasis including a high abundance of MCOs (Multicopper oxidases) and cupredoxins, which has been found previously in other soil dwelling *Nitrosocosmicus* species (Abiola, *et al*., 2025; Reyes, *et al*., 2020b). Furthermore, this study identified copper trafficking CopC, membrane proteins (*e.g*. CopD for the import of copper into the cytoplasm) and the copper exporting membrane protein (CopA for exporting copper out from the cytoplasm) in both C4-2 and C2 isolates, which demonstrates the sophisticated mechanisms AOA have adopted for handing copper concentrations within the cell (Reyes, *et al*., 2020b). Cytochrome c oxidase is another copper dependent protein responsible for aerobic respiration (Reyes, *et al*., 2020b), and was also found in both C4-2 and C2 isolates, further corroborating the necessity of copper requirement in these isolates.

Despite best efforts in this study to obtain an axenic (pure) culture of AOA, (*i.e*. by frequent subculturing, use of solid media colony picking and the addition of antibiotics (Chapter 2), the AOA co-existed with other microorganisms, most likely AOB, (albeit in low

abundance based on CARD-FISH and qPCR analysis (See Chapter 2). This may be due to AOA having a growth dependency with the co-cultured microorganisms or the growth conditions were not sufficiently selective to remove faster growing organisms that competed with the AOA for available substrates. Previous studies that isolated soil dwelling AOA like *N. franklandus* that were fully genome sequenced also found *via* fluorescence microscopy that the AOA co-occurred with other microbes (Klein, *et al*., 2022). Despite a low abundance of AOB co-occurring, in this study, high quality MAGs of C4-2 and C2 were obtained, showing a BUSCO score of 90.7% and a MiGA completion of 95.7% with no contaminant identified for both isolates. However, for isolate C4-1, although good quality MAGs were produced, C4-1 had lower genome coverage, resulting in the inability to yield AOA MAG by binning. This may be due to a lower abundance of AOA in culture which may have been outcompeted by the faster growing co-occurring AOB and further work is required to obtain this information.

Techniques such as extinction dilution, whereby AOA is diluted into fresh media for single AOA cell growth could be employed (Ordonez, *et al*., 2025), to obtain axenic or close to axenic AOA cultures. A recent study plating onto solid/semi-solid media under controlled conditions was successfully carried out on the soil dwelling *N. viennensis* (Klein, *et al*., 2022) and a similar approach could be applied herein. Yielding AOA colonies in low nutrient agar and vapor fed ammonia over long incubation periods is another option (Abiola, *et al*., 2024). The genome data obtained herein for isolates C4-2 and C2 was not quite a complete circular genome map, but the sequence information could be enhanced by improving genome completeness of MAGs to obtain a nearly complete or complete genome. This can be done by implementing a combination of both long and short read sequencing technology, rather than just the short read sequenced that was performed in this study. Long read technology such as PacBio and Oxford nanopore can be carried out and a hybrid assembly of short read data and long read data can be used to close the gaps in the MAGs (Vuong, *et al*., 2022). Future work

could also involve meta transcriptomics to identify key genes that are being actively expressed under different conditions, to directly link isolates C4-1, C4-2 and C2 to their functional activity in the environment**.**

Conclusions

In conclusion, this Chapter has provided a greater understanding of the metabolic function and phylogenetic relationships of three AOA isolates (C4-1, C4-2 and C2) originally isolated from agricultural soils that were closely related to *Nitrosocosmicus.* More specifically, isolate C2 was almost identical to *N. franklandus* (with an ANI 99.98%), whilst C4-1 and C4-2 were closely related to potentially novel *Nitrosocosmicus* species, similar to *N. hydrocola*. Further evidence from this study supports the idea that AOA appear to be highly dependent on copper (rather than iron as found in AOBs) for the myriad of copper proteins and enzymes. Despite the absence of any chalkophores, this study found many proteins dependent on copper within the genomes of enrichments C4-2 and C2. For example, both C2 and C4-2 encode several proteins involved in copper handling and homeostasis, including, multicopper oxidases, cupredoxins (as has been found previously in other soil dwelling *Nitrosocosmicus* species) (Abiola, *et al*., 2025; Reyes, *et al*., 2020) and proteins involved in copper transport (*e.g.* CopA, CopC, and CopD) (Reyes, *et al*., 2020). These findings demonstrate the sophisticated mechanisms that AOA (including isolates C4-2 and C2) have for handling copper concentrations within the cell and have further enhanced our knowledge of the important role of copper in AOA in the environment. Future directions for this research would be to obtain complete or near complete axenic AOA cultures using techniques as mentioned above.

Chapter 4 Ammonia-Oxidizing Archaea Growth Response to Copper

4.1 Introduction

Ammonia oxidising microorganisms (AOMs) contain ammonia monooxygenase, a copper-dependent enzyme, that catalyses the oxidation of ammonia to nitrite *via* hydroxylamine (*i.e.* the first and rate limiting step in nitrification) (Han, *et al*., 2024; Yin, *et al*., 2018; Wang, *et al*., 2015). Recent studies investigating marine AOA (*N. maritimus* SCM1) revealed the presence of numerous copper rich enzymes such as plastocyanin (Shafiee, *et al*., 2021; Santoro, *et al*., 2015; Walker, *et al*., 2010). In soils, genome, and proteome analysis of the AOA *N. viennensis* (EN76) identified several copper-dependent enzymes such as multicopper oxidases, plastocyanin's and blue copper domain-containing proteins (Reyes, *et al*., 2020). Indeed, AOA contain a Cu-rich electron transfer network, consisting of cupredoxins, multicopper oxidases (MCOs) and nitrite reductases (NirK), to drive ammonia oxidation through the Cu-dependent ammonia monooxygenase (AMO) (Diamond, *et al*., 2022; Jacquot, *et al*., 2014; Stahl, *et al*., 2012; Schleper and Nicol, 2010). Further evidence herein also supports the idea that AOA appear to be highly dependent on copper., For example, the identification of copper handling proteins revealed a high abundance of multicopper oxidases and cupredoxins found within the MAGs of C4-2 and C2 enrichments (Chapter 3, **3.3.3**) (and in other soil dwelling *Nitrosocosmicus* species) (Abiola, *et al*., 2025; Reyes, *et al*., 2020); together with Cytochrome c oxidase, which is a copper dependent protein responsible for aerobic respiration and proteins involved in copper transport (*e.g*. CopA, CopC, and CopD) (Reyes, *et al*., 2020). This demonstrates the sophisticated mechanisms that AOA (including isolates C4-2 and C2) have for handling copper concentrations within the cell. In contrast to AOA, the AOB (such as the soil dwelling *Nitrosomonas europaea*) rely predominantly on heme containing proteins that bind Fe, especially multiheme c-type cytochromes, such as,

hydroxylamine oxidoreductase (HAO) to funnel electrons to the respiratory chain (Diamond, *et al*., 2022; Gorman-Lewis, *et al*., 2019; Arp, *et al*., 2002).

In terrestrial and marine environments, the bioavailability of free Cu^{2+} is highly dependent on the extent in which copper is complexed by organic matter. Previous studies looking at seawater and freshwater showed how Cu^{2+} speciation is highly dominated through the complexion of dissolved organic matter, clays and inorganic ligands (SO_4^{2-}, Cl^-, PO_4^{3-}) (Reyes, *et al*., 2020; Maurer, *et al*., 2013). These studies revealed how copper complexation can consequentially lower the concentration of bioavailable Cu^{2+}. Soluble Cu^{2+} is used as a proxy for bioavailable Cu due to Cu^{1+} being unstable in aqueous solution, which is ultimately converted back to Cu^{2+} (Reyes, *et al*., 2020), (Maurer, *et al*., 2013). Therefore, it is assumed that only soluble Cu^{2+} is available for biological uptake, unless a high affinity mechanism using chalkopohores is utilised to scavenge strongly bound or unavailable copper in soil particles (clays, oxides), and organic matter (humic substances) (Kenney, *et al*., (2018). However, in our study copper-binding chalkophores from the genomes of AOA isolates C2 and C4-2 were not identified (Chapter 3, **3.3.3**).

One study by Shafiee, *et al*., (2021) showed that the marine AOA *N. maritimus* (SCM1) had the highest growth rate at 1 pM free Cu^{2+} compared to the marine AOB *Nitrosococcus oceani* (C-107) which had optimal growth at 0.1 pM free Cu^{2+}, with heavy reliance on bioavailable Cu^{2+} compared to AOBs (Gwak, *et al*., 2020). The same experiment investigating the threshold for copper limitation on *N. viennensis* (EN76) demonstrated NH_4^+ oxidation at free Cu^{2+} concentrations of 10^{-14} M and growth, although slower, was even measured in concentrations of free Cu^{2+} as low as 10^{-15} M (Reyes, *et al*., 2020a). This threshold is considerably lower than *N. maritimus* (SCM1) under very similar conditions. One explanation could be environmental adaptations in the soil dwelling AOA containing high affinity uptake systems, possibly by chalkophore-like mechanisms or highly efficient copper transporters

(CopC, CopD-like systems), thus, making it possible to survive in low free Cu^{2+} environments compared to the marine AOA. In this study, AOA isolate (C4-2), cultivated from an agricultural soil (Chapter 2, **2.2.2**), with high sequence identity to *N. hydrocola* (Chapter 3, **3.3.2**) together with *N. franklandus* (C13) (Lehtovirta-Morley et al., 2016), were evaluated to determine the effects of bioavailable Cu^{2+} on growth.

4.1.1 Aims

This chapter aims to investigate the effect of copper concentration on the growth of AOA enrichment C4-2 and *N. franklandus* C13 to better understand the role of copper in the metabolic processes from two different AOA species that inhabit agricultural soils.

4.2 Methods

4.2.1 Metal Speciation Analysis

Metal speciation analysis was carried out using Visual MINTEQ (version 3.1) software (Gustafsson, 2014) to measure total Cu, unchelated Cu and cupric copper (Cu^{2+}). Components of FWM (**Table A.1**) were added to Visual MINTEQ (version 3.1) software at carrying Copper concentrations and Visual MINTEQ job was executed to identify the following: total copper (Cu_T nM), chelated copper (Cu_u pM) and bioavailable cupric (Cu^{2+} pM) copper concentrations.

4.2.2 Effect of copper on AOA growth.

The effect of copper on the growth AOA (isolate C4-2, See Chapter 2, **2.2.2**) and *N. franklandus* (C13) (Lehtovirta-Morley et al., 2016) was investigated. Duran bottles (250 mL) were rinsed in 10% *(v/v)* HCl and sterilised by autoclaving. Prior to inoculation, cells were washed in sterile 1x PBS to minimise copper carry-over from previous sub-cultures. Triplicate cultures were inoculated into sterile bottles containing FWM (100 mL) and 2 mM NH_4^+, but

with addition of 120 µM EDTA final concentration and containing different copper concentrations (*i.e.* 0, 1, 2.5, 5, 7.5, 10, 12.5, 15, 25, 40 and 100 nM final concentration). Abiotic controls consisted of uninoculated FWM prepared at each copper concentration and 2 mM NH_4^+. Both abiotic controls and AOA cultures were incubated statically at 37°C in the dark for a total of 49 days with sub-samples (2 mL) removed every 7 days. Growth was determined by measuring NO_2^- and NH_4^+ concentration using a SEAL AA3 Continuous Flow Analyser (See, Chapter 2, **2.2.4**). Abiotic controls were checked throughout for NO_2^- accumulation and no NO_2^- was observed, confirming the absence of contamination.

Background Cu^{2+} in sterile FWM, Cu^{2+} stocks, inoculum and trace metals (without the addition of Cu^{2+}) was measured (and the quantity of Cu^{2+} adjusted accordingly, to ensure correct final Cu^{2+} concentrations) by using inductively coupled plasma mass spectrometry (ICP-MS, Agilent ICP-MS 7900) run with no gas and helium (He) collision mode and a micro mist nebuliser. ICP-MS analysis of copper stocks is presented in Appendix **Table A.5**. Samples were diluted to a final volume of 5 mL using sterile Milli-Q water containing less than 0.2% total dissolved solids (TDS), so a low matrix plasma program was selected. RF (radio frequency) power was set to 1550 W, carrier gas was set to 1.05 mL min^{-1} with extraction lens one set to 0 V and extraction lens two set to 2-225 V. Helium (He) collision mode was used with a cell gas flow rate of 4 mL min^{-1}, samples were introduced at 0.4 revolutions per second for 30 seconds, stabilised for 30 seconds and probe was rinsed (30 seconds) and washed (40 seconds) between measurements. Nickel sampler cone with Ni X-lens skimmer cone was used and an external copper standard (SigmaAldrich) was used together with an yttrium internal standard (SigmaAldrich). ICP-MS calibration curve is presented in Appendix **Figure B.4**, and shows the total Cu_T is within the detection limits.

4.2.3 qPCR analysis of *amoA* genes

Cells from sub-samples (2 mL) were pelleted by centrifugation at 20000 x g for 10 mins at 4°C. DNA was extracted using the DNAeasy PowerSoil Pro kit (Qiagen) according to the manufacturer's instructions. Quantification of *amoA* gene abundance was performed using using a SensiFAST SYBR No-ROX (Bioline) on a CFX Opus 384 Real-Time PCR System (BioRad) and the primers CrenamoA-23F/CrenamoA-616R (Tourna, *et al*., 2008) and amoA-1F/amoA-2R (Rotthauwe, *et al*., 1997) as previously described (Chapter 2, **2.2.3**). DNA standards of each target consisted of 10^1 to 10^8 copies in 10 µL reactions consisting of 1 µL of DNA template and 200 nM primers. Archaeal and bacterial *amoA* quantification efficiency was between 81.62-90.32%, a slope value between -3.58 and -3.85, with a Y-intercept of 37.05-39.52 and a R^2 of 0.99.

4.2.4 Statistical Analysis

Statistical analysis was performed using a one-way ANOVA followed by Tukey's honestly significant difference HSD post hoc test with the Statistics Kingdom online calculator (https://www.statskingdom.com) (Chapter 2, **2.2.5**). Repeated measures ANOVA was carried out on copper experiments using Statistics Kingdom calculator (https://www.statskingdom.com) (Chapter 2 **2.2.5**).

4.3 Results

4.3.1 Copper bioavailability effects on growth

Uninoculated abiotic controls were carried out with the same additions of copper, and no growth was observed throughout. Copper speciation across a range of copper concentrations (1-100 nM) showed that an increase in total Cu_T resulted in an increase in complexed Cu_u, bioavailable Cu^{2+} and AOA abundance (up to 4-fold) (**Table 4.1**). Specifically, by day 21, AOA *amoA* abundance increased from 1.3 x 10^4 (1 nM Cu_T) to 5.5 x 10^4 (10 nM Cu_T) gene copy mL^{-1} but remained ca. 5 x 10^4 gene copy mL^{-1} (12.5 and 25 nM Cu_T) and at 15 nM Cu_T AOA *amoA* gene abundance decreased to 3.6 x 10^4 gene copy mL^{-1} (± SD) (**Table 4.1**). Highest *amoA* abundance occurred at 40 nM Cu_T (6.8 x 10^4 gene copy mL^{-1} ± SD), and was the lowest at the highest copper concentration (100 nM Cu_T) with 2.5 x 10^4 gene copy mL^{-1} ±SD) (**Table 4.1**).

Table 4.1. Relationship between total copper (Cu_T), unchelated copper (Cu_u), cupric (Cu^{2+}) and AOA abundance (*amoA* gene copies mL^{-1} ± SD, *n*=3) from AOA enrichment C4-2. Gene *amoA* copies derived from stationary phase (day 21) cultures.

Total copper Cu_T (nM)	**Unchelated copper Cu_u (pM)**	**Cupric Cu^{2+} (pM)**	**AOA *amoA* abundance (gene copy mL^{-1} ± SD)**
1	0.00159	4.8 x 10^{-5}	1.3 x 10^4 (± 1169)
2.5	0.00214	1.19 x 10^{-4}	2.4 x 10^4 (± 1969)
5	0.00215	2.38 x 10^{-4}	2.7 x 10^4 (± 6272)
7.5	0.00322	3.57 x 10^{-4}	2.6 x 10^4 (± 3223)
10	0.00421	4.76 x 10^{-4}	5.5 x 10^4 (± 858)
12.5	0.00584	5.95 x 10^{-4}	5.2 x 10^4 (± 1.79 x 10^{-4})
15	0.00701	7.15 x 10^{-4}	3.6 x 10^4 (± 8502)
25	0.01217	0.00119	5.2 x 10^4 (± 1.0 x 10^{-4})
40	0.01870	0.00191	6.8 x 10^4 (± 3401)
100	0.04677	0.00476	2.5 x 10^4 (± 2721)

4.3.2 Growth of *N. franklandus* (C13) at varying copper concentrations

Across the range of copper concentrations (0-100 nM), growth of *N. franklandus* was not inhibited. All cultures showed an exponential decrease in NH_4^+ at days 14-21. At stationary phase (ca. 28-35 days), NH_4^+ was completely oxidised by day 35, with a concomitant increase in NO_2^- (**Figure 4.1A** and **B**). Days 0-28 revealed a significant drop in NH_4^+ difference and increase in NO_2^- for all copper concentrations ($p < 0.05$). NH_4^+ depletion decreased over time with all copper concentrations (**Figure 4.1A** and **C**). During the oxidation of NH_4^+, there was an exponential increase in NO_2^- at days 14-21, shortly followed by no further NO_2^- accumulation during stationary phase at day ca. 28-35 (with ca. 1.8-2. mM) (**Figure 4.1B**). Log plots were largely similar, with NH_4^+ showing linearity at days 0-7 followed by an exponential decrease between days 14-28 (**Figure 4.1C**). The log plot for NO_2^- accumulation revealed an exponential increase at days 7-21, with linearity returning at day 28-35 (**Figure 4.1D**).

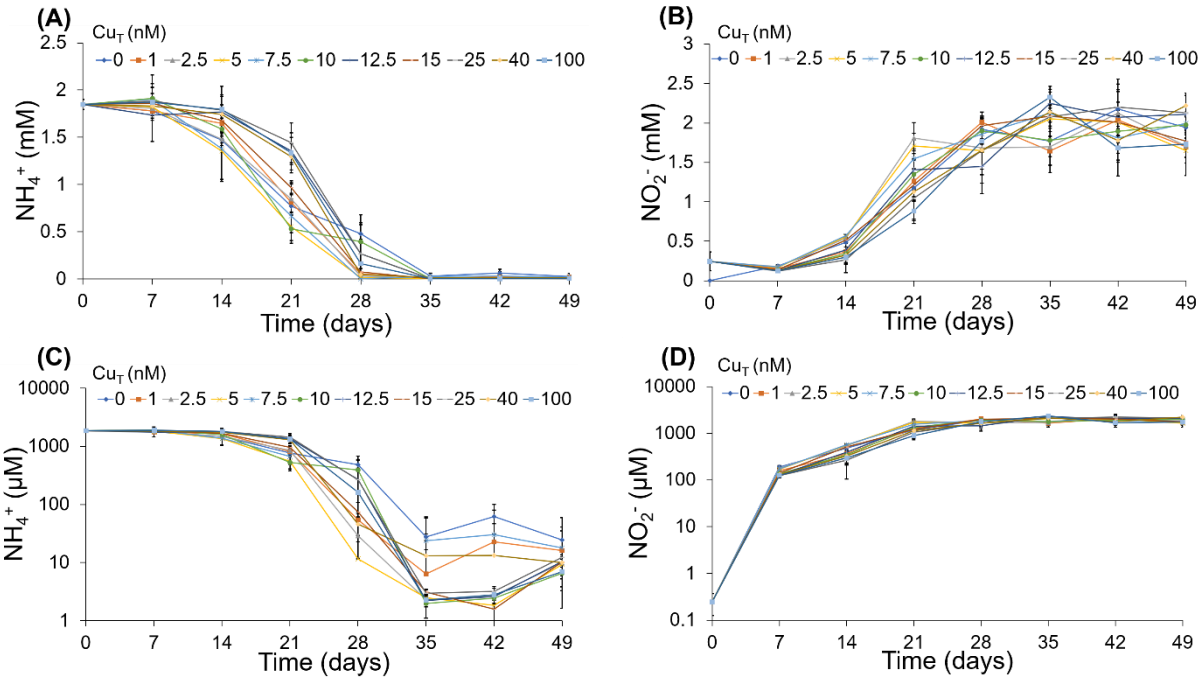


Figure 4.1. Ammonium oxidation (**A**) and nitrite accumulation (**B**) and $\log_{10}$ ammonium (**C**) and nitrite (**D**) over 49 days incubation of *N. franklandus* on varying total copper (Cu_T) concentrations. (Error bars represent SD of the mean, n=3).

4.3.3 Growth Response of AOA isolate C4-2 at varying copper concentrations

Growth of AOA isolate C4-2 at varying copper concentrations was investigated (**Figure 4.2)**. At lower Cu_T concentrations (*i.e.* 0-1 nM) incomplete oxidation of NH_4^+ occurred (with a decrease from ca. 1.95 mM at day 0 to 0.7 mM of NH_4^+ at day 49). Contrastingly, Cu_T concentrations at ≥ 2.5 nM all exhibited a complete oxidation of NH_4^+ (from ca. 1.95 at day 0 to 0 mM of NH_4^+ by day 21), with exponential phase reached ay day 14, followed by stationary phase at day 21 (**Figure 4.2A**). Similarly at 0-1 nM Cu_T, NO_2^- accumulation was lower (*i.e.* < ca. 800-1000 at day 49) compared to ≥ 2.5 nM Cu_T, which showed an increase in NO_2^- from 0 to ca. 2000 μM NO_2^- at day 49 (**Figure 4.2B**). Interestingly, a decline on NO_2^- was observed at day 49 for Cu_T concentrations between 5-100 nM (**Figure 4.2B**). Repeated measures ANOVA revealed a significant decrease in NH_4^+ between days 0-21 and a significant increase of NO_2^- between days 0-28 ($p < 0.05$), suggesting isolate C4-2 expressed growth at all (0-100 nM Cu_T concentrations). However, with lower copper Cu_T (*i.e.* at 0-1 nM), a decrease in NH_4^+ depletion and accumulation of NO_2^- rates can be observed, compared to higher Cu_T concentrations (**Figure 4.2**), the results suggest that even Cu_T concentrations as low as 0-1 nM does have an impact on C4-2 enrichment growth. Logarithmic plots reveal steepest slopes at Cu_T ranges 2.5-100 nM, with a shallower slope seen for 0-1 nM Cu_T for oxidation of NH_4^+ and NO_2^- accumulation respectively (**Figure 4.2C** and **D**). Similarly to the growth of *N. franklandus*, C4-2 isolate only shows growth at copper treatments 0-1 nM, despite growth being slower at 0-1 nM (**Figure 4.2 A** and **B**).

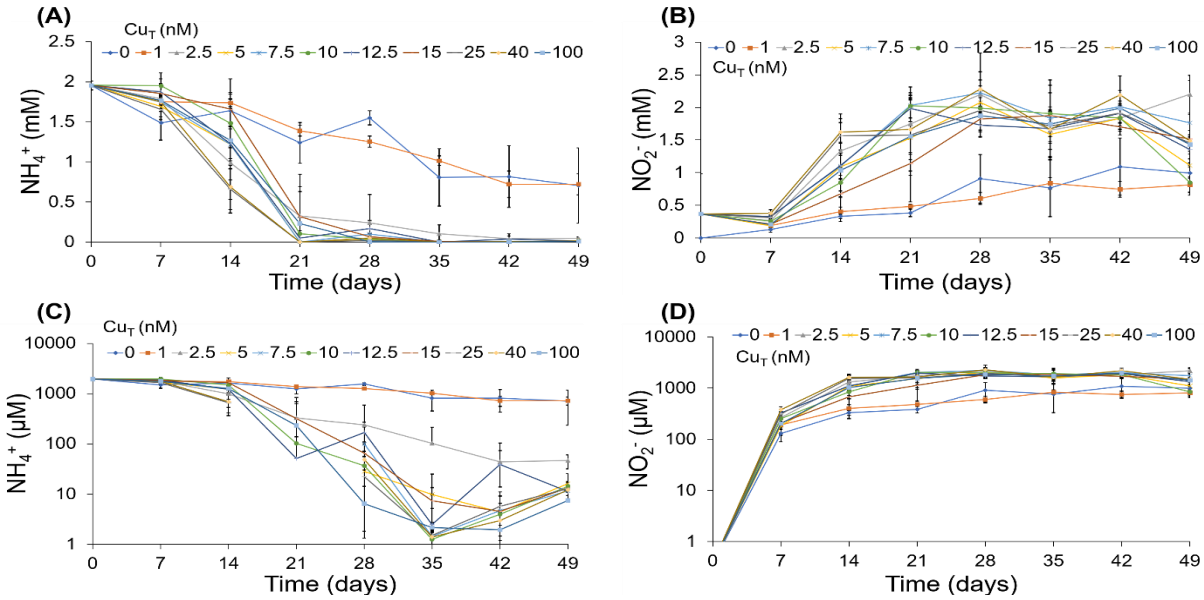


Figure 4.2. Ammonium oxidation (**A**) and nitrite accumulation (**B**) and $\log_{10}$ ammonium (**C**) and nitrite (**D**) over 49 days incubation of enrichment C4-2 grown on varying total copper (Cu_T) concentrations. (Error bars represent SD of the mean, *n*=3).

Table 4.2. Total (Cu_T) and bioavailable (Cu^{2+}) copper concentrations used for AOA copper experiments.

Total Copper CuT (nM)	Bioavailable Cu^{2+} (pM)
1	4.8×10^{-5}
2.5	1.19×10^{-4}
5	2.38×10^{-4}
7.5	3.57×10^{-4}
10	4.76×10^{-4}
12.5	5.95×10^{-4}
15	7.15×10^{-4}
25	0.00119
40	0.00191
100	0.00476

4.3.4 Changes in AOA and AOB amoA gene abundance during growth at varying copper concentrations

qPCR analysis of archaeal *amoA* genes was also performed during the growth of AOA isolate C4-2 on various copper concentrations (**Figure 4.3A**). AOA *amoA* gene abundance increased, with incremental increases found with copper concentrations between 2.5-100 nM, further corroborating the ammonia and nitrite data that copper is a limiting factor for AOA growth (**Figure 4.3**). Across all Cu_T concentrations, AOA abundance was the lowest levels during the lag phase (by day 7), but increased during the exponential phase (at day 14), with a further increase at stationary phase (day 21) for Cu_T concentrations (10-40 nM) (**Figure 4.3A**). Little difference is revealed between exponential and stationary phase for lower Cu_T concentrations of 2.5 nM and 5 nM, and at the highest Cu_T concentration of 100 nM (**Figure 4.3A**). As expected, Cu_T concentration of 1 nM revealed very little increase in AOA abundance (ca. 1600 at day 7 to -3700 at day 14 *amoA* gene copies $mL^{-1)}$, and 13,000 *amoA* gene copies mL^{-1} at day 21 (**Figure 4.3A**). The highest *amoA* gene copies mL^{-1} was in the presence of 40 nM Cu_T concentration (ca. 68,500 *amoA* gene copies mL^{-1} at day 21), *i.e.* 5-fold greater than at 1 nM Cu_T concentration (**Figure 4.3A**).

At day 7 (lag phase) Cu_T treatments at 2.5, 7.5, 10 and 15 nM yielded significantly higher AOA *amoA* gene copies mL^{-1} than 1 nM ^{1} Cu_T treatment ($p < 0.05$) (**Figure 4.3A**). At day 14, $Cu_T \geq 2.5$ nM resulted in significantly higher AOA *amoA* gene copies mL^{-1} than 1 nM ($p < 0.05$), but with no significant difference observed among ≥ 2.5 nM Cu_T treatments ($p > 0.05$) (**Figure 4.3A**). Day 21 also showed significantly higher AOA *amoA* gene copies mL^{-1} at 40 nM Cu_T ($p < 0.05$), with uniquely labelled "c" (**Figure 4.3A**). At day 21, lower Cu_T treatments did not differ significantly from each other ($p > 0.05$), while 10, 12.5, 25 and 100 nM treatments were significantly higher than 1-2.5 nM treatments ($p < 0.05$) (**Figure 4.3A**). AOB *amoA* gene copies mL^{-1} were typically 100-fold lower than AOA, suggesting AOB co-

contaminants was low with no significant increase in AOB *amoA* gene copies mL^{-1} by day 21 (all $p > 0.05$) (**Figure 4.3B**). AOA *amoA* gene copies mL^{-1} in the inoculum were ca. 5.3 x 10^4 copies, which was ca. 3500-fold higher than AOB *amoA* gene copies mL^{-1} (ca. 15 copies) (**Figure 4.3**), suggesting very low AOB abundance during the copper experiment. Moreover, AOA abundances in the copper treatments 10, 12.5, 25 and 40 nM Cu_T were similar to that found in the inoculum (*i.e.* between 5.1 x 10^4 – 6.8 x10^4 *amoA* gene copies mL^{-1}).

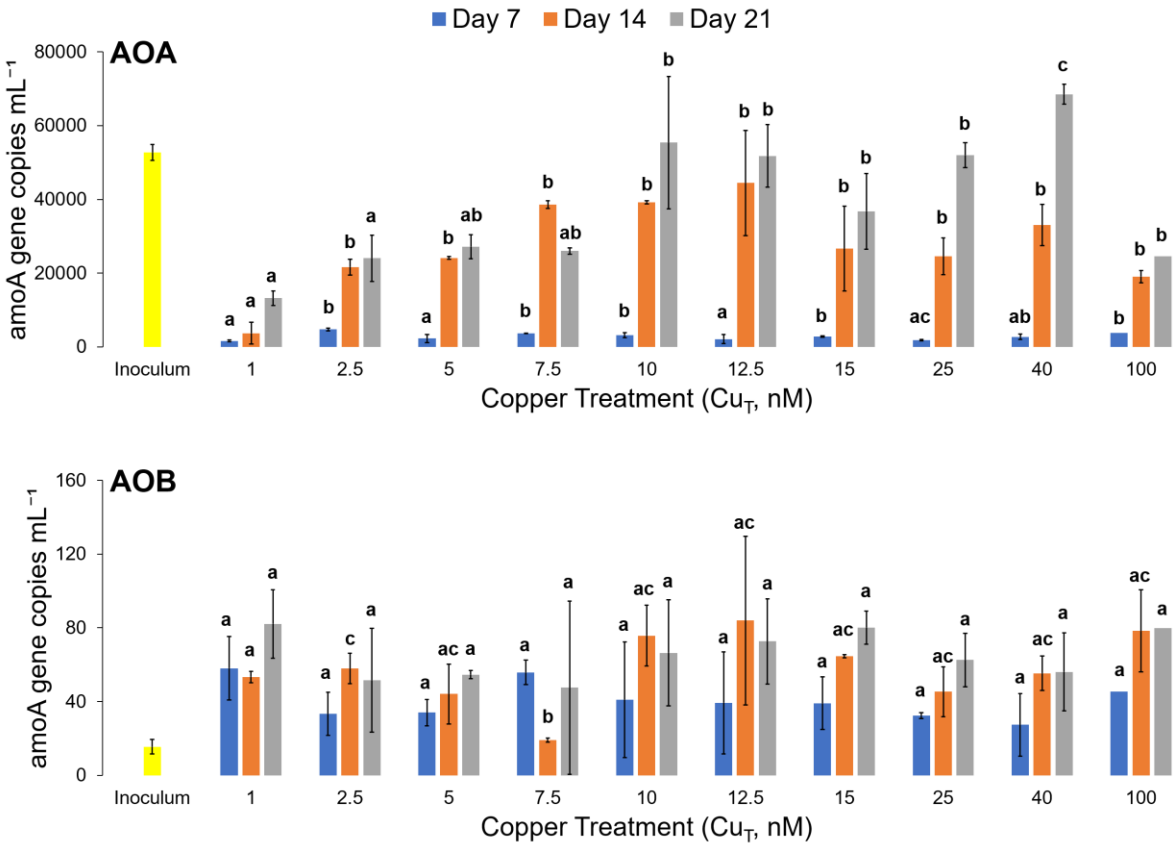


Figure 4.3. Abundance of AOA **(A)** and AOB (**B**) during growth on varying Cu_T concentrations over 21 days. Yellow represents original inoculum. (Error bars represent SD of the mean, n=3). Letters denote significant difference.

4.4 Discussion

This study investigated the growth of AOA isolate C4-2 and *N. franklandus* (C13) at various concentrations of bioavailable Cu^{2+}. This study demonstrated that growth of *N. franklandus* was not inhibited by copper and was able to completely oxidise NH_4^+ to NO_2^- across all bioavailable copper Cu^{2+} concentrations and in the absence of added copper. In contrast to *N. franklandus*, when AOA isolate C4-2 was grown in the absence of copper or at low bioavailable copper concentrations (*i.e.* 1 nM) the oxidation of NH_4^+ to NO_2^- was inhibited somewhat compared to when grown at higher Cu_T concentrations (*i.e.* $\geq$ 2.5 nM) which all exhibited complete oxidation of NH_4^+. Similar to AOA isolate C4-2, other soil dwelling AOA have also exhibited slower growth rates at reduced bioavailability of Cu^{2+} (*i.e.* at concentrations at 1.19 x 10^{-4} pM, **Table 4.1**), including *N. oleophilus* (MY3), *Nitrosotenuis chungbukensis* (MY2), *Nitrosoarchaeum koreense* (MY1), *N. viennensis* (EN76) and AOB *N. europaea* (ATCC 19718) with bioavailable Cu^{2+} concentrations of 10^{-16} M (Gwak, *et al*., 2020; Reyes, *et al*., 2020a). Contrastingly, *N. franklandus* was not affected even at bioavailable Cu^{2+} of 10^{-17} M. Furthermore, another study which investigated copper speciation of *N. viennensis (EN76*) using copper-chelating TETA (Triethylenetetramine) instead of EDTA, also showed growth at extremely low free Cu^{2+} of 10^{-15} M (Reyes, *et al*., 2020a).

Interestingly bioavailable Cu(II) as low as 4.8 x 10^{-5} pM displayed growth but at a lower rate, with significantly lower *amoA* gene copies (**Table 4.1**). Previous studies measuring the abundance of AOA, *N. viennensis* (EN76) showed a stimulation in growth (ca. 75 – 95%) when amended with copper, compared to no copper addition (Gwak, *et al*., 2020). The same study also noted strains *N. koreense* (MY1) and *N. oleophilus* (MY3) all exhibits growth over time at bioavailable Cu(II) at 10^{-17} M, but at a slower rate, this is also seen for C4-2 isolate at bioavailable Cu(II) of 4.8 x 10^{-5} pM (**Table 4.1**).

The results from this study vastly differ from marine AOA such as *N. maritimus* (SCM1) where growth was impeded at free Cu^{2+} levels below 10^{-12} M (Shafiee, *et al*., 2021). A potential explanation for the differing Cu-limiting thresholds between some terrestrial and marine AOA could be distinct mechanisms in Cu-uptake with dissimilar copper proteomes. For example, previous studies using transcriptomic analysis of *N. viennensis* (EN76) identified differences in a large quantity of copper transporters and regulatory genes (Kerou, *et al*., 2016). Comparing *N. viennenis* (EN76) and *N. maritimus* (SCM1), there was a vast difference in the upregulation and downregulation of genes such as multicopper oxidase (MCO) and CopC/D proteins (Reyes, *et al*., 2020b). *N. viennensis* (EN76) displayed impaired growth and ammonia oxidation at free Cu^{2+} limiting conditions (10^{-16} M), with significant upregulation of genes responsible for copper binding and uptake, such as CopC/D transporter proteins and multicopper oxidases (MCOs) (Reyes, *et al*., 2020b). This same stress response may be carried out for members of soil dwelling *Nitrosocosmicus* such as *N. franklandus* and *N. hydrocola* related to isolate C4-2. An abundant network of copper managing proteins such as, putative periplasmic copper chaperone proteins (CopC) (Zhang, *et al*., 2006), cytoplasmic membrane proteins for copper intake (CopD) (Orell, *et al*., 2010), together with copper storage proteins CopA, CopT and CopZ have all been observed in soil dwelling AOA genomes (Kerou, *et al*., 2016; González-Guerrero, *et al*., 2008). Contrastingly, *N. maritimus* (SCM1) only contains one CopC/D gene copy, with no copper storage proteins (Amin, *et al*., 2013; Walker, *et al*., 2010). The copper proteome network in isolate C4-2 could be similar to *N. viennenis* (EN76) which contains a vast copper network, thus resulting in growth even at levels as low as 10^{-16} M free Cu^{2+} (Offre, *et al*., 2014). High quality metagenome assemblies (MAGs) from isolate C4-2 also revealed a vast network of copper binding/managing proteins (Chapter 3, **3.3.3**). Specifically, a total of 10 MCOs, 29 cupredoxins, 2 copies of CopC, 3 copies of CopD, heavy metal-associated protein, cytochrome c oxidase and P-type ATPase (CopA) were discovered from

MAGs. The high abundance of copper managing proteins could also explain the ability of isolate C4-2 to grow under Cu^{2+} limited conditions, like *N. viennensis* (EN76) compared to marine AOA *N. maritimus* (SCM1).

Interestingly, *N. franklandus* (C13) was also able to oxidise ammonia even without the addition of any copper, suggesting it may be able to scavenge and potentially retain copper from previous sub-culturing and/or the soil environment. The genome sequence of *N. franklandus* (C13) revealed the presence of multiple copies of CopD genes with a high abundance of periplasmic multicopper oxidases and copper-transporting P-type ATPases for tight control of copper homeostasis (Shafiee, *et al*., 2022; Nicol, *et al*., 2019). Further study using copper additions at even lower concentrations to achieve free Cu^{2+} concentrations $<10^{-16}$ M could reveal the copper threshold required for ammonia oxidation for *N. franklandus*. Additionally, another reason could be due to copper proteins in the proteome possessing different copper binding affinities for inorganic complexes of Cu and free Cu^{2+} (Shafiee, *et al*., 2021). However, there is no experimental data to back up this potential reason (Jenkins, *et al*., 2016).

The adaptations of AOA isolate C4-2 and *N. franklandus* (C13) have developed from environments that have low bioavailability of copper, despite the original soils that isolate C4-2 was derived from contained ca. 0.54 μM of total copper, which was significantly higher than copper concentrations used in this study. However, in soil environments, only a small fraction of copper is bioavailable, in an accessible form for the microorganism (Jez, *et al*., 2023; Fu, *et al*., 2020). This is largely as a result of soils possessing a high abundance of ligands which chelates Cu^{2+} with a high affinity, and these ligands are usually found in inorganic oxide minerals, clays, inorganic molecules such sulphates, phosphate and organic matter such as humic/fulvic acids (Mengel, *et al*., 2001). Alkaline conditions also result in higher formation of complexed Cu (Gorman-Lewis, *et al*., 2019), which suggests that both isolate C4-2 that had

soil pH of 6.9 and *N. franklandus* from soils at pH 7.5 can adapt to growth in lower Cu^{2+} levels such as in alkaline soils. Moreover, the evidence from this study suggests that *N. franklandus* may be better adapted than isolate C4-2 given that complete ammonia oxidation and nitrite production occurred even in the absence of copper, whilst isolate C4-2 was unable to completely oxidise ammonia in either very low copper or in the absence of copper.

One limitation for this study is the potential contribution of heterotrophs towards the bioavailability of Cu(II). Specifically, heterotrophic bacteria may have altered the Cu(II) speciation and reducing the concentration of free Cu(II). A previous study investigating the distribution of copper speciation in biofilms revealed that microbial cells and extracellular polymeric substances (EPS) can bind copper within biofilms (Lin, *et al*., 2020). It was revealed that EPS from terrestrial *Pseudomonas putida* (CZ1) could absorb a large quantity of copper, with (60-67%) found in extracellular fraction of biofilms, (44.7-42.3%) found in the capsular EPS, and 15.5-21.2% found in microbial cell walls and intracellular fraction (Lin, *et al*., 2020). Interestingly, (**Table A.3)** revealed bacteria from the family of *Sphingomonadaceae*, which may be relevant to copper tolerance, as a previous study found that Cu(II) exposure can promote formation of densely packed biofilm structures in *Sphingomonadaceae*, with the development of a thick EPS matrix enveloping cells (Vílchez, *et al*., 2007). *Planctomycetota* (**Table A.3)** are also commonly associated with biofilm formation, potentially making them indirect contributors to Cu(II) bioavailability through matrix and aggregate formation (Bengtsson and Øvreås, 2010).

Another limitation of this study is the use of glass bottles which can absorb Cu^{2+} and alter the bioavailable Cu^{2+} concentration for the AOA cultures, despite vigorous acid washing (Lago, *et al*., 2025; Fuhrmann, *et al*., 2004). Furthermore, these experiments should be conducted in metal-free laboratories (Reyes *et al*., 2020b; Shafiee, *et al*., 2021). Previous studies investigating AOA growth at varying Cu(II) used plastic polycarbonate/polypropylene bottles

which were acid-washed (Reyes *et al*., 2020b; Shafiee, *et al*., 2021). Although in the present study the macronutrients and basal salts were treated using a Chelex-100 resin to remove trace metal contaminants, prior to copper experiments, future work should implement the use of plastic polycarbonate bottles, Chelex-100 treated media and conducted in a metal-free laboratory. Transcriptomics studies could also be applied to investigate the adaptability of the AOA C4-2 , *N. franklandus* and other AOA within the *Nitrosocosmicus* clade under limiting free Cu^{2+}, to identify key genes involved in the copper uptake systems. Such information would provide a better understanding of the copper proteome and stress responses of AOA from within the *Nitrosocosmicus* genera in copper limited environments. This would provide vital information to uncover specific copper-management strategies and further broaden our understanding of how members of *Nitrosocosmicus* clade adapt to copper in different environments.

Conclusions

In conclusion, this study demonstrated the ability of AOA C4-2 to grow and oxidise ammonium at varying bioavailable Cu^{2+} concentrations (as low as 10^{-16} M), together with *N. franklandus* (C13) which also grew at even lower concentrations (10^{-17} M) of bioavailable Cu^{2+} and in the absence of copper. This result aligns with previous studies carried out on *N. viennenis* (EN76), which exhibited growth at 10^{-16} M of bioavailable Cu^{2+} and further supports the growing evidence that a distinct niche of copper dependent soil dwelling AOA occur that are different to the marine AOA (*N. maritimus* (SCM1). The application of genome-wide transcriptomic analysis at limited bioavailable Cu^{2+} would provide further insights in the strategies implemented by AOA members of the *Nitrosocosmicus* genus in terms of key copper-related genes involved in upregulation and downregulation.

Chapter 5 Molecular biology and over-expression of putative Archaeal multicopper oxidases

5.1 Introduction

AOA are ubiquitous with a globally impactful metabolism, they are difficult to obtain in pure culture, in turn, reactions underlying AOA metabolism remains largely uncharacterised (Wright and Lehtovirta-Morley, 2023). The genome for marine *N. maritimus* (SCM1) (Walker, *et al*., 2010) and soil dwelling *N. franklandus* (C13) (Nicol, *et al*., 2019), reveals many differences compared to better-characterised ammonia oxidising bacteria (AOB). Notably, AOA lack genes encoding the multi c-type heme enzyme hydroxylamine oxidoreductase (HAO), an essential enzyme for abstracting electrons from hydroxylamine (NH_2OH), thus oxidising NH_2OH to form NO_2^- (Voland, R. *et al*., 2025; Kerou, *et al*., 2016). Further discovery of more recent AOA genomes supports the lack of canonical c-type heme biosynthesis pathways, suggesting a distinct mechanism for NH_2OH oxidation.

Like other MCOs, AOA MCOs accept electrons at their type 1 copper site (T1Cu) and pass them to the trinuclear type 2 and type 3 (T2Cu/T3Cu) cluster that performs the four-electron reduction of molecular oxygen (O_2) to H_2O (Gräff, *et al*., 2020; Lawton, *et al*., 2009; Sakurai and Kataoka, 2007). Several MCOs residing from AOA have been grouped by lineages with AOA producing numerous homologs of copper nitrite reductases (NirK) (Voland, *et al*., 2025). NirKs contain a T1Cu center for electron transfer and T2Cu centres for catalysis (Horrell, *et al*., 2017). Additionally, AOA encode many 2 domain (2dMCO), homotrimeric MCOs containing a trinuclear T2Cu/T3Cu center comprising a T2Cu site and binuclear type 3 copper site (T3Cu) (Voland, *et al*., 2025; Zhalnina, *et al*., 2014). A recent study has reported the X-ray crystal structure of a 2dMCO from *N. maritimus,* Nmar_1354, which showed ability for NH_2OH oxidation and bears a unique structural feature termed the “dangler” cupredoxin domain (**Figure 5.1**). Nmar_1354 resides from the MCO4 lineage, which is unique to AOA

from the order Nitrososphaerales and genera *Nitrosopumilus* and *Nitrosarchaeum* (Voland, *et al*., 2025). At the start of this PhD project several putative MCOs in *N. maritimus* were identified as targets for characterisation (Nmar_1131, Nmar_1354 and Nmar_1667), and at the time, the report had not yet been published by Voland *et al*., (2025). Nmar_1131 and Nmar_1354 are both homotrimer 2dMCOs, both consisting of a T1Cu and trinuclear T2Cu/T3Cu copper site. Substrate oxidation occurs at the T1Cu site, with electrons being relayed (ca. 12-13 Å) to the trinuclear T2Cu/T3Cu copper site through a His-Cys-His bridge (Voland *et al*., 2025) (**Figure 5.1**). At the T2Cu/T3Cu site, O_2 binds and is reduced to 2 H_2O, carried out using four electrons and four protons ($O_2 + 4\ H^+ + 4\ e^- \rightarrow 2\ H_2O$) (Sekretaryova, *et al*., 2019).

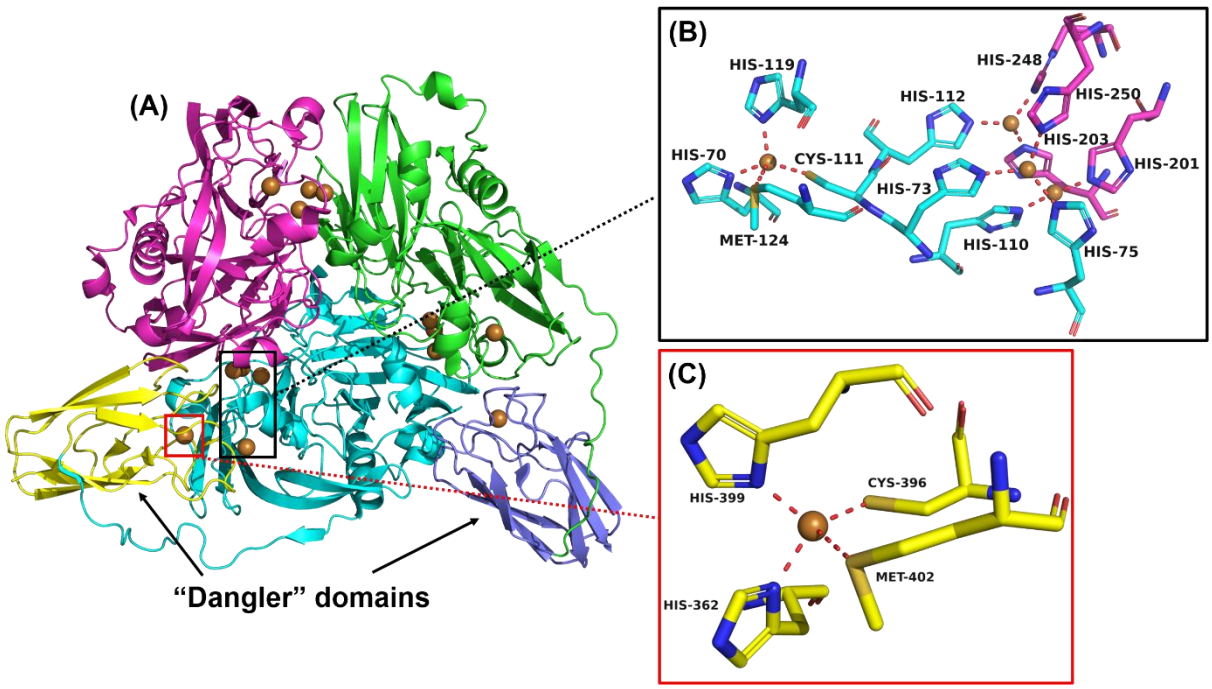


Figure 5.1. Nmar_1354 MCO (PDB 9CCK). **(A)** Quaternary structure of Nmar_1354 revealing canonical 2dMCO, with additional cupredoxin domains, with each colour representing a monomeric subunit. **(B)** Mononuclear T1Cu site with coordinated by two His, a Met and Cys residues, together with a T3Cu/T2Cu trinuclear copper site, containing three copper ions coordinated by eight histidine residues. **(C)** "Dangler" T1Cu cupredoxin domain, two His, Met and Cys coordinating residues.

Target Nmar_1667, is a homotrimeric, 2-domain NirK protein. Subunits of copper NirKs contain two copper sites, a T1Cu electron transfer site and a T2Cu copper catalytic site. Similarly, the T1Cu sites seen in Nmar_1131, Nmar_1354 and NFRAN_2427, the T1Cu site in Nmar_1667 also possess two His, a Met and Cys residue coordinating copper, with T2Cu site possessing four His coordinating residues (**Figure 5.2**). NirK catalysis is initiated at the T1Cu site, where an electron is accepted from substrate and, similarly to 2dMCO, 3dMCO (**Figure 5.1** – **5.2**) electrons are relayed (ca. 12-13 Å) by the His-Cys-His bridge to the T2Cu site. T2Cu site binds to NO_2^- and is reduced to NO and H_2O by a proton-coupled mechanism ($NO_2^- + e^- + 2\ H^+ \rightarrow NO + H_2O$) (Fukuda, *et al*., 2016) (**Figure 5.2**).

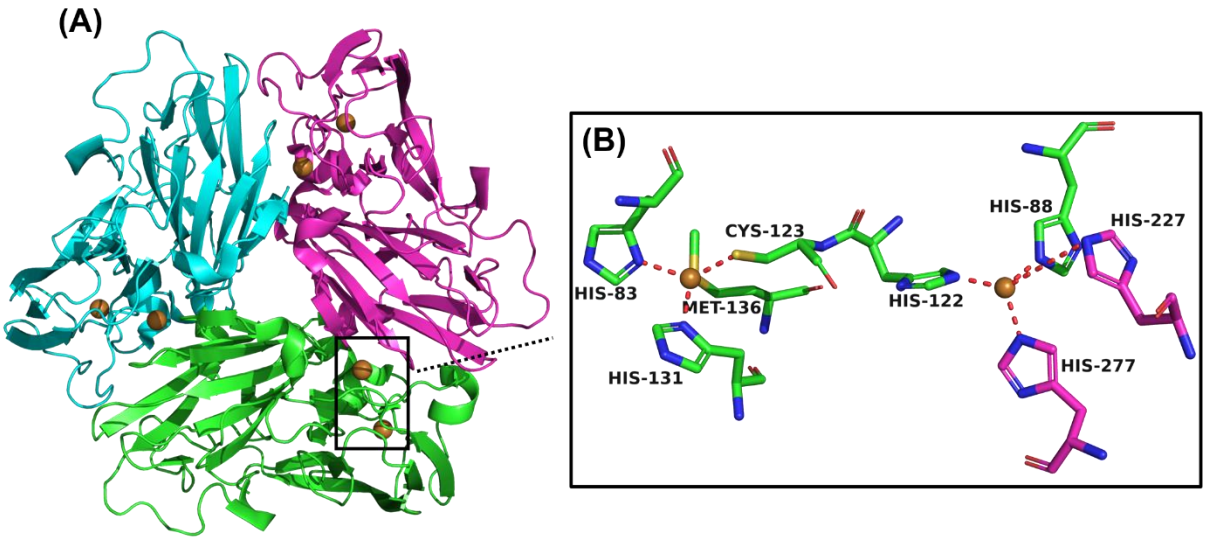


Figure 5.2. 2dMCO NirK from *Nitrosomonas europaea* (PDB 4KNU) (Lawton, *et al*., 2013). **(A)** Quaternary structure of revealing monomeric 2dMCO, with different colour representing monomeric subunit. **(B)** Mononuclear T1Cu site with coordinated by two His, a Met and Cys residues, together with a T2Cu mononuclear copper site, containing four His resides coordinating the copper ion.

In addition, a terrestrial AOA MCO was also targeted. The three-domain (3dMCO) NFRAN_2427 is a monomeric MCO and carries out the same catalysis as its 2dMCO counterparts, with substrate oxidation at T1Cu copper site being relayed (ca. 12-13 Å) by the His-Cys-His bridge, followed by the reduction of O_2 to 2 H_2O using four electrons and four protons ($O_2 + 4\ H^+ + 4\ e^- \rightarrow 2\ H_2O$).

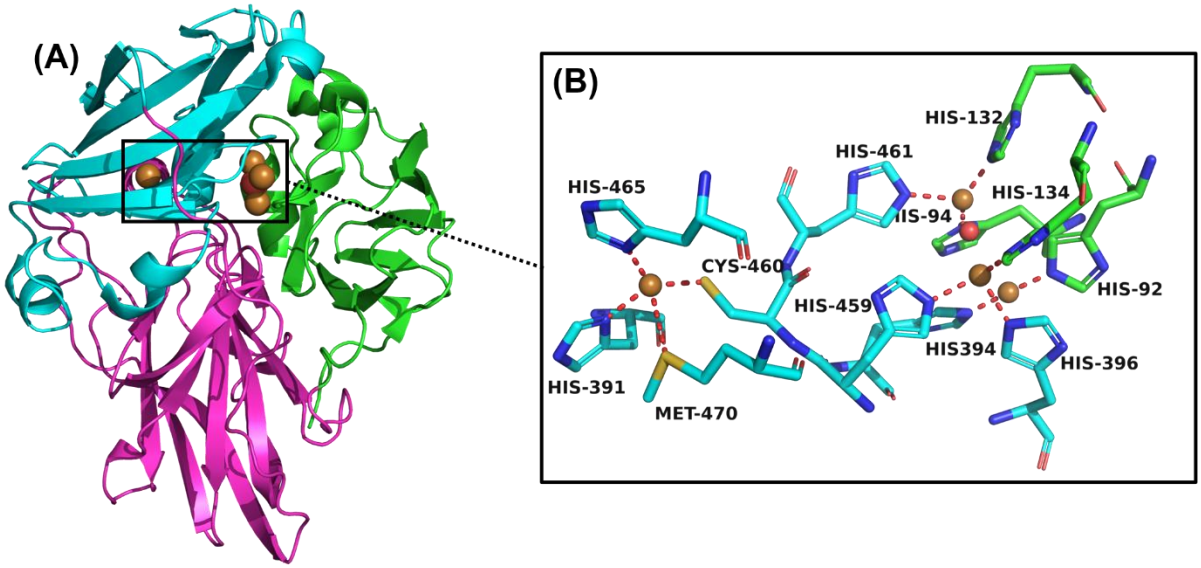


Figure 5.3. 3dMCO from hyperthermophilic archaeon *Pyrobaculum aerophilum* (PDB 3AW5) (Sakuraba, *et al*., 2011). **(A)** Quaternary structure of revealing monomeric 3dMCO, with different colour representing each domain. **(B)** Mononuclear T1Cu site with coordinated by two His, a Met and Cys residues, together with a T3Cu/T2Cu trinuclear copper site, containing three copper ions coordinated by eight histidine residues.

Producing recombinant MCOs in heterologous expression systems such as *E. coli* is an essential strategy to enable detailed functional and structural studies (Hayat, *et al*., 2018; Rosano and Ceccarelli, 2014; Chen, R. 2012). Benefits from successful heterologous expression is that it provides enough protein for characterisation, allowing identification of enzyme sequences or metal cofactor content, unveiling structure-function relationships that would be difficult to carry out from native isolates in small quantities (Rosano and Ceccarelli,

2014). Recombinant production of pure MCOs is instrumental for obtaining high protein yields, if protein is crystallisable, X-ray crystallography studies can be carried forward to determine three-dimensional structure. In turn, providing further understanding of copper-site geometry and oxygen reaction mechanisms, using spectroscopic techniques on crystals or in solution (Handing, *et al*., 2018; Goulding and Jeanne Perry, 2003). In addition, having a recombinant system allows for site-directed mutagenesis to be conducted. MCOs are also of high interest in industrial applications as environmentally friendly biocatalysts, *e.g.* laccases are a widely used MCOs biocatalysts in biotechnology (Satomura, *et al*., 2021; Takamura, *et al*., 2021; Sherif, *et al*., 2013).

5.1.1 Aims

The aim of this Chapter was to identify MCOs from *N. maritimus* (SCM1) and *N. franklandus* (C13), which could be tested for heterologous over-expression in *Escherichia coli* and then used to characterise their biochemical properties and determine their three-dimensional structures by X-ray crystallography.

5.2 Methods

5.2.1 Expression of Recombinant Proteins

The *Escherichia coli* codon optimised MCO gene for each target was provided by Professor Jonathan Worrall. Expression of all target MCO proteins were carried out in *E. coli* BL21(DE3)-RIL cells, cloned into pET28a construct containing a T7 promoter with a 6xhis affinity tag. Enabling selective and rapid purification of expressed MCO. Prior to cloning, N-terminal signal sequences of protein targets were identified and removed using designed primers that start downstream of the signal sequence. An NdeI site was added to the forward primer and a HindIII was added to the reverse primer, followed by ligation into the

NdeI/HindIII site of pET28a using TOP10 *E. coli* cells. Primers used for targets can be found in appendix **Table A.6**. pET28a plasmid was isolated using a mini prep kit (MACGEREY-NAGEL), then transformed in BL21 (DE3) RIL cells for protein purification. Mini prepped constructs were sent for sequencing (Eurofins) to ensure target protein sequence was located at NdeI/HindIII sites prior to overexpression. *E. coli* BL21 (DE3) RIL cells containing pET28a plasmid were inoculated into (1.5% *(w/v)* LB agar and left at 37°C overnight. Single colony transformants were picked and inoculated into sterile 5 mL overnight cultures using a sterile toothpick and left overnight for 16-20 hours (37°C, 180 rpm). Contents of overnights were inoculated into 2 L sterile flasks containing 1 L of LB for Nmar_1131/Nmar_1354/Nmar_1667 and 2xYT for NFRAN_2427 media with kanamycin (50 µg mL^{-1}) and chloramphenicol (50 µg mL^{-1}) (37°C, 180 rpm). Optimal expression conditions were screened for each protein target, such as a range of isopropyl-β-D-thiogalactopyranoside (IPTG) (Melford) concentrations (0.1 mM, 0.3 mM, 0.5 mM and 1 mM) and at varying incubation temperatures after induction (16°C, 18°C, 20°C, 25°C, 30°C and 37°C) with shaking as well as static incubation. All protein targets showed best expression at 0.5 – 1 mM IPTG (Melford) at 16°C - 20°C with shaking at 180 rpm. At OD_{600} 0.8, the temperature was lowered to 16 °C, followed by the addition of 1 mM $CuSO_4$ and 0.5 M IPTG (Melford). After ca. 16-18 h cells were harvested at 3501g for 20 min at 4°C.

5.2.2 Purification from pET28a constructs

Expression from pET28a construct underwent cell lysis using an Emulsiflex-C3, with 3 cycles at 15,000 – 20,000 psi. Cells were pelleted at 38,724g for 20 min at 4 °C and samples were collected for 12.5% *(w/v)* SDS-PAGE analysis. Supernatant was diluted in equilibration buffer (50 mM Tris pH 7.5, 500 mM NaCl, 20 mM imidazole), loaded to a HisTrap nickel column (Cytiva) and eluted on an AKTA pure, using a linear imidazole gradient in elution

buffer (50 mM Tris pH 7.5, 500 M NaCl, 500 mM imidazole). Samples from peaks were selected for 12.5% *(w/v)* SDS-PAGE gel analysis, identified target proteins were pooled and centrifuged in 10 kDa cutoff Centricon (Vivaspin) at 4°C. Once volume reached 2 mL, sample underwent size-exclusion chromatography (SEC) using a HiLoad 16/600 Superdex-200 column (Cytiva) in buffer (50 mM Tris pH 7.5, 150 mM NaCl), ATKA peak sample were collected and analysed with 12.5% *(w/v)* SDS-PAGE gel.

5.2.3 Protein Purification using Ammonium Sulfate "Cut" precipitation

Ammonium sulfate (AS) cuts were carried out to selectively purify NFRAN_2427 MCO. NFRAN_2427 was expressed using pET28a His_6 tag vector and underwent AS cut purification after Ni-NTA purification (See Section **5.2.3**). A concentration of 2 mg mL^{-1} NFRAN_2427 MCO was incubated in 0.5 M, 1.0 M, 1.5 M, 2.0 M, 2.5 M and 3.0 M of AS and left at 4°C for 30 minutes. Following incubation, each AS concentration sample was centrifuged at 15,000g for 10 minutes (4°C), forming a pellet in each concentration. 20 µL of supernatant was mixed with cracking buffer with DTT and pellets from each AS concentration was resuspended in 20 µL of cracking buffer containing DTT.

5.2.4 Extraction and Refolding of Inclusion Bodies

Attempts were made to solubilise inclusion bodies formed from expression trials; expression and cell lysis was carried out as previously described in **5.2.2**/**5.2.3** (Kim *et al*., 2015). After cell lysis, supernatant was discarded and insoluble fraction containing target protein was resuspended in 20 mL of ice-cold buffer (20 mM Tris pH 7.5, 500 mM NaCl, 2 M urea and 2% *(v/v)* Triton X-100). The insoluble fraction was washed twice and separated by centrifugation between washes, at 38,724g for 10 minutes (4°C). Washed pellets were resuspended in 50 mL binding buffer (20 mM Tris pH 7.5, 500 mM NaCl, 6 M guanidine HCl,

5 mM imidazole and 1 mM DTT) and stirred for 1 h at RT. Stirred sample was then centrifuged at 38,724g for 10 minutes (4°C) to remove un-solubilised material and dialysed in 5 L of buffer with no guanidine HCl (20 mM Tris pH 7.5, 500 mM NaCl, 5 mM imidazole). Simultaneous purification and refolding were carried out on target protein, to identify whether inclusion body extraction was successful. Supernatant from previous centrifuge step was loaded onto a HisTrap HP nickel column (Cytiva). Column was washed with 10 column volumes of 20 mM Tris pH 7.5, 500 mM NaCl, 6 M guanidine HCl, 5 mM imidazole and 1 Mm DTT buffer, followed by 10 column volumes of washing with 20 mM Tris pH 7.5, 500 mM NaCl, 6 M urea, 20 mM imidazole and 1 Mm DTT buffer. Target protein was refolded on column using AKTA in buffer (20 mM Tris pH 7.5, 500 mM NaCl, 20 mM imidazole), using a linear gradient. Samples were collected and analysed on a 12.5% SDS-PAGE gel.

5.2.5 SDS-PAGE and Western Blot

Cell samples from expressions trails were obtained by collecting 100 µL and spinning at 15,000g for 1 minute to collect pellet, the pellet was resuspended in 50 µL of cracking buffer with DTT and boiled for 10 minutes. Supernatant and flow-through from expression trails was mixed with 20 µL of cracking buffer with DTT and boiled for 10 minutes, 20 µL of fraction samples from purification was collected and mixed with 20 µL of cracking buffer. Sample was loaded onto 12.5% *(w/v)* SDS-PAGE gel and ran using electrophoresis buffer (50 mM Tris pH 8.3 1.92 M glycine, 1% *(w/v)* SDS). Gel was stained using a Coomassie brilliant blue stain for ca. 4 h and left to destain overnight in water, ready for analysis. For anti-6xHis western blot analysis, the gel was transferred to a nitrocellulose transfer membrane (ThermoScientific) inside an iblot 2 dry blotting device (ThermoScientific) for 7 minutes. Membrane was transferred to a 50 mL falcon tube and blocked using 5% dry milk powder in TBS-T buffer (50 mM Tris, 150 mM NaCl, 0.05% Tween-20, pH 7.4) for >1 hour shaking at RT. After blocker

was removed and membrane was incubated in primary anti-his antibody (Abcam) 1:2000 for 1 hour at room temperature (1 µL added to 2 mL TBS-T). Membrane was washed 3 times using ca. 5 mL of TBS-T for 10 minutes each time on a rocker. Lastly, membrane was incubated with secondary antibody (Abcam) 1:2000 for >1 hour in TBS-T buffer and incubated at room temperature. One more wash with TBS-T was carried forward as previously described and 800 µL of chemiluminescent substrate (BioRad) was added to membrane and mixed thoroughly before visualisation on the iBright (Invitrogen).

5.3 Results

5.3.1 Target Protein Characteristics

Prior to expression and purification trials, key biochemical and structural characteristics of all protein targets were considered. A total of 6 MCOs exist in *N. maritimus* (SCM1) and 8 MCOs exist in *N. franklandus* (C13). **Table 5.1** reveals characteristics to bear in mind during expression and purification trials. During expression Nmar targets are expected to assemble a homotrimeric assembly in their native size (kDa), therefore, SDS and reducing agent (DTT) from SDS-PAGE gel will denature and dissociate polypeptide chains into monomer size.

Table 5.1. Summary of target protein characteristics.

Protein	**Domain**	**Oligomeric State**	**Copper Sites**	**Native Size (kDa)**	**SDS-PAGE Size (kDa)**
Nmar_1131	2-Domain MCO	Homotrimer	T1Cu + T3/T2Cu	ca. 112.4	ca. 36.4
Nmar_1354	2-Domain MCO	Homotrimer	T1Cu + T3/T2Cu	ca. 143.5	ca. 47
Nmar_1667	2-Domian NirK	Homotrimer	T1Cu + T2Cu	ca. 145.6	ca. 47.4
NFRAN_2427	3-Domain MCO	Monomer	T1Cu + T3/T2Cu	ca. 154.6	ca. 50.2

Nmar_1131 was a selected target as it has been predicted to be a laccase-like two-domain MCO, Nmar_1354 was also selected as it contains a unique dangler cupredoxin domain at the C-terminal and Nmar_1667 was selected as it is a NO-forming nitrite NirK. Additionally, NFRAN_2427 was selected as another target as it belongs to terrestrial AOA *N. franklandus* (C13), enabling comparison of terrestrial MCOs encoded by AOA with marine MCO.

NFRAN_2427 contains a three-domain structure for comparison with two-domain Nmar constructs.

5.3.2 Protein expression and purification of targets

The expected molecular weight of target pET28a/Nmar_1131 is approximately 36 kDa. **Figure 5.1A** shows no inducible band at pre-IPTG, with a visible band ca. 36 kDa clearly seen at post-IPTG and overnight (O/N) samples, suggesting successful expression of Nmar_1131. To assess whether Nmar_1131 was expressing in a soluble form, lysate was fractionated into soluble and insoluble components (**Figure 5.**1**A**). SDS-PAGE gel revealed majority of target size (ca. 36 kDa) Nmar_1131 accumulating in the insoluble (pellet) fraction (**Figure 5.5A**), indicative of inclusion body formation. A faint band ca. 36 kDa was observed in the soluble fraction, revealing limited soluble Nmar_1131 expression (**Figure 5.1A**). The band at ca. 36 kDa can be observed in the flow-through (FT), indicating poor recovery of soluble Nmar_1131 and incomplete binding to HisTrap HP nickel resin. In addition, a prominent band can be seen at a higher weight of ca. 40 kDa, seen across Ni-NTA fractions (**Figure 5.4B**). Pooled Nmar_1131 eluate was subjected to SEC and fractions were analysed by 12.5% SDS-PAGE (**Figure 5.4C**). A band running higher than expected can be seen in SEC fractions (ca. 40 kDa), when compared with the estimated ca. 36 kDa size of Nmar_1131. Notably, an improvement in purity is displayed after SEC purification, despite the improved purity, other faint bands at >40 kDa can be seen, suggesting potential co-eluting contaminant proteins (**Figure 5.4C**). To clarify whether the band running ay ca. 40 kDa is Nmar_1131 target protein, peptide identification can be carried out by mass spectrometry on purified product.

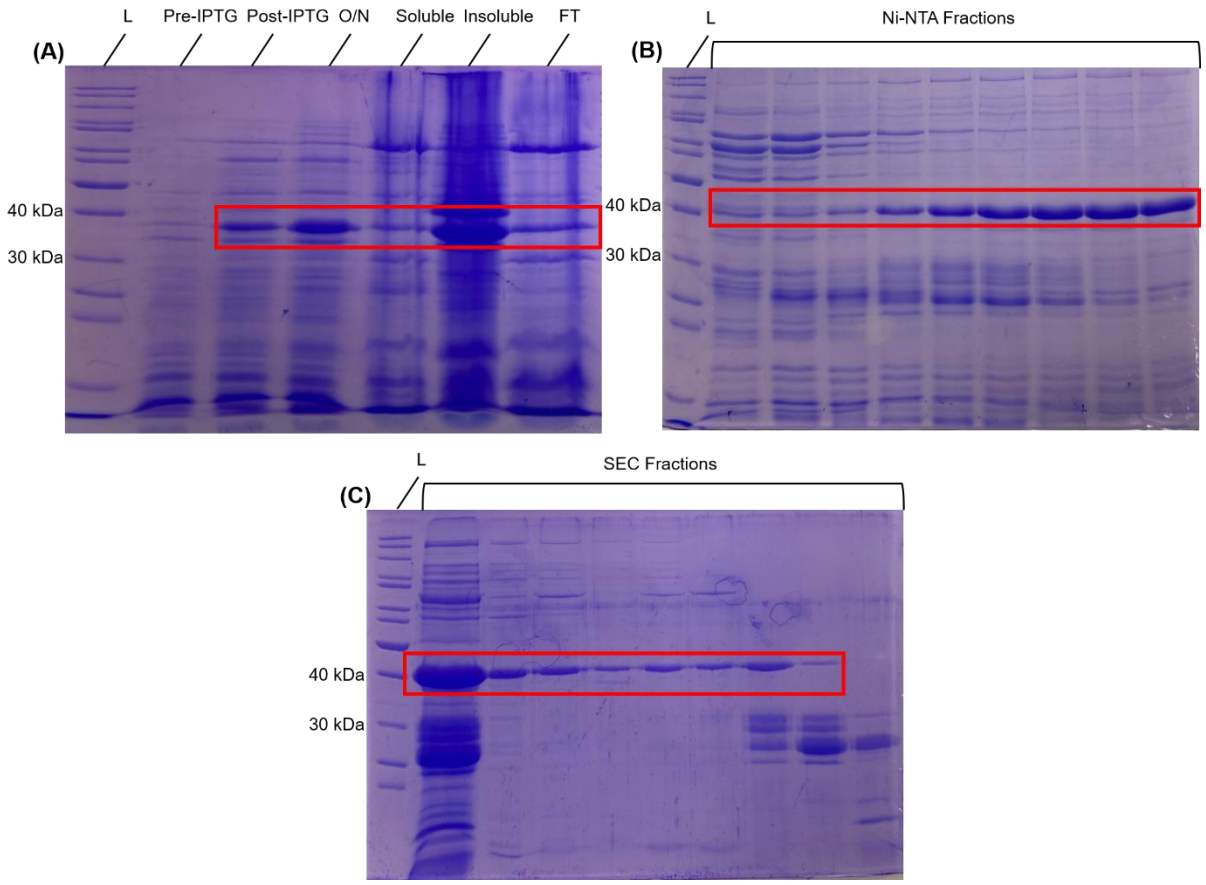


Figure 5.4. 12.5% *(w/v)* SDS-PAGE gel of Nmar_1131. **(A)** = Cell samples at different stages of expression, **(B)** = Ni-NTA fractions and **(C)** = size-exclusion chromatography (SEC) fractions. MCO Nmar_1131 contains a MW ca. 36.3 kDa. L = 200 kDa PageRuler protein ladder (ThermoScientific), O/N = overnight and FT = flow-through. Red box outlines potential target Nmar_1131.

The expression trials of Nmar_1354, which has a molecular weight of approximately 47 kDa, did not yield any bands on SDS-PAGE at the expected weight of ca. 47 kDa (**Figure 5.5**). In contrast to Nmar_1131, where expression was associated with inclusion body formation (**Figure 5.4**), Nmar_1354 does not reveal successful expression in either the soluble fraction or insoluble fraction (**Figure 5.5A**). Ni-NTA fractions align with this finding as no target bands at ca. 47 kDa is revealed (**Figure 5.5B**).

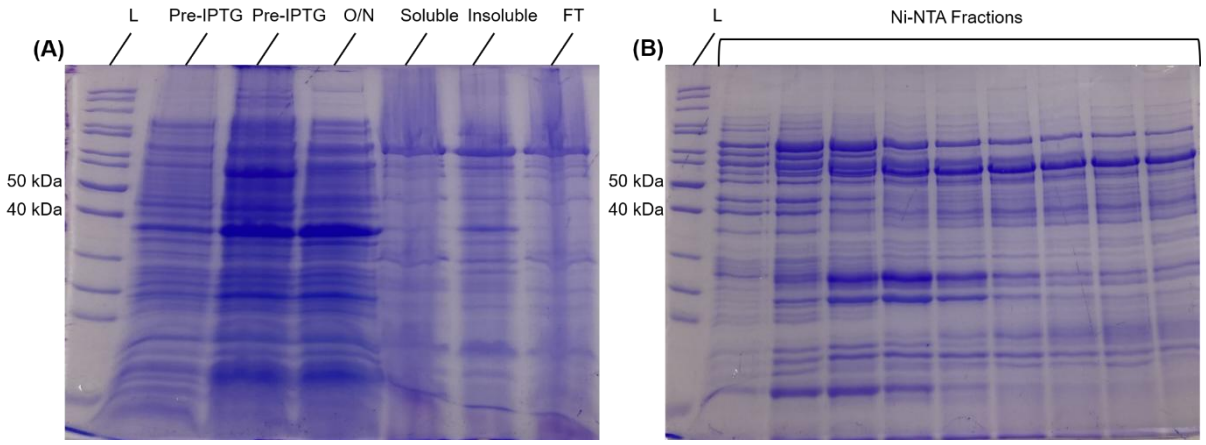


Figure 5.5. 12.5% *(w/v)* SDS-PAGE gel of Nmar_1354. **(A)** = Cell samples at different stages of expression, **(B)** = Ni-NTA fractions and **(C)** = size-exclusion chromatography (SEC) fractions. MCO Nmar_1354 contains a MW ca. 47 kDa. L = 200 kDa PageRuler protein ladder (ThermoScientific), O/N = overnight and FT = flow-through.

Expression of Nmar_1667, which has a molecular weight of approximately 47 kDa, produced a faint band at ca. 47 kDa (**Figure 5.6A**) in the post-IPTG and overnight (O/N) lanes. A prominent band at ca. 47 kDa can also be observed in the flow-through of Ni-NTA flow through (**Figure 5.6A**), suggesting incomplete binding of Nmar_1667 to nickel resin. Despite the soluble fraction showing no clear evidence of substantial Nmar_1667, nevertheless Nmar_1667 soluble fraction was subjected to Ni-NTA purification. However, eluate fractions revealed no clear target band at ca. 47 kDa (**Figure 5.6B**).

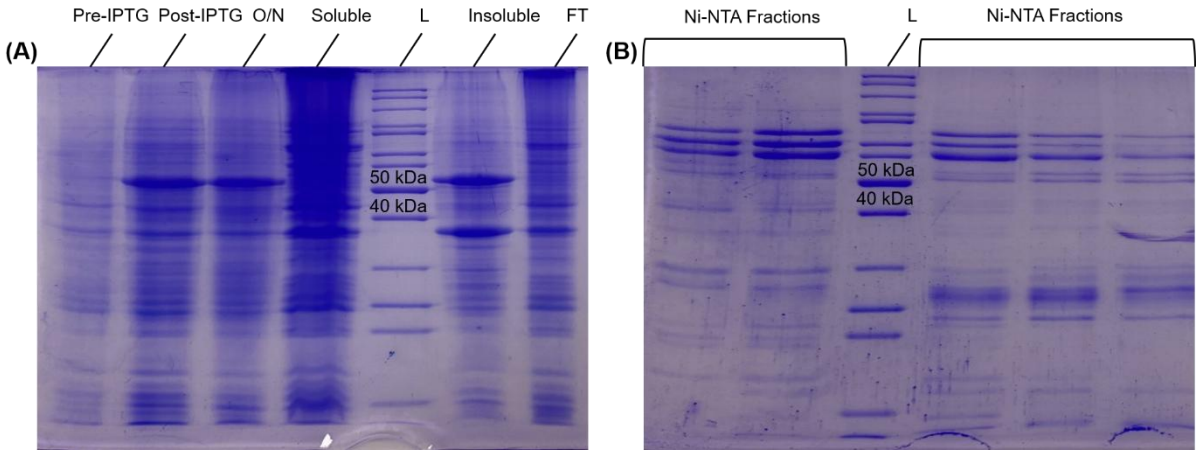


Figure 5.6. 12.5% *(w/v)* SDS-PAGE gel of Nmar_1667. **(A)** = Cell samples at different stages of expression, **(B)** = Ni-NTA fractions and **(C)** = size-exclusion chromatography (SEC) fractions. MCO Nmar_1667 contains a MW ca. 47 kDa. L = 200 kDa PageRuler protein ladder (ThermoScientific), O/N = overnight and FT = flow-through.

Expression of NFRAN-2427 MCO, with a molecular weight of approximately 50 kDa in 2xYT media led to higher yields compared to LB media, thus SDS-PAGE analysis for cell samples were overloaded and distinguishing bands proved impossible. However, a distinct band at ca. 50 kDa can be seen post Ni-NTA and SEC purification (**Figure 5.7**), suggesting soluble target NFRAN_2427 MCO was present in soluble fraction. Despite, successful capture of target NFRAN_2427 MCO using IMAC (**Figure 5.7A**), SEC purification reveals the presence of many bands (**Figure 5.7B**), thus sample was not sufficiently pure for downstream functional and structural experiments.

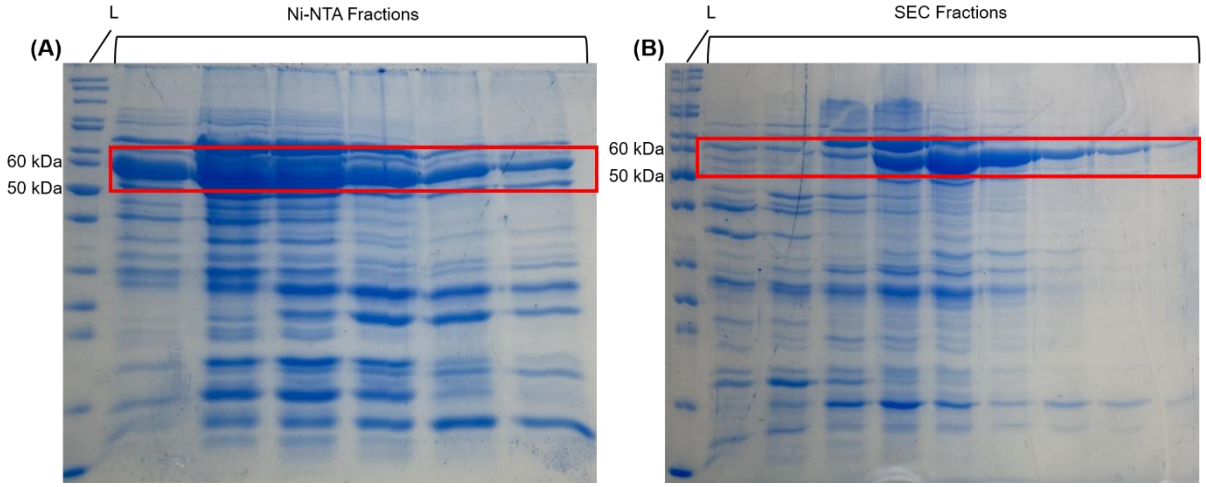


Figure 5.7. 12.5% *(w/v)* SDS-PAGE gel of NFRAN_2427. **(A)** = Ni-NTA fractions and **(B)** = size-exclusion chromatography (SEC) fractions. MCO NFRAN_2427 contains a MW ca. 50 kDa. L = 200 kDa PageRuler protein ladder (ThermoScientific). Red box outlines potential target NFRAN_2427.

An AS cut strategy was implemented for NFRAN_2427 MCO, as protein precipitation was observed when NFRAN_2427 MCO was buffer exchanged into buffer containing ≥1.5 M AS. Screen of AS at varying concentration was used against 2 mg mL^{-1} NFRAN_2427, with 2.0 M AS displaying highest intensity band at ca. 50 kDa (**Figure 5.8A**). NFRAN_2427 was buffer exchanged into buffer containing ≥2.0 M AS, post SEC purification, leading to a relatively pure target band sitting between 60-50 kDa (**Figure 5.8B**). However, AS cut led to very low yields of collected NFRAN_2427, with <1 mg mL^{-1} in 100 µL of final sample.

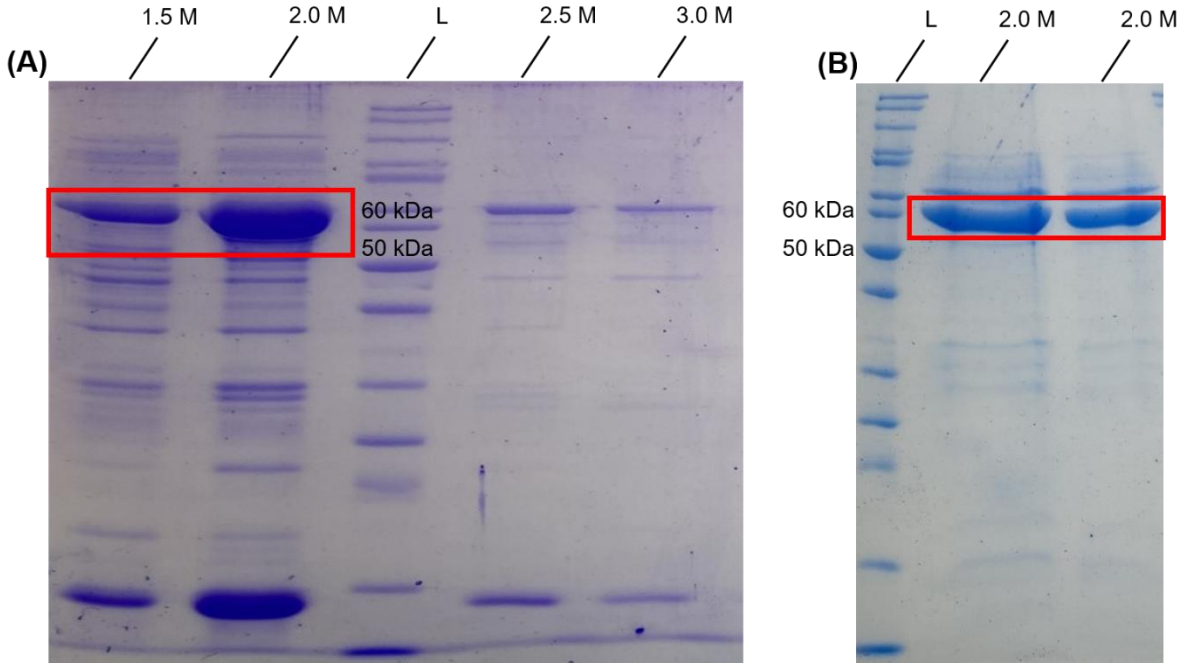


Figure 5.8. 12.5% *(w/v)* SDS-PAGE gel of NFRAN_2427 supernatant from AS cuts. **(A)** = AS cut of NFRAN_2427 MCO at varying AS concentrations (M) and **(B)** = 2.0 M AS concentration that yielded purest sample. MCO NFRAN_2427 contains a MW ca. 50 kDa. L = 200 kDa PageRuler protein ladder (ThermoScientific). Red box outlines potential target NFRAN_2427.

5.3.3 Western Blot Analysis

To clarify whether target proteins were expressing from the pET28a vector and gain further insight into where target protein reside at different stages of expression/purification, an anti-6xHis western blot was carried out. Western blots from Nmar_1131 and Nmar_1667 reveal a prominent band in the insoluble fraction around expected molecular weights (ca. 36 kDa) for Nmar_1131 (**Figure 5.9**) and (ca. 47 kDa) for Nmar_1667 (**Figure 5.9**). An intense band in Ni-NTA fractions can be seen for Nmar_1131, however, multiple bands at lower size are shown in SDS-PAGE gel (< 36 kDa) (**Figure 5.9**), indicative of possible protein breakdown with products having the N-terminus intact. Nmar_1667 displays a fainter band at around ca. 47 kDa , with multiple bands seen at < 47 kDa (**Figure 5.9**), similar to Nmar_1131 containing possible protein breakdown. Contrastingly, Nmar_1354 does not show any distinguishable expected molecular weight bands (ca. 47 kDa) (**Figure 5.9**). NFRAN_2427 shows very faint bands between 55-35 kDa from Ni-NTA fractions, aligning with low yields obtained post AS cut. Bands seem to sit ca. 50 kDa, indicative of successful expression of target NFRAN_2427 (**Figure 5.9).**

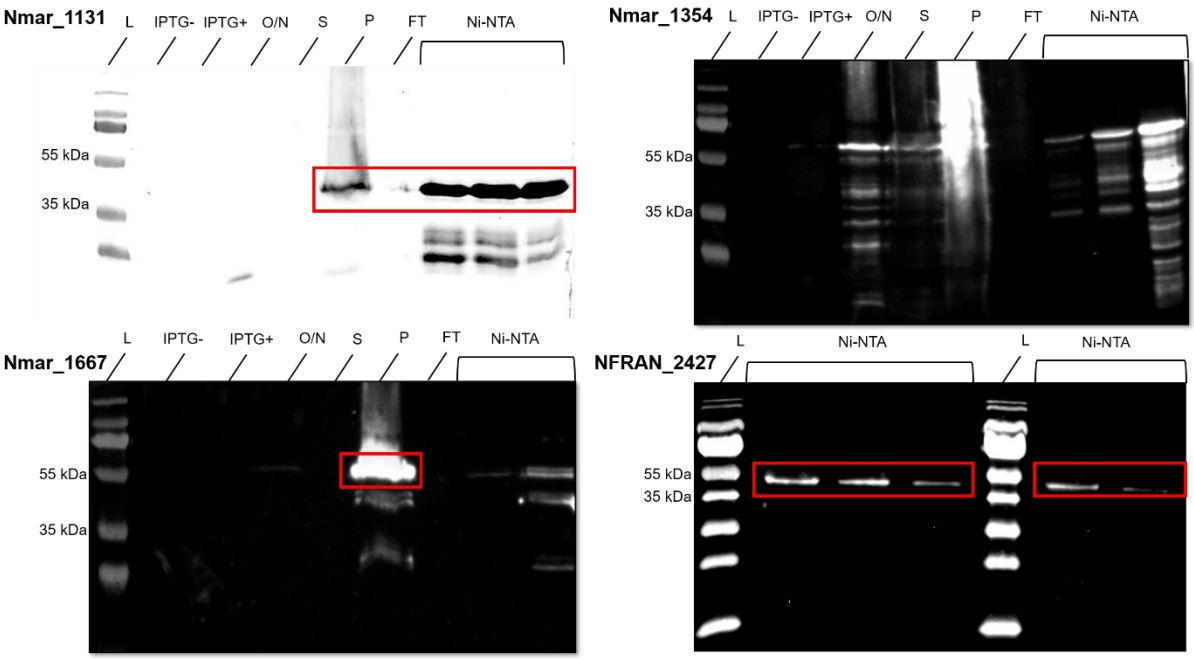


Figure 5.9. 12.5% anti-6xHis Western blots of Nmar_1131, Nmar_1354, Nmar_1667 and NFRAN_2427. IPTG- = Pre-IPTG, IPTG+ = Post-IPTG, O/N = overnight, S = soluble fraction, P = insoluble fraction (pellet), FT = flow-through and L = 250 kDa pre-stained PageRuler protein ladder (ThermoScientific). Red box outlines potential target proteins.

5.3.4 Inclusion body extraction of Nmar_1131

An attempt to extract inclusion bodies from Nmar_1131 was carried out. SDS-PAGE analysis (**Figure 5.4A**) and anti-6xHis western blot (**Figure 5.9A**) reveals this protein to remain in insoluble fraction (pellet) after cell lysis and centrifugation. A protocol on inclusion body extraction and refolding was replicated from Kim *et al*., (2015). However, attempts were unsuccessful as Nmar_1131 MCO remined in all insoluble fractions throughout the protocol (**Figure 5.10**). SDS-PAGE reveals no target bands of ca. 36 kDa or in S (soluble) fraction post dialysis and centrifugation, on the other hand, the pellet from centrifuged dialysate revealed target band ca. 36 kDa (**Figure 5.10**).

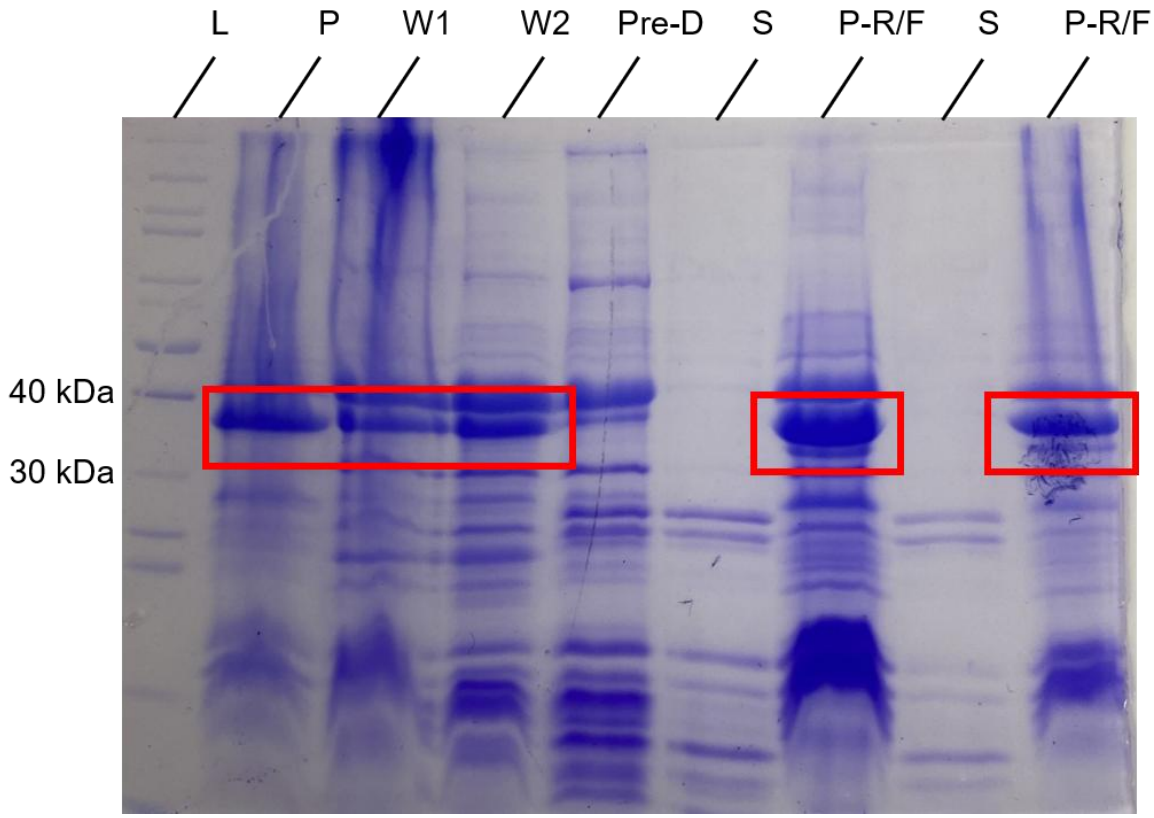


Figure 5.10. 12.5% *(w/v)* SDS-PAGE gel from Nmar_1131 extraction of inclusion bodies and refolding attempt. P = insoluble fraction (pellet), W1/W2 = washes 1 and 2, Pre-D = cells before dialysis, S = soluble fraction and P-R/F = insoluble pellet fraction after inclusion body extraction and refolding. L = 250 kDa pre-stained PageRuler protein ladder (ThermoScientific).

5.4 Discussion

During this study a wide range of inducing (IPTG) concentrations (0.1, 0.25, 0.5, 1.0 mM), temperatures (16°C, 20°C, 25°C, 30°C, 37°C), and shaking speeds (static, 80 rpm, 100 rpm, 120 rpm, 150 rpm and 180 rpm) were used. Despite this, inclusion bodies from Nmar1131 persisted, together with no clear SDS-PAGE bands seen for Nmar_1354 and Nmar_1667. NFRAN_2427, continued to yield very low quantities of protein from a AS cut, with SDS-PAGE revealing an extra band, suggesting inadequate purity for crystallisation (**Figure 5.7B**).

Nmar_1131 and NFRAN_2427 yielded the most promising results, despite total yield after purification coming to < 1mg mL^{-1}. Inclusion body issues were persistent in Nmar_1131 as shown in SDS-PAGE/anti-6xHis western blots (**Figure 5.4A** and **5.9**) and proposedly NFRAN_2427 had issues with inclusion body formation, due to its very low yields and no clear band being obtained in insoluble fraction in SDS-PAGE gel. The formation of inclusion bodies is driven by hydrophobic interactions, which acts as shielding from surrounding aqueous solution (Bhatwa, *et al*., 2021; Winkler *et* al., 2010). This results in protein aggregation and promotes nucleation by acting as seeds for aggregation of similar proteins (Morell *et al*., 2008). This is commonly found in *E. coli* hosts, as they do not possess the proper protein-modification machinery, this coupled with high expression results in hydrophobic residue exposure and inclusion body formation (Bhatwa, *et al*., 2021; Schramm, *et al*., 2020). A variety of factors influences inclusion body formation in *E. coli*, factors include, environmental conditions, target protein properties, host cell metabolism, protein synthesis and modification machinery (Bhatwa, *et al*., 2021; Donovan, *et al*., 1996; Strandberg and Enfors, 1991). High rates of protein synthesis, commonly applied to recombinant protein production, using expression vectors with strong promoters, such as the T7 promoter for this study (pET28a) with high copy numbers and codon optimised in engineered *E. coli* for fast growth (Bhatwa, *et al*., 2021; Zeng and Yang, 2019). This fast growth rate could exceed the ability of the *E. coli* host cells to

manage target protein post-translational modifications and folding, resulting in target protein misfolding and aggregate formation as target protein hydrophobic residues, which is usually buried in native protein is exposed on its surface. (Walsh, *et al*., 2010). Another factor is the lack of proper *E. coli* host machinery for target protein post-translational modifications (PTMs), leading to misfolding, particularly when expressing mammalian, plant, archaeal and fungal proteins in *E. coli* (Bhatwa, *et al*., 2021). To achieve native, biologically active target protein, proper machinery for PTMs is essential and can also alter other target protein characteristics, such as hydrophobicity, solvent access and protein charge (Walsh, *et al*., 2010).

Inclusion body formation from protein aggregation is also dictated by the structural and physiochemical features of the target proteins themselves (Bhatwa, *et al*., 2021). This included target protein molecular weight (kDa), number of low complexity regions and abundance of contiguous hydrophobic residues. Previous studies investigated the relationship of protein molecular weight and abundance of soluble proteins produce from *E. coli* (Dyson, *et al*., 2004; Goh, *et al*., 2004). Overall single-domain target proteins with a lower molecular weight were soluble and maintained protein functionality. The reason for this, is partly due to larger, multi-domain proteins requiring greater numbers of folding intermediates for protein folding pathway (Wilkinson and Harrison, 1991). Multi-domains proteins, increase the likelihood of protein aggregation, as abundance of disordered and large stretches of hydrophobic regions increases the risk of inclusion body formation (Bhatwa, *et al*., 2021).

Various strategies can be implemented to minimise formation of recombinant protein inclusion bodies in *E. coli*. Previous studies trying to minimise inclusion body formation include, introducing a short heat shock (45°C) prior to induction (Oganesyan, *et al*., 2007). Bacterial defence mechanisms include the releasee of heat shock proteins, which are molecular chaperones, some of them actively assist in protein folding. Target proteins may be assisted by native bacterial chaperones to prevent the formation of inclusion bodies and improve solubility

(Baneyx and Mujacic, 2004). Concentration of inducer (IPTG) could be further reduced to 0.01-0.05 mM or adding 1-2% of glucose to lower the rate of protein expression *via* catabolic repression (Sina, *et al*., 2015; Jhamb and Sahoo, 2012; Grossman, *et al*., 1998). Osmotic stress using sorbitol (0.5-1.0 M) and NaCl (0.2-0.8 M) can be used on *E. coli*, similarity to heat shock, this promotes native bacterial chaperone synthesis to assist target protein folding (Diamant, *et al*., 2001; Blackwell and Horgan, 1991). Additionally, ethanol (3% *(v/v)*) can be used to elicit head shock for chaperone production (Oganesyan, *et al*., 2007). Co-factors, specifically for metalloenzymes such as MCOs, incubation in copper may aid in proper protein folding (Bushmarina, *et al*., 2006). Co-expressing chaperones may aid in proper protein folding in *E. coli*, including GroEL, GroES, DnaK, DnaJ, ClpB, etc. (Jhamb and Sahoo, 2012; Marco, *et al*., 2007). A weaker promoter to slower expression rate, such as pBAD promoter, which also aids in the recovery of target protein in soluble fractions (Hjelm, *et al*., 2017). A fusion tag that promotes protein solubility could be useful, such as maltose binding protein (MBP) or Spheroplast protein Y (Spy), which has previously increased recombinant protein solubility ca. 30 fold (Ruan, *et al*., 2020; Vu, *et al*., 2014).

Despite the challenges faced during overexpression and purification of Nmar_1131, Nmar_1354, Nmar_1667 and NFRAN_2427, a recent study managed to solve the X-ray crystal structure for 2dMCO Nmar_1354 containing “dangler” cupredoxin unit at 1.84 Å resolution (Voland, *et al*., 2025) (**Figure 5.1**). Similar to the expression trials in this study, Voland, *et al* 2025 induced cells at OD600 ca. 1.0 and induced *E. coli* at final concentration of 0.4 mM IPTG. Cultures were left to grow for 18 h prior to harvesting and cells were lysed using osmotic shock. The *Nmar-1354* gene was cloned into a pET22(b)+ vector, consisting of a pelB leader sequence and 6xHis tag at the C-terminal. As the pelB sequence leads target protein to *E. coli* periplasm, osmotic shock can yield fairly pure target protein before purification protocols. Interestingly, in the study by Voland *et al* 2025., Nmar_1354 was not stable in SDS-PAGE conditions, with

pure Nmar_1354 revealing multiple bands in SDS-PAGE gel despite it being pure (Voland, *et al*., 2025). This may explain the SDS-PAGE gels of Nmar constructs displaying multiple bands in SDS-PAGE and anti-6xHis western blots (**Figure 5.5B** and **5.9**). Nmar_1354 showed ability to oxidise NH_2OH to generate HNO, which then dimerizes to N_2O using O_2, aligning with copper dependence of AOA to drive the nitrification pathway. Nmar_1131, Nmar1667 and NFRAN_2427 may function in a similar manner and these finding offer an explanation as to why AOA lack canonical heme-based HAO found in AOBs.

Conclusions

In conclusion, although this study proved unsuccessful in obtaining adequate MCO yields for functional and structural studies, many strategies are available to counteract inclusion body formation and low expression levels. Carrying out expression trails using pET26b or pET22(b)+ vector with pelB leader sequence on Nmar_111, Nmar_1667 and NFRAN_2427 could be a good starting point.

Chapter 6: Spectroscopic and structural characterization of two-domain LPMOs

6.1.1 Polysaccharide Degradation for Industrial Biotechnology

Complex carbohydrates are found throughout all domains of life and represent some of the most abundant renewable organic materials on Earth (Hemsworth, *et al*., 2016). Polysaccharides found in plants, such as cellulose, hemicellulose and pectin form major components of the cell wall (Pauly and Keegstra, 2008). Additionally, other polysaccharides such as chitin provides structural support in fungal cell walls, insect exoskeletons and crustacean shells (Qiu, *et al*., 2022). Currently, petroleum remains the dominant carbon source for the industrial production of fuels and a wide range of chemical products (Hemsworth, *et al*., 2016). However, increasing global energy demands, the finite nature of fossil fuels and concerns surrounding climate change has intensified the need to develop sustainable forms of energy (Himmel, *et al*., 2007). Consequently, research efforts have focussed on identifying renewable carbon sources to reduce fossil fuel reliance to support a transition towards sustainable, chemical, fuel and material production (Isikgor, *et al*., 2015).

Polysaccharides such as cellulose and chitin represent a rich carbon source, making them attractive for biotechnological conversion into fuels (Hemsworth, *et al*., 2016; Hahn-Hägerdal, *et al*., 2006). The goal is to break down the complex material into the individual carbohydrate component monomers, which are later converted into ethanol and other compounds (Hahn-Hägerdal, *et al*., 2006). These fuels are significantly more environmentally friendly than fossil fuels as CO_2 released during combustion can be reabsorbed from the atmosphere through photosynthesis as new crops are grown (Drewniak, *et al*., 2015; Slade, *et al*., 2009). Research has pushed for the enzymatic degradation of polysaccharides, as this method makes it possible to use highly selective reactions that can be performed by naturally designed proteins for biotechnological applications (Poshina, *et al*., 2018). Furthermore, high performance of enzymatic degradation of insoluble natural polysaccharides (cellulose and

chitin) that contain a highly ordered structure is hard to achieve. The highly ordered structures and high crystallinity of recalcitrant cellulose and chitin make them considerably difficult for accessibility of the substrate for enzymatic breakdown (Poshina, *et al*., 2018). To overcome this, Lytic polysaccharide monooxygenases (LPMOs) can be used as early-stage disruption of crystalline polysaccharide surfaces using copper dependant oxidative chemistry to introduce chain breaks in resistant polysaccharide surfaces (Quinlan, *et al*., 2011; Vaaje-Kolstad, *et al*., 2010). This oxidative cleavage generates new access points for glycoside hydrolases, cellulases and chitinases, thereby enhancing polysaccharide depolymerisation (Calderaro, *et al*., 2021; Vaaje-Kolstad, *et al*., 2010).

6.1.2 Lytic Polysaccharide Monooxygenases

Lytic polysaccharide monooxygenases (LPMOs) are copper-dependent enzymes and are responsible for the oxidation of recalcitrant polysaccharides, such as chitin and cellulose (Kumar, *et al*., 2024; Bissaro and Eijsink, 2023; Forsberg and Courtade, 2023; Rieder and Sørlie, 2023). First discovered in the 1950s, Reese, *et al*., 1950 discovered the concept of non-hydrolytic cleavage of cellulose, increasing accessibility for cellulases. LPMOs are co-cultured with other glycoside hydrolases (GH) to boost the breakdown of recalcitrant/crystalline polysaccharides (Hemsworth, 2023; Walton, *et al*., 2023; Moon, *et al*., 2022). In 2010, CBP21 showed capabilities of glycosidic cleavage of chitin, revealing increased potential for biomass breakdown (Vaaje-Kolstad *et al*., 2010). LPMOs are classified by their auxiliary activities (AA) (AA9, AA10, AA11, AA13-AA17) (Tamburrini, *et al*., 2021), which are grouped by organism and LPMO activity. The auxiliary activity families are listed in the CAZy (Carbohydrate-Active enZYmes) database (Levasseur, *et al*., 2013), that categorises carbohydrate-active enzymes based on their sequence similarity. This makes it a key tool for understanding how carbohydrate-active enzymes function in carbohydrate metabolism (Cantarel, *et al*., 2009).

LPMOs bind a single Cu ion through a motif known as the histidine brace (Quinlan, *et al*., 2011) (**Figure 6.1**) and use this metal centre to perform oxidative chemistry on the polysaccharide substrate (Meier, *et al*., 2018). This action enhances accessibility of GHs, therefore boosting efficiency of biomass degradation. The enzymatic depolymerisation of long chain polysaccharides in biomass degradation is highly beneficial in biotechnological applications in biorefineries containing a cocktail of LPMOs and GHs; to produce commodity chemicals, biofuels and biomaterials (Hemsworth, 2023; Moon, *et al*., 2022; Vandhana, *et al*., 2022; Ipsen, *et al*., 2021; Walton, *et al*., 2023).

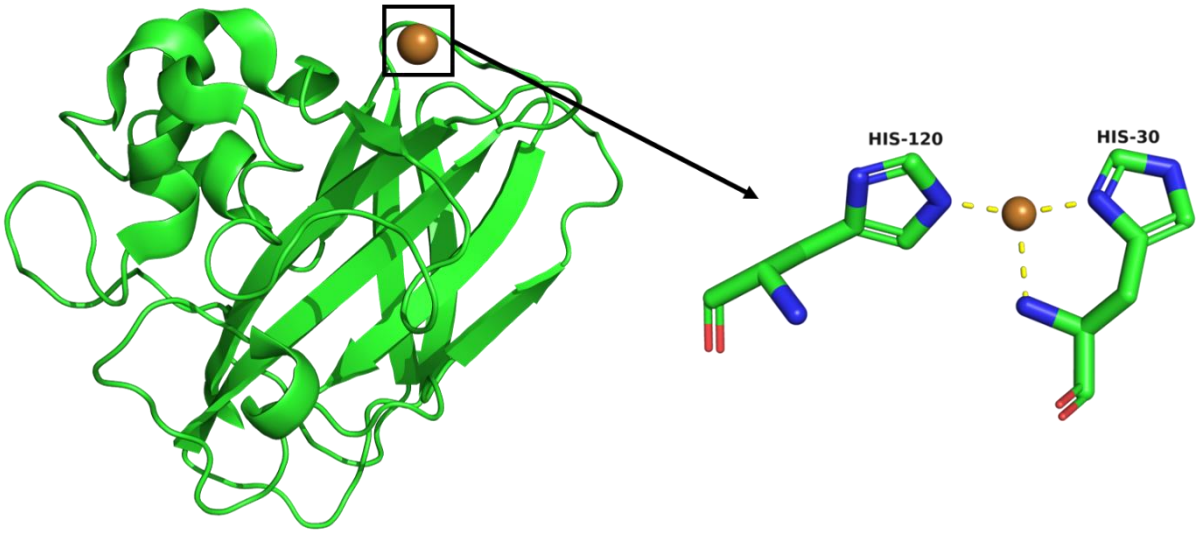


Figure 6.1 Crystal structure of AA10 *Sli*LPMO10E and its copper active site. *Sli*LPMO10E crystal structure shown on the left in cartoon representation showing β-strands and α-helical regions. Histidine brace (blue sticks) can be seen on the right, coordinated by the nitrogen's on Nε of His-31, the Nδ of His-106 and the nitrogen of His-31 side chain. PDB 5ftz (Chaplin *et al*., 2016).

Understanding the catalytic cycle of LPMOs is an ongoing active area of research. It is widely accepted that LPMOs can carry out their oxidative chemistry through activation of either molecular oxygen (O_2) or using the two-electron reduced form of hydrogen peroxide (H_2O_2) (Meier, *et al*., 2018; Walton and Davies, 2016; Tandrup *et al*., 2018). To start the catalytic cycle the resting state Cu(II)-LPMO needs to be reduced to Cu(I)-LPMO by an external reductant and then bind to the polysaccharide substrate (R-H), positioning the active Cu site near to the target glycosidic bond. Two mechanisms have been proposed for O_2 activation based on the active oxidant: the superoxo mechanism and oxyl mechanism (Batka *et al*., 2024), (Hagemann and Hedegård 2023). In the superoxo mechanism, Cu(II)-superoxo LPMO undergoes hydrogen atom abstraction (HAA), resulting in hydroxylated polysaccharide substrate (**Figure 6,2** (3-4)). In the oxyl mechanism, Cu(II)-superoxo is protonated and further reduced to produce a Cu(II)-oxyl, in turn, again resulting in a hydroxylating polysaccharide

substrate (**Figure 6.2** (5-6), (8-9)) (Batka *et al*., 2024). Alternatively, Cu(II)-oxyl can be directly formed in the presence of H_2O_2 (**Figure 6.2** (7)). The instability of the hydroxylated product results in spontaneous decomposition and cleavage of glycosidic bonds.

As expected for metalloenzymes, catalytic competency depends critically on the electronic structure of the active site as illustrated above. In the case of LPMOs this is highly influenced by the histidine brace active site. In their resting state, the Cu(II) centre coordinated by the N-terminal amino group and two histidine nitrogen's, generates an axial ligand field that produces a characteristic $d_{x^2-y^2}$-based ground state (Lindley, *et al*., 2022; Ciano, *et al*., 2020). Subtle perturbations to this geometry arising from substrate binding, reductive activation, or interaction with exogenous ligands can alter the distribution of unpaired spin density and thus the reactivity of the copper centre (Courtade, *et al*., 2020). Electron paramagnetic resonance (EPR) spectroscopy provides a direct probe of these changes by reporting on the g-tensor, hyperfine coupling, and spin-delocalisation pathways associated with the Cu(II) ion and its coordinating residues and has thus been used as valuable experimental tool in LPMO research (Lindley, *et al*., 2022; Ciano, *et al*., 2020; Courtade, *et al*., 2020).

Streptomyces lividans is a Gram-positive, filamentous, soil dwelling bacterium, responsible for aerobic biomass decomposition and plays a pivotal role in the decomposition of chitin and cellulose. They are also the producers of a wide variety of bioactive natural products, including more than half of all clinical antibiotics (AnnÃ and Mellaert, 1993), (Ahmed, *et al*., 2020). A total of seven LPMOs are present in *S. lividans*, of which three exist as a single LPMO domain (*Sli*LPMO10A, *Sli*LPMO10B and *Sli*LPMO10E), three as two-domain LPMOs (*Sli*LPMO10C, *Sli*LPMO10F and *Sli*LPMO10G) with a CBM5/2 appended *via* a linker and one, *Sli*LPMO10D, which contains a C-terminal transmembrane helix. These LPMOs are all secreted to the extracellular environment leading to a role in biomass breakdown to be proposed. *S. lividans* displays a distinct dependence on Cu for morphological development (Zhong, *et al*., 2022). Recent studies have identified a role of *Sli*LPMO10E working in tandem with a cellulose-synthase to deposit glycans at hyphal tips in the production of aerial hyphae during the life cycle (Zhong, *et al*., 2022). This work therefore illustrates a house-keeping role for an *Sli*LPMO10E as opposed to biomass recycling. Although not described in bacteria, this is found in copper containing X325 proteins from ectomycorrhizal fungus *Laccaria bicolor*, which are evolutionarily related to LPMOs without the ability to oxidatively cleave polysaccharide bonds, However X325s have been linked to hyphal associated biological processes rather than biomass degradation (Labourel, *et al*., 2020).

species can lead to the dissociation of H_2O_2 (**Figure 6.2** (10)) (Forsberg *et al*., 2019), (Batka *et al*., 2024).

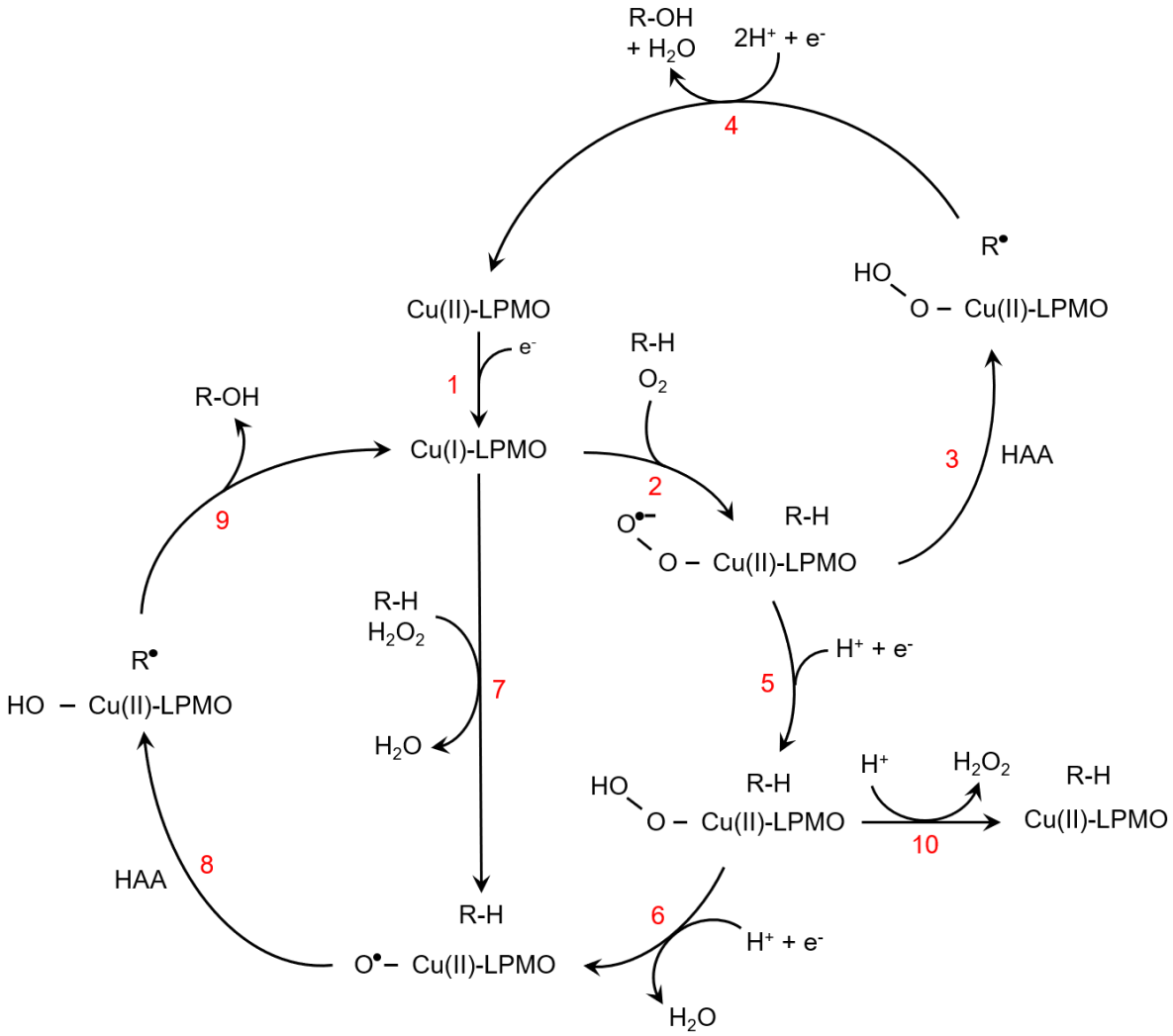


Figure 6.2 Proposed mechanisms for the oxidation of glycosidic bonds in polysaccharides (R-H) (Batka, *et al*., 2024).

The AA10 family (formerly CBM33) (Isaksen, *et al*., 2014) encompasses LPMOs that are found in archaea, bacteria, fungi and viruses. Current experimental evidence for the LPMOs from the AA10 family, reveals strict C1/C4 glycosidic bond oxidation of chitin and cellulose. Chitin-active AA10 LPMOs can either carry out C1, or contrastingly, cellulose-active LPMOs can oxidise only C1 or oxidise both C1 and C4 bonds (**Figure 6.3**). There is no evidence of cellulose-active LPMOs oxidising just C4 carbon in the AA10 family (Li, *et al*., 2021). Transcriptomic studies reveal LPMOs, *Sc*LPMO10E and *Sc*LPMO10G from *S. coelicolor* were significantly upregulated when *S. coelicolor* was grown with chitin as a sole carbon source. Six out of the seven total LPMOs that reside in *S. coelicolor* show activity towards chitin, suggesting a preference for chitin over cellulose in *S. coelicolor* (Votvik, *et al.,* 2023).

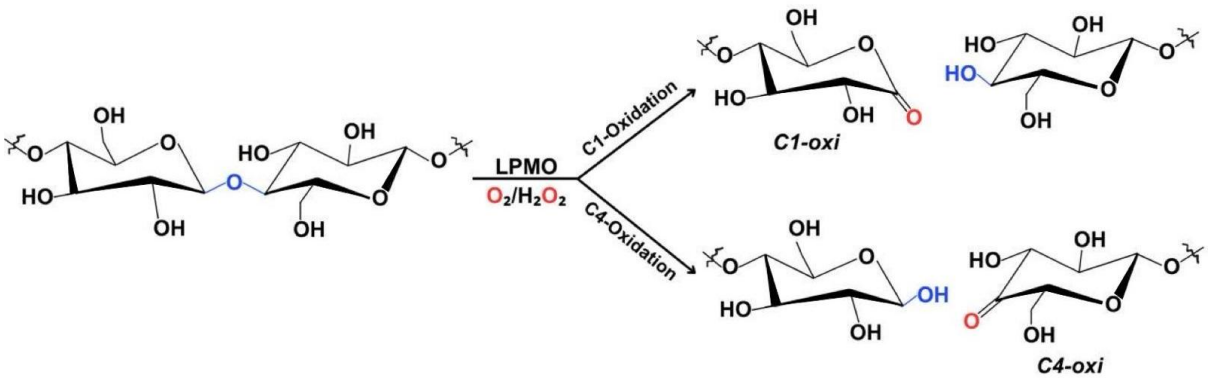


Figure 6.3 Oxidised products formed from cellulose during LPMO catalysis.

6.1.3 Carbohydrate-Binding Modules in Polysaccharide Degradation

Carbohydrate-binding modules (CBMs) are small, non-catalytic protein domains which are part of modular enzymes, particularly carbohydrate-active enzymes (CAZymes). CBMs have proven to be highly effective in enhancing polysaccharide conversion in GHs, and this has led to extensive studies being carried out on CBMs (Forsberg and Courtade, 2023; Guillén, *et al*., 2010). The main role of CBMs in polysaccharide degrading enzymes such as LPMOs and GHs is to bind and direct enzymes to substrate (polysaccharide), thereby increasing the concentration of enzyme on substrate surface (Forsberg and Courtade, 2023). The catalytic

domains and CBMs are connected by a flexible linker which vary in length and composition depending on the type of CBM. The complex nature of CBMs means there is an extensive network of different types, which are divided based on their sequence similarity. In the last 20 years the CAZy database has doubled, with 93 distinct families currently known (Forsberg and Courtade, 2023; Drula, *et al*., 2022). Currently, CBMs are grouped into three types based on their function and structure, "exo-binding" CBMs that bind to glycan chain termini, "endo-type" CBMs that bind internally on glycan chains and "surface-binding" CBMs. The majority of CBMs in LPMOs are "surface-binding" CBMs which are optimal for binding of crystalline surfaces. Despite this, "exo-binding" have been observed in fungal starch oxidising AA13s, together with "endo-type" chitin/chito-oligosaccharide-binding in CBM14 and CBM18 from the AA15 family (Gilbert, *et al*., 2013).

In the AA10 family of LPMOs, multimodular members display greater performance in chitin/cellulose oxidation with tighter substrate binding, compared to the catalytic domain only LPMOs (Forsberg and Courtade, 2023). Recent studies investigating truncated LPMO (catalytic domain only) and full-length (LPMO domain + CBM2) *Sc*LPMO10C discovered diminished efficiency of cellulose oxidation in the truncated version compared to the naturally occurring counterpart (Courtade, *et al*., 2018). The CBM2 of *Sc*LPMO10C provided enhanced binding affinity to cellulose, in turn, minimizing diversion into off-pathway reactions which would ultimately deactivate the enzyme (Forsberg, *et al*., 2023) (**Figure 6.4**). Off-pathway reactions refer to reactions of unbound, reduced LPMO with H_2O_2 or O_2 which does not lead to productive cleavage of polysaccharide. This is followed by oxidative damage of LPMO in the histidine brace coordination site, causing Cu leakage and irreversible deactivation (Stepnov, *et al*., 2022; Paradisi, *et al*., 2019).

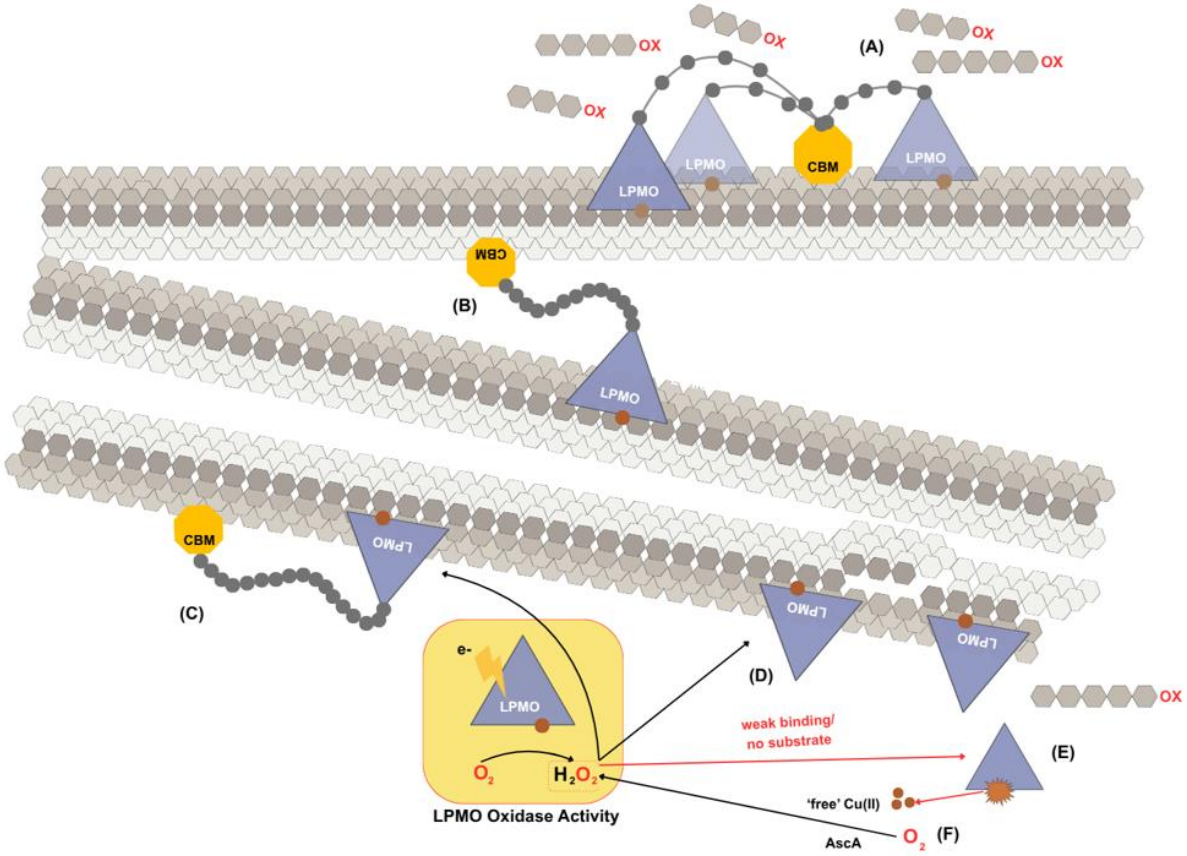


Figure 6.4 Dynamics of LPMOs that contain CBM connected by a flexible linker. **(A)** shows LPMO domain (blue) connected to a CBM (yellow) by a flexible linker (grey), illustrating how the CBM can dock the substrate in different locations, increasing chance to oxidise polysaccharide chain in close proximity and increases release of shorter polysaccharide chains. **(B)** Illustrates how CBM can bind to one polysaccharide chain while the LPMO can oxidise another chain. LMPO with no linker spend longer time in solution, which leads to higher oxidase activity and greater generation of H_2O_2 and can fuel LPMO reactions **(C, D)**. Alternatively, LPMO Cu active site can undergo oxidative damage followed by enzyme inactivation **(E)**. This ultimately leads to leakage of Cu(II) from LPMO active site followed by acceleration of H_2O_2 production.

6.1.4 Structural Dynamics of Linkers in Two-domain LPMOs

Intrinsically disordered linker regions in proteins adopt a heterogeneous array of conformations and therefore elucidating conformational ensembles of multimodular proteins is critical for a mechanistic understanding of their functional dynamics. Despite the role of linkers in connecting the Cu-binding LPMO domain with CBMs, linkers are also increasingly recognised as crucial functional elements which modulate structural and functional interactions between folded domains (Tuveng, *et al*., 2020). Characteristics such as amino acid sequence and length determine linker functionality. Different residues in the linker may promote extended conformations or promote flexibility with conformational freedom. For example, serine-rich linkers are thought to be more flexible than pro-rich linkers (Forsberg, *et al*., 2023), charged residues (arginine, lysine, aspartate and glutamate) can promote extended conformations while residues such as glycine may promote flexibility and greater conformational freedom in linkers (Chen, *et al*., 2013), (van Rosmalen, *et al*., 2017).

Despite linkers playing an important role in regulating interactions between the Cu-binding LPMO/CBM domain and the overall shape of two-domain LPMOs, there is limited studies regarding the functional and structural role of these linkers. Previous studies on LPMO linkers include linker truncations, where shortening of the linker in fungal *Bc*LPMO9C resulted in lower cellulose-binding affinity (Srivastava, *et al*., 2022). Previously mentioned studies carried out by Courtade, *et al*., 2018, revealed NMR data that showed extended conformation of linker region, allowing LPMO domain and CBM to move independently from each other. Small-angle X-ray scattering (SAXS) data of *O*-glycosylated linker from fungal *Hj*LPMO9A shows how the linker can be both extended as well as highly flexible, to tightly regulate binding of the linker to the LPMO domain (Higasi, *et al*., 2021).

The flexible nature of these linkers presents challenges for structural characterization of two-domain LPMOs. These challenges arise as conformational heterogeneity prevents

uniform packing into a crystal lattice hindering crystallization, making X-ray crystallography increasingly difficult. On the other hand, truncation to LPMO or CBM units on their own, increases the chances of obtaining crystals for diffraction studies, providing an alternative way to obtain a high-resolution X-ray crystal structures of the individual domain structures (Ma, *et al*., 2022), (Méndez-Líter, *et al*, 2021). Combined biophysical and computational approaches can be used to obtain insight into the structure of two-domain LPMOs. One such approach is size-exclusion chromatography small-angle X-ray scattering (SEC-SAXS), to determine average dimensions, molecular envelope data and flexible conformations. Experimental SEC-SAXS data (Gräwert, *et al*., 2020) can be used together with SAXS-based computational tools such as Ensemble Optimisation methods (EOM) (Sagar, *et al*., 2021) or MultiFoXS (Schneidman-Duhovny, *et al*., 2016) to generate a pool of thousands of possible conformations to select best-fit models.

6.1.4.1 Aims

The aim of this chapter was to obtain structural and active site information on the three two-domain LPMOs present in *S. lividans*, *Sli*10C, *Sli*10F and *Sli*10G. Constructs that were designed for over-expression in *E. coli* were provided and the purified two-domain proteins were tested for Cu(II) binding using fluorescence spectroscopy. In addition, electron paramagnetic resonance (EPR) spectroscopy has been used as it can directly probe the Cu(II) active site in LPMOs, allowing to determine the geometry, ligand environment, protonation state, and coordination number. By constructing an LPMO domain only construct for the *Sli*10G protein, then a comparison of the EPR properties of the Cu(II) site in the presence and absence of a CBM can be made. Finally, SEC-SAXS experiments with carried out on the two-domain samples at the beamline BL21 at Diamond Light Source. These data were analysed alongside SEC-SAXS data for the LPMO domain only of *Sli*LPMO10G, for which an X-ray crystal structure was determined.

6.2 Methods

6.2.1 LPMO Cloning and Site-directed Mutagenesis

A modified pET26b (Novagen) plasmid was used to clone into the genes for the two-domain LPMOs as well as the truncated LPMO domains, so that following cleavage of the pelB leader sequence the N-terminal His residue of the histidine brace was the first residue (Chaplin, *et al*., 2016). Multimodular *Sli*LPMO10C, *Sli*LPMO10F and *Sli*LPMO10G, in the pET26b vector were provided by Dr Erik Vijgenboom (University of Leiden, NL). The LPMO only domain of CBD (Catalytic Binding Domain) - *Sli*LPMO10G was provided by Professor Jonathan Worrall. Forward and reverse primers (**Table A.6)** with NcoI and HindIII restriction sites, respectively, were deigned to create *Sli*LPMO10F construct. A PCR mix consisting of 0.2 µM of forward and reverse primer, 0.2 mM dNTPs, 1 µL of template DNA, 1x taq DNA

polymerase buffer (ThermoScientific), 1 µL of taq DNA polymerase (ThermoScientific) and nuclease-free water to make up to a final volume of 50 µL. All PCR reactions were carried out using a Veriti Model 384 well thermocycler (Applied Biosystems). Gel-extraction of the PCR product was carried using the NucleoSpin gel and PCR Clean-up kit (MACHEREY-NAGEL) followed by ligation using T4 DNA ligase (ThermoScientific) at room temperature for 16 hours, followed by transformation to TOP10 (*Escherichia coli*) competent cells, with the resulting transformants, mini-prepped using a NucleoSpin plasmid mini-prep kit (MACHEREY-NAGEL) and DNA sequenced (Eurofins) using T7 forward and reverse sequencing primers.

6.2.2 Recombinant Overexpression and Purification

Expression of all LPMOs were carried out in *E. coli* BL21(DE3)-RIL cells. pET26b plasmid containing target proteins were plated onto (1.5% *(w/v)* LB agar plates with 50 µg mL^{-1} of kanamycin and 20 µg mL^{-1} chloramphenicol and left at 37°C overnight. Single colony transformants were picked and inoculated into sterile 5 mL overnight cultures (LB 50 µg mL^{-1} of kanamycin and 20 µg mL^{-1} chloramphenicol) using a sterile toothpick and grown overnight for 16-20 h (37°C, 180 rpm). Cells from the overnight cultures were inoculated into 2 L sterile flasks containing 500 mL of LB media with 50 µg mL^{-1} of kanamycin and 20 µg mL^{-1} chloramphenicol (37°C, 180 rpm). At an OD_{600} of 0.4, the temperature was lowered to 16 °C with the addition of 0.5 M isopropyl-β-D-thiogalactopyranoside (IPTG) (Melford) on reaching an OD_{600} of 0.6. After 16-18 h cells were harvested at 3501 *g* for 20 min at 4°C. The pellet was combined, weighed and resuspended in $1/30^{th}$ of the total culture volume in cold (4 °C) 50 mM Tris/HCl, pH 7.5, 1 mM EDTA and 20% sucrose. The cell pellet was stirred for 1 h at 4 °C with 60 µL addition of $MgSO_4$ (1 M) for every gram of cell pellet for another 30 min (4°C). The smooth cell suspension was then centrifuged at 38,724 *g* for 20 min at 4 °C and the supernatant

collected and stored at 4 °C, followed by cell pellet resuspension in 1/30th volume of original culture of Milli-Q water (4°C). After stirring for 1 h at 4°C and centrifugated (38,724 *g* for 20 min at 4°C), both supernatants were combined. and dialysed in 5 L of 5 mM Tris/HCl, pH 7.5, 1 mM EDTA (two-domain LPMOs) overnight at 4°C. Dialysate was diluted in buffer A (10 mM Tris/HCl, pH 7.5), DEAE column was equilibrated in buffer A and bound LPMO was eluted with a linear salt gradient using buffer B (10 mM Tris/HCl, 1 M NaCl, pH 7.5). Fractions eluting with a UV absorbance at 280 nm were checked using a 15% *(w/v)* SDS-PAGE gel and fractions containing a band on the SDS-gel with the expected molecular weight for the respective LPMO pooled followed by concentrating using a 5 kDa cutoff Centricon (Vivaspin) at 4°C. Once concentrated to < 2 mL, LPMO and two-domain proteins were loaded onto a HiLoad 16/600 Superdex-75 column (Cytiva) equilibrated in buffer C (50 mM Tris/HCl, pH 7.5, 150 mM NaCl). Fractions showing high purity after checking using 15% SDS-PAGE were pooled, further concentrated for experimentation, or stored at -20 ºC.

6.2.3 Concentration determination and preparation of copper solutions.

Concentration of each purified protein was determined using a Cary60 UV-Vis spectrophotometer (Agilent Technologies), using protein extinction coefficients (ε) determined at 280 nm using ProtParam (Gasteiger *et al*., 2005) with the respective protein sequence as the input and are reported in **Table 6.2**. A100 mM $CuSO_4$ stock solution was prepared in buffer A for further dilution and subsequent addition of Cu(II) ions to the as isolated samples. Removal of excess Cu(II) was carried out by using a desalting PD-10 column (Cytiva). For experiments investigating the binding of Cu(II) binding protein samples were first incubated with a final concentration of 10 mM EDTA overnight at 4°C, followed by desalting using a PD-10 column (Cytiva) and LPMOs were buffer exchanged into 50 mM Tris/HCl, pH 7, 150 mM NaCl prior to further experimentation.

6.2.4 Fluorescence Spectroscopy

Fluorescence spectra were measured at 20 °C using a LS 50B fluorimeter (PerkinElmer Life Sciences). Emission was measured between 300-400 nm, following excitation at 330 nm. A 0.2 mM stock solution of $CuSO_4$ was used for titration of Cu(II) ions into protein samples, through the addition of 6 µL aliquots to 3 mL of protein at 2 µM concentration, with each Cu(II) addition corresponding to 0.2 equivalents of Cu(II), Samples were mixed thoroughly after each addition before spectral acquisition.

6.2.5 Substrate Binding Assay

Two-domain LPMOs (10 µM) in a total volume of 20 µL in 50 mM Tris/HCl pH 7.5, 150 mM NaCl were incubated with ca. 5 mg total weight of Whatman filter paper (cellulose) and β-chitin (squid pen ink) at 18 °C overnight. The supernatant was removed by centrifugation (15,000 *g* for 3 min) and kept as the unbound fraction (20 µL). The solid substrate was washed in 2 mL of buffer C followed by centrifugation and repeated three times to obtain the solid bound fraction sample. 20 µL of SDS-PAGE loading buffer with DTT was added to unbound and bound samples. Samples were heated to 90°C for 10 min and loaded to a 15% *(w/v)* SDS-PAGE gel for visualisation.

6.2.6 Electron Paramagnetic Resonance Spectroscopy

Low temperature EPR spectra were measured on a Bruker E500 CW spectrometer equipped with an Oxford Instruments liquid helium temperature controller. Each experimental spectrum was an average of 4 scans collected on 2048 data points. Other instrumental parameters were:,sampling time 25 ms, modulation amplitude 5 G, modulation frequency 100 kHz. Ranges of temperatures (10 K – 35 K) and microwave power values (12.62 mW (12 dB) – 0.0126 mW (42 dB)) were used to find optimal conditions for comparative analysis of

different Cu^{2+} spectra, targeting a trade-off of the EPR signal intensity and resolution of spectral features). The temperature of 20 K and two MW power values of 3.17 mW (18 dB) and 0.2 mW (30 dB) were chosen as the conditions common for all proteins.

The baseline spectra to be subtracted from the EPR spectra of the proteins were obtained as averaged spectra measured of three water samples, at appropriate temperatures and powers. All spectra were shifted along magnetic field axis in accordance with the microwave frequencies used in measurements of different EPR samples, to align all spectra to common g-values. Spectra subtraction with variable coefficient (Svistunenko, *et al*., 2006) was used in deconvolution complex EPR spectra into individual EPR signals. The Cu^{2+} EPR line shape simulations were performed by EasySpin (Stoll and Schweiger, 2006), primarily under the constrains listed below. Some of the constraints have been lifted for simulation of the Component II spectrum (Figure 6.7B).

The constraints:

a) The tensor for the hyperfine interaction of the electron spin with copper nucleus A_{Cu} ($I_{Cu} = 3/2$) was assumed axial, with the $A_{Cux} = A_{Cuy} = 0.1\ A_{Cuz}$ (small variations around empirically picked factor of 0.1 were found to produce a minimal effect on the simulated spectra).
b) All A_{Cu} tensors were assumed co-linear with the *g*-tensor (all Euler angles were assumed zero).
c) The three nitrogen atoms ($I_N = 1$) coordinating the Cu^{2+} ion were assumed to produce isotropic hyperfine interactions ($A_{Nx} = A_{Ny} = A_{Nz} = A_{Niso}$) of the same magnitude for all three atoms.

For the Component II EPR spectrum simulation, constraint a) was lifted: all three A_{Cu} values were considered different: $A_{Cux} = K_x * A_{Cuz}$ and $A_{Cuy} = K_y * A_{Cuz}$, where the coefficients

K_x and K_y were varied during the fitting procedures starting from those which ensured the relative magnitudes of the three A-values close to what was reported form the DFT calculations for the LPMOs from *P. luminescens* (Gómez-Piñeiro, *et al*., 2022) and *B. licheniformis* (without substrate bound) (Courtade, *et al*., 2020).

Similarly, to modify constraint c) in the Component II spectrum simulation, we adopted the approach that the A-values for the three nitrogen atoms, while still approximated as isotropic for each, were different keeping, however, their relative values close to the proportions determined from the DFT calculation (Gómez-Piñeiro, *et al*., 2022) (DFT model B2).

6.2.7 Simulated to experimental spectra fitting methodology

The values of g_z, ΔH_z (the z-component of the individual line width) and A_{Cuz} have been determined with a good confidence directly from experimental spectra (Figure SX01). The esfit EasySpin routine (Stoll and Schweiger, 2006) was used when different number of parameters were varied at different stages of the fitting. Rather than varying g_x and g_y independently, two other values, $g_\perp$ and $\Delta g_\perp$, were varied, where $g_\perp = (g_x + g_y)/2$ and $\Delta g_\perp = (g_y - g_x)/2$ thus yielding $g_x = g_\perp - \Delta g_\perp$ and $g_y = = g_\perp + \Delta g_\perp$. At first step, $g_\perp$ and $\Delta g_\perp$ were varied. The optimal values were then set as the starting values for fitting of four parameters, $g_\perp$, $\Delta g_\perp$, K_x and K_y. Finally, the two line-widths parameters, ΔH_x and ΔH_y, were varied.

6.2.8 Crystallization and X-ray Crystallography of Copper Binding Domain

Prior to crystallisation screening, purified CBM - *Sli*LPMO10G catalytic domain was buffer exchanged into (50 mM Tris/HCl, 150 mM NaCl, pH 8) using a PD-10 column (Cytiva) and concentrated to ca. 13 mg mL^{-1} using a 5 kDa cutoff Centricon (Vivaspin) at 4°C.

Crystallisation conditions were screened for the Cu(II)-loaded CBD - *Sli*LPMO10G LPMO domain (ca. 13 mg mL^{-1}) using an ARI-Gryphon crystallisation robot in 96-well sitting drop MRC plates (Molecular Dimensions). Crystal were found in 0.1 M Tris/HCl, pH 8, 32 % *(w/v)* PEG 6000 (JCSG Plus screen), at a 1:1 protein:screen condition ratio at 0.4 µL. Crystal conditions were tested in batches at different ratios of protein:precipitant solution (1:1, 1:2 and 1:3) at 100 µL final volume, with crystals formed in the 1:2 batches after 4-5 weeks at 18°C. A single large crystal (ca. 40 µm) was fished from crystal batch and soaked in 20 % glycerol cryoprotectant prior to flash-freezing by plunging into liquid nitrogen. Crystallographic data was measured at Diamond Light Source, beamline I04, at 100 K using a 0.953738 Å X-ray wavelength and a EIGER X 16M detector. Data were processed using the xia2.multiplex pipeline, which uses DIALS for spot finding, indexing, integration and scaling of multiplex datasets. Data was scaled and merged using AIMLESS (Evans, 2011) in CCP4i2 (Agirre, *et al*., 2023) and molecular replacement in MrBUMP (Keegan and Winn, 2008) to identify suitable homologous search models from the Protein Data Bank (PDB). Cycles of refinement and model building were carried out using Refmac5 (Murshudov, *et al*., 2011) and Coot (Emsley, *et al*., 2010), respectively in CCP4i2. Riding hydrogens were added during refinement. The structure was validated using tools within Coot. A summary of data collection and refinement statistics are given in Table 6.1. Map files for viewing in PyMOL were generated from mtz files in Phenix version 1.21-5207-000 (Liebschner, *et al*., 2019).

Table 6.1. Crystallographic processing and refinement statistics for the LPMO domain of CBD - *Sli*LPMO10G. Values in parenthesis are for the highest resolution shell.

Space group	P1 21 1
Unit cell (Å,°)	a = 60.7, b = 31.0, c = 70.5, 90.0° 106.6° 90°
Resolution (Å)	52.06 - 1.20 (1.22 – 1.20)
Unique reflections	78,238 (3347)
Multiplicity	33.6 (11.8)
Completeness (%)	98.7 (86.1)
Mn(I/sd)	26.6 (5.1)
R_{merge}	0.162 (1.167)
Wilson B-factor (Å^2)	5.618
$CC_{1/2}$	1.0 (0.92)
R_{work}	0.113
R_{free}	0.132
RMSD bond lengths (Å)	0.0123
RMSD bond angles (°)	1.806
Ramachandran most favoured (%)	98.6
PDB accession code	**TBC**

6.2.9 Size-Exclusion Small Angle X-ray Scattering (SEC-SAXS)

Purified two-domain LPMO samples (*Sli*LPMO10C, *Sli*LPMO10F and *Sli*LPMO10G) together with the CBD - *Sli*LPMO10G were shipped to Diamond Light Source, beamline BL21 for SEC-SAXS experiments (Cowieson, *et al*., 2020). All samples were prepared in 50 mM Tris/HCl pH 7.5, 150 mM NaCl and 2% glycerol, with the addition of glycerol and NaCl to help reduce radiation damage and interparticle effects, respectively. Samples were prepared in triplicate to volumes of 60 µL and at three different concentrations, 5 mg mL^{-1}, 7.5 mg mL^{-1} and 10 mg mL^{-1}. A 10x buffer stock was prepared for matching size-exclusion chromatography using a Shodex. KW-402.5 column and for buffer subtraction. Samples were passed down a Shodex KW-402.5 column equilibrated in 50 mM Tris/HCl pH 7.5, 150 mM NaCl and 2% *(v/v)* glycerol buffer at a flow rate of 0.16 mL/min. X-rays of 12.4 keV (λ = 1.00 Å) with a photon flux of ca. 10^{12} photons per second were focused to a beam size < 75 µm at the sample. Scattering X-rays were measured using a Eiger 4M D detector and data was collected continually across elution.

6.2.10 Processing and Analysis of SEC-SAXS Data

ATSAS 3.2.1 software package (Manalastas-Cantos *et al*., 2021) was used for SEC-SAXS data processing and ensure data quality, prior to ensemble modelling. Elution profiles were quality checked to ensure a well-defined elution peak and buffer subtracted using CHROMIX (Panjkovich and Svergun, 2018) automatic sample/buffer frame selection. X-ray scattering data following buffer subtraction generated from buffer subtraction was analysed in Primus (Manalastas-Cantos *et al*., 2021). Beam stop bleed-through at low q-region and high-q noise at high-q regions were trimmed prior to further analysis. Guinier analysis (ln I(q) versus q^2) was performed using the lowest-q region with $q \cdot Rg \leq 1.3$. The fit was linear and free of low-q upturns or downturns, consistent with negligible aggregation and interparticle

interference. Distance distribution analysis was carried out for the P(r) curve, which was examined to ensure profile was consistent with expected shape of either catalytic domain only LPMO and two-domain LPMOs. Kratky plots ($q^2I(q)$ vs. q) were generated to assess flexibility and folding state of LPMOs, to confirm the plots matched expected shapes for compact catalytic domains and multimodule LPMOs that possess a flexible linker. Molecular weight (MW) analysis was implemented to ensure estimated MW values were consisted with expected LPMO MW. Following confirmation of data quality, an *ab initio* molecular envelope was generated using DAMMIF (Franke and Svergun, 2009). Eight DAMMIF runs were averaged and aligned using DAMAVER (Volkov, and Svergun, 2003) to obtain a consensus model. Generated model was further refined using DAMMIN (Svergun, 1999) to improve agreement between experimental scattering data and generated model. Model generation was carried out in slow mode, followed by 20 repetitions and generated SAXS envelopes were later utilised for ensemble modelling.

6.2.11 Ensemble Optimisation Methods and AlphaFold Model Alignment to Experimental SAXS Envelope

Ensemble modelling of multimodule LPMOs were carried out using Ensemble Optimisation Method (EOM) (https://www.embl-hamburg.de/biosaxs/eom.html) *via* ATSAS web interface (Manalastas-Cantos, *et al*., 2021), (Tria, *et al*., 2015). Rigid body models of the LPMO domain and CBM5 or CBM2 were produced using AlphaFold3 (Abramson, *et al*., 2024). Domain boundaries, including those of the flexible linker region, were defined based on annotation from the NCBI conserved Domain Database. AlphaFold3 models were produced of just the catalytic domain and CBM5 or CBM2 of each construct and were added as rigid bodies to the EOM interface. Linker region residues were specified as disordered and the rigid LPMO domain and CBM5 or CBM2 were listed as structures. A pool of 50,000 conformations was

generated using the RANCH module. GAJOE parameters included, 2000 generations and 100 number of ensembles, prior to job submission.

6.3 Results

6.3.1 Protein purification

All two-domain LPMOs were typically overexpressed from 4 L of LB media and exported to the periplasmic space of *E. coli*. Following disruption of the periplasm using osmotic shock procedures, the pI of each protein determined the type of ion-exchange column used **(Table 6.2).** Following ion-exchange chromatography (IEX) all proteins were subjected to size-exclusion chromatography (SEC), with fractions from the major eluting peak at the retention volumes reported in **Table 6.2** and checked by SDS-PAGE (**Figure 6.5**). Unfortunately, truncated *Sli*LPMO10F failed to express despite best efforts in expression troubleshooting.

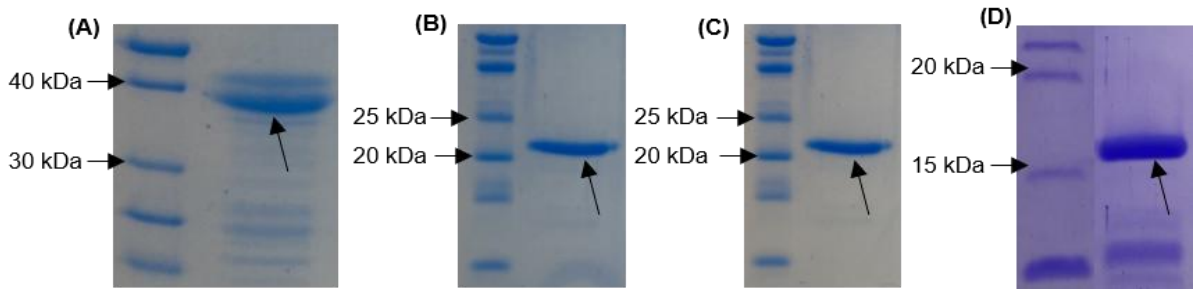


Figure 6.5. 15% *(w/v)* SDS-PAGE images with protein samples loaded after SEC. Two-domain *Sli*LPMO10C **(A)**, *Sli*LPMO10F **(B)**, *Sli*LPMO10G **(C)**. The LPMO domain of CBD - *Sli*LPMO10G **(D).** The first gel lane is a molecular weight ladder (10 - 200 kDa PageRuler™) with bands corresponding to certain molecular weights labelled.

It was notable that for some of the proteins low intensity bands were present as well as highly stained bands running at the expected molecular weight for each protein **(Table 6.2).** This would suggest the presence of some minor contaminants. Calculated yields for each protein

are reported in **Table 6.2**. The LPMO domain of truncated CBD - *Sli*LPMO10G was also successfully purified (**Table 6.2** and **Figure 6.4**).

Table 6.2. Summary of LPMO properties and purification parameters for the two-domain *Sli*LPMO10C + CBM2, *Sli*LPMO10F + CBM5 and *Sli*LPMO10G + CBM5 and the LPMO domain only of *Sli*LPMO10G.

	*Sli*LPMO10C + CBM2	*Sli*LPMO10F + CBM5	*Sli*LPMO10G + CBM5	CBD - *Sli*LPMO10G
Molecular Weight	34,539 Da	22,517 Da	22,234 Da	15,064 Da
Extinction Coefficients (ε)	75775 M^{-1} cm^{-1}	76345 M^{-1} cm^{-1}	66835 M^{-1} cm^{-1}	32220 M^{-1} cm^{-1}
pI	4.34	4.78	4.85	6.13
IEX resin	DEAE	DEAE	DEAE	HiTrap Q HP
SEC resin	HiLoad 16/600 Superdex-75	HiLoad 16/600 Superdex-75	HiLoad 16/600 Superdex-75	HiLoad 16/600 Superdex-75
Retention Volume from SEC	60 mL	82 mL	82 mL	100 mL
Yield	ca. 12.4 mg/ml	ca. 8.7 mg/mL	ca. 22.3 mg/mL	ca. 13.4 mg/mL
Trp (%)	3.33	6.07	5.14	3.55
No. of Trp	11	13	11	5

6.3.2 Tryptophan quenching of LPMOs upon Cu(II) binding

All three two-domain LPMOs (*Sli*LPMO10C, *Sli*LPMO10F and *Sli*LPMO10G) were successfully over-expressed in *E. coli* and purified. Typical yields from 4 L of growth media ranged from ~10 mg/mL for *Sli*LPMO10F to ~25 mg/mL for *Sli*LPMO10G. Titrating Cu(II) into the as purified two-domain LPMOs resulted in quenching of Trp fluorescence at $\lambda = 330$ nm (**Figure 6.6**), with a break point reached between 0.8 and 1.0 equivalents of copper addition, indicating a binding stoichiometry close to 1:1 (**Figure 6.6**). Similarly, a binding stoichiometry of 1:1 was obtained for the LPMO only domain of *Sli*LPMO10G, thus indicating that the appended CBMs do not interfere with copper uptake into the active site.

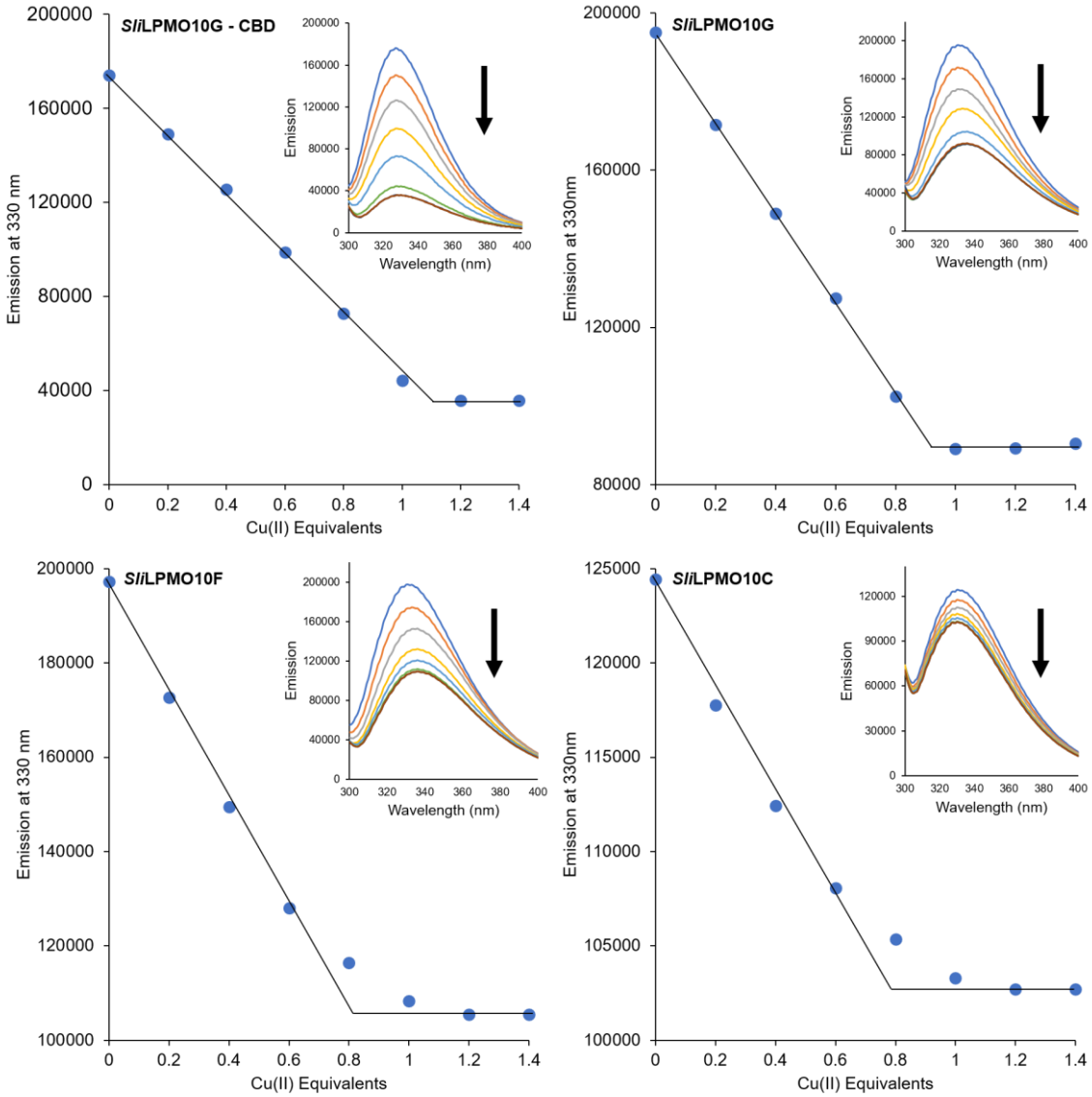


Figure 6.6. Cu(II) binding at pH 7.5 and 18°C. Changes in the Trp emission spectrum upon titration of the various proteins (2 μM) with addition of stock 200 mM Cu(II)SO_4 aliquots and plotted as function of [Cu(II)] equivalents. The intersection of the straight lines is an estimation of stoichiometry, and the arrows indicate the direction of the emission change.

6.3.3 Chitin and Cellulose Binding Assays

Substrate binding experiments of two-domain LPMOs using β-chitin and cellulose were monitored by SDS-PAGE. From **Figure 6.7** it is apparent that there is a clear specificity preference for β-chitin over cellulose, for which a band is present in all UB fractions for LPMOs containing β-chitin binding CBM5. For *Sli*LPMO10F and *Sli*LPMO10G both contain a chitin specific CBM5 whereas *Sli*LPMO10C possesses a CBM2 which has specificity for both cellulose and chitin. Interestingly, for the cellulose binding assays, despite using the same volumes of protein/substrate and loading equal volumes of sample on the SDS-PAGE gel (15 µL), difference intensity bands in the UB fraction for *Sli*LPMO10F and *Sli*LPMO10G were observed, which suggests higher cellulose binding affinity in *Sli*LPMO10F compared to *Sli*LPMO10G, whereas cellulose binding affinity in *Sli*LPMO10C shows equal quantity in bound and unbound fractions.

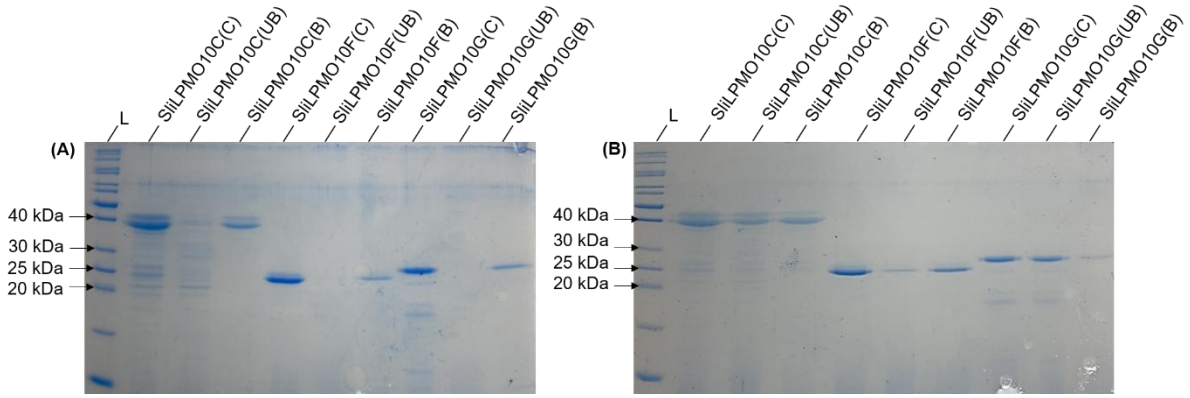


Figure 6.7. LPMO substrate binding assay visualised on a 15% SDS-PAGE gel against a 10-200 kDa unstained protein ladder (PageRuler™). β-chitin substrate binding assay **(A)** and cellulose substrate binding assay **(B)**. with control = (C), unbound fraction = (UB) and bound fraction = (B).

6.3.4 EPR spectroscopy on the two-domain LPMOs

Previous EPR spectroscopy studies with AA10 LPMO members (Gómez-Piñeiro, *et al*., 2022; Lindley, *et al*., 2022; Chaplin, *et al*., 2016) have inferred that a high degree of plasticity exists within the first coordination sphere geometry of the copper active site. Geometric configurations of the copper site may be deduced from experimentally determined *g* and hyperfine coupling, *A*, tensors obtained from EPR measurements. The X-band EPR spectra at 20 K for the Cu^{2+} bound two-domain LPMOs are reported in **Figure 6.8A**. The *Sli*LPMO10C spectrum shows well resolved low-field hyperfine components consistent with a single Cu^{2+} species. Simulation gives the EPR parameters reported in **Table 6.3** and the corresponding simulated spectrum shown in **Figure 6.8A** immediately under the experimental *Sli*LPMO10C spectrum. These parameters are consistent with a rhombic EPR signature ($g_z \neq g_y \neq g_x$), indicative of distorted copper geometry that lies between square base pyramidal and trigonal bipyramidal.

The interpretation of the *Sli*LPMO10F and *Sli*LPMO10G EPR spectra (**Figure 6.8A**) is somewhat more complex. In the case of the *Sli*LPMO10F protein, the EPR spectrum can be deconvoluted into two components (I and II), made possible due to the dramatic change in line shape upon increased concentration of Cu^{2+} added (**Figure 6.8B**). (This effect was not found for *Sli*10G.) Component I (**Figure 6.8B**) has been simulated when using the EPR parameters reported in **Table 6.3**. These parameters are consistent with an almost axial EPR signature ($g_z > g_y \approx g_x$), suggesting the copper site geometry has changed relative to the rhombic EPR signature.

Component II in the *Sli*LPMO10F spectrum is practically identical to the spectrum of *Sli*LPMO10G (**Figure 6.8A**). The complex line shape of this EPR spectrum, directly detected in *Sli*LPMO10G and extracted as a component from the *Sli*LPMO10F spectrum, is difficult to

simulate within the constraints adopted in our other Cu^{2+} spectra simulation (see Methods section). Specifically, the two assumptions: a) $A_{Cux} = A_{Cuy} = 0.1\ A_{Cuz}$ and b) the A_N-values are all isotropic and have the same value for all three coordinating nitrogen atoms, have to be abandoned for Component II simulation. Parameters used to simulate Component II spectrum (**Figure 6.8B**) are reported in **Table 6.3**.

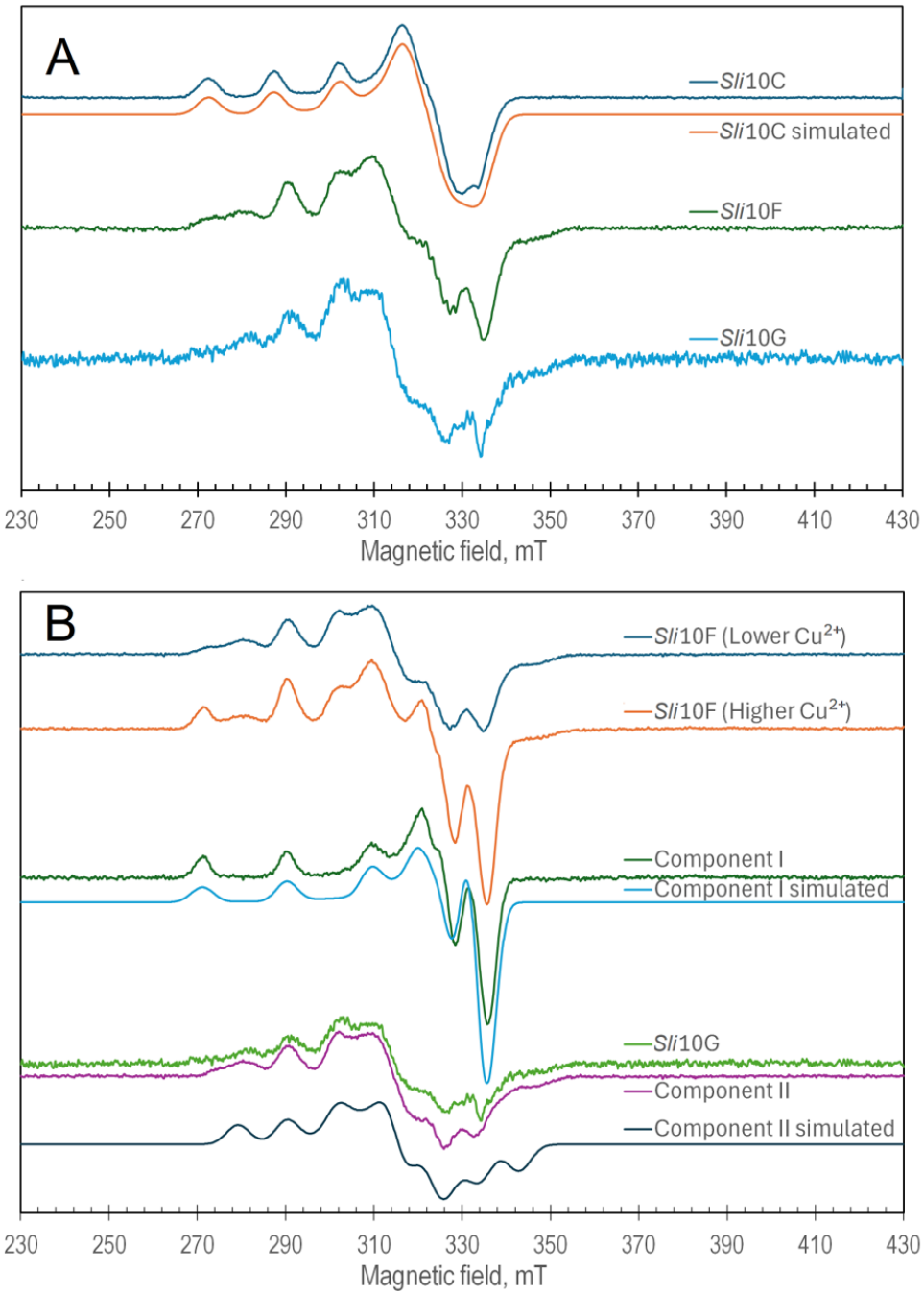


Figure 6.8. EPR spectroscopy of the three two-domain LPMOs (*Sli*LPMO10C, *Sli*LPMO10F and *Sli*LPMO10G). **A** – The 20 K EPR spectra of the three proteins measured at 0.2 mW (30 dB). The spectra have been normalised to a similar intensity for convenient comparison. The *Sli*LPMO10C spectrum has been simulated using the parameters reported in **Table 6.3**.

B - Deconvolution of the *Sli*LPMO10F EPR spectra into two components – I and II. The two components were obtained by spectral subtraction with variable coefficient (Svistunenko, *et al*., 2006) as follows: Component I = *Sli*LPMO10F (Higher [Cu^{2+}]) – 0.870 x *Sli*LPMO10F (Lower [Cu^{2+}]). Component II = *Sli*LPMO10F (Lower Cu^{2+}]) – 0.228 x *Sli*LPMO10F (Higher [Cu^{2+}]). The line shape of Component II is very close to the *Sli*LPMO10G spectrum (the two shown next to each other, the *Sli*LPMO10G being the same as in A). The Component I and Component II spectra were simulated when using the parameters reported in **Table 6.3**.

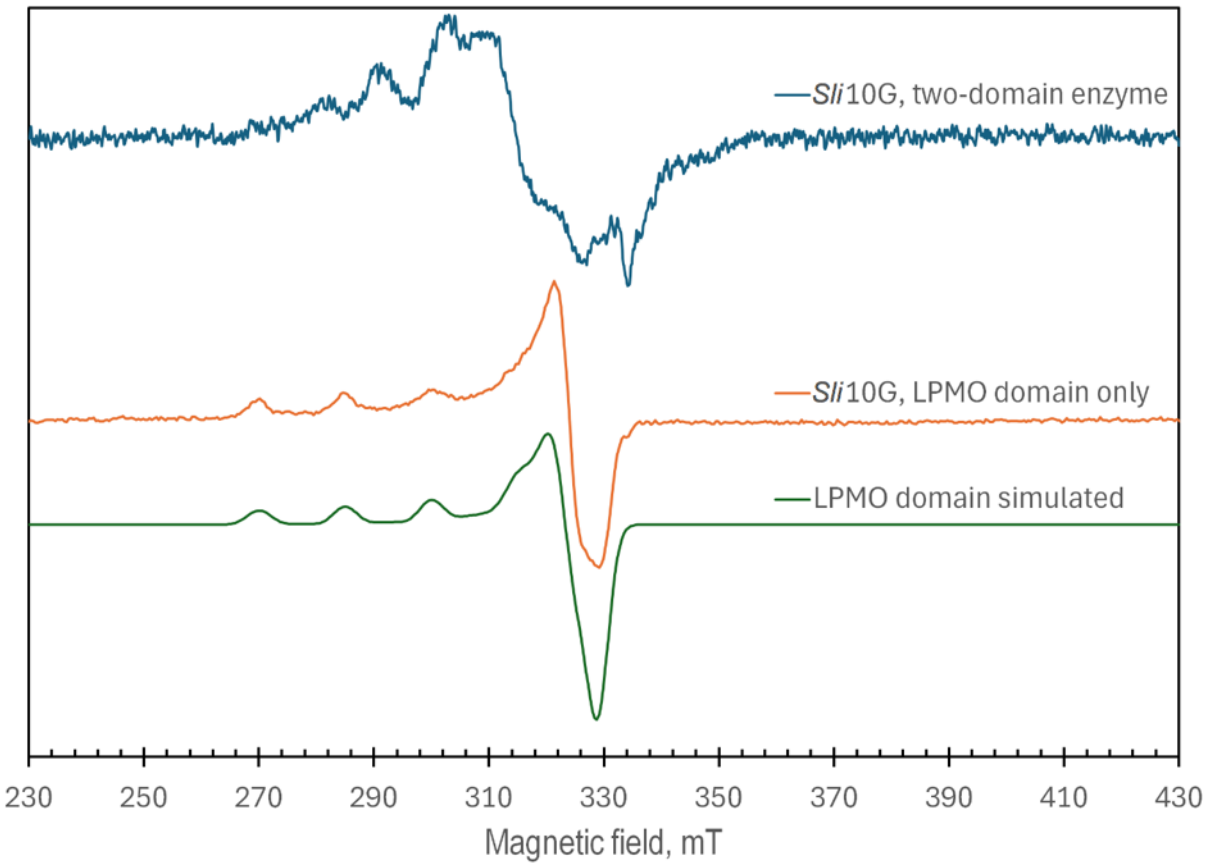


Figure 6.9. The 20 K EPR spectra of the *Sli*LPMO10G protein – in the two-domain and single (LPMO) domain forms, measured at 0.2 mW (30 dB). The parameters of the LPMO domain only EPR spectrum simulation are reported in **Table 6.3**.

Table 6.3. The 20 K EPR spectra of the *Sli*LPMO10G protein – in the two-domain and single (LPMO) domain forms, measured at 0.2 mW (30 dB). The parameters of the LPMO domain only EPR spectrum simulation are also reported.

	Δg	g_x	g_y	g_z	ΔH_x, MHz	ΔH_y, MHz	ΔH_z, MHz	A_{Cux}, MHz	A_{Cuy}, MHz	A_{Cuz}, MHz
*Sli*LPMO10C	0.248	2.020	2.1	2.268	140	190	130	46	46	460
*Sli*LPMO10F Component I	0.180	2.049	2.051	2.229	60	133	120	58.5	58.5	585
*Sli*LPMO10F Component II (also applicable to two-domain *Sli*LPMO10G)	0.2303	2.0247	2.1073	2.255	140	260	160	241	123	350
*Sli*LPMO10G (LPMO domain only)	0.228	2.057	2.073	2.285	120	130	80	47	47	470

See the Methods section for details on the following:

Dg = g_z - g_x – the parameter used in (Gómez-Piñeiro, *et al*., 2022).

The g_z, DH_z and A_{Cuz} values were determined with a high confidence directly from experimental spectrum (see **Figure B.7**). For the values of the hyperfine constants for the three nitrogen atoms in the four proteins see **Table A.7**.

6.3.5 X-ray Crystal Structure of the LPMO Domain of *Sli*LPMO10G

The X-ray crystal structure of the LPMO domain of *Sli*LPMO10G was solved to a resolution of 1.20 Å by molecular replacement, using MrBump (Keegan and Winn, 2008) (**Table 6.1**) with Cu atoms omitted from the search model. Two protein monomers were found in the crystallographic asymmetric unit, which adopt a head-to-head arrangement involving the substrate binding surfaces with one monomer rotated 180° with respect to the other (**Figure 6.10**). This arrangement is not biologically relevant, and SEC as well as SEC-SAXS data (see below) are consistent with the LPMO domain existing in solution as a monomer. Three strong (> 30 σ F_o-F_c) omit electron density features are present in the head-to-head region of the monomers of the asymmetric unit. Two of these features in the $2F_o$-F_c map are consistent with the expected location of the Cu active site and have each been modelled and refined with a full-occupancy Cu atom. The third is located at the interface of the two monomers, and is best modelled and refined as a Cu atom with 60% occupancy. Occupancy was determined based on the appearance of difference map features after refinement.

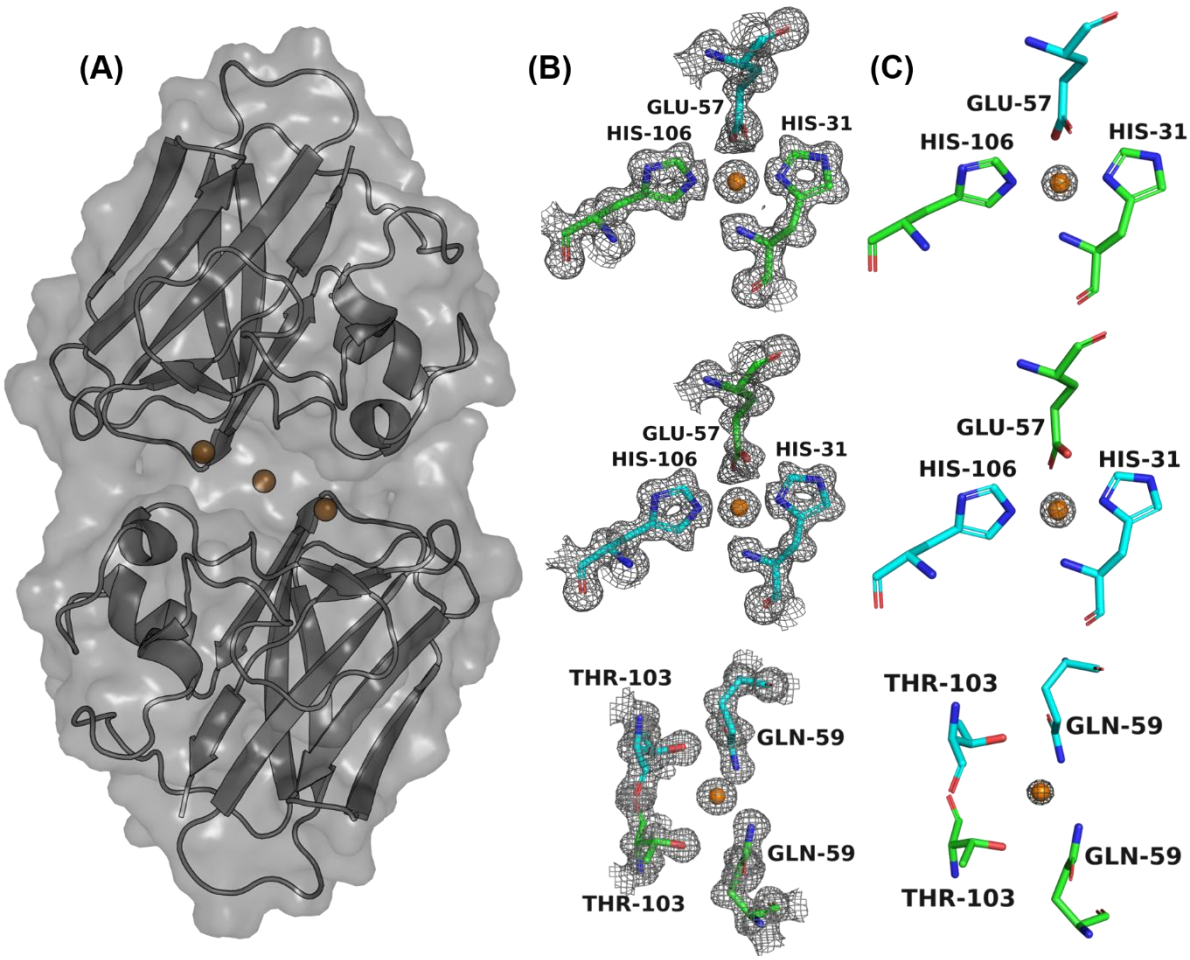


Figure 6.10. Asymmetric unit of truncated *Sli*LPMP10G dimer represented in a transparent surface, showing the three orange copper ion spheres shared between both LPMO chains **(A)**. $2F_o$-F_c electron density map contoured at 1 σ at the copper active site in the dimer conformation **(B)**. $2F_o$-F_c electron density map, contoured at 10 σ **(C)**. The strong spherical density around the copper at this level supports confidence in placement of copper ion, which also confirms its high occupancy even in the absence of surrounding residues.

Continuous electron density was observed from the N-terminal His31 to the C-terminal Asn169 in both crystallographic monomers. The structure contained six anti-parallel β-strands arranged in two β-sheets; one side contains S1, S3 and S6 and the other S4, S5 and S7. In between the connecting loops, smaller β-stands are found which are 3 to 4 resiues long S2, S8

and S9. This β-sandwich fold is common across all AA10 structures (Munzone, *et al*., 2020; Fowler, *et al*., 2019) including the previously determined X-ray structure of the LPMO domain of *Sli*LPMO10E (Chaplin *et al*, 2016). Embellishing the β-sandwich fold are a total of three α-helices (H1, H2 and H3) with α-helices H1 and H2 sitting between β-strands S1 and S2 are present; H3 sits beween S2 and S3 which are all separated in the loop region. Two disulfide bonds are present in the strtucture, Cys14/Cys21 and Cys40/Cys135. Disulfide bond Cys14/Cys21 is bonded between a H1 α-helix and loop region, whereas Cys40/Cys135 is bonded at S9 β-stands to another loop region.

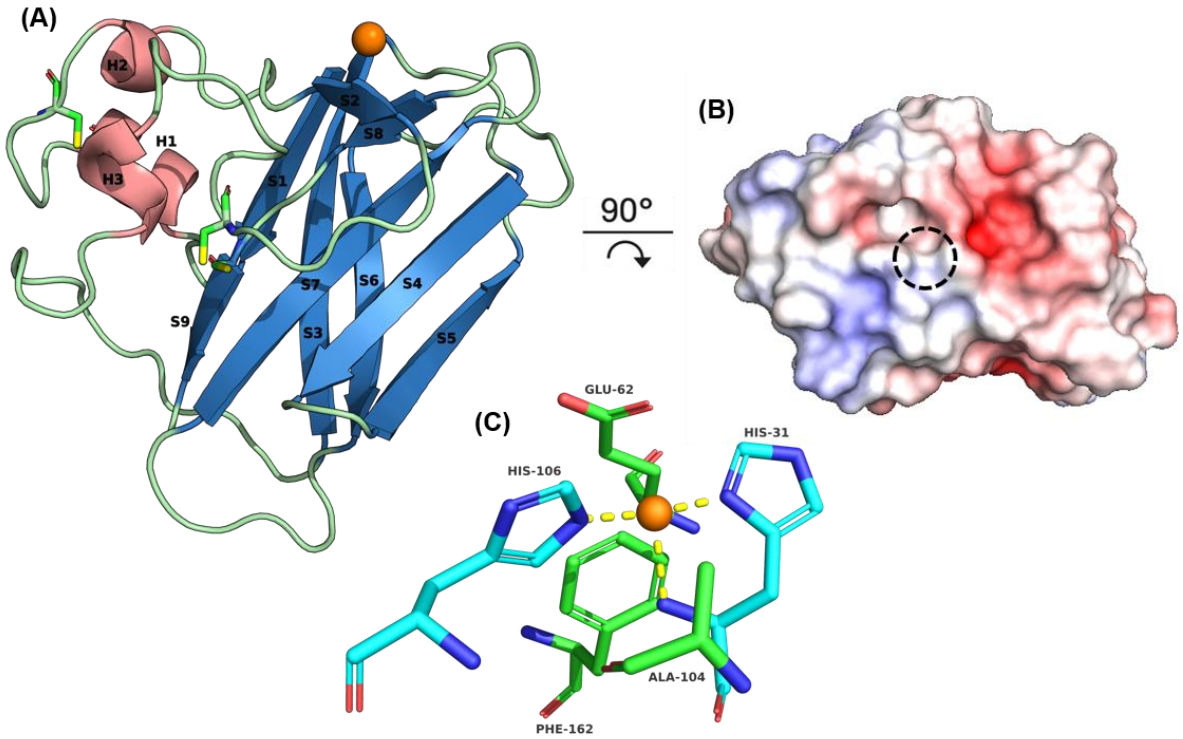


Figure 6.11. X-ray crystal structure of truncated CBD - *Sli*LPMO10G. β-sheets shown in blue, α-helices displayed in pink, loop regions shown in green and disulfide bonds displayed as sticks and copper represented as an orange sphere. Histidine brace active site represented is shown represented as sticks **(A)**. The electrostatic surface potential (calculated by (APBS)), shows a dashed circle at the copper active site with region showing mix of positive (blue) and negative (red) surface charges **(B)**. Copper active site showing histidine brace and second coordination sphere residues **(C)**.

The active site Cu is coordinated by the N-terminal amino group of His-31 and the Nδ1 and Nε2 atoms of His-31 and His-106, respectively (**Figure 6.11**). This results in the T-shaped His-brace coordination geometry. No water molecules occupy the first coordination sphere; however, electron density is visible for the side chain of Glu-57 from the opposing monomer that can coordinate the Cu ion in a chelating manner (**Figure 6.11 B - C**). As mentioned, in solution, this LPMO domain is a monomer and thus this Glu coordination arising from opposing monomers is likely considered an artefact of crystallisation. The third Cu located at the centre of the dimer interface displays coordination by the side chains of Gln-59 and Thr-103 from each chain (**Figure 6.11 B - C**). However, the Oγ1(Thr-103) and Oε1(Gln-59) are positioned at distances of 3.1 Å and 3.4 Å from the Cu ion, these distances are longer than typically found for metal-ligand coordination bonds (Shah, *et al*., 2023). Given these features, it is likely that the third Cu site represents a crystallographic artefact with no biological relevance.

6.3.6 SEC-SAXS of LPMOs

*Sli*LPMO10C, *Sli*LPMO10G and *Sli*LPMO10F together with the isolated LPMO domain of CBD - *Sli*LPMO10G were studied by SEC-SAXS to determine effect of LPMO linker on overall conformation. Compactness and flexibility of LPMOs were examined through Kratky plots (**Figure 6.15**) and Quality of SEC-SAXS data was assessed through a Guinier analysis. ln I(q) versus q^2 was performed over the lowest-q region with $q \cdot Rg \leq 1.3$. The linear fit was consistent with negligible interparticle effects. Distance distribution analysis (**Figure 6.16**) further validated data quality to look at overall shape and to estimate D_{max} in P(r) curves and resembled the expected shape of the target LPMO (globular or elongated). Elution peak shows the largest (*Sli*LPMO10C (34.5 kDa)) being eluted first at lower frame number compared to smaller LPMOs (**Figure 6.12**) (Graewert, *et al*., 2020). CBD-*Sli*LPMO10G, full

length *Sli*LPMO10G and *Sli*LPMO10F all show elution at later frames. Buffer/sample frames selection for buffer subtraction include, CBD - *Sli*LPMO10G (342-280)/(456-426), *Sli*LPM10G (373-311)/(445-415), *Sli*LPMO10F (94-32)/(438-408) and *Sli*LPMO10C (137-69)/(375-342). Selected regions are shown in **Figure 6.12**).

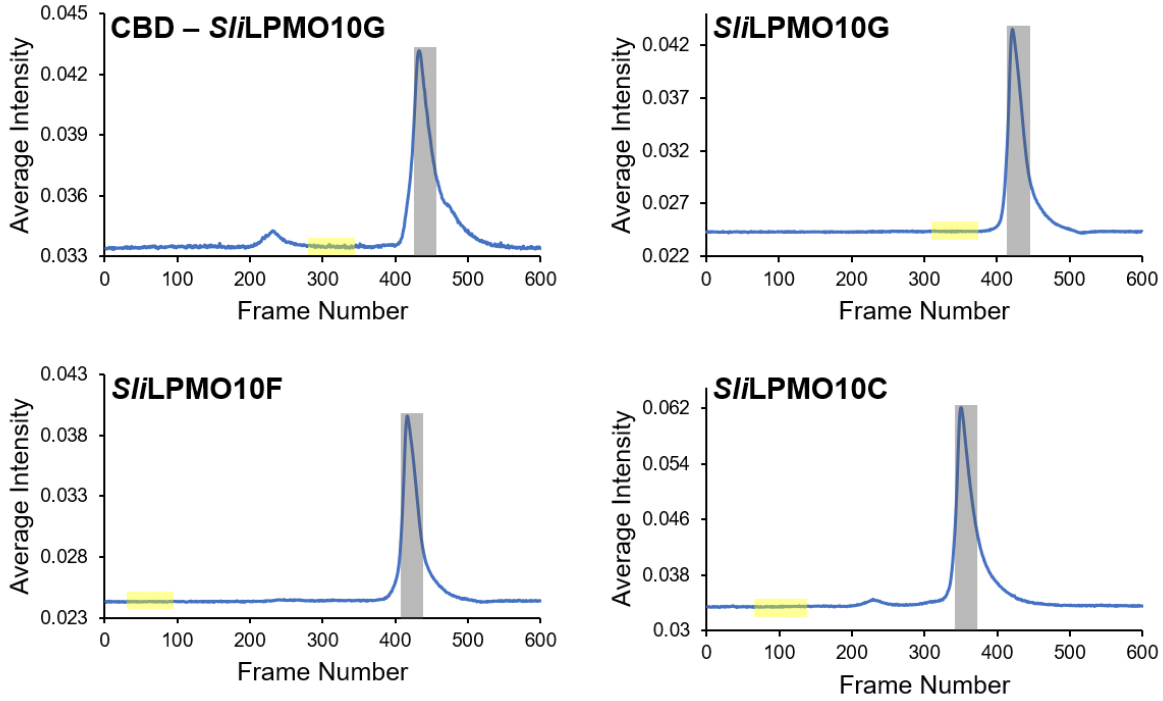


Figure 6.12. Size-exclusion chromatography elution profiles of LPMOs sent for SAXS analysis. Buffer frames (yellow) and sample frames (grey) that were selected for buffer subtraction.

6.3.7 Scattering Profiles of LPMOs

Prior to analyses, frames corresponding to buffer (frames distant from elution peaks) and sample regions (largest elution peak) were selected for frame averaging and buffer subtraction (selecting frames distant from elution peaks). The scattering intensity (log Iq)) was plotted against the scattering vector (q) to obtain scattering profiles to gain insight on protein shape and data quality. All scatter profiles show a smooth decay in log l(q) at increasing q, and no upturn was observed at low-q, indicating an absence of aggregation or interparticle effect (**Figure 6.13**). The monomeric CBD – *Sli*LPMO10G contains a steep decay at increasing q, which is typical for a compact, globular protein. In contrast, two-domain *Sli*LPMO10G, *Sli*LPMO10F and *Sli*LPMO10C, displays a less steep decay with a more gradual transition, especially at the mid-q region. Although these differences are subtle, these deviations are characteristic of multimodule, elongated proteins structures that are connected by a flexible linker, this increases conformational heterogeneity (Kikhney and Svergun, 2015). While *Sli*LPMO10C is still consistent with a multimodule structure, a slightly steeper decay at mid-q region can be observed (**Figure 6.13**). This characteristic could suggest a more compact protein which is less flexible to its other two-domain counterparts (*Sli*LPMO10G and *Sli*LPMO10F).

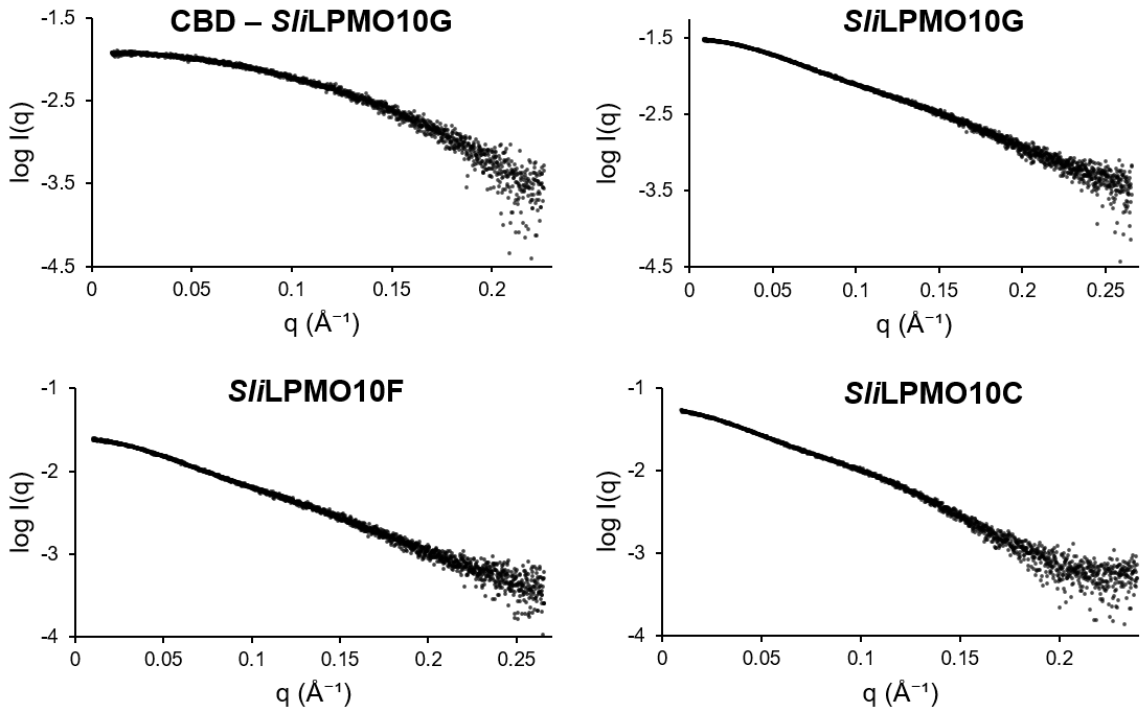


Figure 6.13. Logarithmic scale scattering profiles of LPMOs. Scattering intensity (log l(q)) plotted against the scattering vector (q).

6.3.8 Guinier Analysis of SEC-SAXS Data

A radius of gyration (Rg) analysis was carried out to assess monodispersity and overall size of LPMOs in solution. Guinier analysis was performed by plotting the natural logarithm of scattering intensity (ln I(q)) against the square of the scattering vector (q^2) at low q-region. CBD – *Sli*LPMO10G shows a consistent linear Guinier fit throughout low q-regions (**Figure 6.14**) (red). Further suggesting a well-folded and monomeric structure with an Rg of 14.35 ±0.05. Two-domain *Sli*LPMO10F and *Sli*LPMO10G also show consistent linearity in the Guinier region; slopes are significanty steeper than that of CBD – *Sli*LPMO10G, suggesting significantly larger Rg values of 23.54 ±0.05 for *Sli*LPMO10G and 23.97 ±0.05 for *Sli*LPMO10F (Kikhney and Svergun, 2015). Interestingly, *Sli*LPMO10C has a slightly steeper slope compared to *Sli*LPM10F and *Sli*LPMO10G (**Figure 6.14**) with Rg value 29.68 ±0.09, further suggesting a structured and more compact domain with less flexibility comparred to *Sli*LPM10F and *Sli*LPMO10G. (Matsui, *et al*., 2024). Residual plots in **Figure 6.14** shows even

residual distribution around zero for CBD – *Sli*LPMO10G, *Sli*LPMO10G and *Sli*LPMO10F, with mimimal deviation. Slight curvature, reflecting minor curvature in Guiniear region, can be seen for *Sli*LPMO10C, suggesting subtle deviation from ideal Guiniear region.

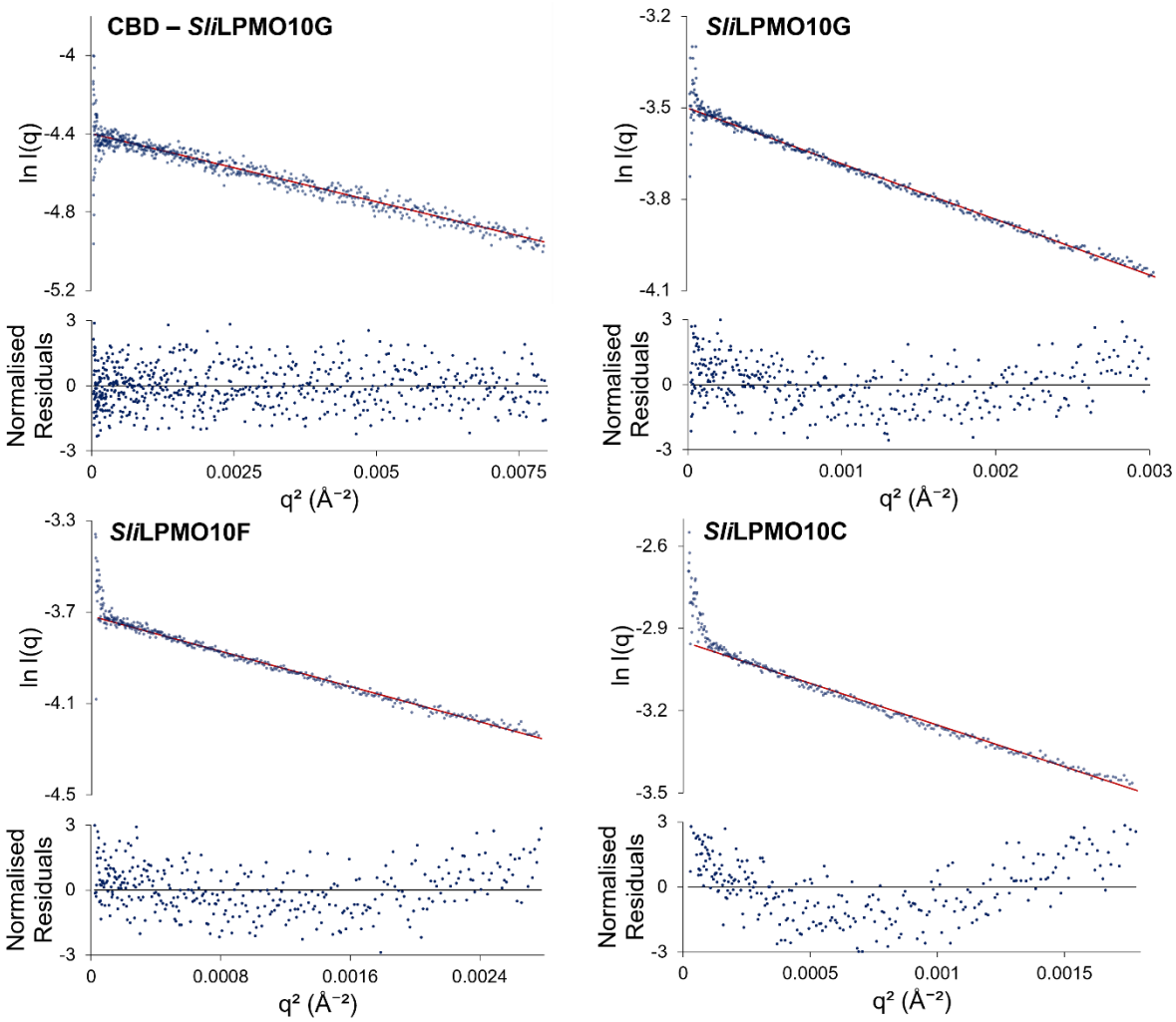


Figure 6.14. Guinier plot of LPMOs. The natural logarithm of the scattering intensity (In I(q)) is plotted against the squared root of the scattering vector (q^2), with corresponding normalised residuals for Guinier line quality fit evaluation.

6.3.9 Kratky Profiles Analysis of LPMOs

To gain further insight on the structural confomation of two-domain/CBD LPMOs, a kratky plot was produced with a function of a scattering vector (q), to highlight folding/flexibility (q^2)·I(q)) against the scattering vector (q). Monomeric CBD – *Sli*LPMO10G contains a bell-shaped curve peak which decays back to y=0 (**Figure 6.15**), this is a characteristic of rigid and globular conformation. This particular krakty plot scatter shows that there is a lack of flexibility, suggesting absence of a flexible linker and confirming that the domain behaves as a rigid body in solution. However, two-domain *Sli*LPMP10G and *Sli*LPMO10F contain a significantly broader peak with a slower rate of decay at increasing q, compared to monomeric CBD – *Sli*LPMO10G. This characteristic is canonical for proteins that have conformational flexibilty, and shows that these proteins are multidomain connected by a flexible linker (Kikhney and Svergun, 2015). On the other hand, *Sli*LPMO10C shows a higher, sharper peak and rapid decay at increasing q, compared to *Sli*LPMO10F and *Sli*LPMO10G. These features are consistent with a more compact/rigid conformation, likely reflecting differences in linker length and/or rigidity.

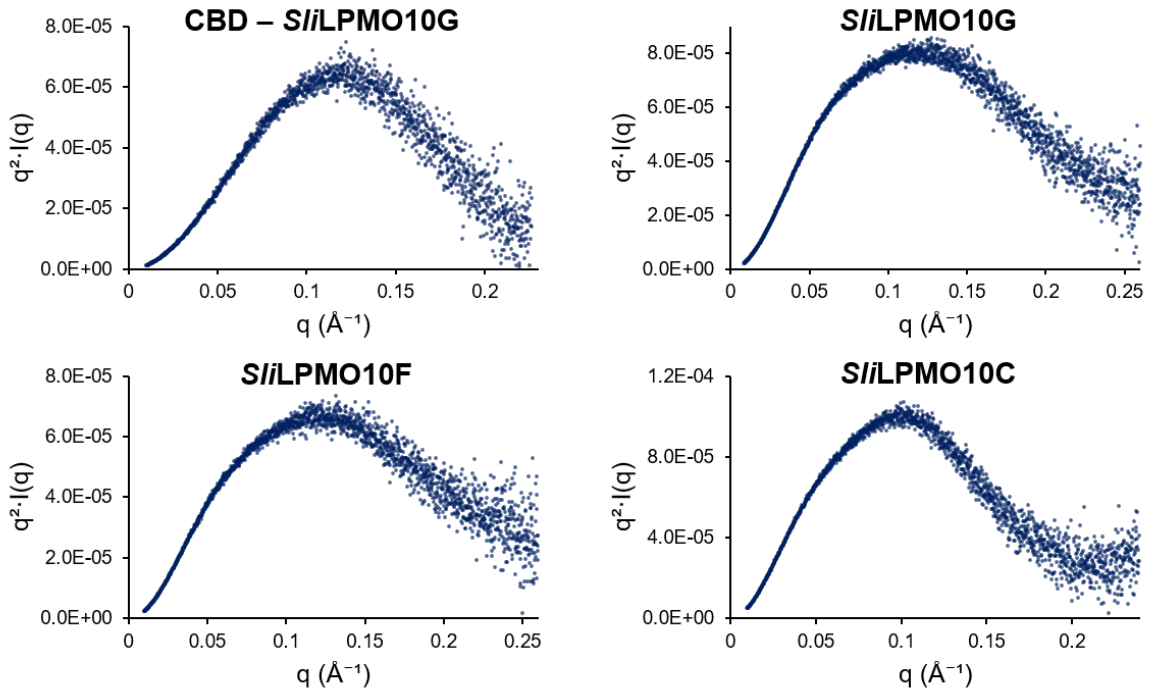


Figure 6.15. Kratky plot of LPMOs. Arbitrary (q2·I(q)) to highlight structural features and folding states, plotted against the scattering vector (q).

6.3.10 Pair-Distribution Analysis of SEC-SAXS Data

Pair-distance distribution (P(r)) functions were calculated from experimental SEC-SAXS data to gain insight on LPMO size and shape. A D_{max} (maximum dimension) calculated from P(r) reveals the largest intraparticle distances within a molecule, giving information on overall protein size. P(r) is the relative frequency of pairwise distances, which is plotted against (r), representing the distribution of distances between all pairs of scattering centers within the particle (in Å). As expected, P(r) curves for the monomeric CBD – *Sli*LPMO10G contain a symetric bell-shaped distribution and a D_{max} value of 44.2 Å (**Figure 6.16**), consistent with characteristics of a globular/monomeric protein. In contrast, *Sli*LPMO10C, *Sli*LPMO10F and *Sli*LPMO10G exhibited asymmetric and broader P(r) distributions, indicating the presence of larger protein structure. The secondary shoulder exhibited by the two-domain LPMOs suggest the presence of multidomain structures such as the flexible linker and CBM5/CBM2. P(r)

analysis calculates a D_{max} which is the furthest distance between any two points within a particle, providing information on overall size of target proteins (Kikhney and Svergun, 2015). Similar D_{max} (maximum intra-particle distance) values are observed in smaller (22.5 kDa) *Sli*LPMO10F (D_{max} = 74.3 Å) and *Sli*LPMO10G (D_{max} = 73.6 Å) (**Figure 6.16**). The broadest distribution can be seen in *Sli*LPM10C with a D_{max} of 93.5 Å, as expected as *Sli*LPMO10C is the largest LPMO at 34.5 kDa. The P(r) curves support conclusions drawn from the Kratky plots (**Figure 6.15**) which were consistent of a monomeric domain for CBD - *Sli*LPMO10G (Putnam, *et al*., 2013). D_{max} results from P(r) curves are also consistent with sizes of proteins from 15 kDa in CBD – *Sli*LPMO10G compared to the larger two-domain LPMOs.

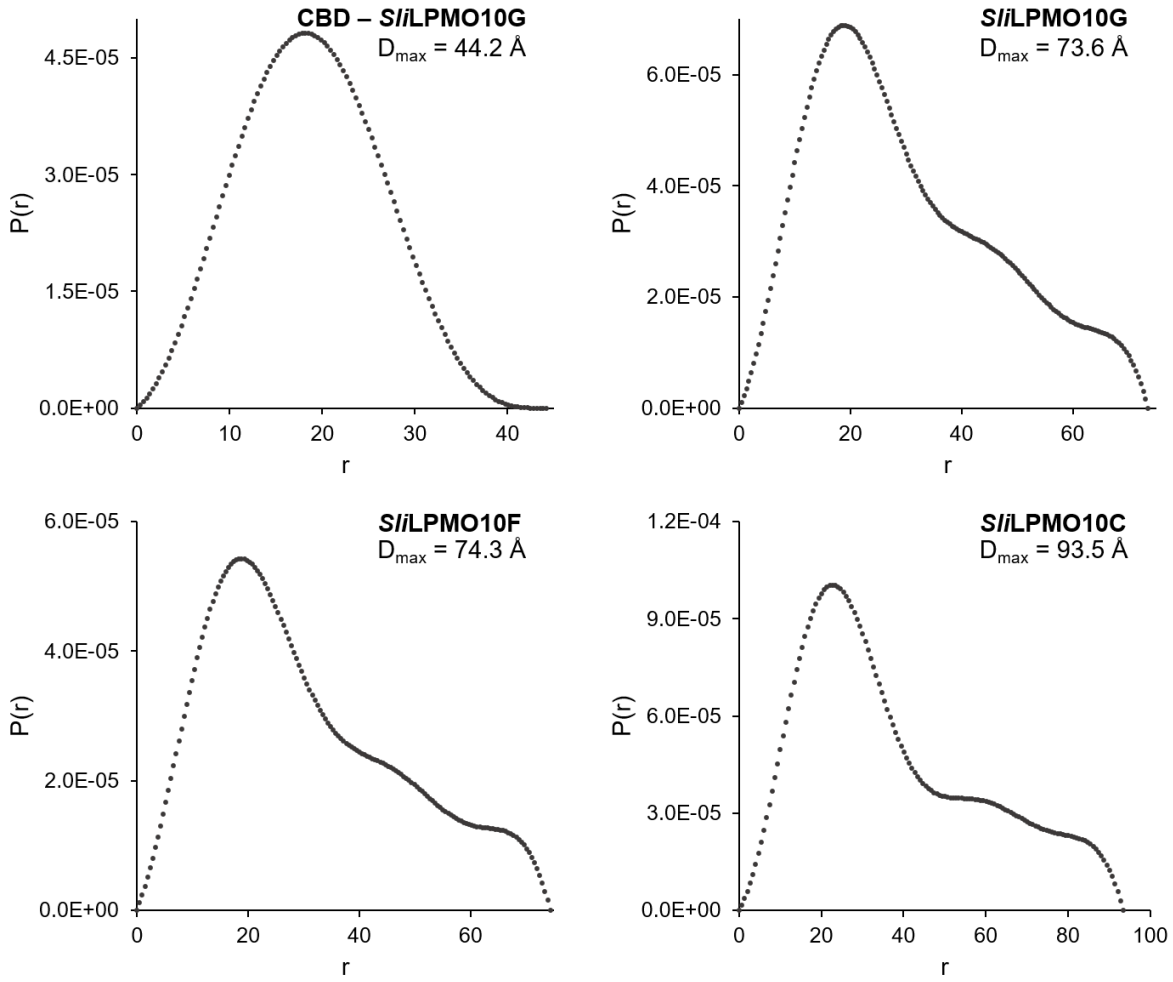


Figure 6.16. Pair-distance distribution curves (P(r)) curves of LPMOs. Relative frequency of pairwise distances P(r), plotted against distance between pairs of scattering centres.

6.3.11 *Ab Initio* Reconstruction of SEC-SAXS Envelopes

Ab initio reconstruction was performed using DAMMIF, models were averaged with DAMMIN and filtered/aligned using DAMAVER. Experimental scattering curves and fits to reconstructed model can be seen along with corresponding molecular envelopes (**Figure 6.17**). *Sli*LPMO10G generates a compact, globular shaped envelope consistent of a single-domain structure with an excellent fit to data (X^2 = 0.918). Contrastingly, the two-domain proteins generated elongated/modular shapes, consistent with the presence of catalytic domain, flexible linker and CBM5/2. *Sli*LPMO10G and *Sli*LPMO10F both show envelopes of similar size showing a two-domain structure with excellent fits to data (X^2 = 0/961/1.003). *Sli*lPMO10C envelope also shows a clear two-domain elongated structure, but with a larger space in the LPMO domain region. The LPMO/CBM2 domain of *Sli*LPMO10C is larger in size compared to multimodular *Sli*LPMO10F and *Sli*LPMO10G, this is shown in **Figure 6.17**. A good fit to data can be seen for *Sli*LPMO10C (X^2 =1.139).

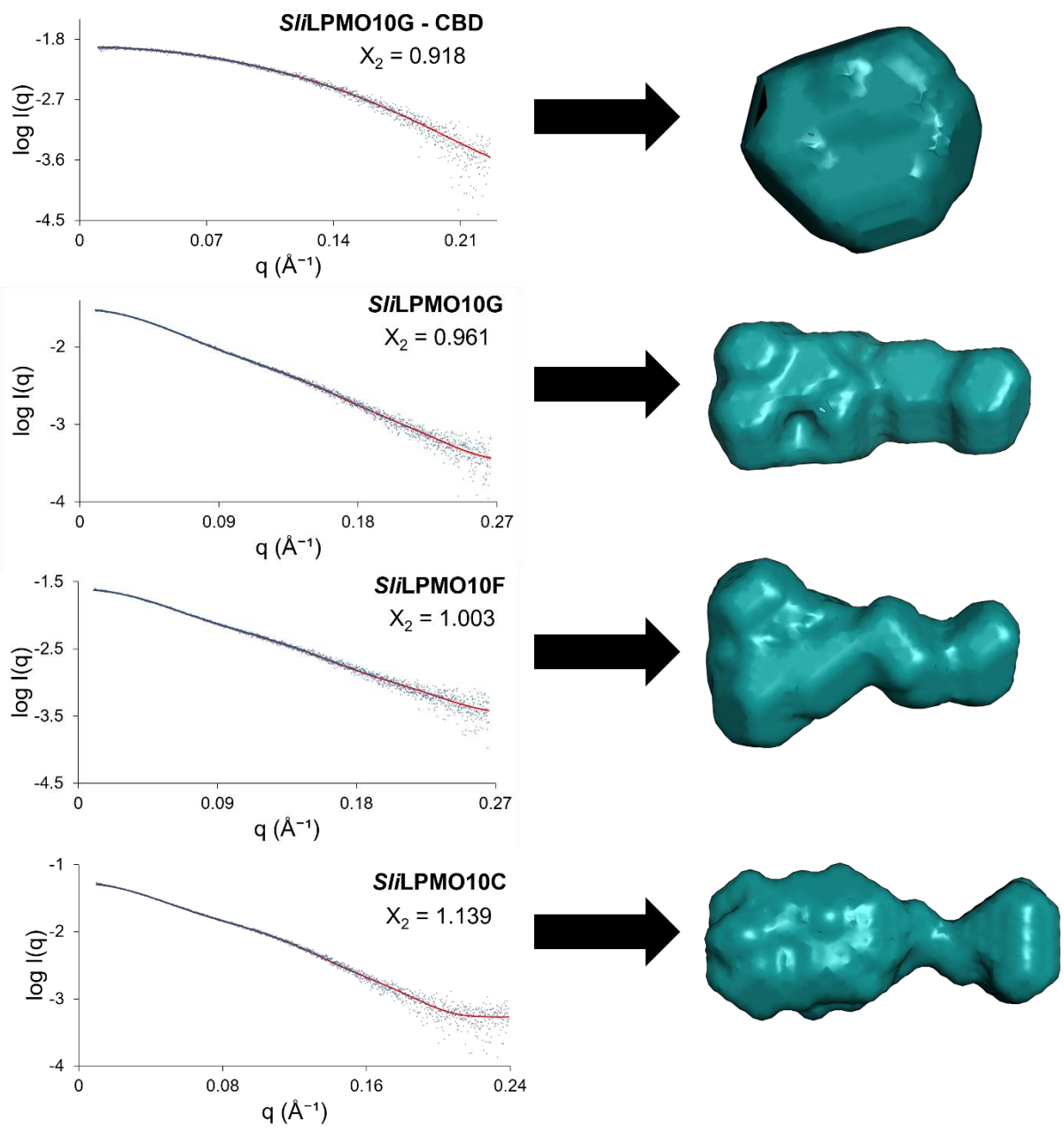


Figure 6.17. *Ab initio* reconstruction of catalytic domain only (*Sli*LPMO10G – CBD) and full length (*Sli*LPMO10G, *Sli*LPMO10F and *Sli*LPMO10C. SEC-SAXS scattering (blue) data with fits to data (red) for model generation.

Molecular weight (MW) estimation from SEC-SAXS data provides an essential check for data quality and for determining the oligomeric state of the protein in solution. In PRIMUS (Manalastas-Cantos *et al.*, 2021), the MW derived from experimental SEC-SAXS were in reasonable agreement with theoretical MW values, which was estimated using Bayes estimation (**Table 6.4**).

Table 6.4. Theoretical and SAXS-derived molecular weights of LPMO constructs, SAXS-derived MW were calculated using primus (ATSAS) (Manalastas-Cantos, *et al*., 2021).

LPMO Construct	Theoretical (kDa)	SEC-SAXS MW Bayes Estimate (kDa)
*Sli*LPMO10G – CBD	15	11.6
*Sli*LPMO10G + CBM5	22.5	18.5
*Sli*LPMO10F + CBM5	22.5	16.8
*Sli*LPMO10C + CBM2	33	31.1

6.3.12 Ensemble Optimization Method Fitting of SEC–SAXS Data

Experimental SAXS data from two-domain LPMOs underwent Ensemble Optimisation Method (EOM) (Tria, *et al*., 2015) to investigate conformational variability. Rigid body models (copper-binding domain + CBM5/CBM2) produced from Alphafold3 were inputted into the EOM pipeline, along with the sequence for the LPMO domain, CBM5/CBM2 and disordered flexible linker sequence. A large pool of 50,000 conformers was then generated to identify a best-fit model for alignment with SAXS envelope. **Figure 6.18** shows experimental SAXS scattering (orange) against best-fit curve produced from EOM job (purple) (Burger, *et al*., 2016). All EOM jobs showed an excellent fit between experimental SAXS scatter and EOM curves across q-region, with $X_2 = 0.889$ (*Sli*LPMO10G), $X_2 = 1.025$ (*Sli*LPMO10F) and $X_2 =$

0.923 (*Sli*LPMO10C). A slight upwards curve at high q can be seen in *Sli*LPMO10C compared to *Sli*LPMO10F and *Sli*LPMO10G.

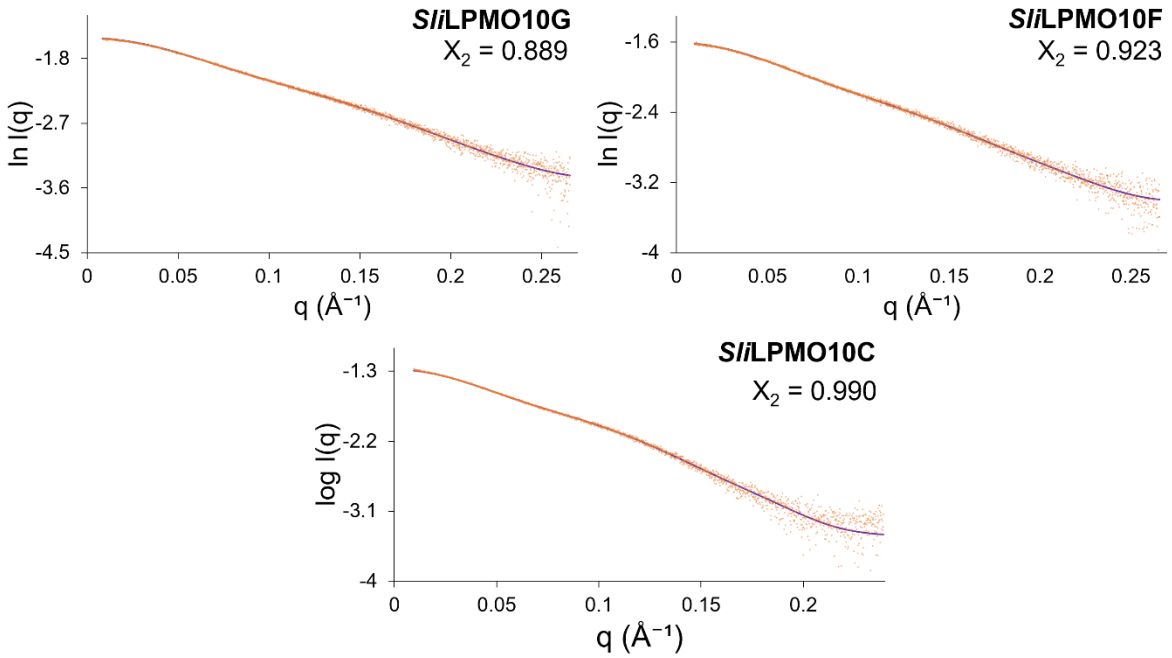


Figure 6.18. Ensemble Optimisation Method (EOM) of multidomain LPMOs. Scattering curve from experimental SAXS data (blue) against best-fit EOM line (red).

6.3.13 Comparison of Experimental SEC-SAXS and EOM-Derived Rg and D_{max}

Rg/D_{max} values obtained through EOM analysis were compared to the experimental SEC-SAXS derived Rg/D_{max} values. SEC-SAXS and EOM derived values display close agreement (**Table 6.5**), thus, confirming that generated ensembles reflect true conformational space of full length LPMOs in solution. Furthermore, the close agreement provides further validation of experimental scattering data and model reliability.

Table 6.5. Comparison of Rg ad D_{max} values obtained from experimental SEC-SAXS analysis and EOM modelling for LPMO constructs.

LPMO Construct	SEC-SAXS Rg (Å)	SEC-SAXS D_{max} (Å)	EOM Average Rg (Å)	EOM Average D_{max} (Å)
*Sli*LPMO10G + CBM5	23.54	73.6	24.33	78.78
*Sli*LPMO10F + CBM5	23.97	74.3	24.96	80.17
*Sli*LPMO10C + CBM2	29.68	93.5	30.89	96.10

6.3.14 Alignment of LPMO Models to SEC-SAXS Envelopes

The SAXS *ab initio* envelope for monomeric CBD – *Sli*LPMO10G was aligned with its AlphaFold-predicted model using CIFSUP (ATSAS) to validate its conformation. **Figure 6.19** shows an excellent fit with no significant unoccupied regions, further supporting a globular and compact structure for CBD – *Sli*LPMO10G. This is consistent with SAXS plots such as Kratky plots (**Figure 6.15**) and P(r) distributions (**Figure 6.16**), which suggest a monomeric and single well-folded domain with no flexible linker.

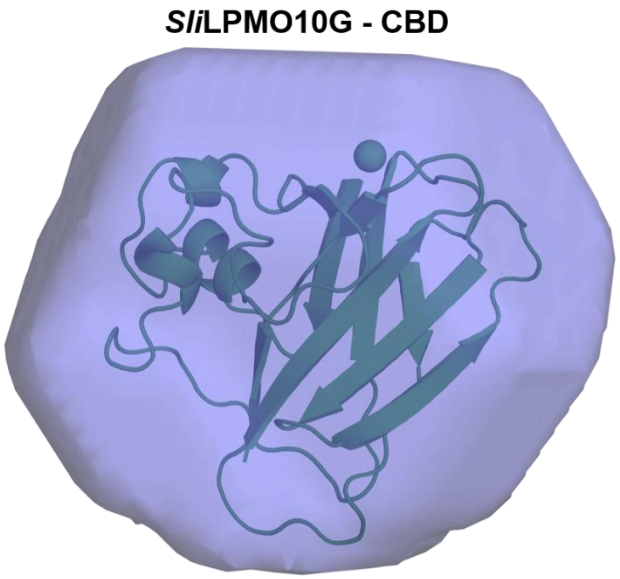


Figure 6.19. *Sli*LPMO10G CBD – Alphafold3 model with experimental SAXS envelope. Transparent surface represents SAXS envelope and purple cartoon represents CBD – *Sli*LPMO10G Alphafold3 model.

Experimental SAXS envelopes for the two-domain LPMOs were aligned with their corresponding EOM outputs using CIFSUP (ATSAS). **Figure 6.20** shows *Sli*LPMO10C fit which has significantly higher unoccupied space within the SAXS envelope compared to *Sli*LPMO10F and *Sli*LPMO10G. These may be due to differences in linker sequence giving *Sli*LPMO10C a more rigid linker compared to *Sli*LPMO10F and *Sli*LPMO10G. On the other hand, *Sli*LPMO10F and *Sli*LPMO10G both show an excellent fit to experimental SAXS envelope with no significant unoccupied space. CIFSUP alignments further reinforce insights gained from SAXS plots such as Kratky plot (**Figure 6.15**) and P(r) plot (**Figure 6.16**) which suggest a multimodule structure connected by a flexible linker

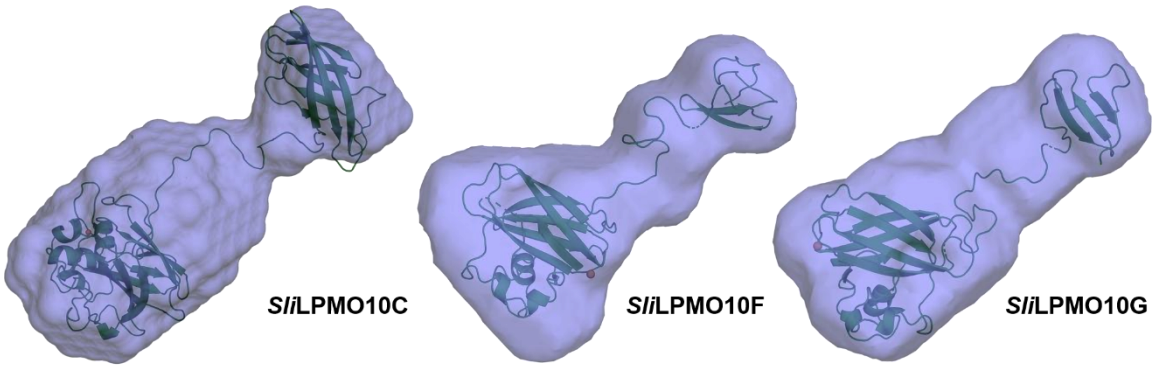


Figure 6.20. *Sli*LPMO10C, *Sli*LPMO10F and SliLPMO10G Alphafold3 model aligned with experimental SAXS envelope, using EOM. Transparent surface represents SAXS envelope and purple cartoon represents CBD – *Sli*LPMO10G Alphafold3 model.

6.3.15 Multiple Sequence Alignment of LPMO Linkers

A multiple sequence alignment using clustalW in Jalview viewer (Waterhouse, *et al*., 2009) was carried out to analyse amino acid composition within the flexible linkers to gain further insight on experimental SAXS plots. Constructs *Sli*LPMO10F and *Sli*LPMO10G both share a shorter Gly-rich linker at 2.2 kDa (*Sli*LPMO10F) and 1.8 kDa (*Sli*LPMO10G). In contrast, *Sli*LPMO10C with significantly larger linker at 4.6 kDa is abundant in Pro, Thr and Ser. This different composition in *Sli*LPMO10C has been linked with reduced backbone flexibility in linkers rich in Pro (Courtade, *et al*., (2018)) resulting in greater rigidity and extended conformation exhibited by linker. On the other hand, *Sli*LPMO10F and *Sli*LPMO10G contain a high percentage of Gly residues which are very small and flexible. Despite sequences length of 24 AA for *Sli*LPMO10G, 28 AA for *Sli*LPMO10F and 29 AA for *Sli*LPMO10C, the D_{max} value derived from P(r) plot (**Figure 6.16**) is significantly greater for *Sli*LPMO10C, compared to *Sli*LPMO10F and *Sli*LPMO10G. High sequence linker divergence across *Sli*LPMO10C was discovered against less divergent *Sli*LPMO10F and *Sli*LPMO10G. A percentage identity of 27.78% was calculated when compared *Sli*LPMO10C and *Sli*LPMO10G 26.67% for *Sli*LPMO10C compared to *Sli*LPMO10F and 70.83% when comparing

*Sli*LPMO10F to *Sli*LPMO10G. The conservation and consensus in **Figure 6.21** further show the divergence between *Sli*LPMO10C compared to *Sli*LPMO10F and *Sli*LPMO10G, compared to Gly dominant *Sli*LPMO10F/*Sli*LPMO10G.

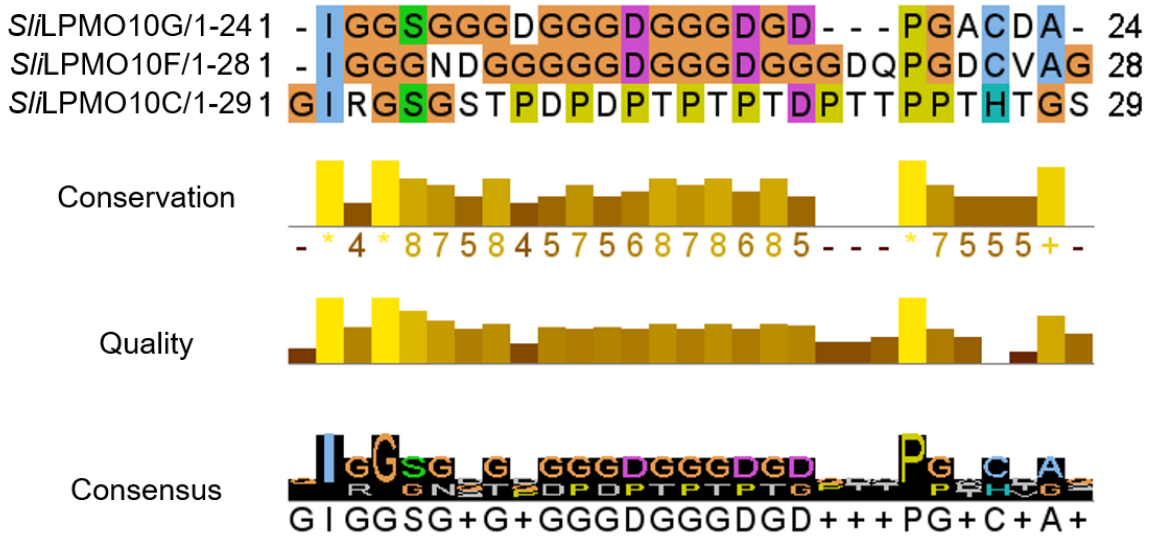


Figure 6.21. Multiple sequence alignment of linker regions from *Sli*LPMO10C, *Sli*LPMO10F and SliLPMO10G using ClustalW in Jalview.

6.4 Discussion

The main focus of this Chapter was to over-express and purify all three two-domain LPMOs that are secreted in *S. lividans*, with the aim of investigating solution-state structural properties using SEC-SAXS. In this respect all two-domain LPMO constructs were successfully over-expressed in *E. coli* and purified in the apo-form. For all LPMO constructs produced in this Chapter addition of Cu(II) resulted in a quenching of Trp fluorescence which is consistent with copper binding to the active site with high affinity and stoichiometries close to one (Chaplin, *et al*., 2026). Thus, the appended CBMs do not appear to influence copper uptake, based on the fluorescent measurements Although the substrate binding assays (**Figure 6.7**) give some insight into possible variations in substrate specificity between the two-domain

proteins, no activity assays with the two-domain constructs were carried out. Future work could investigate this in detail and where possible activities could be compared between two-domain proteins and the LPMO domain only, so as to assess the CBM effect.

The EPR data portray a complex mix of copper coordination species that are accessible across the three two-domain LPMOs. Heterogeneity within the His brace coordination sphere has been previously recognised and leads to multiple species being observed in the EPR spectra (Lindley, *et al*., 2022; Chaplin, *et al*., 2016). A recent study has used quantum chemical approaches to compute Cu(II) g-tensors and hyperfine coupling constants for a suite of core Cu(II) coordination models that can be populated starting from the AA10 resting state coordination sphere, CuN_3O_2, where N is the N-terminal amine group and the His brace, and the O represents two H_2O molecules (Gómez-Piñeiro, *et al*., 2022). This distorted square pyramidal geometry can interconvert between eight different coordinate states, which include loss of water ligands, deprotonation and dissociation of the N-terminal amino group (Gómez-Piñeiro, *et al*., 2022). For each of these coordinating states the Δg and A_{max}^{Cu} have been computed and thus can act as benchmarks to predict from experimental EPR data the coordination geometries present in a LPMO sample.

For *Sli*LPMO10C the EPR spectrum and simulation is consistent with a single copper species with a rhombic EPR signature, with the Δg and A_{CuZ} values reported in **Table 6.3** closely aligned with the predicted parameters of Δg = 0.201 and A_{max}^{Cu} = 446 for a square planar CuN_3O species, with a single H_2O ligand. Such a coordination species is also consistent based on the EPR parameters reported in **Table 6.3** for the LPMO domain only of *Sli*LPMO10G. For *Sli*LPMO10F, component I displayed an axial EPR signature which based on the experimentally determined Δg and A_{CuZ} values reported in **Table 6.3** are a close match to the predicted, Δg = 0.184 and A_{max}^{Cu} = 569 values, for a square planar CuN_3O complex, in which the H_2O ligand is now deprotonated to from a hydroxide ligand. For component II the

simulated parameters reported in **Table 6.3** align with those reported for a five-coordinate copper centre with two H_2O molecules coordinating the Cu(II) ion (Δg = 0.225 and A_{CuZ} = 370). This is also the case for the *Sli*LPMO10G (**Table 6.3**). **Figure 6.22** summaries the various copper geometries for the various *Sli*LPMOs that have been identified based on the comparison of the predicted and experimental data. This reveals that there is heterogeneity in the copper coordination sphere between the various LPMOs, with *Sli*LPMO10F demonstrating coordination geometries of four and five within the sample. It is noted that none of the EPR parameters reported are consistent with a copper state whereby the N-terminal amino group is dissociated. Such a state has been inferred from Cu(II) binding kinetic studies with *Sli*LPMO10E (Chaplin *et al.* 2016). It should be remembered however that the heterogeneity of copper coordination geometries identified for the Cu(II) state may be of little functional significance as it is the reduced Cu(I) state which, inherently has a lower coordination number (loss of H_2O/OH^- ligand) and thus assisting in minimising coordination heterogeneity, is the functionally active state.

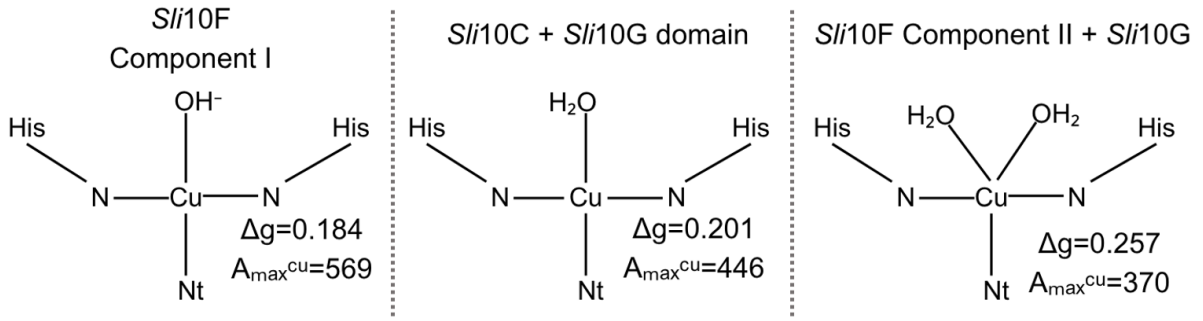


Figure 6.22. Schematic of coordination environment for LPMO copper sites of *Sli*10G-domain only, *Sli*10G, *Sli*10F and *Sli*10C, with the predicted EPR parameters reported

Crystallisation of multidomain enzymes that possess intrinsically disordered linkers remains a very difficult challenge. Flexible linkers can adopt many inter-domain orientations, thus, preventing the formation of a well-ordered lattice for crystallisation and poor diffraction quality of crystals (Powell, 2016). To overcome this issue and be able to examine the structure

of two-domain LPMOs from *S. lividans*, the low-resolution solution technique SEC-SAXS was implemented. This combined with ensemble optimisation method (EOM) modelling can reveal how linker composition can dictate conformation in solution.

Disordered/flexible linkers found in multimodule LPMOs has proven to be significant in structural flexibility and appears to have functional importance improved yield of oxidised sugars at low substrate concentrations (Courtade, *et al*., *2018*), (Méndez-Líter, *et al*., 2021), (Srivastava, *et al*., 2022). A previous study looking at substrate binding affinity of *Sc*LPMO10C from *Streptomyces coelicolor* (analogous to *Sli*LPMO10C), revealed how the Pro-rich 34 AA long linker can adopt a disordered or extended conformation (Forsberg, *et al*., 2023), allowing the catalytic copper-binding domain and CBM2 to move independently of each other (Courtade, *et al*., 2018). This aligns with the SAXS data obtained with similar length (29 AA) Pro-rich linker in *Sli*LPMO10C, where unoccupied space can be observed in the catalytic copper-binding domain (**Figure 6.20**). Pro residues display greater rigidity and rotational restrictions, in turn, this exhibits an extended conformation with domains moving independently from each other. As SAXS is a solution-based techniques, it captures a broad conformational range, this can explain the large unoccupied space within SAXS envelope and AlphaFold3 model.

On the other hand, Gly-rich *Sli*LPMO10F and *Sli*LPMO10G (**Figure 6.20**) shows a tight fit with experimental SAXS envelope. As glycine possesses high levels of flexibility due to a very small sidechain, this maximises backbone freedom. *Sli*LPMO10F and *Sli*LPMO10G contains Gly/Asp repeats (**Figure 6.21**) ("GGGDGGGD"); Asp is another disorder promoting residue and promotes disorder through charge repulsion to keep the flexible linker extended without creating rigidity (van Rosmalen, *et al*., 2017). The highly flexible nature of prevalent Gly/Asp repeats suggest a confinement of motion in catalytic copper-binding domain and CBM5, compared to flexible linker that moves more freely**. Figure 6.20** containing EOM

alignment with experimental SAXS envelope shows higher volume of unoccupied space in *Sli*LPMO10F and *Sli*LPMO10G, compared to the linker of *Sli*LPMO10C. The flexible linker of *Sli*LPMO10F and *Sli*LPMO10G could be adopting a coil-like shape which coils back on themselves (Ceballos-Alcantarilla and Merkx, 2021), explaining the unoccupied space within the linker.

Despite difficulties crystallising two-domain LPMOs, many studies have solved the X-ray crystal structure of the catalytic copper-binding domain and CBM of the LPMOs; due to their high rigidity (Ipsen, *et al*., 2021; Frandsen and lo Leggio, L. 2016; Boraston, *et al*., 2004). The crystal structure of copper-binding domain (CBD - *Sli*LPMO10G) was solved at 1.2 Å, with copper ion coordinated with canonical histidine brace motif. Overall structure of CBD – *Sli*LPMO10G shows high similarity with other chitin active AA10s, such as the highly conserved Glu-68, Ala-125, Glu-196 and Phe-198 (*Sc*LPMO10D) from *S. coelicolor* (Votvik, *et al*., 2023), which is highly similar to Glu-62, Ala-104 and Phe-162 which sits at close proximity of the histidine brace. Another example includes chitin active AA10s from *Bacillus thuringiensis* (*Bt*LPMO10A and *Bt*LPMO10B) (Zhang, *et al*., 2023) which also possess a Glu, Phe and Ala in close proximity of the histidine brace. AA10 LPMOs of *Sli*LPMO10E contains an active site with Glu, Phe and Ala residues at close proximity with very similar bond length and angles.

Conclusions

In conclusion, a complex mixture of copper coordination geometries amongst the three two-domain *Sli*LPMOs have been identified experimentally. Based on a recent theoretical study that benchmarked EPR parameters for a set of Cu(II) coordination complexes that may exist in LPMOs, at least four distinct Cu(II) coordination complexes are present across the *Sli*LPMOs investigated under the experimental conditions employed **(Figure 6.22).** From the

structural data, the combination of SEC-SAXS envelope modelling combined with EOM, has provided further insights of the two-domain structure of *Sli*LPMOs in solution.

Chapter 7 Serial X-ray crystallography of *Sli*LPMO10E: Towards Time-Resolved Studies with a Molecular Oxygen Releasing Photocage

7.1 Introduction

Lytic polysaccharide monooxygenases (LPMOs) accelerate the depolymerisation of recalcitrant polysaccharides by oxidative cleavage of glycosidic C-H bonds. LPMOs are catalytically competent towards molecular oxygen O_2 or hydrogen peroxide H_2O_2 (See, Chapter 6, **6.1**) (Stepnov, *et al*., 2022; Rieder, *et al*., 2021; Wang, *et al*., 2020). Studies have suggested a mechanism employing O_2 as a co-substrate, followed by O_2 activation (Wang, *et al*., 2019a; Beeson, *et al*., 2015), initiated by the reduction of resting Cu(II) to active Cu(I) state, leading to the formation of reactive O_2 species CuO_2^+ and CuO^+ (See, Chater 6 **Figure 6.1**) for hydrogen atom abstraction from polysaccharide substrates (Batka, *et al*., 2024; Hedegård and Ryde, 2018). These intermediates can be captured as snapshots by implementing time-resolved X-ray crystallography, making it possible to capture enzymes in motion (Wilson, 2022).

Characterising enzyme reactions and catching structural snapshots of reaction intermediates, remains a significant challenge in structural biology (Smyth, *et al*., 2025; Caramello and Royant, 2024). Techniques to solve these issues have significantly advanced through the development of serial crystallography. This method involves reaction initiation, quickly followed by measurement of a single still diffraction image after a set time delay. Variation of time delays makes it possible to capture a stop-motion series of protein structures during catalysis (Smyth, *et al*., 2025). The main method used to initiation enzyme catalysis can

be carried out by using light pulses or soaking protein microcrystals (10 – 50 μm in size) in reductant/substrate as well as other reagents (Wolff *et al*.,2023). However, the majority of enzymatic reaction are not naturally activated by a light source, such as rhodopsin's or flavins (Smyth *et al*., 2025; Nango, *et al*., 2016; Kupitz, *et al*., 2014)

To investigate O_2 activation in LPMOs, the experiment must be carried out under strict anaerobic conditions to prevent active Cu(I) oxidising back to resting state Cu(II) after soaking in reductant. One possible approach to explore is the implementation of a O_2-photocage (oxycage), utilising light to initiate release of O_2 from the oxycage soaked in LPMO microcrystals that have been reduced to the Cu(I) form (Monteiro, *et al*., 2021; Ludovici *et al.*, 2002).

To capture intermediate steps of O_2 activation at the metal site of a metalloenzyme such as an LPMO, femtosecond time-resolved X-ray crystallography at an X-ray free-electron laser (XFEL) can be employed (de Wijn, *et al*., 2022). Single-hit serial data collection minimises radiation damage of on protein micro-crystals by avoiding cumulative dose on a single crystal, preventing radiation damage and makes it possible to captures events on sub millisecond time scales (Hunter, *et al*., 2014). This is particularly important for metalloenzymes, where preserving a redox state and the structural changes or rearrangements associated, are important for deciphering mechanism. At room temperature the likelihood of capturing subtle Cu-O_2 binding events in an LPMO and geometrical arrangements at the Cu site following the initial O_2 binding is important. Further understanding of O_2 activation in LPMOs provides opportunity for protein engineering, site-directed mutagenesis on histidine-brace for beneficial geometries and even extend catalytic lifetime under harsh conditions (Hall, *et al*., 2025).

7.1.1 Aims

The aims of this study were to capture structural snapshots of LPMO (AA10) from *Streptomyces lividans* (*Sli*LPMO10E) following O_2 activation at the copper histidine-brace site, using time-resolved X-ray crystallography. This will be carried out by producing reduced Cu(I) state *Sli*LPMO10E micro-crystals in anaerobic conditions, initiating O_2 binding by laser excitation to release O_2 from a cobalt oxycage [(μ-peroxo)(μ-hydroxo)bis[bis(bipyridyl)cobalt(III)] (HPBC) and capturing structural snapshots with single-hit femtosecond time-resolved crystallography.

7.2 Methods

7.2.1 *Sli*LPMO10E Cloning and Site-directed Mutagenesis

A modified pET26b (Novagen) plasmid was used to clone into the genes for the LPMO *Sli*LPMO10E, so that following cleavage of the pelB leader sequence the N-terminal His residue of the His-brace was the first residue. *Sli*LPMO10E in the pET26b vector was provided by Professor Jonathan Worrall (Chaplin, *et al*., 2016). *Sli*LPMO10E was selected as microcrystals yields were highest with this target.

7.2.2 Recombinant Overexpression and Purification of *Sli*LPMO10E

Expression of all *Sli*LPMO10E were carried out in *E. coli* BL21(DE3)-RIL cells. The pET26b plasmid containing the gene for *Sli*LPMO10E was inoculated into (1.5% *(w/v)* LB agar and left at 37°C overnight. Single colony transformants were picked and inoculated into sterile 5 mL overnight cultures using a sterile toothpick and left overnight for 16-20 hours (37°C, 180 rpm). Contents of overnights were inoculated into 2 L sterile flasks containing 500 mL of LB media with 50 μg mL^{-1} of kanamycin and chloramphenicol (37°C, 180 rpm). At

OD_{600} 0.4, the temperature was lowered to 16 °C with the addition of 0.5 M isopropyl-β-D-thiogalactopyranoside (IPTG) (Melford) on reaching OD_{600} 0.6. After 16-18 h cells were harvested and a sucrose shock procedure as described in Chapter 6, **6.2.2** was carried out. Dialysate of *Sli*LPMO10E (pI 8.42) was diluted in buffer A (10 mM Tris/HCl, pH 7), HiTrap SP HP cation exchange column (Cytiva) was equilibrated in buffer A and the dialysate loaded and the bound *Sli*LPMO10E eluted with a linear salt gradient using buffer B (10 mM Tris/HCl pH 7, 1 M NaCl). Fractions eluting with a UV absorbance at 280 nm were checked using a 15% SDS-PAGE gel and fractions containing a band on the SDS-gel with the expected molecular weight for *Sli*LPMO10E pooled followed by concentrating using a 5 kDa cutoff Centricon (Vivaspin) at 4°C. Once concentrated to <2 mL, *Sli*LPMO10E was loaded onto a HiLoad 26/600 Superdex-75 column (Cytiva) equilibrated in buffer C (50 mM Tris/HCl pH 7, 150 mM NaCl). Fractions showing high purify after checking using a 15% SDS-PAGE gel, pooled and concentrated to required concentration using a 5 kDa cutoff Centricon (Vivaspin) at 4°C.

7.2.3 Concentration determination of *Sli*LPMO10E and Copper Solution Preparation

Concentration of purified *Sli*LPMO10E was determined using a Cary60 UV-Vis spectrophotometer (Agilent Technologies), using protein extinction coefficients determined at 280 nm (ε) (36,690 M^{-1} cm^{-1} for *Sli*LPMO10E) using ProtParam (Gasteiger *et al*., 2005) with the respective protein sequence as the input. A 100 mM $CuSO_4$ stock solution was prepared in buffer A for stoichiometric addition of $CuSO_4$, followed by removal of excess Cu(II) using a desalting PD-10 column (Cytiva).

7.2.4 Crystallization and Cryo-X-ray Crystallography of *Sli*LPMO10E

*Sli*LPMO10E was concentrated to ca. 12 mg mL^{-1} using a 5 kDa cutoff Centricon (Vivaspin) at 4°C in buffer C. Crystallisation conditions were screened on the Cu(II)-loaded *Sli*LPMO10E (ca. 12 mg mL^{-1}) using an ARI-Gryphon 96-well crystallisation robot, using screens (JCSG Plus, PACT premier, PGA, ProPlex, 3D structure screen and ammonium sulfate screen). Crystals were found in 0.1 M MES, pH 6, 2.4 M ammonium sulfate (Grid Screen Ammonium Sulfate) at a 1:1 and 1:2 protein:screen condition ratio in 0.4 µL volume drops. Crystal conditions were then tested using batches at different ratios of protein:precipitant solution, with large crystals being formed in the 1:1 and 1:2 drop ratios at a final volume of 100 µL after 4-5 weeks at 18°C. A single large crystal (ca. 50 µm) was fished from the batch and soaked in 20 % sucrose cryoprotectant prior to flash-freezing by plunging into liquid nitrogen. Crystallographic data was measured at Diamond Light Source, beamline I04, at 100 K using a 0.953738 Å X-ray wavelength and a EIGER X 16M detector. Data was processed using the xia2.multiplex pipeline, which uses DIALS for spot finding, indexing, integration and scaling of multi-sweep datasets. Data was scaled and merged using AIMLESS (Evans, 2011) in CCP4i2 (Agirre, *et al*., 2023) and molecular replacement in PHASER (McCoy, *et al*., 2007), using PDB code 5ftz from a previously solved *Sli*LPMO10E X-ray structure. Cycles of refinement and model building were carried out using Refmac5 (Murshudov, *et al*., 2011) and Coot (Emsley, *et al*., 2010), respectively in CCP4i2. Riding hydrogens were added during refinement. The structure was validated using tools within Coot. Map files for viewing in PyMOL were generated from mtz files in Phenix version 1.21-5207-000 (Liebschner, *et al*., 2019). Data collection and refinement statistics are reported in **Table 7.1.**

Table 7.1 Cryo-Crystallographic processing and refinement statistics for *Sli*LPMO10E. Values in parenthesis are for the highest resolution shell.

Space group	P1
Unit Cell (Å)	a = 33.5, b = 36.1, c = 61.5, 81.4° 78.6° 64.1°
Resolution (Å)	60.18 - 1.25 (1.27 – 1.25)
Unique Reflections	66903 (3041)
Multiplicity	3.6 (3.1)
Completeness (%)	95.1 (87.8)
Mn(I/sd)	17.5 (8.3)
R_{merge}	0.04 (0.089)
Wilson B-factor (Å^2)	5.697
$CC_{1/2}$	0.999 (0.987)
R_{work}	0.108
R_{free}	0.139
RMSD bond lengths (Å)	0.0104
RMSD bond angles (°)	1.806
Ramachandran most favoured (%)	98.2
PDB accession code	**TBC**

7.2.5 Crystallisation of *Sli*LPMO10E microCrystals

Microcrystals of *Sli*LPMO10E (ca. 12 mg mL^{-1}) in buffer C were obtained in batches, containing a 1:7 protein:condition ratio in 0.1 M MES pH 6 and 2.4 M ammonium sulfate, at final volume of 200 µL. Batches were left for 5-6 weeks at 18°C, with ca. 70% of batches forming micro-crystals at ≤15 µm in size. Crystal batches that yielded larger crystals were resuspended and broken down into micro-crystals by repeated up-and-down pipetting using

P200 tip, followed by vortex mixing to further reduce crystal size. Crystal slurries were checked on a Hirox digital microscope (MXG-2500REZ lens) resulting in micro-crystal sizes of ≤10-25 μm

7.2.6 Spectroscopic Validation of HPBC on Bovine Myoglobin

HPBC oxycage was synthesised at the University of Essex using a published method (Ludovici, *et al*., 2002). HPBC powder, 0.1 M MES pH 6 and 2.4 M ammonium sulfate buffer and bovine myoglobin powder was left in the anaerobic chamber for 24 h, under dark conditions to prevent photolysis of HPBC. HPBC powder was dissolved in 0.1 M MES pH 6 and 2.4 M ammonium sulfate buffer to make stock solution (10 mg mL^{-1}). 100 μL of dissolved HPBC was mixed with 900 μL of 0.1 M MES pH 6 and 2.4 M ammonium sulfate buffer (1 mg mL^{-1}) in a cuvette and a spectrum was taken using Cary60 UV-Vis spectrophotometer (Agilent Technologies). 10 μM final concentration of bovine myoglobin in 0.1 M MES pH 6 and 2.4 M ammonium sulfate buffer was reduced by mixing dithionite powder with a pipette tip, until solution was bright red (indicative of reduced state). Bovine myoglobin was then passed down a desating column (PD-10) to remove excess dithionite. 20 μL of bovine myoglobin was thoroughly mixed into a cuvette containing HPBC oxycage dissolved in buffer, with an air-tight cap to prevent O_2 exposure and a deoxy spectrum was measured. Cuvette was flashed three times and another spectrum was measured to identify shift to oxy spectrum (**Figure 7.7**). This was repeated until no further changes were identified in spectrum.

7.2.7 Sample Preparation and Crystallographic Data Measurements at SACLA XFEL

Microcrystal batches were transferred into the anaerobic chamber (<0.1 ppm (parts per million) O_2 reading) at the SACLA experimental hall and tubes were left open for one-hour intervals to remove O_2 before use. Eppendorf tubes containing HPBC oxycage powder,

reductant powder (ascorbic acid, cysteine) and stock crystal condition (0.1 M MES pH 6, 2.4 M ammonium sulfate) were also left open to remove O_2; all handing was done in the dark under red light. Protein *Sli*LPMO10E microcrystals were handled under anaerobic conditions and concentrated down by centrifugation at 15,000g for 3 minutes, 100 µL of slurry was removed from the top and 100 µL of crystal slurry was loaded onto a silicon nitride chip at aperture size of 15 or 10. Vacuum pump was used to orientate micro-crystals and chip was placed into a holder (**Figure 7.1**). Loaded chips were sealed on each side using EVAL [poly(ethylene-co-vinyl alcohol)] film, allowing high transmission of 308 nm laser light while maintaining a humid environment, preventing microcrystals from drying out and importantly forming a barrier to ambient oxygen contamination. A 10 mg mL^{-1} stock of HPBC oxycage was made from crystal condition buffer and stocks of reductants, ascorbic acid (25 mM) and cysteine (50 mM) were made in crystal condition (0.1 M MES pH 6, 2.4 M ammonium sulfate). 100 µL additions of HPBC oxycage (5 mg mL^{-1} final concentration) was added to *Sli*LPMO10E micro-crystals, together with final concentration of reductants, ascorbic acid (5 mM) and cysteine (1.5 mM), additions of oxycage and reductants were left to soak for 20 minutes prior to chip loading.

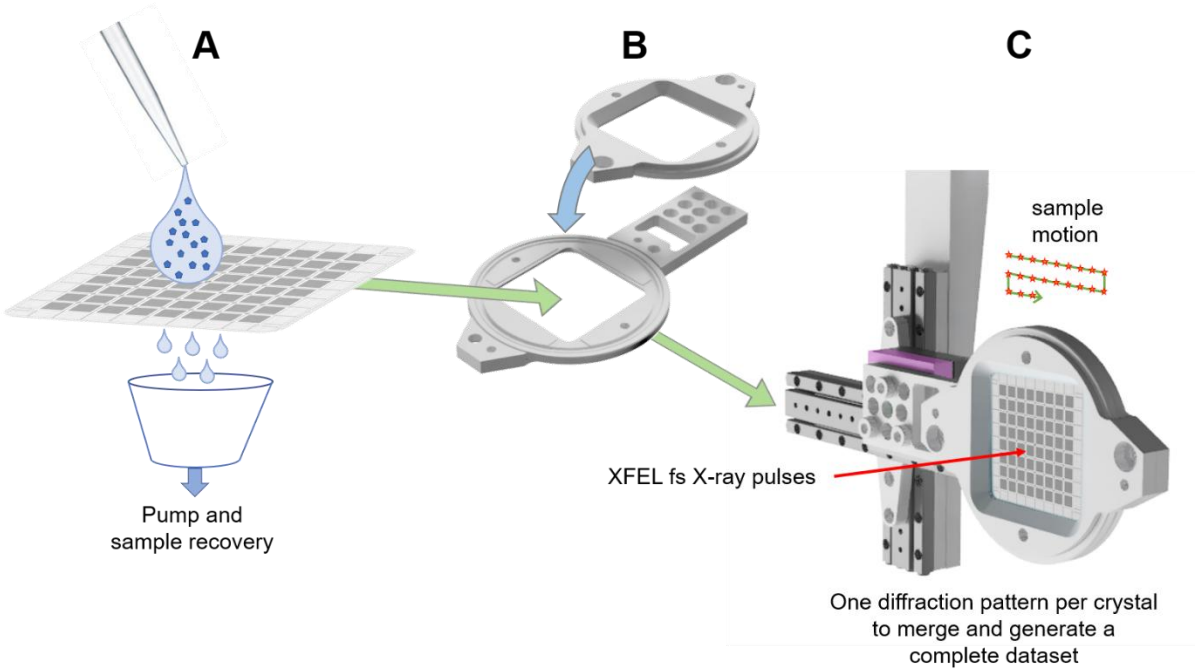


Figure 7.1. Chip delivery system at XFEL for room temperature time-resolved X-ray crystallography. **(A)** loading of silicon chip with 1000s of micro-crystals. **(B)** Assembly of chip mount. **(C)** Direction of chip motion. during X-ray pulses. Figure from (Lučić, *et al*., 2021).

7.2.8 Data processing at XFEL

SACLA beamline BL3EH2 was used with an X-ray energy of 11 KeV, 10 fs pulses delivered at 30 Hz and a beam size of 1.4 x 1.4 (h) x (v) microns. Data were measured on a Rayonix detector and were processed on site at SACLA using CrystFEL. Crystals were delivered to the beam using the fixed target silicon chip approach as described previously. Final data was processed by Dr Amy Thompson (Diamond Light Source) using xia2.ssx (https://doi.org/10.1016/bs.mie.2024.10.004). Standard settings were used for all parameters. A mask was generated for shadows on the detector from the beam-stop and the fixed target chips, as well as a ring from 1.655 – 1.625 Å where broad silicon diffraction was occurring. Resolution cut offs were applied where $CC_{1/2}$ fell smoothly to 0.3.

Table 7.2. Serial femtosecond crystallography (XFEL) data-collection and merging statistics.

	Cu(II)-LPMO aerobic	Cu(I)-LPMO anaerobic (asc)	Cu(I)-LPMO anaerobic (cys)	Cu(I)-LPMO anaerobic cage/dark + (cys)
Resolution range (Å)	32.97 – 1.30 (1.32 – 1.30)	32.92 – 1.32 (1.34 – 1.32)	32.93 – 1.30 (1.32 – 1.30)	32.97 – 1.30 (1.32 – 1.30)
Space group	P1	P1	P1	P1
Unit cell	33.62, 36.42, 62.21, 82.55, 78.15, 65.02	33.63, 36.35, 62.20, 82.60, 78.02, 65.01	33.65, 36.37, 62.23, 82.60, 78.03, 65.02	33.52, 36.42, 62.20, 82.59, 78.21, 65.00
No. crystals merged	17590	9831	11558	13542
Total reflections	3299315 (35102)	1840978 (25055)	2148236 (22600)	2551160 (26978)
Unique reflections	62686 (3153)	59735 (3128)	62614 (3152)	62582 (3149)
Multiplicity	52.6 (11.1)	30.8 (8.0)	34.3 (7.2)	40.8 (8.6)
Completeness (%)	97.5 (100)	97.4 (100)	97.4 (100)	97.4 (100)
I/sigma(I)	9.6 (1.8)	6.9 (1.1)	7.1 (1.0)	8.8 (1.4)
R-split	0.167 (0.489)	0.215 (0.874)	0.203 (0.907)	0.183 (0.706)
$CC_{1/2}$	0.929 (0.606)	0.912 (0.311)	0.924 (0.305)	0.940 (0.425)

7.3 Results

7.3.1 Microscopy of *Sli*LPMO10E Micro-Crystals

Light microscopy images of the *Sli*LPMO10E microcrystals used in this study are shown in **Figure 7.2**. **Figure 7.2 (A)** reveals a dense slurry of *Sli*LPMO10E microcrystals as x100 magnification and **Figure 7.2 (B)** illustrates a higher (x400) magnification of microcrystals.

Higher magnification reveals needle-shaped crystals which range of 10-25 µm in size. Microcrystals were employed for serial x-ray crystallography experiments at SACLA (XFEL) facility (Japan)

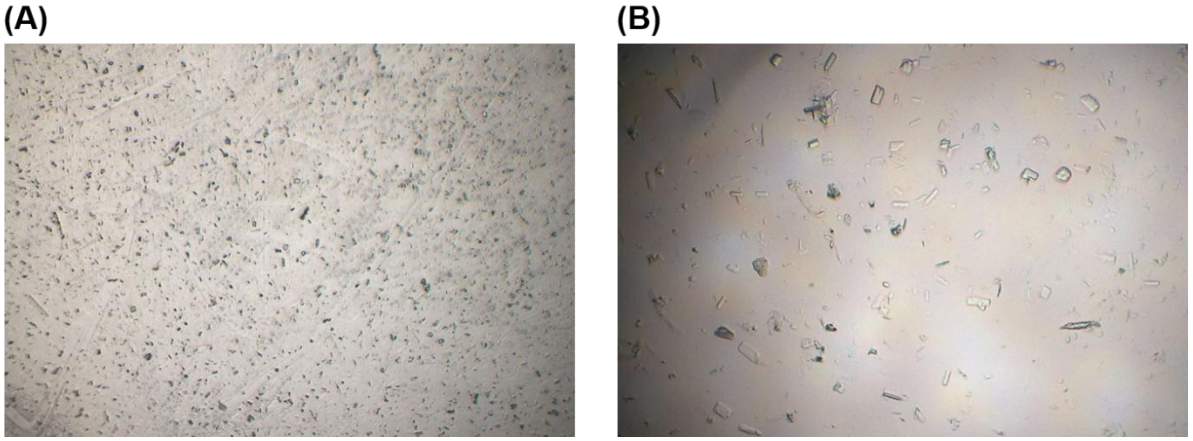


Figure 7.2. Light microscopy images of *Sli*LPMO10E microcrystals used for serial crystallography. Low magnification (x100) shows a distributed and dense slurry of microcrystals **(A)** and higher magnification (x400) reveals individual crystals with rod-like shape **(B)**.

7.3.2 Cryo and Room Temperature X-ray Structures of *Sli*LPMO10E

The cryo crystal structure (100 K) of *Sli*LPMO10E was determined at a resolution of 1.25 Å at I04 beamline, using anisotropic refinement. The copper active site of both *Sli*LPMO10E chains, reveals $2F_o$-F_c density for a single coordinating axial water ligand with a distance of ca. 2.1-2.2 Å from copper ion (**Figure 7.3**). Similarly to the truncated *Sli*LPMO10G structure solved at 100 K (See, Chapter 6, **6.3.5**), *Sli*LPMO10E exists as a monomer in solution but two protein molecules are found in the crystallographic asymmetric unit and are arranged, as shown in **Figure 7.4**. The asymmetric unit reveals the LPMO *Sli*LPMO10E monomers do not adopt the copper face-to-face orientation, seen for *Sli*LPMO10G.

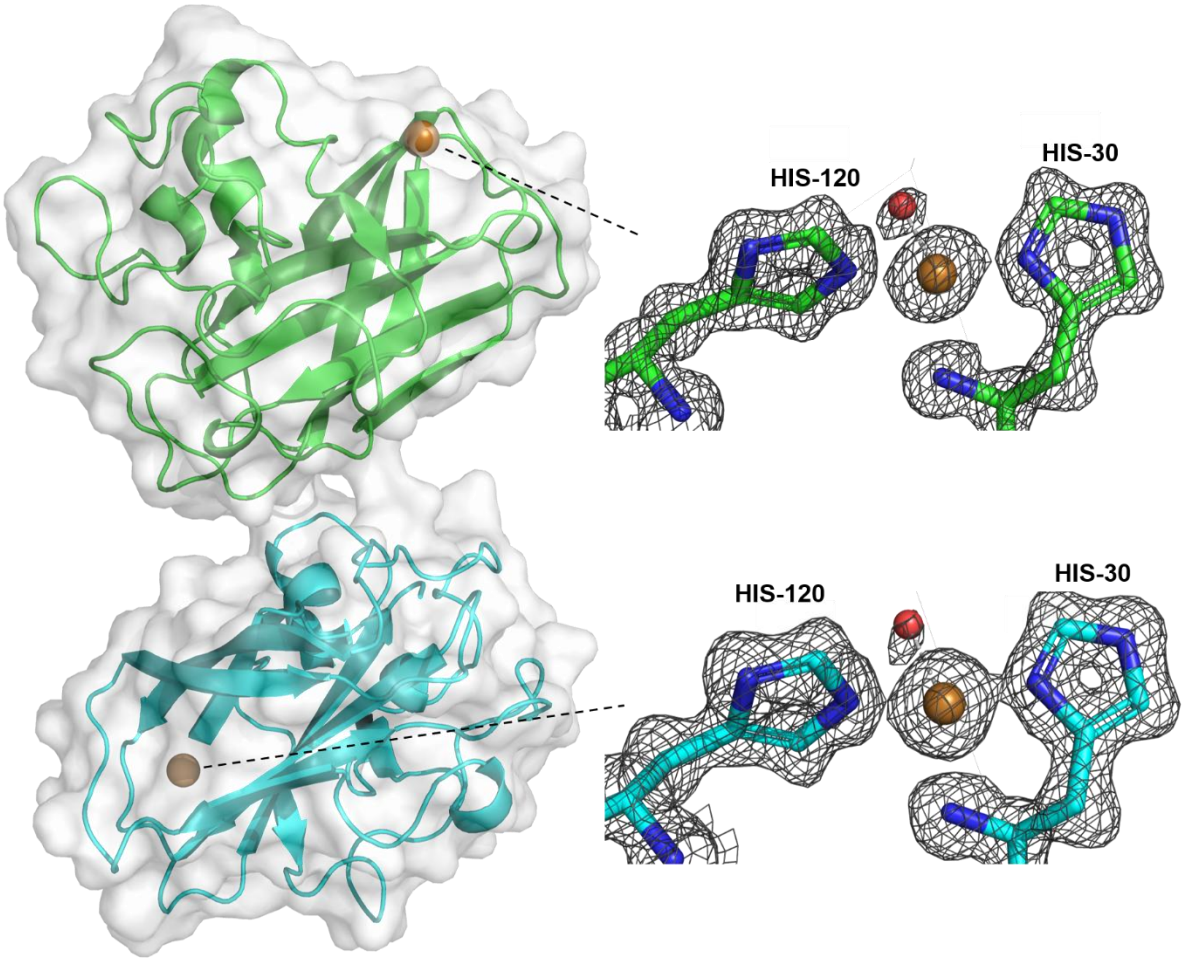


Figure 7.3 Asymmetric unit of truncated *Sli*LPMO10E dimer represented in a transparent surface, showing the two orange copper ion spheres in both LPMO chains **(A)**. A $2F_o$-F_c electron density map contoured at 1 σ at the copper active site in the dimer conformation. Orange sphere = copper and red sphere = water.

The serial room temperature (RT) structure of Cu(II) *Sli*LPMO10E (aerobic conditions) was solved to 1.4 Å resolution by XFEL (SACLA), using single-hit serial crystallography (**Figure 7.4**). Interestingly, the dimer arrangement in the crystallographic asymmetric unit for the RT *Sli*LPMO10E structure (**Figure 7.4**) is very different compared to the solved cryo structure (**Figure 7.3**). Furthermore, whereas the cryo *Sli*LPMO10E structure revealed only one strong water shaped electron density in close proximity to the Cu ion, the RT structure of *Sli*LPMO10E revealed two water shaped electron density features close to the Cu site (**Figure**

7.4). The utilisation of single-hit serial crystallography methods in RT *Sli*LPMO10E could explain the presence of two water electron densities, with mitigation of radiation damage, compare to rotation method used for cryo *Sli*LPMO10E structure.

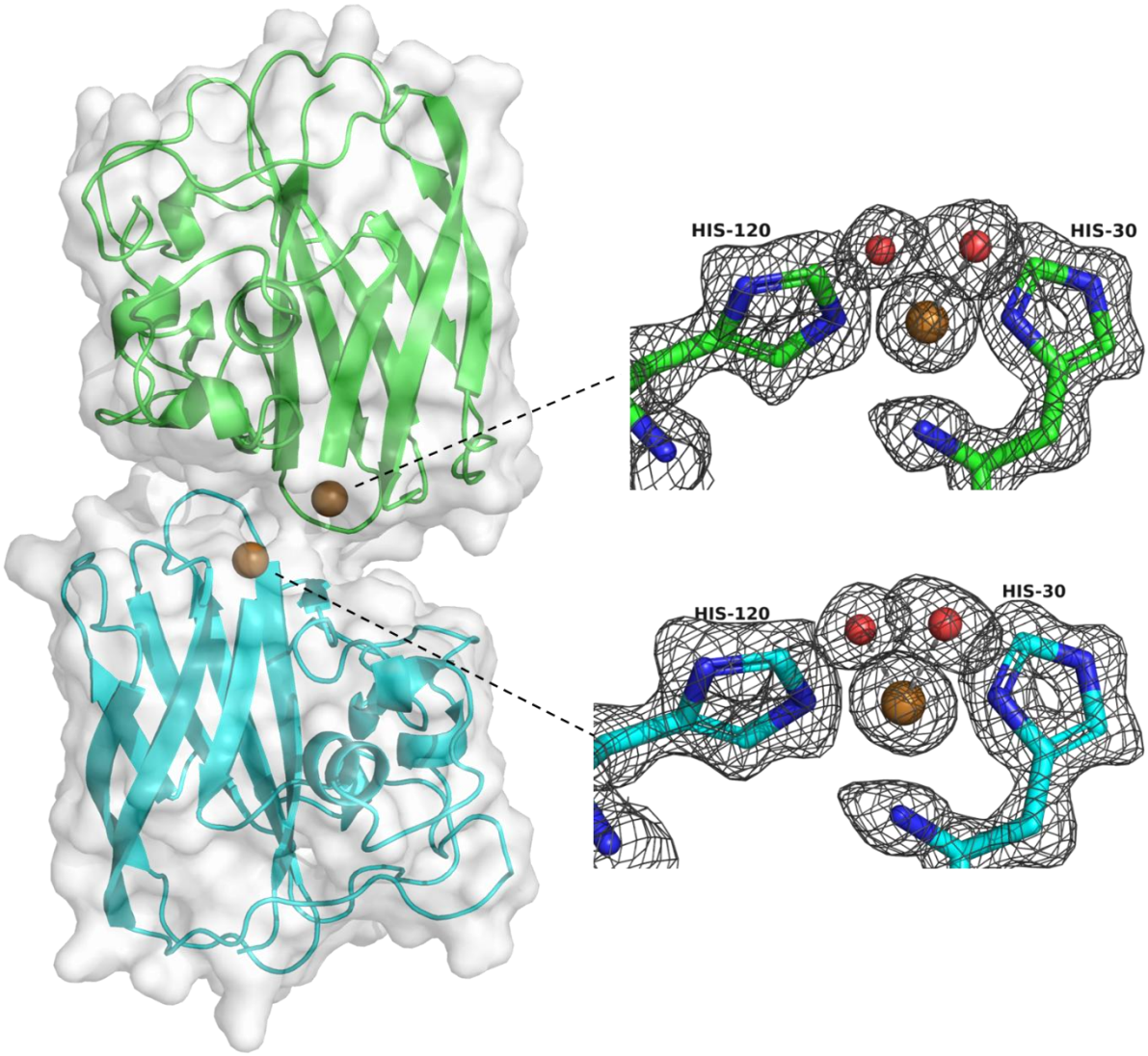


Figure 7.4. Asymmetric unit of truncated *Sli*LPMO10E dimer represented in a transparent surface, showing the two orange copper ion spheres in both LPMO chains **(A)**. A $2F_o$-F_c electron density map contoured at 1 σ at the copper active site in the dimer conformation. Orange sphere = copper and red sphere = water.

A continuous and well-defined electron density map was found for residues 30-201. Each *Sli*LPMO10E monomer consisted of seven anti-parallel β-strands, with one side consisting of β-strands (S1, S3 and S6) and the other side consisting of (S4, S5, S7 and S8), with a small β-strand (S2). Five α-helices are shown (H1-H5) and two disulfide bonds are revealed in *Sli*LPMO10E structure (Cys-43/Cys-55) and (Cys-78/Cys-196) (**Figure 7.5A**). Cryo and RT copper sites show distances of 1.9-2.2 Å between copper ion and N of histidine rings for the cryo structure and 1.9-2.1 Å for between copper ion and N of histidine rings in the RT structure (**Figure 7.5B**).

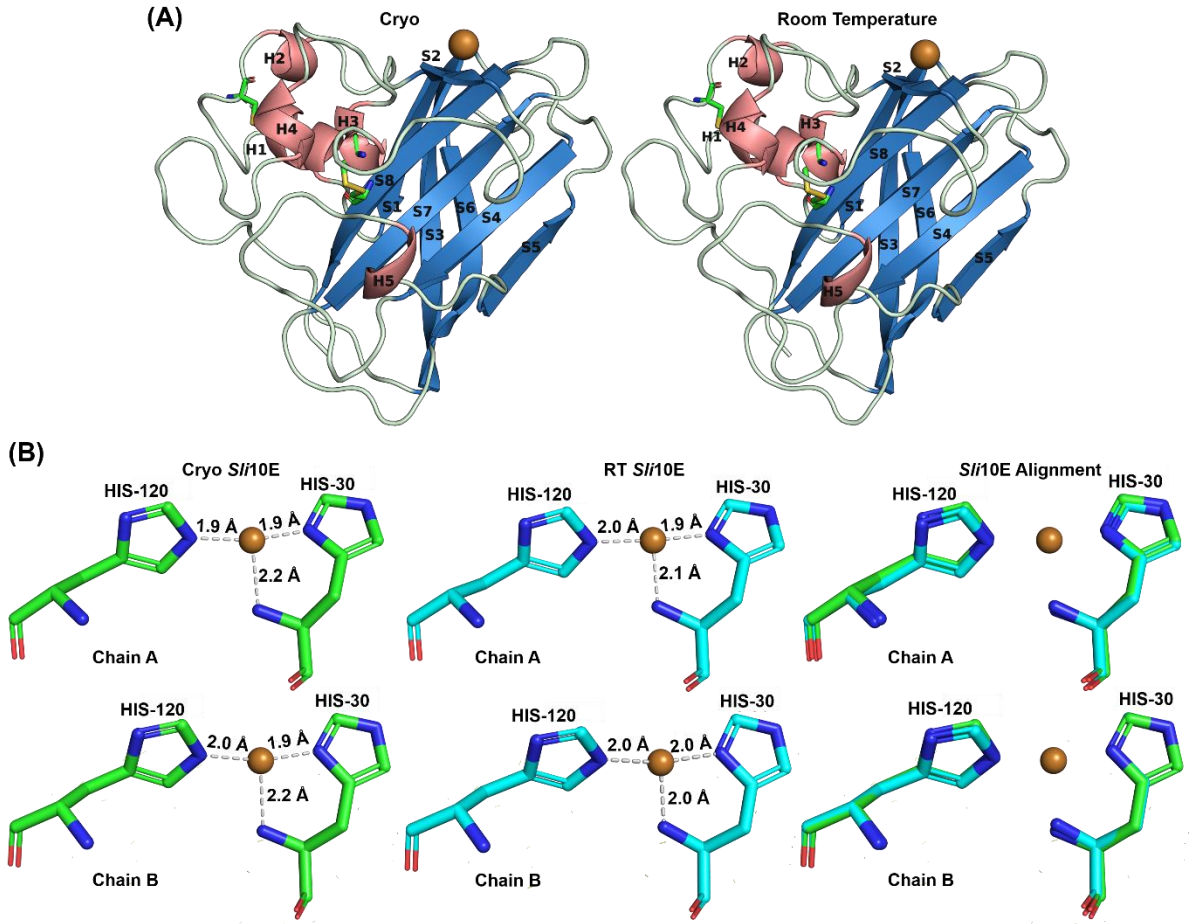


Figure 7.5 (A) Cryo and RT X-ray crystal structure of SliLPMO10E. β-sheets shown in blue, α-helices displayed in pink, loop regions shown in white and disulfide bonds displayed as sticks and copper represented as an orange sphere. **(B)** Copper site comparisons of both copper centres in the asymmetric unit. Green sticks = *Sli*LPMO10E cryo structure and cyan sticks = *Sli*LPMO10E RT XFEL structure.

7.3.3 Spectroscopic Validation of HPBC on Bovine Myoglobin

UV-Vis spectrum of the HPBC oxycage reveals a spectrum that is consistent with literature (Ludovici, *et al*., 2002) (**Figure 7.6A**). Measuring anaerobically prepared deoxy bovine myoglobin mixed with HPBC oxycage prior to flashing confirms a deoxy spectrum (**Figure 7.6B**). After flashing the cuvette using a photographic camera flash a shift in the spectrum occurs. This is indicative of O_2 being released from the photocage and forming oxy-myoglobin, *i.e.* a Fe^{2+}–His → O_2•–Fe^{3+}–His transition (**Figure 7.6D**).

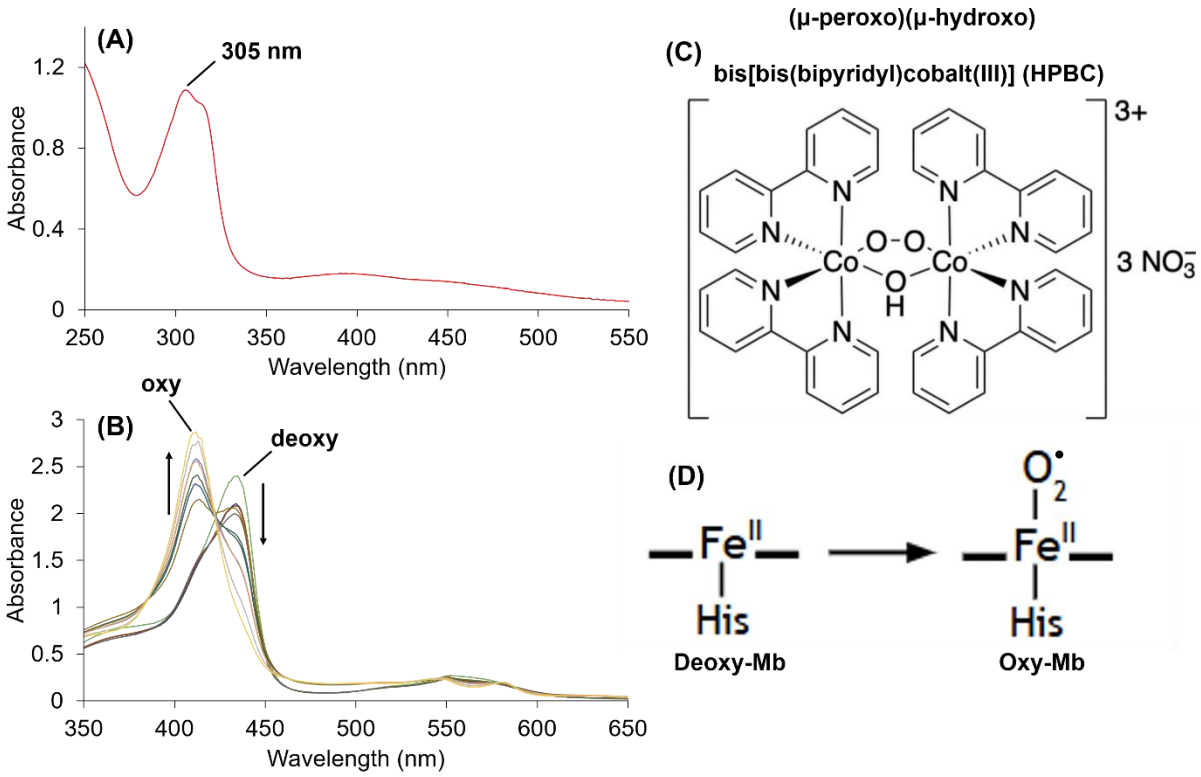


Figure 7.6. **(A)** Spectra of HPBC in crystallisation buffer. **(B)** Spectra of HPBC mixed with bovine myoglobin (10 μM), revealing HPBC uncaging between flashing (deoxy → oxy). **(C)** Structure of HPBC **(D)** Schematic of the deoxy → oxy myoglobin transition.

7.3.4 Active Site Copper Coordination under Reduced States and HPBC Soaks

Under strictly dark and anaerobic conditions, 20 minute soaks of *Sli*LPMO10E micro-crystals in oxycage and reductants were measured using the fixed target chip system using the XFEL at SACLA., five chips for Cu(I)-*Sli*LPMO10E reduced (asc) and Cu(I)-*Sli*LPMO10E soaked with oxycage and six chips for Cu(I)-*Sli*LPMO10E (cys) reduced. The addition of reductants ascorbic acid (asc) and cysteine (cys) under anaerobic condition yielded *Sli*LPMO10E with no electron density features for coordinating water molecules at the the copper ion, indicative of a fully reduced Cu(I) *Sli*LPMO10E structure (**Figure 7.7**). Contrastingly, when microcrystals were soaked in cys reductant and HPBC oxycage in dark anaerobic conditions, the copper site reveals the return of two spherical densities coordinating the copper ion (**Figure 7.7**). Because of the oxycage soaks compromising the reduced states of *Sli*LPMO10E, data collection using the laser to initiate O_2 release from the photocage was not pursued.

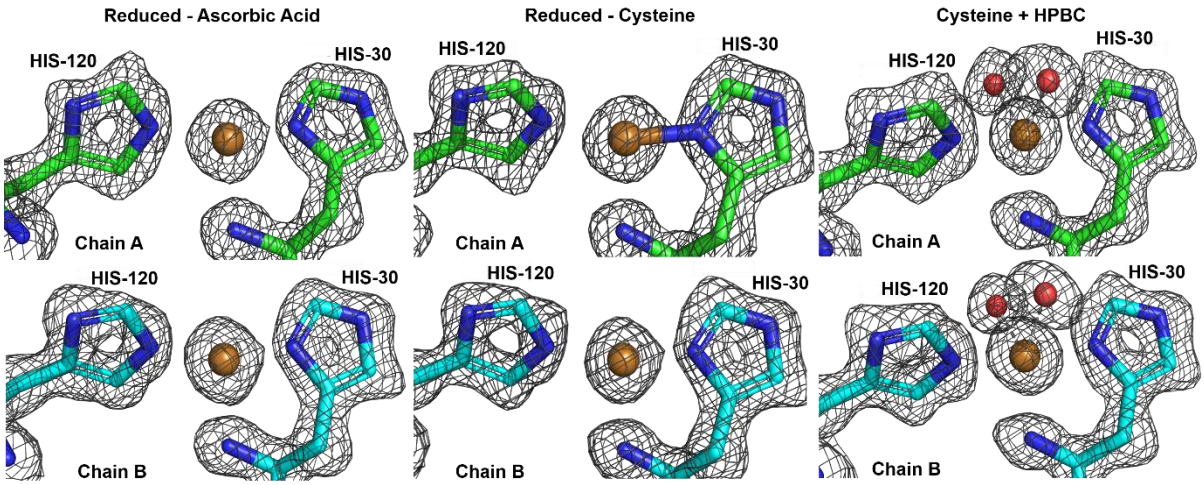


Figure 7.7 histidine-brace site for *Sli*LPMO10E showing both chains from asymmetric unit with a 2Fo-Fc electron density map contoured at 1 σ at the copper active sites. Dark resting state structure, and reduced ascorbic acid/cysteine structures reveal differences at copper site. Orange sphere = copper and red spheres = water.

.

7.4 Discussion

The cryo *Sli*LPMO10E structure (**Figure 7.3**) reveals copper geometry similar to EPR data collected in (Chapter 6, **6.3.4**), exhibiting geometry similar to full length *Sli*LPMO10C and truncated catalytic domain only *Sli*LPMO10G, with copper ion coordinating a single water. The use of rotational methods to obtain cryo *Sli*LPM10E structure indicates beam driven reduction as only one water density is seen at copper site. This feature is well documented when working with metalloproteins, where X-ray dose accumulates causing radiation damage and reduction of the Cu(II) site (Weik and Colletier, 2010; Ravelli and Garman, 2006). Cryo conditions can reduce solvent mobility, with flash-cooling leading to lattice contracting and locking structure into one conformer (Fraser, *et al*., 2011). Contrastingly, single-hit serial crystallography at room temperature used a single dose on each microcrystal, avoiding dose accumulation (Lučić, *et al*., 2022; Lučić, *et al*., 2020). Microcrystals diffract prior to radiation damage, making it possible to capture the *Sli*LPMO10E structure in its native state (two water coordination under aerobic conditions) (**Figure 7.4**).

The observation from the time-resolved experiments indicates how an active Cu(I) state can readily bind O_2 for activation and the form the first oxygenated intermediate. The aerobic RT *Sli*LPMO10E represents a pre-catalytic Cu(II) (**Figure 7.4**), coordinating two H_2O molecules in its resting state. The anaerobic, reduced Cu(I) shows a primed state, ready to bind and activate O_2 to drive the catalytic cycle (Vandhana, *et al*., 2022; Courtade, *et al*., 2020). However, **Figure 7.8** displays the return of resting Cu(II) state after soaking with cys reductant and HPBC oxycage.

An explanation for this is not attributed to O_2 leakage. Small quantities of cysteine mixed with HPBC under the presence of O_2 can lead to the formation of H_2O_2 and drive Cu(I) state back to the resting Cu(II) state (MacArthur, *et al*., 1995). It could also be that in the

presence of Cys, Cu(I)-LPMO and dinuclear Co(III) complex that some uncontrolled redox chemistry is occurring that is oxidising the Cu(I) state of the LPMO back to the Cu(II) form.

Optimisation of methods to obtain O_2 intermediates using serial crystallography would be beneficial to obtain a pre-catalytic Cu(I) state mixed with reductant and HPBC oxycage. Once this is achieved, laser excitation can be implemented in attempts to obtain *Sli*LPMO10E intermediates at the copper site during catalysis. In this respect, further solution studies using spectroscopy and possibly electrochemical measurements will be needed to understand these processes.

The canonical Cu(II) resting state in the *Sli*LPMO10E copper active site with 2 coordinating H_2O ligands can be observed in other resting state LPMOs, such as fungal AA9 LPMO from *Neurospora crassa* (PDB 4EIR) (O'Dell, *et al*., 2017). (See, Chapter 6**, 6.1.2**). Previous studies working on AA10 (*Jd*LPMO10A) with neutron and atomic resolution work, managed to capture densities with elongated structure that is assignable to an O_2 at the copper site (Bacik, *et al*., 2017). Further evidence of the formation of superoxol/hydroperoxo intermediates exist, cryo-trapping techniques on an AA9 (*Nc*LPMO9D) revealed densities at the copper site that are consistent with a mixture of superoxo and hydroperoxo intermediates (Schröder, *et al*., 2021).

Conclusions

In Conclusion, this study has showed the possibility of obtaining a chemically reduced Cu(I) state from *Sli*LPMO10E using single hit serial crystallography at XFEL facilities. Method optimisation to ensure pre-catayltic Cu(I) state of *Sli*LPMO10E with mixed HPBC oxycage, prior to laser excitation would be the next steps for this work.

Chapter 8 Concluding Remarks

This thesis set out to isolate AOA using a combination of liquid and solid media cultivation techniques and perform physiological experiments with a particular focus on the bioavailability of copper, in order to gain further insights into soil dwelling AOA and the role of copper. A total of three AOA enrichments were obtained from agricultural soils (C2, C4-2 and C4-2), with C4-2 undergoing further physiological characterisation. AOA C4-2 was found to grow at concentrations up to 50 mM NH_4^+, but with growth inhibited in the presence of higher levels of NO_2^- (*i.e.* at 50 mM NO_2^-). The optimal growth temperature range for C4-2 was between 30 - 40°C and at pH between pH 6-8. However, both C2 and C4-2 were not axenic and the presence of co-occurring AOB were detected (albeit in low abundance) by a combination of microscopy and CARD-FISH and qPCR methods (Chapter 2, **2.3.2**). If axenic cultures of a new *Nitrosocosmicus* isolate could be obtained, physiology experiments can be carried out without potential interference from co-cultured AOB or other heterotrophic contaminants.

A combination of *amoA* amplicon and whole genome sequencing was performed on the enriched AOA (Chapter 3). The *amoA* amplicons and *amoA*/16 S rRNA gene sequences mined from whole-genome sequencing revealed that AOA C2, C4-1 and C4-2 had high similarity to other high NH_4^+ tolerant AOA (*i.e. N. oleophilus* MY3 and *N. franklandus* (C13)). High quality MAGs were successfully obtained for C2 and C4-2 isolates, making it possible to mine for key AOA genes for carbon fixation (3HP/4HB), NH_4^+ oxidation/transport, ROS defence, urea metabolism and copper handling/binding proteins. MAGs analysis of C2 showed an ANI % of (99.98%) for *N. franklandus* (C13), suggesting C2 is likely related to *N. franklandus* (C13). However, AOA C4-2 showed an ANI % of 97.06% for an unknown *Nitrosocosmicus* species, and may be a potentially novel species within the *Nitrosocosmicus* clade (Chapter 3, **3.3.2**). In this study, genome reconstruction relied on MAGs from enrichment cultures, therefore genome

completeness and genome contamination was influenced by the presence of co-cultured microorganisms. Axenic cultures of C4-2 and C2 would make it possible to obtain a complete closed genome of a potentially novel *Nitrosocosmicus* species, using a combination of long read sequencing (Oxford Nanopore or Pacbio) and short read Illumina sequencing. This would resolve repetitive regions, identify genomic islands and identify plasmids from the complete genome.

This study also compared the tolerance of copper limiting conditions for AOA C4-2 with *N. franklandus* (C13). Whilst C4-2 grew optimally at bioavailable Cu(II) concentrations of 1.19 x 10^{-4} pM, with growth rate reduced at 4.8 x 10^{-5} pM; *N. franklandus* (C13) did not exhibit any change of growth including at the lowest bioavailable Cu(II) concentration (4.8 x 10^{-5} pM). This suggests that AOA C4-2 and *N. franklandus* (C13) occupy distinct copper niches, which was also distinct from *N. viennensis* (EN76) which showed growth at 10^{-16} M bioavailable Cu(II) concentrations (Reyes *et al*. 2020a) (See, Chapter 4, **4.3.1**). The copper experiments performed herein, would have also benefitted from transcriptomic and proteomic analysis under copper-limited conditions of axenic cultures of C4-2, to further characterise copper handling and copper storage genes in relation to NH_4^+ oxidation. This would confirm whether copper niches between C4-2 and *N. franklandus* (C13) reveal differences in copper demand, copper uptake, copper storage and copper-dependant NH_4^+ oxidation.

Attempts to over-express selected (2dMCO) multicopper oxidases from marine AOA (*N. maritimus* (SCM1), Nmar_1131, Nmar_1354, NirK-like Nmar_1667 and (3dMCO) from terrestrial *N. franklandus* (C13) (NFRAN_2427) did not yield soluble proteins for functional structural studies. Issues with no clear expression for Nmar_1354 and very low proteins yield, with inclusion body formation for Nmar_1131, Nmar_1667 and NFRAN_2427 were main challenges during this study. Nonetheless, routes forward can be implemented to extract target protein from inclusion body to obtain higher yields for functional/structural studies. Methods

including periplasmic expression with pET26b or pET22b(+) with pelB leader sequence, co-expression with chaperones, and copper supplementation can be trialled for future work (Chapter 5). If archaeal MCOs and NirK proteins could be over-expressed and purified in soluble form, structural and biochemical characterisation of their role in AOA copper biology could be investigated. Structural studies using X-ray crystallography coupled to EPR spectroscopy would likely unearth rich insight into the bioinorganic chemistry of the copper active site and would form the basis of mechanistic studies. The latter would commence with aerobic assays using hydroxyl amine (NH_2OH) and investigating reaction pathways intermediates and products produced.

Issues from overexpression of AOA MCOs/NirKs steered the project towards working with LPMOs residing in *S. lividans*. Spectroscopic techniques (EPR) and structural studies using SEC-SAXS on multimdular LPMOs (*Sli*LPMO10C, *Sli*LPMO10F and *Sli*LPMO10G) revealed a mixture of Cu(II) coordination geometries. The same techniques were applied to truncated version of *Sli*LPMO10G copper binding-domain, with the addition of X-ray crystallography of *Sli*LPMO10G coper-binding domain using cryo SSX. The *Sli*LPMO10G copper-binding domain reveals a dimer asymmetric unit artifact with the canonical histidine-brace, coordinated by His-31 N from histidine ring, N from His-31 side chain and N from His-106 histidine ring. SEC-SAXS of multimodular *Sli*LPMOs gave further insight on the multimodular architecture of *Sli*LPMO proteins in solution (Chapter 6, **6.3.14**).

Finally, this study explored the structural mechanisms of O_2 activation in the *Sli*LPMO10E copper-binding site. Single-hit serial crystallography at an XFEL (SACLA, Japan) on *Sli*LPMO10E microcrystals showed how a reduced Cu(I) state can be achieved for the first oxygenated intermediate, with the addition of reductant (ascorbic acid/cysteine) in strictly anaerobic conditions. The aims of the study were to implement an oxygen releasing photocage (HPBC) into reduced *Sli*LPMO10E microcrystals to release oxygen under laser

excitation and reveal O_2 intermediates, as structural snapshots. However, upon mixing HPBC photocage with reductant (cysteine), a resting state Cu(II) showing spherical density for 2 coordinating waters was observed due to photocage leakage. Optimisation of methods is required to obtain O_2 intermediates and obtain a pre-catalytic Cu(I) when mixing reductant with HPBC. Once achieved laser excitation can be implemented to obtain *Sli*LPMO10E O_2 intermediates during its catalytic cycle. The same study also showed how radiation damage can be eliminated using SFX techniques at XFEL as microcrystals can diffract prior to radiation damage. SSX technique displays radiation damage with the absence of 2 coordinating water around the copper ion, revealing a Cu(I) state, induced by radiation damage (Chapter 7, **7.3.2**).

Future directions of this work could be to perform additional AOA isolations to obtain axenic cultures using the techniques discussed in Chapter 2 (**2.4**). Further physiological and copper experiments could be carried out on the other AOA enrichments recovered in this study (C2, C4-1) and compared with C4-2. The copper experiment performed on AOA C4-2 could also be repeated using < 1nM Cu concentration, as growth was still observed for both *N. franklandus* (C13) and C4-2 (Chapter 4 **4.3.2**/**4.3.3**). Further genomic analysis could also be performed to carry out a hybrid assembly of long-read sequencing together with short-read sequencing, to obtain a fully closed circular genome and potentially identify other novel *Nitrosocosmicus* species from the AOA enrichments (Chapter 3, **3.3.2**) for further study. Numerous alternative approaches can be employed to obtain soluble MCOs/Nirk from AOA using recombinant overexpression for functional/structural studies, discussed in Chapter 5, **5.4**. SEC-SAXS could be repeated on *Sli*LPMO10C using a different buffer, removing potential aggregation. This could potentially improve agreement at low-q range in Guinier plot and generate higher quality envelope (Chapter 6). Future work for Chapter 7 includes method optimisation using HPBC photocage soaked into reduced *Sli*LPMO10E microcrystals to obtain

pre-catalytic Cu(I), prior to laser excitation. This would make it possible to capture structural snapshots of *Sli*LPMO10E during catalysis.

References

Abiola, C., Gwak, J.-H., Lee, U.-J., Adigun, A. O., & Rhee, S.-K. (2025). Genomic profiling of soil nitrifying microorganisms enriched on floating membrane filter. *Journal of Microbiology*, *63*(4), e2502002.

Abiola, C., Gwak, J.-H., Lee, U.-J., Awala, S. I., Jung, M.-Y., Park, W., & Rhee, S.-K. (2024). Growth of soil ammonia-oxidizing archaea on air-exposed solid surface. *ISME Communications*, *4*(1).

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *630*(8016), 493–500.

Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Baslé, A., Beilsten-Edmands, J., Borges, R. J., Brown, D. G., Burgos-Mármol, J. J., Berrisford, J. M., Bond, P. S., Caballero, I., Catapano, L., Chojnowski, G., Cook, A. G., Cowtan, K. D., Croll, T. I., Debreczeni, J. É., Devenish, N. E., … Yamashita, K. (2023). The *CCP* 4 suite: integrative software for macromolecular crystallography. *Acta Crystallographica Section D Structural Biology*, *79*(6), 449–461.

Ahmed, Y., Rebets, Y., Estévez, M. R., Zapp, J., Myronovskyi, M., & Luzhetskyy, A. (2020). Engineering of Streptomyces lividans for heterologous expression of secondary metabolite gene clusters. *Microbial Cell Factories*, *19*(1), 5.

Aigle, A., Prosser, J. I., & Gubry-Rangin, C. (2019). The application of high-throughput sequencing technology to analysis of amoA phylogeny and environmental niche specialisation of terrestrial bacterial ammonia-oxidisers. *Environmental Microbiome*, *14*(1), 3.

Alves, R. J. E., Kerou, M., Zappe, A., Bittner, R., Abby, S. S., Schmidt, H. A., Pfeifer, K., & Schleper, C. (2019). Ammonia Oxidation by the Arctic Terrestrial Thaumarchaeote Candidatus Nitrosocosmicus arcticus Is Stimulated by Increasing Temperatures. *Frontiers in Microbiology*, *10*.

Alves, R. J. E., Minh, B. Q., Urich, T., von Haeseler, A., & Schleper, C. (2018). Unifying the global phylogeny and environmental distribution of ammonia-oxidising archaea based on amoA genes. *Nature Communications*, *9*(1), 1517.

Amin, S. A., Moffett, J. W., Martens-Habbena, W., Jacquot, J. E., Han, Y., Devol, A., Ingalls, A. E., Stahl, D. A., & Armbrust, E. V. (2013). Copper requirements of the ammonia-oxidizing archaeon *Nitrosopumilus maritimus* SCM1 and implications for nitrification in the marine environment. *Limnology and Oceanography*, *58*(6), 2037–2045.

Anas, M., Liao, F., Verma, K. K., Sarwar, M. A., Mahmood, A., Chen, Z.-L., Li, Q., Zeng, X.-P., Liu, Y., & Li, Y.-R. (2020). Fate of nitrogen in agriculture and environment: agronomic, eco-physiological and molecular approaches to improve nitrogen use efficiency. *Biological Research*, *53*(1), 47.

AnnÃ©, J., & Mellaert, L. (1993). *Streptomyces lividans* as host for heterologous protein production. *FEMS Microbiology Letters*, *114*(2), 121–128.

Armstrong, F. A. J., Stearns, C. R., & Strickland, J. D. H. (1967). The measurement of upwelling and subsequent biological process by means of the Technicon Autoanalyzer® and associated equipment. *Deep Sea Research and Oceanographic Abstracts*, *14*(3), 381–389.

Arp, D., Sayavedra-Soto, L., & Hommes, N. (2002). Molecular biology and biochemistry of ammonia oxidation by Nitrosomonas europaea. *Archives of Microbiology*, *178*(4), 250–255.

Bacik, J.-P., Mekasha, S., Forsberg, Z., Kovalevsky, A. Y., Vaaje-Kolstad, G., Eijsink, V. G. H., Nix, J. C., Coates, L., Cuneo, M. J., Unkefer, C. J., & Chen, J. C.-H. (2017). Neutron and Atomic Resolution X-ray Structures of a Lytic Polysaccharide Monooxygenase Reveal Copper-Mediated Dioxygen Binding and Evidence for N-Terminal Deprotonation. *Biochemistry*, *56*(20), 2529–2532.

Baneyx, F., & Mujacic, M. (2004). Recombinant protein folding and misfolding in Escherichia coli. *Nature Biotechnology*, *22*(11), 1399–1408.

Banning, N. C., Maccarone, L. D., Fisk, L. M., & Murphy, D. v. (2015). Ammonia-oxidising bacteria not archaea dominate nitrification activity in semi-arid agricultural soil. *Scientific Reports*, *5*(1), 11146.

Baskaran, V., Patil, P. K., Antony, M. L., Avunje, S., Nagaraju, V. T., Ghate, S. D., Nathamuni, S., Dineshkumar, N., Alavandi, S. v., & Vijayan, K. K. (2020). Microbial community profiling of ammonia and nitrite oxidizing bacterial enrichments from brackishwater ecosystems for mitigating nitrogen species. *Scientific Reports*, *10*(1), 5201.

Batka, A. E., Thomas, W. C., Tudorica, D. A., Sayler, R. I., & Marletta, M. A. (2024). Second-Sphere Histidine Catalytic Function in a Fungal Polysaccharide Monooxygenase. *Biochemistry*, *63*(23), 3136–3146.

Bayer, B., Hansman, R. L., Bittner, M. J., Noriega-Ortega, B. E., Niggemann, J., Dittmar, T., & Herndl, G. J. (2019). Ammonia-oxidizing archaea release a suite of organic compounds potentially fueling prokaryotic heterotrophy in the ocean. *Environmental Microbiology*, *21*(11), 4062–4075.

Bayer, B., McBeain, K., Carlson, C. A., & Santoro, A. E. (2023). Carbon content, carbon fixation yield and dissolved organic carbon release from diverse marine nitrifiers. *Limnology and Oceanography*, *68*(1), 84–96.

Beaudor, M., Vuichard, N., Lathière, J., Evangeliou, N., van Damme, M., Clarisse, L., & Hauglustaine, D. (2023). Global agricultural ammonia emissions simulated with the ORCHIDEE land surface model. *Geoscientific Model Development*, *16*(3), 1053–1081.

Beeckman, F., Motte, H., & Beeckman, T. (2018). Nitrification in agricultural soils: impact, actors and mitigation. *Current Opinion in Biotechnology*, *50*, 166–173.

Beeson, W. T., Vu, V. v., Span, E. A., Phillips, C. M., & Marletta, M. A. (2015). Cellulose Degradation by Polysaccharide Monooxygenases. *Annual Review of Biochemistry*, *84*(1), 923–946.

Bei, Q., Reitz, T., Schädler, M., Hodgskiss, L. H., Peng, J., Schnabel, B., Buscot, F., Eisenhauer, N., Schleper, C., & Heintz-Buschart, A. (2024). Metabolic potential of *Nitrososphaera* -associated clades. *The ISME Journal*, *18*(1).

Bengtsson, M. M., & Øvreås, L. (2010). Planctomycetes dominate biofilms on surfaces of the kelp Laminaria hyperborea. *BMC Microbiology*, *10*(1), 261.

Bernadó, P., Mylonas, E., Petoukhov, M. v., Blackledge, M., & Svergun, D. I. (2007). Structural Characterization of Flexible Proteins Using Small-Angle X-ray Scattering. *Journal of the American Chemical Society*, *129*(17), 5656–5664.

Bhatwa, A., Wang, W., Hassan, Y. I., Abraham, N., Li, X.-Z., & Zhou, T. (2021). Challenges Associated With the Formation of Recombinant Protein Inclusion Bodies in Escherichia coli and Strategies to Address Them for Industrial Applications. *Frontiers in Bioengineering and Biotechnology*, *9*.

Bissaro, B., & Eijsink, V. G. H. (2023). Lytic polysaccharide monooxygenases: enzymes for controlled and site-specific Fenton-like chemistry. *Essays in Biochemistry*, *67*(3), 575–584.

Blackwell, J. R., & Horgan, R. (1991). A novel strategy for production of a highly expressed recombinant protein in an active form. *FEBS Letters*, *295*(1–3), 10–12.

Blum, M., Andreeva, A., Florentino, L. C., Chuguransky, S. R., Grego, T., Hobbs, E., Pinto, B. L., Orr, A., Paysan-Lafosse, T., Ponamareva, I., Salazar, G. A., Bordin, N., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunic, I., Llinares-López, F., … Bateman, A. (2025). InterPro: the protein sequence classification resource in 2025. *Nucleic Acids Research*, *53*(D1), D444–D456.

Bollmann, A., Bullerjahn, G. S., & McKay, R. M. (2014). Abundance and Diversity of Ammonia-Oxidizing Archaea and Bacteria in Sediments of Trophic End Members of the Laurentian Great Lakes, Erie and Superior. *PLoS ONE*, *9*(5), e97068.

Bollmann, A., French, E., & Laanbroek, H. J. (2011). *Isolation, Cultivation, and Characterization of Ammonia-Oxidizing Bacteria and Archaea Adapted to Low Ammonium Concentrations* (pp. 55–88).

Boraston, A. B., Bolam, D. N., Gilbert, H. J., & Davies, G. J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochemical Journal*, *382*(3), 769–781.

Bouvier, T., & del Giorgio, P. A. (2003). Factors influencing the detection of bacterial cells using fluorescence in situ hybridization (FISH): A quantitative review of published reports. *FEMS Microbiology Ecology*, *44*(1), 3–15.

Burger, V. M., Arenas, D. J., & Stultz, C. M. (2016). A Structure-free Method for Quantifying Conformational Flexibility in proteins. *Scientific Reports*, *6*(1), 29040.

Bushmarina, N. A., Blanchet, C. E., Vernier, G., & Forge, V. (2006). Cofactor effects on the protein folding reaction: Acceleration of α-lactalbumin refolding by metal ions. *Protein Science*, *15*(4), 659–671.

Cabezas, M. P., Fonseca, N. A., & Muñoz-Mérida, A. (2024). MIMt: a curated 16S rRNA reference database with less redundancy and higher accuracy at species-level identification. *Environmental Microbiome*, *19*(1), 88.

Calderaro, F., Bevers, L. E., & van den Berg, M. A. (2021). Oxidative Power: Tools for Assessing LPMO Activity on Cellulose. *Biomolecules*, *11*(8), 1098.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research*, *37*(Database), D233–D238.

Caramello, N., & Royant, A. (2024). From femtoseconds to minutes: time-resolved macromolecular crystallography at XFELs and synchrotrons. *Acta Crystallographica Section D Structural Biology*, *80*(2), 60–79.

Ceballos-Alcantarilla, E., & Merkx, M. (2021). *Understanding and applications of Ser/Gly linkers in protein engineering* (pp. 1–22).

Ciano, L., Paradisi, A., Hemsworth, G. R., Tovborg, M., Davies, G. J., & Walton, P. H. (2020). Insights from semi-oriented EPR spectroscopy studies into the interaction of lytic polysaccharide monooxygenases with cellulose. *Dalton Transactions*, *49*(11), 3413–3422.

Chaplin, A. K., Wilson, M. T., Hough, M. A., Svistunenko, D. A., Hemsworth, G. R., Walton, P. H., Vijgenboom, E., & Worrall, J. A. R. (2016). Heterogeneity in the Histidine-brace Copper Coordination Sphere in Auxiliary Activity Family 10 (AA10) Lytic Polysaccharide Monooxygenases. *Journal of Biological Chemistry*, *291*(24), 12838–12850.

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2022). GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics*, *38*(23), 5315–5316.

Che, J., Zhao, X. Q., Zhou, X., Jia, Z. J., & Shen, R. F. (2015). High pH-enhanced soil nitrification was associated with ammonia-oxidizing bacteria rather than archaea in acidic soils. *Applied Soil Ecology*, *85*, 21–29.

Chen, R. (2012). Bacterial expression systems for recombinant protein production: E. coli and beyond. *Biotechnology Advances*, *30*(5), 1102–1107.

Chen, S. (2023). Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta*, *2*(2).

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890.

Chen, X., Zaro, J. L., & Shen, W.-C. (2013). Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews*, *65*(10), 1357–1369.

Chicano, T. M., Dietrich, L., de Almeida, N. M., Akram, Mohd., Hartmann, E., Leidreiter, F., Leopoldus, D., Mueller, M., Sánchez, R., Nuijten, G. H. L., Reimann, J., Seifert, K.-A., Schlichting, I., van Niftrik, L., Jetten, M. S. M., Dietl, A., Kartal, B., Parey, K., & Barends, T. R. M. (2021). Structural and functional characterization of the intracellular filament-forming nitrite oxidoreductase multiprotein complex. *Nature Microbiology*, *6*(9), 1129–1139.

Cholet, F., Agogué, H., Ijaz, U. Z., Lachaussée, N., Pineau, P., & Smith, C. J. (2024). Low-abundant but highly transcriptionally active uncharacterised Nitrosomonas drive ammonia-oxidation in the Brouage mudflat, France. *Science of The Total Environment*, *946*, 174312.

Clark, D. R., McKew, B. A., Dong, L. F., Leung, G., Dumbrell, A. J., Stott, A., Grant, H., Nedwell, D. B., Trimmer, M., & Whitby, C. (2020). Mineralization and nitrification: Archaea dominate ammonia-oxidising communities in grassland soils. *Soil Biology and Biochemistry*, *143*, 107725.

Courtade, G., Ciano, L., Paradisi, A., Lindley, P. J., Forsberg, Z., Sørlie, M., Wimmer, R., Davies, G. J., Eijsink, V. G. H., Walton, P. H., & Aachmann, F. L. (2020). Mechanistic basis of substrate–O $_2$ coupling within a chitin-active lytic polysaccharide monooxygenase: An integrated NMR/EPR study. *Proceedings of the National Academy of Sciences*, *117*(32), 19178–19189.

Courtade, G., Forsberg, Z., Heggset, E. B., Eijsink, V. G. H., & Aachmann, F. L. (2018). The carbohydrate-binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose oxidation. *Journal of Biological Chemistry*, *293*(34), 13006–13015.

Cowieson, N. P., Edwards-Gayle, C. J. C., Inoue, K., Khunti, N. S., Doutch, J., Williams, E., Daniels, S., Preece, G., Krumpa, N. A., Sutter, J. P., Tully, M. D., Terrill, N. J., & Rambo, R. P. (2020). Beamline B21: high-throughput small-angle X-ray scattering at Diamond Light Source. *Journal of Synchrotron Radiation*, *27*(5), 1438–1446.

Daims, H., Brühl, A., Amann, R., Schleifer, K.-H., & Wagner, M. (1999). The Domain-specific Probe EUB338 is Insufficient for the Detection of all Bacteria: Development and Evaluation of a more Comprehensive Probe Set. *Systematic and Applied Microbiology*, *22*(3), 434–444.

Daims, H., Lebedeva, E. v., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A., Kirkegaard, R. H., von Bergen, M., Rattei, T., Bendinger, B., Nielsen, P. H., & Wagner, M. (2015). Complete nitrification by Nitrospira bacteria. *Nature*, *528*(7583), 504–509.

Daims, H., Lücker, S., & Wagner, M. (2016). A New Perspective on Microbes Formerly Known as Nitrite-Oxidizing Bacteria. *Trends in Microbiology*, *24*(9), 699–712.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2).

de La Torre, J. R., Walker, C. B., Ingalls, A. E., Könneke, M., & Stahl, D. A. (2008). Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environmental Microbiology*, *10*(3), 810–818.

de Marco, A., Deuerling, E., Mogk, A., Tomoyasu, T., & Bukau, B. (2007). Chaperone-based procedure to increase yields of soluble recombinant proteins produced in E. coli. *BMC Biotechnology*, *7*(1), 32.

de Vries, W. (2021). Impacts of nitrogen emissions on ecosystems and human health: A mini review. *Current Opinion in Environmental Science & Health*, *21*, 100249.

de Wijn, R., Melo, D. V. M., Koua, F. H. M., & Mancuso, A. P. (2022). Potential of Time-Resolved Serial Femtosecond Crystallography Using High Repetition Rate XFEL Sources. *Applied Sciences*, *12*(5), 2551.

DeLong, E. F. (1992). Archaea in coastal marine environments. *Proceedings of the National Academy of Sciences*, *89*(12), 5685–5689.

DeLong, E. F. (1998). Everything in moderation: Archaea as 'non-extremophiles.' *Current Opinion in Genetics & Development*, *8*(6), 649–654.

Destan, E., Yuksel, B., Tolar, B. B., Ayan, E., Deutsch, S., Yoshikuni, Y., Wakatsuki, S., Francis, C. A., & DeMirci, H. (2021). Structural insights into bifunctional thaumarchaeal crotonyl-CoA hydratase and 3-hydroxypropionyl-CoA dehydratase from Nitrosopumilus maritimus. *Scientific Reports*, *11*(1), 22849.

Di, H. J., Cameron, K. C., Shen, J.-P., Winefield, C. S., O'Callaghan, M., Bowatte, S., & He, J.-Z. (2010). Ammonia-oxidizing bacteria and archaea grow under contrasting soil nitrogen conditions. *FEMS Microbiology Ecology*, *72*(3), 386–394.

Diamant, S., Eliahu, N., Rosenthal, D., & Goloubinoff, P. (2001). Chemical Chaperones Regulate Molecular Chaperones in Vitro and in Cells under Combined Salt and Heat Stresses. *Journal of Biological Chemistry*, *276*(43), 39586–39591.

Diamond, S., Lavy, A., Crits-Christoph, A., Matheus Carnevali, P. B., Sharrar, A., Williams, K. H., & Banfield, J. F. (2022). Soils and sediments host Thermoplasmata archaea encoding novel copper membrane monooxygenases (CuMMOs). *The ISME Journal*, *16*(5), 1348–1362.

Djoko, K. Y., Chong, L. X., Wedd, A. G., & Xiao, Z. (2010). Reaction Mechanisms of the Multicopper Oxidase CueO from *Escherichia coli* Support Its Functional Role as a Cuprous Oxidase. *Journal of the American Chemical Society*, *132*(6), 2005–2015.

Donovan, R. S., Robinson, C. W., & Glick, B. R. (1996). Review: Optimizing inducer and culture conditions for expression of foreign proteins under the control of thelac promoter. *Journal of Industrial Microbiology*, *16*(3), 145–154.

Drewniak, B. A., Mishra, U., Song, J., Prell, J., & Kotamarthi, V. R. (2015). Modeling the impact of agricultural land use and management on US carbon budgets. *Biogeosciences*, *12*(7), 2119–2129.

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., & Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Research*, *50*(D1), D571–D577.

Dyson, M. R., Shadbolt, S. P., Vincent, K. J., Perera, R. L., & McCafferty, J. (2004). Production of soluble mammalian proteins in Escherichia coli: identification of protein features that correlate with successful expression. *BMC Biotechnology*, *4*(1), 32.

Eady, R. R., & Samar Hasnain, S. (2022). New horizons in structure-function studies of copper nitrite reductase. *Coordination Chemistry Reviews*, *460*, 214463.

Ebeling, W., Hennrich, N., Klockow, M., Metz, H., Orth, H. D., & Lang, H. (1974). Proteinase K from *Tritirachium album* Limber. *European Journal of Biochemistry*, *47*(1), 91–97.

Eijsink, V. G. H., Petrovic, D., Forsberg, Z., Mekasha, S., Røhr, Å. K., Várnai, A., Bissaro, B., & Vaaje-Kolstad, G. (2019). On the functional characterization of lytic polysaccharide monooxygenases (LPMOs). *Biotechnology for Biofuels*, *12*(1), 58.

Elling, F. J., Könneke, M., Mußmann, M., Greve, A., & Hinrichs, K.-U. (2015). Influence of temperature, pH, and salinity on membrane lipid composition and TEX86 of marine planktonic thaumarchaeal isolates. *Geochimica et Cosmochimica Acta*, *171*, 238–255.

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of *Coot*. *Acta Crystallographica Section D Biological Crystallography*, *66*(4), 486–501.

Erguder, T. H., Boon, N., Wittebolle, L., Marzorati, M., & Verstraete, W. (2009). Environmental factors shaping the ecological niches of ammonia-oxidizing archaea. *FEMS Microbiology Reviews*, *33*(5), 855–869.

Erisman, J. W., Galloway, J. N., Seitzinger, S., Bleeker, A., Dise, N. B., Petrescu, A. M. R., Leach, A. M., & de Vries, W. (2013). Consequences of human modification of the global

nitrogen cycle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1621), 20130116.

Evans, P. R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D Biological Crystallography*, *67*(4), 282–292.

Feng, H., Guo, J., Peng, C., Kneeshaw, D., Roberge, G., Pan, C., Ma, X., Zhou, D., & Wang, W. (2023). Nitrogen addition promotes terrestrial plants to allocate more biomass to aboveground organs: A global meta-analysis. *Global Change Biology*, *29*(14), 3970–3989.

Finn, R. D., Clements, J., & Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research*, *39*(suppl), W29–W37.

Forsberg, Z., & Courtade, G. (2023). On the impact of carbohydrate-binding modules (CBMs) in lytic polysaccharide monooxygenases (LPMOs). *Essays in Biochemistry*, *67*(3), 561–574.

Forsberg, Z., Nelson, C. E., Dalhus, B., Mekasha, S., Loose, J. S. M., Crouch, L. I., Røhr, Å. K., Gardner, J. G., Eijsink, V. G. H., & Vaaje-Kolstad, G. (2016). Structural and Functional Analysis of a Lytic Polysaccharide Monooxygenase Important for Efficient Utilization of Chitin in Cellvibrio japonicus. *Journal of Biological Chemistry*, *291*(14), 7300–7312.

Forsberg, Z., Sørlie, M., Petrović, D., Courtade, G., Aachmann, F. L., Vaaje-Kolstad, G., Bissaro, B., Røhr, Å. K., & Eijsink, V. G. (2019). Polysaccharide degradation by lytic polysaccharide monooxygenases. *Current Opinion in Structural Biology*, *59*, 54–64.

Forsberg, Z., Stepnov, A. A., Tesei, G., Wang, Y., Buchinger, E., Kristiansen, S. K., Aachmann, F. L., Arleth, L., Eijsink, V. G. H., Lindorff-Larsen, K., & Courtade, G. (2023). The effect of linker conformation on performance and stability of a two-domain lytic polysaccharide monooxygenase. *Journal of Biological Chemistry*, *299*(11), 105262.

Fowler, Claire. A., Sabbadin, F., Ciano, L., Hemsworth, G. R., Elias, L., Bruce, N., McQueen-Mason, S., Davies, G. J., & Walton, P. H. (2019). Discovery, activity and characterisation of an AA10 lytic polysaccharide oxygenase from the shipworm symbiont Teredinibacter turnerae. *Biotechnology for Biofuels*, *12*(1), 232.

Fowler, D., Coyle, M., Skiba, U., Sutton, M. A., Cape, J. N., Reis, S., Sheppard, L. J., Jenkins, A., Grizzetti, B., Galloway, J. N., Vitousek, P., Leach, A., Bouwman, A. F., Butterbach-Bahl, K., Dentener, F., Stevenson, D., Amann, M., & Voss, M. (2013). The global nitrogen cycle in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1621), 20130164.

Frandsen, K. E. H., & lo Leggio, L. (2016). Lytic polysaccharide monooxygenases: a crystallographer's view on a new class of biomass-degrading enzymes. *IUCrJ*, *3*(6), 448–467.

Franke, D., & Svergun, D. I. (2009). *DAMMIF* , a program for rapid *ab-initio* shape determination in small-angle scattering. *Journal of Applied Crystallography*, *42*(2), 342–346.

Fraser, J. S., van den Bedem, H., Samelson, A. J., Lang, P. T., Holton, J. M., Echols, N., & Alber, T. (2011). Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proceedings of the National Academy of Sciences*, *108*(39), 16247–16252.

French, E., Kozlowski, J. A., & Bollmann, A. (2021). Competition between Ammonia-Oxidizing Archaea and Bacteria from Freshwater Environments. *Applied and Environmental Microbiology*, *87*(20).

French, E., Kozlowski, J. A., Mukherjee, M., Bullerjahn, G., & Bollmann, A. (2012). Ecophysiological Characterization of Ammonia-Oxidizing Archaea and Bacteria from Freshwater. *Applied and Environmental Microbiology*, *78*(16), 5773–5780.

Fu, C., Tu, C., Zhang, H., Li, Y., Li, L., Zhou, Q., Scheckel, K. G., & Luo, Y. (2020). Soil accumulation and chemical fractions of Cu in a large and long-term coastal apple orchard, North China. *Journal of Soils and Sediments*, *20*(10), 3712–3721.

Fuhrmann, M., & Fitts, J. P. (2004). Adsorption of Trace Metals on Glass Fiber Filters. *Journal of Environmental Quality*, *33*(5), 1943–1944.

Fukuda, Y., Tse, K. M., Nakane, T., Nakatsu, T., Suzuki, M., Sugahara, M., Inoue, S., Masuda, T., Yumoto, F., Matsugaki, N., Nango, E., Tono, K., Joti, Y., Kameshima, T., Song, C., Hatsui, T., Yabashi, M., Nureki, O., Murphy, M. E. P., … Mizohata, E. (2016). Redox-coupled proton transfer mechanism in nitrite reductase revealed by femtosecond crystallography. *Proceedings of the National Academy of Sciences*, *113*(11), 2928–2933.

Galloway, J. N., & Cowling, E. B. (2021). Reflections on 200 years of Nitrogen, 20 years later. *Ambio*, *50*(4), 745–749.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook* (pp. 571–607). Humana Press.

Geisseler, D., Smith, R., Cahn, M., & Muramoto, J. (2021). Nitrogen mineralization from organic fertilizers and composts: Literature survey and model fitting. *Journal of Environmental Quality*, *50*(6), 1325–1338.

Ghimire-Kafle, S., Weaver, M. E., & Bollmann, A. (2023). Ecophysiological and Genomic Characterization of the Freshwater Complete Ammonia Oxidizer *Nitrospira* sp. Strain BO4. *Applied and Environmental Microbiology*, *89*(2).

Giguere, A. T., Taylor, A. E., Myrold, D. D., Mellbye, B. L., Sayavedra-Soto, L. A., & Bottomley, P. J. (2018). Nitrite-oxidizing activity responds to nitrite accumulation in soil. *FEMS Microbiology Ecology*, *94*(3).

Gilbert, H. J., Knox, J. P., & Boraston, A. B. (2013). Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Current Opinion in Structural Biology*, *23*(5), 669–677.

Goh, C.-S., Lan, N., Douglas, S. M., Wu, B., Echols, N., Smith, A., Milburn, D., Montelione, G. T., Zhao, H., & Gerstein, M. (2004). Mining the Structural Genomics Pipeline: Identification of Protein Properties that Affect High-throughput Experimental Analysis. *Journal of Molecular Biology*, *336*(1), 115–130.

Gómez-Piñeiro, R. J., Drosou, M., Bertaina, S., Decroos, C., Simaan, A. J., Pantazis, D. A., & Orio, M. (2022). Decoding the Ambiguous Electron Paramagnetic Resonance Signals in the Lytic Polysaccharide Monooxygenase from *Photorhabdus luminescens*. *Inorganic Chemistry*, *61*(20), 8022–8035.

Gong, C., Tian, H., Liao, H., Pan, N., Pan, S., Ito, A., Jain, A. K., Kou-Giesbrecht, S., Joos, F., Sun, Q., Shi, H., Vuichard, N., Zhu, Q., Peng, C., Maggi, F., Tang, F. H. M., & Zaehle, S. (2024). Global net climate effects of anthropogenic reactive nitrogen. *Nature*, *632*(8025), 557–563.

González-Guerrero, M., & Argüello, J. M. (2008). Mechanism of Cu^{+}-transporting ATPases: Soluble Cu^{+} chaperones directly transfer Cu^{+} to transmembrane transport sites. *Proceedings of the National Academy of Sciences*, *105*(16), 5992–5997.

Gopalasingam, C. C., Johnson, R. M., Chiduza, G. N., Tosha, T., Yamamoto, M., Shiro, Y., Antonyuk, S. v., Muench, S. P., & Hasnain, S. S. (2019). Dimeric structures of quinol-dependent nitric oxide reductases (qNORs) revealed by cryo–electron microscopy. *Science Advances*, *5*(8).

Gorman-Lewis, D., Martens-Habbena, W., & Stahl, D. A. (2019). Cu(II) adsorption onto ammonia-oxidizing bacteria and archaea. *Geochimica et Cosmochimica Acta*, *255*, 127–143.

Goulding, C. W., & Jeanne Perry, L. (2003). Protein production in Escherichia coli for structural studies by X-ray crystallography. *Journal of Structural Biology*, *142*(1), 133–143.

Graewert, M. A., da Vela, S., Gräwert, T. W., Molodenskiy, D. S., Blanchet, C. E., Svergun, D. I., & Jeffries, C. M. (2020). Adding Size Exclusion Chromatography (SEC) and Light

Scattering (LS) Devices to Obtain High-Quality Small Angle X-Ray Scattering (SAXS) Data. *Crystals*, *10*(11), 975.

Gräff, M., Buchholz, P. C. F., le Roes-Hill, M., & Pleiss, J. (2020). Multicopper oxidases: modular structure, sequence space, and evolutionary relationships. *Proteins: Structure, Function, and Bioinformatics*, *88*(10), 1329–1339.

Grant, J. R., Enns, E., Marinier, E., Mandal, A., Herman, E. K., Chen, C., Graham, M., Van Domselaar, G., & Stothard, P. (2023). Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Research*, *51*(W1), W484–W492.

Gräwert, T. W., & Svergun, D. I. (2020). Structural Modeling Using Solution Small-Angle X-ray Scattering (SAXS). *Journal of Molecular Biology*, *432*(9), 3078–3092.

Grossman, T. H., Kawasaki, E. S., Punreddy, S. R., & Osburne, M. S. (1998). Spontaneous cAMP-dependent derepression of gene expression in stationary phase plays a role in recombinant expression instability. *Gene*, *209*(1–2), 95–103.

Gubry-Rangin, C., Hai, B., Quince, C., Engel, M., Thomson, B. C., James, P., Schloter, M., Griffiths, R. I., Prosser, J. I., & Nicol, G. W. (2011). Niche specialization of terrestrial archaeal ammonia oxidizers. *Proceedings of the National Academy of Sciences*, *108*(52), 21206–21211.

Gubry-Rangin, C., Nicol, G. W., & Prosser, J. I. (2010). Archaea rather than bacteria control nitrification in two agricultural acidic soils. *FEMS Microbiology Ecology*, *74*(3), 566–574.

Gubry-Rangin, C., Novotnik, B., Mandič-Mulec, I., Nicol, G. W., & Prosser, J. I. (2017). Temperature responses of soil ammonia-oxidising archaea depend on pH. *Soil Biology and Biochemistry*, *106*, 61–68.

Gubry-Rangin, C., Williams, W., & Prosser, J. I. (2018). Approaches to understanding the ecology and evolution of understudied terrestrial archaeal ammonia-oxidisers. *Emerging Topics in Life Sciences*, *2*(4), 619–628.

Guo, C., Liu, X., & He, X. (2022). A global meta-analysis of crop yield and agricultural greenhouse gas emissions under nitrogen fertilizer application. *Science of The Total Environment*, *831*, 154982.

Gwak, J.-H., Jung, M.-Y., Hong, H., Kim, J.-G., Quan, Z.-X., Reinfelder, J. R., Spasov, E., Neufeld, J. D., Wagner, M., & Rhee, S.-K. (2020). Archaeal nitrification is constrained by copper complexation with organic matter in municipal wastewater treatment plants. *The ISME Journal*, *14*(2), 335–346.

Hagemann, M. M., & Hedegård, E. D. (2023). Molecular Mechanism of Substrate Oxidation in Lytic Polysaccharide Monooxygenases: Insight from Theoretical Investigations. *Chemistry – A European Journal*, *29*(7).

Hahn-Hägerdal, B., Galbe, M., Gorwa-Grauslund, M. F., Lidén, G., & Zacchi, G. (2006). Bio-ethanol – the fuel of tomorrow from the residues of today. *Trends in Biotechnology*, *24*(12), 549–556.

Hall, K. R., Elisa Rønnekleiv, S., Gautieri, A., Lilleås, H., Skaali, R., Rieder, L., Nikoline Englund, A., Landsem, E., Emrich-Mills, T. Z., Ayuso-Fernández, I., Kjendseth Røhr, Å., Sørlie, M., & Eijsink, V. G. H. (2025). Structure–Function Analysis of an Understudied Type of LPMO with Unique Redox Properties and Substrate Specificity. *ACS Catalysis*, *15*(12), 10601–10617.

Han, S., Kim, S., Sedlacek, C. J., Farooq, A., Song, C., Lee, S., Liu, S., Brüggemann, N., Rohe, L., Kwon, M., Rhee, S.-K., & Jung, M.-Y. (2024). Adaptive traits of *Nitrosocosmicus* clade ammonia-oxidizing archaea. *mBio*, *15*(11).

Handing, K. B., Niedzialkowska, E., Shabalin, I. G., Kuhn, M. L., Zheng, H., & Minor, W. (2018). Characterizing metal-binding sites in proteins with X-ray crystallography. *Nature Protocols*, *13*(5), 1062–1090.

Harhangi, H. R., le Roy, M., van Alen, T., Hu, B., Groen, J., Kartal, B., Tringe, S. G., Quan, Z.-X., Jetten, M. S. M., & Op den Camp, H. J. M. (2012). Hydrazine Synthase, a Unique

Phylomarker with Which To Study the Presence and Biodiversity of Anammox Bacteria. *Applied and Environmental Microbiology*, *78*(3), 752–758.

Hasan, A. A., Kalinina, E., Tatarskiy, V., & Shtil, A. (2022). The Thioredoxin System of Mammalian Cells and Its Modulators. *Biomedicines*, *10*(7), 1757.

Hatzenpichler, R. (2012). Diversity, Physiology, and Niche Differentiation of Ammonia-Oxidizing Archaea. *Applied and Environmental Microbiology*, *78*(21), 7501–7510.

Hatzenpichler, R., Lebedeva, E. v., Spieck, E., Stoecker, K., Richter, A., Daims, H., & Wagner, M. (2008). A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proceedings of the National Academy of Sciences*, *105*(6), 2134–2139.

Hayat, S. M. G., Farahani, N., Golichenari, B., & Sahebkar, A. (2018). Recombinant Protein Expression in Escherichia coli (E.coli): What We Need to Know. *Current Pharmaceutical Design*, *24*(6), 718–725.

Hayatsu, M., Tago, K., Uchiyama, I., Toyoda, A., Wang, Y., Shimomura, Y., Okubo, T., Kurisu, F., Hirono, Y., Nonaka, K., Akiyama, H., Itoh, T., & Takami, H. (2017). An acid-tolerant ammonia-oxidizing γ-proteobacterium from soil. *The ISME Journal*, *11*(5), 1130–1141.

Hedegård, E. D., & Ryde, U. (2018). Molecular mechanism of lytic polysaccharide monooxygenases. *Chemical Science*, *9*(15), 3866–3880.

Hemsworth, G. R. (2023). Revisiting the role of electron donors in lytic polysaccharide monooxygenase biochemistry. *Essays in Biochemistry*, *67*(3), 585–595.

Hemsworth, G. R., Déjean, G., Davies, G. J., & Brumer, H. (2016). Learning from microbial strategies for polysaccharide degradation. *Biochemical Society Transactions*, *44*(1), 94–108.

Heylen, K., Vanparys, B., Gevers, D., Wittebolle, L., Boon, N., & de Vos, P. (2007). Nitric oxide reductase (*norB*) gene sequence analysis reveals discrepancies with nitrite reductase (*nir*) gene phylogeny in cultivated denitrifiers. *Environmental Microbiology*, *9*(4), 1072–1077.

Higasi, P. M. R., Velasco, J. A., Pellegrini, V. O. A., de Araújo, E. A., França, B. A., Keller, M. B., Labate, C. A., Blossom, B. M., Segato, F., & Polikarpov, I. (2021). Light-stimulated T. thermophilus two-domain LPMO9H: Low-resolution SAXS model and synergy with cellulases. *Carbohydrate Polymers*, *260*, 117814.

Hiis, E. G., Vick, S. H. W., Molstad, L., Røsdal, K., Jonassen, K. R., Winiwarter, W., & Bakken, L. R. (2024). Unlocking bacterial potential to reduce farmland N2O emissions. *Nature*, *630*(8016), 421–428.

Himmel, M. E., Ding, S.-Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., & Foust, T. D. (2007). Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science*, *315*(5813), 804–807.

Hink, L., Gubry-Rangin, C., Nicol, G. W., & Prosser, J. I. (2018). The consequences of niche and physiological differentiation of archaeal and bacterial ammonia oxidisers for nitrous oxide emissions. *The ISME Journal*, *12*(4), 1084–1093.

Hink, L., Nicol, G. W., & Prosser, J. I. (2017). Archaea produce lower yields of N $_2$ O than bacteria during aerobic ammonia oxidation in soil. *Environmental Microbiology*, *19*(12), 4829–4837.

Hird, K., Campeciño, J. O., & Hegg, E. L. (2025). From genes to function: regulation, maturation, and evolution of cytochrome *c* nitrite reductase in nitrate reduction to ammonium. *Applied and Environmental Microbiology*, *91*(7).

Hjelm, A., Karyolaimos, A., Zhang, Z., Rujas, E., Vikström, D., Slotboom, D. J., & de Gier, J.-W. (2017). Tailoring *Escherichia coli* for the <scp>l</scp> -Rhamnose P $_{BAD}$ Promoter-Based Production of Membrane and Secretory Proteins. *ACS Synthetic Biology*, *6*(6), 985–994.

Hoang, H. G., Thuy, B. T. P., Lin, C., Vo, D.-V. N., Tran, H. T., Bahari, M. B., Le, V. G., & Vu, C. T. (2022). The nitrogen cycle and mitigation strategies for nitrogen loss during organic waste composting: A review. *Chemosphere*, *300*, 134514.

Hodgskiss, L. H., Melcher, M., Kerou, M., Chen, W., Ponce-Toledo, R. I., Savvides, S. N., Wienkoop, S., Hartl, M., & Schleper, C. (2023). Unexpected complexity of the ammonia monooxygenase in archaea. *The ISME Journal*, *17*(4), 588–599.

Horrell, S., Kekilli, D., Strange, R. W., & Hough, M. A. (2017). Recent structural insights into the function of copper nitrite reductases. *Metallomics*, *9*(11), 1470–1482.

Hoshino, T., Yilmaz, L. S., Noguera, D. R., Daims, H., & Wagner, M. (2008). Quantification of Target Molecules Needed To Detect Microorganisms by Fluorescence In Situ Hybridization (FISH) and Catalyzed Reporter Deposition-FISH. *Applied and Environmental Microbiology*, *74*(16), 5068–5077.

Hossain, S., Chow, C. W. K., Cook, D., Sawade, E., & Hewa, G. A. (2022). Review of Nitrification Monitoring and Control Strategies in Drinking Water System. *International Journal of Environmental Research and Public Health*, *19*(7), 4003.

Hu, L., Dong, Z., Wang, Z., Xiao, L., & Zhu, B. (2022). The contributions of ammonia oxidizing bacteria and archaea to nitrification-dependent N2O emission in alkaline and

neutral purple soils. *Scientific Reports*, *12*(1), 19928. https://doi.org/10.1038/s41598-022-23084-1

Hu, L., Wang, Z., Wang, Z., Wang, L., Fang, J., & Liu, R. (2024). Community Composition and Functional Characterization of Microorganisms in Surface Sediment of the New Britain Trench. *Current Microbiology*, *81*(9), 282.

Hu, Y., Wang, Y., Wang, R., Wang, X., & Liu, S.-J. (2024). Dirammox-dominated microbial community for biological nitrogen removal from wastewater. *Applied Microbiology and Biotechnology*, *108*(1), 389.

Huang, L., Chakrabarti, S., Cooper, J., Perez, A., John, S. M., Daroub, S. H., & Martens-Habbena, W. (2021). Ammonia-oxidizing archaea are integral to nitrogen cycling in a highly fertile agricultural soil. *ISME Communications*, *1*(1).

Huang, Z., Jiang, Y., Song, X., Hallerman, E., Peng, L., Dong, D., Ma, T., Zhai, J., & Li, W. (2018). Ammonia-oxidizing bacteria and archaea within biofilters of a commercial recirculating marine aquaculture system. *AMB Express*, *8*(1), 17.

Hunter, M. S., Segelke, B., Messerschmidt, M., Williams, G. J., Zatsepin, N. A., Barty, A., Benner, W. H., Carlson, D. B., Coleman, M., Graf, A., Hau-Riege, S. P., Pardini, T., Seibert,

M. M., Evans, J., Boutet, S., & Frank, M. (2014). Fixed-target protein serial microcrystallography with an x-ray free electron laser. *Scientific Reports*, *4*(1), 6026.

Hutchins, D. A., & Capone, D. G. (2022). The marine nitrogen cycle: new developments and global change. *Nature Reviews Microbiology*, *20*(7), 401–414.

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*(1), 119.

Ipsen, J. Ø., Hallas-Møller, M., Brander, S., lo Leggio, L., & Johansen, K. S. (2021). Lytic polysaccharide monooxygenases and other histidine-brace copper proteins: structure, oxygen activation and biotechnological applications. *Biochemical Society Transactions*, *49*(1), 531–540.

Isaksen, T., Westereng, B., Aachmann, F. L., Agger, J. W., Kracher, D., Kittl, R., Ludwig, R., Haltrich, D., Eijsink, V. G. H., & Horn, S. J. (2014). A C4-oxidizing Lytic Polysaccharide Monooxygenase Cleaving Both Cellulose and Cello-oligosaccharides. *Journal of Biological Chemistry*, *289*(5), 2632–2642.

Ishii, K., MuÃŸmann, M., MacGregor, B. J., & Amann, R. (2004). An improved fluorescence in situ hybridization protocol for the identification of bacteria and archaea in marine sediments. *FEMS Microbiology Ecology*, *50*(3), 203–213.

Isikgor, F. H., & Becer, C. R. (2015). Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polymer Chemistry*, *6*(25), 4497–4559.

Jacquot, J. E., Horak, R. E. A., Amin, S. A., Devol, A. H., Ingalls, A. E., Armbrust, E. V., Stahl, D. A., & Moffett, J. W. (2014). Assessment of the potential for copper limitation of ammonia oxidation by Archaea in a dynamic estuary. *Marine Chemistry*, *162*, 37–49.

Jenkins, S. N., Murphy, D. v., Waite, I. S., Rushton, S. P., & O'Donnell, A. G. (2016). Ancient landscapes and the relationship with microbial nitrification. *Scientific Reports*, *6*(1), 30733.

Jez, E., Pellegrini, E., & Contin, M. (2023). Copper Bioavailability and Leaching in Conventional and Organic Viticulture under Environmental Stress. *Applied Sciences*, *13*(4), 2595.

Jhamb, K., & Sahoo, D. K. (2012). Production of soluble recombinant proteins in Escherichia coli: Effects of process conditions and chaperone co-expression on cell growth and production of xylanase. *Bioresource Technology*, *123*, 135–143. https://doi.org/10.1016/j.biortech.2012.07.011

Jiang, H., Huang, L., Deng, Y., Wang, S., Zhou, Y., Liu, L., & Dong, H. (2014). Latitudinal Distribution of Ammonia-Oxidizing Bacteria and Archaea in the Agricultural Soils of Eastern China. *Applied and Environmental Microbiology*, *80*(18), 5593–5602.

Johansen, K. S. (2016). Discovery and industrial applications of lytic polysaccharide mono-oxygenases. *Biochemical Society Transactions*, *44*(1), 143–149.

Jon Petter Gustafsson. (2025). *https://vminteq.com/*.

Jones, S. M., & Solomon, E. I. (2015). Electron transfer and reaction mechanism of laccases. *Cellular and Molecular Life Sciences*, *72*(5), 869–883.

Jung, M., Kim, J., Sinninghe Damsté, J. S., Rijpstra, W. I. C., Madsen, E. L., Kim, S., Hong, H., Si, O., Kerou, M., Schleper, C., & Rhee, S. (2016). A hydrophobic ammonia-oxidizing archaeon of the *Nitrosocosmicus* clade isolated from coal tar-contaminated sediment. *Environmental Microbiology Reports*, *8*(6), 983–992.

Jung, M.-Y., Gwak, J.-H., Rohe, L., Giesemann, A., Kim, J.-G., Well, R., Madsen, E. L., Herbold, C. W., Wagner, M., & Rhee, S.-K. (2019). Indications for enzymatic denitrification to N2O at low pH in an ammonia-oxidizing archaeon. *The ISME Journal*, *13*(10), 2633–2638.

Jung, M.-Y., Park, S.-J., Kim, S.-J., Kim, J.-G., Sinninghe Damsté, J. S., Jeon, C. O., & Rhee, S.-K. (2014). A Mesophilic, Autotrophic, Ammonia-Oxidizing Archaeon of Thaumarchaeal Group I.1a Cultivated from a Deep Oligotrophic Soil Horizon. *Applied and Environmental Microbiology*, *80*(12), 3645–3655.

Jung, M.-Y., Park, S.-J., Min, D., Kim, J.-S., Rijpstra, W. I. C., Sinninghe Damsté, J. S., Kim, G.-J., Madsen, E. L., & Rhee, S.-K. (2011). Enrichment and Characterization of an Autotrophic Ammonia-Oxidizing Archaeon of Mesophilic Crenarchaeal Group I.1a from an Agricultural Soil. *Applied and Environmental Microbiology*, *77*(24), 8635–8647.

Jung, M.-Y., Sedlacek, C. J., Kits, K. D., Mueller, A. J., Rhee, S.-K., Hink, L., Nicol, G. W., Bayer, B., Lehtovirta-Morley, L., Wright, C., de la Torre, J. R., Herbold, C. W., Pjevac, P., Daims, H., & Wagner, M. (2022). Ammonia-oxidizing archaea possess a wide range of cellular ammonia affinities. *The ISME Journal*, *16*(1), 272–283.

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., & Wang, Z. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, *7*, e7359.

Katoh, K., Asimenos, G., & Toh, H. (2009). *Multiple Alignment of DNA Sequences with MAFFT* (pp. 39–64).

Keegan, R. M., & Winn, M. D. (2008). *MrBUMP* : an automated pipeline for molecular replacement. *Acta Crystallographica Section D Biological Crystallography*, *64*(1), 119–124.

Kenney, G. E., & Rosenzweig, A. C. (2018). Chalkophores. *Annual Review of Biochemistry*, *87*(1), 645–676.

Kerou, M., Offre, P., Valledor, L., Abby, S. S., Melcher, M., Nagler, M., Weckwerth, W., & Schleper, C. (2016). Proteomics and comparative genomics of *Nitrososphaera viennensis* reveal the core genome and adaptations of archaeal ammonia oxidizers. *Proceedings of the National Academy of Sciences*, *113*(49).

Kérouel, R., & Aminot, A. (1997). Fluorometric determination of ammonia in sea and estuarine waters by direct segmented flow analysis. *Marine Chemistry*, *57*(3–4), 265–275.

Kikhney, A. G., & Svergun, D. I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Letters*, *589*(19PartA), 2570–2577.

Kikuchi, S., Fujitani, H., Ishii, K., Isshiki, R., Sekiguchi, Y., & Tsuneda, S. (2023). Characterisation of bacteria representing a novel *Nitrosomonas* clade: Physiology, genomics and distribution of missing ammonia oxidizer. *Environmental Microbiology Reports*, *15*(5), 404–416.

Kim, H.-W., Lee, S.-Y., Park, H., & Jeon, S.-J. (2015). Expression, refolding, and characterization of a small laccase from Thermus thermophilus HJ6. *Protein Expression and Purification*, *114*, 37–43.

Kim, J., Jung, M., Park, S., Rijpstra, W. I. C., Sinninghe Damsté, J. S., Madsen, E. L., Min, D., Kim, J., Kim, G., & Rhee, S. (2012). Cultivation of a highly enriched ammonia-oxidizing archaeon of thaumarchaeotal group I.1b from an agricultural soil. *Environmental Microbiology*, *14*(6), 1528–1543.

Kim, J.-G., Park, S.-J., Sinninghe Damsté, J. S., Schouten, S., Rijpstra, W. I. C., Jung, M.-Y., Kim, S.-J., Gwak, J.-H., Hong, H., Si, O.-J., Lee, S., Madsen, E. L., & Rhee, S.-K. (2016). Hydrogen peroxide detoxification is a key mechanism for growth of ammonia-oxidizing archaea. *Proceedings of the National Academy of Sciences*, *113*(28), 7888–7893.

Klein, T., Poghosyan, L., Barclay, J. E., Murrell, J. C., Hutchings, M. I., & Lehtovirta-Morley, L. E. (2022). Cultivation of ammonia-oxidising archaea on solid medium. *FEMS Microbiology Letters*, *369*(1).

Komori, H., & Higuchi, Y. (2015). Structural insights into the O_2 reduction mechanism of multicopper oxidase. *Journal of Biochemistry*, *158*(4), 293–298.

Könneke, M., Bernhard, A. E., de la Torre, J. R., Walker, C. B., Waterbury, J. B., & Stahl, D. A. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature*, *437*(7058), 543–546.

Könneke, M., Schubert, D. M., Brown, P. C., Hügler, M., Standfest, S., Schwander, T., Schada von Borzyskowski, L., Erb, T. J., Stahl, D. A., & Berg, I. A. (2014). Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO_2 fixation. *Proceedings of the National Academy of Sciences*, *111*(22), 8239–8244.

Kop, L. F. M., Koch, H., Dalcin Martins, P., Suarez, C., Karačić, S., Persson, F., Wilén, B.-M., Hagelia, P., Jetten, M. S. M., & Lücker, S. (2025). High diversity of nitrifying bacteria and archaea in biofilms from a subsea tunnel. *FEMS Microbiology Ecology*, *101*(5).

Koper, T. E., Stark, J. M., Habteselassie, M. Y., & Norton, J. M. (2010). Nitrification exhibits Haldane kinetics in an agricultural soil treated with ammonium sulfate or dairy-waste compost. *FEMS Microbiology Ecology*, *74*(2), 316–322.

Kubota, K. (2013). CARD-FISH for Environmental Microorganisms: Technical Advancement and Future Applications. *Microbes and Environments*, *28*(1), 3–12.

Kumar, A., Singh, A., Sharma, V. K., Goel, A., & Kumar, A. (2024). The upsurge of lytic polysaccharide monooxygenases in biomass deconstruction: characteristic functions and sustainable applications. *The FEBS Journal*, *291*(23), 5081–5101.

Kumar, M., Kumar, P., Das, P., Solanki, R., & Kapur, M. K. (2020). Potential applications of extracellular enzymes from Streptomyces spp. in various industries. *Archives of Microbiology*, *202*(7), 1597–1615.

Kupitz, C., Basu, S., Grotjohann, I., Fromme, R., Zatsepin, N. A., Rendek, K. N., Hunter, M. S., Shoeman, R. L., White, T. A., Wang, D., James, D., Yang, J.-H., Cobb, D. E., Reeder, B., Sierra, R. G., Liu, H., Barty, A., Aquila, A. L., Deponte, D., … Fromme, P. (2014). Serial time-resolved crystallography of photosystem II using a femtosecond X-ray laser. *Nature*, *513*(7517), 261–265.

Labourel, A., Frandsen, K. E. H., Zhang, F., Brouilly, N., Grisel, S., Haon, M., Ciano, L., Ropartz, D., Fanuel, M., Martin, F., Navarro, D., Rosso, M.-N., Tandrup, T., Bissaro, B., Johansen, K. S., Zerva, A., Walton, P. H., Henrissat, B., Leggio, L. lo, & Berrin, J.-G. (2020). A fungal family of lytic polysaccharide monooxygenase-like copper proteins. *Nature Chemical Biology*, *16*(3), 345–350.

Ladha, J. K., Peoples, M. B., Reddy, P. M., Biswas, J. C., Bennett, A., Jat, M. L., & Krupnik, T. J. (2022). Biological nitrogen fixation and prospects for ecological intensification in cereal-based cropping systems. *Field Crops Research*, *283*, 108541.

Lago, D., Kraxner, J., Galusek, D., & Bernardo, E. (2025). Development of porous membranes by alkali activation of borosilicate glass: effect of wastewater acidity on copper adsorption. *Developments in the Built Environment*, *23*, 100705.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359.

Langmead, B., Wilks, C., Antonescu, V., & Charles, R. (2019). Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics*, *35*(3), 421–432.

Lawton, T. J., Bowen, K. E., Sayavedra-Soto, L. A., Arp, D. J., & Rosenzweig, A. C. (2013). Characterization of a Nitrite Reductase Involved in Nitrifier Denitrification. *Journal of Biological Chemistry*, *288*(35), 25575–25583.

Lawton, T. J., Ham, J., Sun, T., & Rosenzweig, A. C. (2014). Structural conservation of the B subunit in the ammonia monooxygenase/particulate methane monooxygenase superfamily. *Proteins: Structure, Function, and Bioinformatics*, *82*(9), 2263–2267.

Lawton, T. J., & Rosenzweig, A. C. (2011). Detection and characterization of a multicopper oxidase from nitrosomonas Europaea. In *Methods in Enzymology* (Vol. 496).

Lawton, T. J., Sayavedra-Soto, L. A., Arp, D. J., & Rosenzweig, A. C. (2009). Crystal Structure of a Two-domain Multicopper Oxidase. *Journal of Biological Chemistry*, *284*(15), 10174–10180.

Lee, S., Sieradzki, E. T., Nicol, G. W., & Hazard, C. (2023). Propagation of viral genomes by replicating ammonia-oxidising archaea during soil nitrification. *The ISME Journal*, *17*(2), 309–314.

Lehtovirta-Morley, L. E. (2018). Ammonia oxidation: Ecology, physiology, biochemistry and why they must all come together. *FEMS Microbiology Letters*, *365*(9).

Lehtovirta-Morley, L. E., Ge, C., Ross, J., Yao, H., Hazard, C., Gubry-Rangin, C., Prosser, J. I., & Nicol, G. W. (2024). Nitrosotalea devaniterrae gen. nov., sp. nov. and Nitrosotalea sinensis sp. nov., two acidophilic ammonia oxidising archaea isolated from acidic soil, and proposal of the new order Nitrosotaleales ord. nov. within the class Nitrososphaeria of the phylum Nitrososphaerota. *International Journal of Systematic and Evolutionary Microbiology*, *74*(9).

Lehtovirta-Morley, L. E., Ge, C., Ross, J., Yao, H., Nicol, G. W., & Prosser, J. I. (2014). Characterisation of terrestrial acidophilic archaeal ammonia oxidisers and their inhibition and stimulation by organic compounds. *FEMS Microbiology Ecology*, *89*(3), 542–552.

Lehtovirta-Morley, L. E., Ross, J., Hink, L., Weber, E. B., Gubry-Rangin, C., Thion, C., Prosser, J. I., & Nicol, G. W. (2016). Isolation of ‘ *Candidatus* Nitrosocosmicus franklandus’, a novel ureolytic soil archaeal ammonia oxidiser with tolerance to high ammonia concentration. *FEMS Microbiology Ecology*, *92*(5), fiw057.

Lehtovirta-Morley, L. E., Stoecker, K., Vilcinskas, A., Prosser, J. I., & Nicol, G. W. (2011). Cultivation of an obligate acidophilic ammonia oxidizer from a nitrifying acid soil. *Proceedings of the National Academy of Sciences*, *108*(38), 15892–15897.

Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M., & Henrissat, B. (2013). Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnology for Biofuels*, *6*(1), 41.

Levičnik-Höfferle, Š., Nicol, G. W., Ausec, L., Mandić-Mulec, I., & Prosser, J. I. (2012). Stimulation of thaumarchaeal ammonia oxidation by ammonia derived from organic nitrogen but not added inorganic nitrogen. *FEMS Microbiology Ecology*, *80*(1), 114–123.

Li, D., Liu, C.-M., Luo, R., Sadakane, K., & Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. *Bioinformatics*, *31*(10), 1674–1676.

Li, H., Zhuang, L., Cai, H., Ni, Y., Chu, T., Chen, L., Yu, Y., & Wang, Y. (2025). *Nitrosarchaeum haohaiensis* </scp> sp. Nov. <scp> CL1 ^{T} </scp> : Isolation and Characterisation of a Novel Ammonia-Oxidising Archaeon From Aquatic Environments. *Environmental Microbiology Reports*, *17*(3).

Li, J., Solhi, L., Goddard-Borger, E. D., Mathieu, Y., Wakarchuk, W. W., Withers, S. G., & Brumer, H. (2021). Four cellulose-active lytic polysaccharide monooxygenases from Cellulomonas species. *Biotechnology for Biofuels*, *14*(1), 29.

Li, M., Du, C., Liu, J., Quan, X., Lan, M., & Li, B. (2018). Mathematical modeling on the nitrogen removal inside the membrane-aerated biofilm dominated by ammonia-oxidizing archaea (AOA): Effects of temperature, aeration pressure and COD/N ratio. *Chemical Engineering Journal*, *338*, 680–687.

Li, W., Zhen, Y., Yang, Y., Wang, D., & He, H. (2024). Environmental Adaptability and Roles in Ammonia Oxidation of Aerobic Ammonia-Oxidizing Microorganisms in the Surface Sediments of East China Sea. *Journal of Microbiology*, *62*(10), 845–858.

Liebschner, D., Afonine, P. v., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. v., … Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and

electrons: recent developments in *Phenix*. *Acta Crystallographica Section D Structural Biology*, *75*(10), 861–877.

Lin, H., Wang, C., Zhao, H., Chen, G., & Chen, X. (2020). A subcellular level study of copper speciation reveals the synergistic mechanism of microbial cells and EPS involved in copper binding in bacterial biofilms. *Environmental Pollution*, *263*, 114485.

Lin, J. T., Goldman, B. S., & Stewart, V. (1993). Structures of genes nasA and nasB, encoding assimilatory nitrate and nitrite reductases in Klebsiella pneumoniae M5al. *Journal of Bacteriology*, *175*(8), 2370–2378.

Lin, Y., Hu, H.-W., Ye, G., Fan, J., Ding, W., He, Z.-Y., Zheng, Y., & He, J.-Z. (2021). Ammonia-oxidizing bacteria play an important role in nitrification of acidic soils: A meta-analysis. *Geoderma*, *404*, 115395.

Lindley, P. J., Parkin, A., Davies, G. J., & Walton, P. H. (2022). Mapping the protonation states of the histidine brace in an AA10 lytic polysaccharide monooxygenase using CW-EPR spectroscopy and DFT calculations. *Faraday Discussions*, *234*, 336–348.

Liu, L., Liu, M., Jiang, Y., Lin, W., & Luo, J. (2019). *Physiological and genomic analysis of "* Candidatus *Nitrosocosmicus agrestis", an ammonia tolerant ammonia-oxidizing archaeon from vegetable soil*.

Liu, L., Liu, M., Jiang, Y., Lin, W., & Luo, J. (2021). Production and Excretion of Polyamines To Tolerate High Ammonia, a Case Study on Soil Ammonia-Oxidizing Archaeon " *Candidatus* Nitrosocosmicus agrestis." *mSystems*, *6*(1).

Lloyd, K. G., May, M. K., Kevorkian, R. T., & Steen, A. D. (2013). Meta-Analysis of Quantification Methods Shows that Archaea and Bacteria Have Similar Abundances in the Subseafloor. *Applied and Environmental Microbiology*, *79*(24), 7790–7799.

Long, A., Song, B., Fridey, K., & Silva, A. (2015). Detection and diversity of copper containing nitrite reductase genes (nirK) in prokaryotic and fungal communities of agricultural soils. *FEMS Microbiology Ecology*, *91*(2), 1–9.

Lu, X., Bottomley, P. J., & Myrold, D. D. (2015). Contributions of ammonia-oxidizing archaea and bacteria to nitrification in Oregon forest soils. *Soil Biology and Biochemistry*, *85*, 54–62.

Lu, X., Nicol, G. W., & Neufeld, J. D. (2018). Differential responses of soil ammonia-oxidizing archaea and bacteria to temperature and depth under two different land uses. *Soil Biology and Biochemistry*, *120*, 272–282.

Lu, X., Taylor, A. E., Myrold, D. D., & Neufeld, J. D. (2020). Expanding perspectives of soil nitrification to include ammonia-oxidizing archaea and comammox bacteria. *Soil Science Society of America Journal*, *84*(2), 287–302.

Lučić, M., Svistunenko, D. A., Wilson, M. T., Chaplin, A. K., Davy, B., Ebrahim, A., Axford, D., Tosha, T., Sugimoto, H., Owada, S., Dworkowski, F. S. N., Tews, I., Owen, R. L., Hough, M. A., & Worrall, J. A. R. (2020). Serial Femtosecond Zero Dose Crystallography Captures a Water-Free Distal Heme Site in a Dye-Decolorising Peroxidase to Reveal a Catalytic Role for an Arginine in Fe ^{IV} =O Formation. *Angewandte Chemie International Edition*, *59*(48), 21656–21662.

Lučić, M., Wilson, M. T., Svistunenko, D. A., Owen, R. L., Hough, M. A., & Worrall, J. A. R. (2021). Aspartate or arginine? Validated redox state X-ray structures elucidate mechanistic subtleties of FeIV = O formation in bacterial dye-decolorizing peroxidases. *JBIC Journal of Biological Inorganic Chemistry*, *26*(7), 743–761.

Lučić, M., Wilson, M. T., Tosha, T., Sugimoto, H., Shilova, A., Axford, D., Owen, R. L., Hough, M. A., & Worrall, J. A. R. (2022). Serial Femtosecond Crystallography Reveals the Role of Water in the One- or Two-Electron Redox Chemistry of Compound I in the Catalytic Cycle of the B-Type Dye-Decolorizing Peroxidase DtpB. *ACS Catalysis*, *12*(21), 13349–13359.

Ludovici, C., Fröhlich, R., Vogtt, K., Mamat, B., & Lübben, M. (2002). Caged O $_2$. *European Journal of Biochemistry*, *269*(10), 2630–2637.

Luo, Z.-H., Li, Q., Xie, Y.-G., Lv, A.-P., Qi, Y.-L., Li, M.-M., Qu, Y.-N., Liu, Z.-T., Li, Y.-X., Rao, Y.-Z., Jiao, J.-Y., Liu, L., Narsing Rao, M. P., Hedlund, B. P., Evans, P. N., Fang, Y., Shu, W.-S., Huang, L.-N., Li, W.-J., & Hua, Z.-S. (2024). Temperature, pH, and oxygen availability contributed to the functional differentiation of ancient *Nitrososphaeria*. *The ISME Journal*, *18*(1).

Lv, J., Min, D., Cheng, Z., Zhang, J., Li, W., Mu, Y., Liu, S., & Liu, D. (2024). Direct ammonia oxidation (Dirammox) is favored over cell growth in *Alcaligenes ammonioxydans* HO-1 to deal with the toxicity of ammonium. *Biotechnology and Bioengineering*, *121*(3), 980–990.

Ma, L., Li, G., Xu, H., Liu, Z., Wan, Q., Liu, D., & Shen, Q. (2022). Structural and functional study of a novel lytic polysaccharide monooxygenase cPMO2 from compost sample in the oxidative degradation of cellulose. *Chemical Engineering Journal*, *433*, 134509.

Ma, L., Shan, J., & Yan, X. (2015). Nitrite behavior accounts for the nitrous oxide peaks following fertilization in a fluvo-aquic soil. *Biology and Fertility of Soils*, *51*(5), 563–572.

MacArthur, R., Sucheta, A., Chong, F. F., & Einarsdóttir, O. (1995). Photodissociation of a (mu-peroxo)(mu-hydroxo)bis[bis(bipyridyl)-cobalt(III)] complex: a tool to study fast biological reactions involving O2. *Proceedings of the National Academy of Sciences*, *92*(18), 8105–8109.

Manalastas-Cantos, K., Konarev, P. v., Hajizadeh, N. R., Kikhney, A. G., Petoukhov, M. v., Molodenskiy, D. S., Panjkovich, A., Mertens, H. D. T., Gruzinov, A., Borges, C., Jeffries, C. M., Svergun, D. I., & Franke, D. (2021). *ATSAS 3.0* : expanded functionality and new tools for small-angle scattering data analysis. *Journal of Applied Crystallography*, *54*(1), 343–355.

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Molecular Biology and Evolution*, *38*(10), 4647–4654.

Martens-Habbena, W., Qin, W., Horak, R. E. A., Urakawa, H., Schauer, A. J., Moffett, J. W., Armbrust, E. V., Ingalls, A. E., Devol, A. H., & Stahl, D. A. (2015). The production of nitric oxide by marine ammonia-oxidizing archaea and inhibition of archaeal ammonia oxidation by a nitric oxide scavenger. *Environmental Microbiology*, *17*(7), 2261–2274.

Matsui, T., Rajkovic, I., Mooers, B. H. M., Liu, P., & Weiss, T. M. (2024). Adaptable SEC-SAXS data collection for higher quality structure analysis in solution. *Protein Science*, *33*(4).

Maurer, F., Christl, I., Fulda, B., Voegelin, A., & Kretzschmar, R. (2013). Copper Redox Transformation and Complexation by Reduced and Oxidized Soil Humic Acid. 2. Potentiometric Titrations and Dialysis Cell Experiments. *Environmental Science & Technology*, *47*(19), 10912–10921.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). *Phaser* crystallographic software. *Journal of Applied Crystallography*, *40*(4), 658–674.

Meier, K. K., Jones, S. M., Kaper, T., Hansson, H., Koetsier, M. J., Karkehabadi, S., Solomon, E. I., Sandgren, M., & Kelemen, B. (2018). Oxygen Activation by Cu LPMOs in Recalcitrant Carbohydrate Polysaccharide Conversion to Monomer Sugars. *Chemical Reviews*, *118*(5), 2593–2635.

Mekasha, S., Tuveng, T. R., Askarian, F., Choudhary, S., Schmidt-Dannert, C., Niebisch, A., Modregger, J., Vaaje-Kolstad, G., & Eijsink, V. G. H. (2020). A trimodular bacterial enzyme combining hydrolytic activity with oxidative glycosidic bond cleavage efficiently degrades chitin. *Journal of Biological Chemistry*, *295*(27), 9134–9146.

Melillo, J. M. (2021). Disruption of the global nitrogen cycle: A grand challenge for the twenty-first century. *Ambio*, *50*(4), 759–763.

Méndez-Líter, J. A., Ayuso-Fernández, I., Csarman, F., de Eugenio, L. I., Míguez, N., Plou, F. J., Prieto, A., Ludwig, R., & Martínez, M. J. (2021). Lytic Polysaccharide Monooxygenase from Talaromyces amestolkiae with an Enigmatic Linker-like Region: The Role of This Enzyme on Cellulose Saccharification. *International Journal of Molecular Sciences*, *22*(24), 13611.

Mengel, K., Kirkby, E. A., Kosegarten, H., & Appel, T. (2001). Soil Copper. In *Principles of Plant Nutrition* (pp. 599–611). Springer Netherlands.

Miralles-Robledillo, J. M., Martínez-Espinosa, R. M., & Pire, C. (2024). Transcriptomic profiling of haloarchaeal denitrification through RNA-Seq analysis. *Applied and Environmental Microbiology*, *90*(6).

Moin, N. S., Nelson, K. A., Bush, A., & Bernhard, A. E. (2009). Distribution and Diversity of Archaeal and Bacterial Ammonia Oxidizers in Salt Marsh Sediments. *Applied and Environmental Microbiology*, *75*(23), 7461–7468.

Moissl-Eichinger, C., Pausan, M., Taffner, J., Berg, G., Bang, C., & Schmitz, R. A. (2018). Archaea Are Interactive Components of Complex Microbiomes. *Trends in Microbiology*, *26*(1), 70–85.

Molari, M., & Manini, E. (2012). Reliability of CARD-FISH Procedure for Enumeration of Archaea in Deep-Sea Surficial Sediments. *Current Microbiology*, *64*(3), 242–250.

Monks, P. S., Archibald, A. T., Colette, A., Cooper, O., Coyle, M., Derwent, R., Fowler, D., Granier, C., Law, K. S., Mills, G. E., Stevenson, D. S., Tarasova, O., Thouret, V., von Schneidemesser, E., Sommariva, R., Wild, O., & Williams, M. L. (2015). Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer. *Atmospheric Chemistry and Physics*, *15*(15), 8889–8973.

Monteiro, D. C. F., Amoah, E., Rogers, C., & Pearson, A. R. (2021). Using photocaging for fast time-resolved structural biology studies. *Acta Crystallographica Section D Structural Biology*, *77*(10), 1218–1232.

Moon, M., Lee, J.-P., Park, G. W., Lee, J.-S., Park, H. J., & Min, K. (2022). Lytic polysaccharide monooxygenase (LPMO)-derived saccharification of lignocellulosic biomass. *Bioresource Technology*, *359*, 127501.

Morell, M., Bravo, R., Espargaró, A., Sisquella, X., Avilés, F. X., Fernàndez-Busquets, X., & Ventura, S. (2008). Inclusion bodies: Specificity in their aggregation process and amyloid-like structure. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1783*(10), 1815–1825.

Morimoto, S., Hayatsu, M., Takada Hoshino, Y., Nagaoka, K., Yamazaki, M., Karasawa, T., Takenaka, M., & Akiyama, H. (2011). Quantitative Analyses of Ammonia-oxidizing Archaea (AOA) and Ammonia-oxidizing Bacteria (AOB) in Fields with Different Soil Types. *Microbes and Environments*, *26*(3), 248–253.

Morris, J. J., Rose, A. L., & Lu, Z. (2022). Reactive oxygen species in the world ocean and their impacts on marine ecosystems. *Redox Biology*, *52*, 102285.

Mosley, O. E., Gios, E., Close, M., Weaver, L., Daughney, C., & Handley, K. M. (2022). Nitrogen cycling and microbial cooperation in the terrestrial subsurface. *The ISME Journal*, *16*(11), 2561–2573.

Mukhtar, H., Lin, Y.-P., Lin, C.-M., & Lin, Y.-R. (2019). Relative Abundance of Ammonia Oxidizing Archaea and Bacteria Influences Soil Nitrification Responses to Temperature. *Microorganisms*, *7*(11), 526.

Munzone, A., el Kerdi, B., Fanuel, M., Rogniaux, H., Ropartz, D., Réglier, M., Royant, A., Simaan, A. J., & Decroos, C. (2020). Characterization of a bacterial copper-dependent lytic polysaccharide monooxygenase with an unusual second coordination sphere. *The FEBS Journal*, *287*(15), 3298–3314.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., & Vagin, A. A. (2011). *REFMAC* 5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D Biological Crystallography*, *67*(4), 355–367.

Nango, E., Royant, A., Kubo, M., Nakane, T., Wickstrand, C., Kimura, T., Tanaka, T., Tono, K., Song, C., Tanaka, R., Arima, T., Yamashita, A., Kobayashi, J., Hosaka, T., Mizohata, E., Nogly, P., Sugahara, M., Nam, D., Nomura, T., … Iwata, S. (2016). A three-dimensional movie of structural changes in bacteriorhodopsin. *Science*, *354*(6319), 1552–1557.

Neal, A. L., Hughes, D., Clark, I. M., Jansson, J. K., & Hirsch, P. R. (2021). Microbiome Aggregated Traits and Assembly Are More Sensitive to Soil Management than Diversity. *mSystems*, *6*(3).

Nguyen, D.-H., Lin, C., Vu, C.-T., Cheruiyot, N. K., Nguyen, M. K., Le, T. H., Lukkhasorn, W., Vo, T.-D.-H., & Bui, X.-T. (2022). Tropospheric ozone and NO. A review of worldwide variation and meteorological influences. *Environmental Technology & Innovation*, *28*, 102809.

Nicol, G. W., Hink, L., Gubry-Rangin, C., Prosser, J. I., & Lehtovirta-Morley, L. E. (2019). Genome Sequence of " *Candidatus* Nitrosocosmicus franklandus" C13, a Terrestrial Ammonia-Oxidizing Archaeon. *Microbiology Resource Announcements*, *8*(40).

Norton, J. M., Alzerreca, J. J., Suwa, Y., & Klotz, M. G. (2002). Diversity of ammonia monooxygenase operon in autotrophic ammonia-oxidizing bacteria. *Archives of Microbiology*, *177*(2), 139–149.

O'Dell, W. B., Agarwal, P. K., & Meilleur, F. (2017). Oxygen Activation at the Active Site of a Fungal Lytic Polysaccharide Monooxygenase. *Angewandte Chemie International Edition*, *56*(3), 767–770.

Offre, P., Kerou, M., Spang, A., & Schleper, C. (2014). Variability of the transporter gene complement in ammonia-oxidizing archaea. *Trends in Microbiology*, *22*(12), 665–675.

Offre, P., Prosser, J. I., & Nicol, G. W. (2009). Growth of ammonia-oxidizing archaea in soil microcosms is inhibited by acetylene. *FEMS Microbiology Ecology*, *70*(1), 99–108.

Oganesyan, N., Ankoudinova, I., Kim, S.-H., & Kim, R. (2007). Effect of osmotic stress and heat shock in recombinant protein overexpression and crystallization. *Protein Expression and Purification*, *52*(2), 280–285.

Ordonez, A., Hussain, U., Cambon, M. C., Golyshin, P. N., Downie, J., & McDonald, J. E. (2025). Evaluating agar-plating and dilution-to-extinction isolation methods for generating oak-associated microbial culture collections. *ISME Communications*, *5*(1).

Orell, A., Navarro, C. A., Arancibia, R., Mobarec, J. C., & Jerez, C. A. (2010). Life in blue: Copper resistance mechanisms of bacteria and Archaea used in industrial biomining of minerals. *Biotechnology Advances*, *28*(6), 839–848.

Orlando, C., Rizzo, I. C., Arrigoni, F., Zampolli, J., Mangiagalli, M., di Gennaro, P., Lotti, M., de Gioia, L., Marino, T., Greco, C., & Bertini, L. (2024). Mechanism of non-phenolic substrate oxidation by the fungal laccase Type 1 copper site from *Trametes versicolor* : the case of benzo[*a*]pyrene and anthracene. *Dalton Transactions*, *53*(29), 12152–12161.

Ouyang, Y., Evans, S. E., Friesen, M. L., & Tiemann, L. K. (2018). Effect of nitrogen fertilization on the abundance of nitrogen cycling genes in agricultural soils: A meta-analysis of field studies. *Soil Biology and Biochemistry*, *127*, 71–78.

Ouyang, Y., Norton, J. M., & Stark, J. M. (2017). Ammonium availability and temperature control contributions of ammonia oxidizing bacteria and archaea to nitrification in an agricultural soil. *Soil Biology and Biochemistry*, *113*, 161–172.

Palatinszky, M., Herbold, C., Jehmlich, N., Pogoda, M., Han, P., von Bergen, M., Lagkouvardos, I., Karst, S. M., Galushko, A., Koch, H., Berry, D., Daims, H., & Wagner, M. (2015). Cyanate as an energy source for nitrifiers. *Nature*, *524*(7563), 105–108.

Panjkovich, A., & Svergun, D. I. (2018). CHROMIXS: automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data. *Bioinformatics*, *34*(11), 1944–1946.

Paradisi, A., Johnston, E. M., Tovborg, M., Nicoll, C. R., Ciano, L., Dowle, A., McMaster, J., Hancock, Y., Davies, G. J., & Walton, P. H. (2019). Formation of a Copper(II)–Tyrosyl Complex at the Active Site of Lytic Polysaccharide Monooxygenases Following Oxidation by H_2O_2. *Journal of the American Chemical Society*, *141*(46), 18585–18599.

Park, B.-J., Park, S.-J., Yoon, D.-N., Schouten, S., Sinninghe Damsté, J. S., & Rhee, S.-K. (2010). Cultivation of Autotrophic Ammonia-Oxidizing Archaea from Marine Sediments in Coculture with Sulfur-Oxidizing Bacteria. *Applied and Environmental Microbiology*, *76*(22), 7575–7587.

Pauly, M., & Keegstra, K. (2008). Cell-wall carbohydrates and their modification as a resource for biofuels. *The Plant Journal*, *54*(4), 559–568.

Percy Faraday Frankland, & Grace C. Frankland. (1890). V. The nitrifying process and its specific ferment.—Part I. *Philosophical Transactions of the Royal Society of London. (B.)*, *181*, 107–128.

Pereira, O., Qin, W., Galand, P. E., Debroas, D., Lami, R., Hochart, C., Zhou, Y., Zhou, J., & Zhang, C. (2024). Metabolic activities of marine ammonia-oxidizing archaea orchestrated by quorum sensing. *mLife*, *3*(3), 417–429.

Pester, M., Maixner, F., Berry, D., Rattei, T., Koch, H., Lücker, S., Nowka, B., Richter, A., Spieck, E., Lebedeva, E., Loy, A., Wagner, M., & Daims, H. (2014). <scp> *NxrB* </scp> encoding the beta subunit of nitrite oxidoreductase as functional and phylogenetic marker for nitrite-oxidizing <scp> *N* </scp> *itrospira*. *Environmental Microbiology*, *16*(10), 3055–3071.

Pester, M., Rattei, T., Flechl, S., Gröngröft, A., Richter, A., Overmann, J., Reinhold-Hurek, B., Loy, A., & Wagner, M. (2012). *amoA* -based consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of *amoA* genes from soils of four different geographic regions. *Environmental Microbiology*, *14*(2), 525–539.

Pester, M., Schleper, C., & Wagner, M. (2011). The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Current Opinion in Microbiology*, *14*(3), 300–306.

Pi, H.-W., Lin, J.-J., Chen, C.-A., Wang, P.-H., Chiang, Y.-R., Huang, C.-C., Young, C.-C., & Li, W.-H. (2022). Origin and Evolution of Nitrogen Fixation in Prokaryotes. *Molecular Biology and Evolution*, *39*(9).

Poshina, D. N., Raik, S. v., Poshin, A. N., & Skorik, Y. A. (2018). Accessibility of chitin and chitosan in enzymatic hydrolysis: A review. *Polymer Degradation and Stability*, *156*, 269–278.

Powell, D. R. (2016). Review of *X-Ray Crystallography*. *Journal of Chemical Education*, *93*(4), 591–592.

Pratscher, J., Dumont, M. G., & Conrad, R. (2011). Ammonia oxidation coupled to CO_2 fixation by archaea and bacteria in an agricultural soil. *Proceedings of the National Academy of Sciences*, *108*(10), 4170–4175.

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE*, *5*(3), e9490.

Processing serial synchrotron crystallography diffraction data with DIALS (pp. 207–244). (2024).

Prosser, J. I., Hink, L., Gubry-Rangin, C., & Nicol, G. W. (2020). Nitrous oxide production by ammonia oxidizers: Physiological diversity, niche differentiation and potential mitigation strategies. *Global Change Biology*, *26*(1), 103–118.

Prosser, J. I., & Nicol, G. W. (2008). Relative contributions of archaea and bacteria to aerobic ammonia oxidation in the environment. *Environmental Microbiology*, *10*(11), 2931–2941.

Prosser, J. I., & Nicol, G. W. (2012). Archaeal and bacterial ammonia-oxidisers in soil: the quest for niche specialisation and differentiation. *Trends in Microbiology*, *20*(11), 523–531.

Purkhold, U., Pommerening-Röser, A., Juretschko, S., Schmid, M. C., Koops, H.-P., & Wagner, M. (2000). Phylogeny of All Recognized Species of Ammonia Oxidizers Based on Comparative 16S rRNA and *amoA* Sequence Analysis: Implications for Molecular Diversity Surveys. *Applied and Environmental Microbiology*, *66*(12), 5368–5382.

Putnam, D. K., Lowe, E. W., & Meiler, J. (2013). RECONSTRUCTION OF SAXS PROFILES FROM PROTEIN STRUCTURES. *Computational and Structural Biotechnology Journal*, *8*(11), e201308006.

Qiang, R., Wang, M., Li, Q., Li, Y., Li, C., Zhang, J., & Liu, H. (2024). The different responses of AOA and AOB communities to irrigation systems in the semi-arid region of Northeast China. *Frontiers in Microbiology*, *15*.

Qin, W., Amin, S. A., Martens-Habbena, W., Walker, C. B., Urakawa, H., Devol, A. H., Ingalls, A. E., Moffett, J. W., Armbrust, E. V., & Stahl, D. A. (2014). Marine ammonia-

oxidizing archaeal isolates display obligate mixotrophy and wide ecotypic variation. *Proceedings of the National Academy of Sciences*, *111*(34), 12504–12509.

Qiu, J.-G., & Liu, S.-J. (2025). Dirammox (direct ammonia oxidation) to nitrogen (N2): discovery, current status, and perspectives. *Current Opinion in Microbiology*, *83*, 102565.

Qiu, S., Zhou, S., Tan, Y., Feng, J., Bai, Y., He, J., Cao, H., Che, Q., Guo, J., & Su, Z. (2022). Biodegradation and Prospect of Polysaccharide from Crustaceans. *Marine Drugs*, *20*(5), 310.

Quinlan, R. J., Sweeney, M. D., lo Leggio, L., Otten, H., Poulsen, J.-C. N., Johansen, K. S., Krogh, K. B. R. M., Jørgensen, C. I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C. P., Dupree, P., Xu, F., Davies, G. J., & Walton, P. H. (2011). Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the National Academy of Sciences*, *108*(37), 15079–15084.

Raskin, L., Stromley, J. M., Rittmann, B. E., & Stahl, D. A. (1994). Group-specific 16S rRNA hybridization probes to describe natural communities of methanogens. *Applied and Environmental Microbiology*, *60*(4), 1232–1240.

Rasmussen, A. N., & Francis, C. A. (2024). Dynamics and activity of an ammonia-oxidizing archaea bloom in South San Francisco Bay. *The ISME Journal*, *18*(1).

Ravelli, R., & Garman, E. (2006). Radiation damage in macromolecular cryocrystallography. *Current Opinion in Structural Biology*, *16*(5), 624–629.

Redondo, M. A., Jones, C. M., Legendre, P., Guénard, G., & Hallin, S. (2025). Predicting gene distribution in ammonia-oxidizing archaea using phylogenetic signals. *ISME Communications*, *5*(1).

Reese, E. T., Siu, R. G. H., & Levinson, H. S. (1950). The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. Journal of Bacteriology, 59(4), 485–497.

Reyes, C., Hodgskiss, L. H., Baars, O., Kerou, M., Bayer, B., Schleper, C., & Kraemer, S. M. (2020). Copper limiting threshold in the terrestrial ammonia oxidizing archaeon Nitrososphaera viennensis. *Research in Microbiology*, *171*(3–4), 134–142.

Reyes, C., Hodgskiss, L. H., Kerou, M., Pribasnig, T., Abby, S. S., Bayer, B., Kraemer, S. M., & Schleper, C. (2020). Genome wide transcriptomic analysis of the soil ammonia oxidizing archaeon *Nitrososphaera viennensis* upon exposure to copper limitation. *The ISME Journal*, *14*(11), 2659–2674.

Rieder, L., & Sørlie, M. (2023). Recent Advances in Understanding LPMO Catalysis. *Biochemistry*, *62*(22), 3170–3172.

Rieder, L., Stepnov, A. A., Sørlie, M., & Eijsink, V. G. H. (2021). Fast and Specific Peroxygenase Reactions Catalyzed by Fungal Mono-Copper Enzymes. *Biochemistry*, *60*(47), 3633–3643.

Rinke, C., Chuvochina, M., Mussig, A. J., Chaumeil, P.-A., Davín, A. A., Waite, D. W., Whitman, W. B., Parks, D. H., & Hugenholtz, P. (2021). A standardized archaeal taxonomy for the Genome Taxonomy Database. *Nature Microbiology*, *6*(7), 946–959.

Rodrigues-Oliveira, T., Belmok, A., Vasconcellos, D., Schuster, B., & Kyaw, C. M. (2017). Archaeal S-Layers: Overview and Current State of the Art. *Frontiers in Microbiology*, *8*.

Rodriguez, J., Chakrabarti, S., Choi, E., Shehadeh, N., Sierra-Martinez, S., Zhao, J., & Martens-Habbena, W. (2021). Nutrient-Limited Enrichments of Nitrifiers From Soil Yield Consortia of Nitrosocosmicus-Affiliated AOA and Nitrospira-Affiliated NOB. *Frontiers in Microbiology*, *12*.

Rodriguez-R, L. M., Gunturu, S., Harvey, W. T., Rosselló-Mora, R., Tiedje, J. M., Cole, J. R., & Konstantinidis, K. T. (2018). The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. *Nucleic Acids Research*, *46*(W1), W282–W288.

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, *5*.

Rotthauwe, J. H., Witzel, K. P., & Liesack, W. (1997). The ammonia monooxygenase structural gene amoA as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing populations. *Applied and Environmental Microbiology*, *63*(12), 4704–4712.

Ruan, A., Ren, C., & Quan, S. (2020). Conversion of the molecular chaperone Spy into a novel fusion tag to enhance recombinant protein expression. *Journal of Biotechnology*, *307*, 131–138.

Rubio, L. M., Herrero, A., & Flores, E. (1996). A cyanobacterial narB gene encodes a ferredoxin-dependent nitrate reductase. *Plant Molecular Biology*, *30*(4), 845–850.

Rutledge, H. L., & Tezcan, F. A. (2020). Electron Transfer in Nitrogenase. *Chemical Reviews*, *120*(12), 5158–5193.

Rütting, T., Schleusner, P., Hink, L., & Prosser, J. I. (2021). The contribution of ammonia-oxidizing archaea and bacteria to gross nitrification under different substrate availability. *Soil Biology and Biochemistry*, *160*, 108353.

Sagar, A., Jeffries, C. M., Petoukhov, M. v., Svergun, D. I., & Bernadó, P. (2021). Comment on the Optimal Parameters to Derive Intrinsically Disordered Protein Conformational Ensembles from Small-Angle X-ray Scattering Data Using the Ensemble Optimization Method. *Journal of Chemical Theory and Computation*, *17*(4), 2014–2021.

Sagarika, M. S., Parameswaran, C., Senapati, A., Barala, J., Mitra, D., Prabhukarthikeyan, S. R., Kumar, A., Nayak, A. K., & Panneerselvam, P. (2022). Lytic polysaccharide monooxygenases (LPMOs) producing microbes: A novel approach for rapid recycling of agricultural wastes. *Science of The Total Environment*, *806*, 150451.

Sakurai, T., & Kataoka, K. (2007). Basic and applied features of multicopper oxidases, CueO, bilirubin oxidase, and laccase. *The Chemical Record*, *7*(4), 220–229.

Sanjuan, J., Delgado, M. J., & Girard, L. (2020). Editorial: Microbial Control of the Nitrogen Cycle. *Frontiers in Microbiology*, *11*.

Santoro, A. E., Dupont, C. L., Richter, R. A., Craig, M. T., Carini, P., McIlvin, M. R., Yang, Y., Orsi, W. D., Moran, D. M., & Saito, M. A. (2015). Genomic and proteomic characterization of " *Candidatus* Nitrosopelagicus brevis": An ammonia-oxidizing archaeon from the open ocean. *Proceedings of the National Academy of Sciences*, *112*(4), 1173–1178.

Sanz-Luque, E., Chamizo-Ampudia, A., Llamas, A., Galvan, A., & Fernandez, E. (2015). Understanding nitrate assimilation and its regulation in microalgae. *Frontiers in Plant Science*, *6*.

Sarkar, S., Kazarina, A., Hansen, P. M., Ward, K., Hargreaves, C., Reese, N., Ran, Q., Kessler, W., de Souza, L. F. T., Loecke, T. D., Sarto, M. V. M., Rice, C. W., Zeglin, L. H., Sikes, B. A., & Lee, S. T. M. (2025). Metabolism diversification of ammonia-oxidizing archaea and bacteria under different precipitation gradients and land legacies. *Applied Soil Ecology*, *206*, 105831.

Satomura, T., Hirano, T., Inagaki, K., Horinaga, K., Takamura, E., Sakamoto, H., Ohshida, T., Ohshima, T., Sakuraba, H., & Suye, S. (2021). Activity enhancement of multicopper oxidase from a hyperthermophile via directed evolution, and its application as the element of a high performance biocathode. *Journal of Biotechnology*, *325*, 226–232.

Sauder, L. A., Albertsen, M., Engel, K., Schwarz, J., Nielsen, P. H., Wagner, M., & Neufeld, J. D. (2017). Cultivation and characterization of *Candidatus* Nitrosocosmicus exaquare, an ammonia-oxidizing archaeon from a municipal wastewater treatment system. *The ISME Journal*, *11*(5), 1142–1157.

Sauder, L. A., Ross, A. A., & Neufeld, J. D. (2016). Nitric oxide scavengers differentially inhibit ammonia oxidation in ammonia-oxidizing archaea and bacteria. *FEMS Microbiology Letters*, *363*(7), fnw052.

Scarlett, K., Denman, S., Clark, D. R., Forster, J., Vanguelova, E., Brown, N., & Whitby, C. (2021a). Relationships between nitrogen cycling microbial community abundance and composition reveal the indirect effect of soil pH on oak decline. *The ISME Journal*, *15*(3), 623–635.

Schleper, C., & Nicol, G. W. (2010a). *Ammonia-Oxidising Archaea – Physiology, Ecology and Evolution* (pp. 1–41).

Schlesinger, W. H., & Bernhardt, E. S. (2020). The Atmosphere. In *Biogeochemistry* (pp. 51–97). Elsevier.

Schneidman-Duhovny, D., Hammel, M., Tainer, J. A., & Sali, A. (2016). FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Research*, *44*(W1), W424–W429.

Schramm, F. D., Schroeder, K., & Jonas, K. (2020). Protein aggregation in bacteria. *FEMS Microbiology Reviews*, *44*(1), 54–72.

Schröder, G. C., O'Dell, W. B., Swartz, P. D., & Meilleur, F. (2021). Preliminary results of neutron and X-ray diffraction data collection on a lytic polysaccharide monooxygenase under

reduced and acidic conditions. *Acta Crystallographica Section F Structural Biology Communications*, *77*(4), 128–133.

Schulz, E. C., Yorke, B. A., Pearson, A. R., & Mehrabi, P. (2022). Best practices for time-resolved serial synchrotron crystallography. *Acta Crystallographica Section D Structural Biology*, *78*(1), 14–29.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069.

Sekar, R., Pernthaler, A., Pernthaler, J., Warnecke, F., Posch, T., & Amann, R. (2003). An Improved Protocol for Quantification of Freshwater *Actinobacteria* by Fluorescence In Situ Hybridization. *Applied and Environmental Microbiology*, *69*(5), 2928–2935.

Sekretaryova, A., Jones, S. M., & Solomon, E. I. (2019). O_2 Reduction to Water by High Potential Multicopper Oxidases: Contributions of the T1 Copper Site Potential and the Local Environment of the Trinuclear Copper Cluster. *Journal of the American Chemical Society*, *141*(28), 11304–11314.

Shafiee, R. T., Diver, P. J., Snow, J. T., Zhang, Q., & Rickaby, R. E. M. (2021). Marine ammonia-oxidising archaea and bacteria occupy distinct iron and copper niches. *ISME Communications*, *1*(1).

Shafiee, R. T., Snow, J. T., Hester, S., Zhang, Q., & Rickaby, R. E. M. (2022). Proteomic response of the marine ammonia-oxidising archaeon *Nitrosopumilus maritimus* to iron limitation reveals strategies to compensate for nutrient scarcity. *Environmental Microbiology*, *24*(2), 835–849.

Shah, A., Taylor, M. J., Molinaro, G., Anbu, S., Verdu, M., Jennings, L., Mikulska, I., Diaz-Moreno, S., el Mkami, H., Smith, G. M., Britton, M. M., Lovett, J. E., & Peacock, A. F. A. (2023). Design of the elusive proteinaceous oxygen donor copper site suggests a promising future for copper for MRI contrast agents. *Proceedings of the National Academy of Sciences*, *120*(27).

Shen, T., Stieglmeier, M., Dai, J., Urich, T., & Schleper, C. (2013). Responses of the terrestrial ammonia-oxidizing archaeon *Ca* . Nitrososphaera viennensis and the ammonia-oxidizing bacterium *Nitrosospira multiformis* to nitrification inhibitors. *FEMS Microbiology Letters*, *344*(2), 121–129.

Sheng, Y., Abreu, I. A., Cabelli, D. E., Maroney, M. J., Miller, A.-F., Teixeira, M., & Valentine, J. S. (2014). Superoxide Dismutases and Superoxide Reductases. *Chemical Reviews*, *114*(7), 3854–3918.

Sherif, M., Waung, D., Korbeci, B., Mavisakalyan, V., Flick, R., Brown, G., Abou-Zaid, M., Yakunin, A. F., & Master, E. R. (2013). Biochemical studies of the multicopper oxidase

(small laccase) from <*scp*>*S*</*scp*> *treptomyces coelicolor* using bioactive phytochemicals and site-directed mutagenesis. *Microbial Biotechnology*, *6*(5), 588–597.

Shi, M., Li, J., Gao, R., Song, X., Wang, G., Gao, Y., & Yan, S. (2023). Contrasting effects of elevated CO2 on autotrophic prokaryotes with different CO2 fixation strategies in tea plantation soil. *Biology and Fertility of Soils*, *59*(2), 205–215.

Sina, M., Farajzadeh, D., & Dastmalchi, S. (2015). Effects of Environmental Factors on Soluble Expression of a Humanized Anti-TNF-α scFv Antibody in Escherichia coli. *Advanced Pharmaceutical Bulletin*, *5*(4), 455–461.

Singha, A., Sekretareva, A., Tao, L., Lim, H., Ha, Y., Braun, A., Jones, S. M., Hedman, B., Hodgson, K. O., Britt, R. D., Kosman, D. J., & Solomon, E. I. (2023). Tuning the Type 1 Reduction Potential of Multicopper Oxidases: Uncoupling the Effects of Electrostatics and H-Bonding to Histidine Ligands. *Journal of the American Chemical Society*, *145*(24), 13284–13301.

Slade, R., Bauen, A., & Shah, N. (2009). The greenhouse gas emissions performance of cellulosic ethanol supply chains in Europe. *Biotechnology for Biofuels*, *2*(1), 15.

Smyth, P., Jaho, S., Williams, L. J., Karras, G., Fitzpatrick, A., Thompson, A. J., Battah, S., Axford, D., Horrell, S., Lučić, M., Ishihara, K., Kataoka, M., Matsuura, H., Shimba, K.,

Tono, K., Tosha, T., Sugimoto, H., Owada, S., Hough, M. A., … Owen, R. L. (2025). Time-resolved serial synchrotron and serial femtosecond crystallography of heme proteins using photocaged nitric oxide. *IUCrJ*, *12*(5), 582–594.

Soler-Jofra, A., Pérez, J., & van Loosdrecht, M. C. M. (2021). Hydroxylamine and the nitrogen cycle: A review. *Water Research*, *190*, 116723.

Solomon, E. I., Heppner, D. E., Johnston, E. M., Ginsbach, J. W., Cirera, J., Qayyum, M., Kieber-Emmons, M. T., Kjaergaard, C. H., Hadt, R. G., & Tian, L. (2014). Copper Active Sites in Biology. Chemical Reviews, 114(7), 3659–3853.

Song, W., Zhang, S., & Thomas, T. (2022). MarkerMAG: linking metagenome-assembled genomes (MAGs) with 16S rRNA marker genes using paired-end short reads. *Bioinformatics*, *38*(15), 3684–3688.

Srithep, P., Pornkulwat, P., & Limpiyakorn, T. (2018). Contribution of ammonia-oxidizing archaea and ammonia-oxidizing bacteria to ammonia oxidation in two nitrifying reactors. *Environmental Science and Pollution Research*, *25*(9), 8676–8687.

Srivastava, A., Nagar, P., Rathore, S., & Adlakha, N. (2022). The Linker Region Promotes Activity and Binding Efficiency of Modular LPMO towards Polymeric Substrate. *Microbiology Spectrum*, *10*(1).

Stahl, D. A., & de la Torre, J. R. (2012). Physiology and Diversity of Ammonia-Oxidizing Archaea. *Annual Review of Microbiology*, *66*(1), 83–101.

Stein, L. Y. (2019). Insights into the physiology of ammonia-oxidizing microorganisms. *Current Opinion in Chemical Biology*, *49*, 9–15.

Stein, L. Y., & Klotz, M. G. (2016a). The nitrogen cycle. *Current Biology*, *26*(3), R94–R98.

Stempfhuber, B., Engel, M., Fischer, D., Neskovic-Prit, G., Wubet, T., Schöning, I., Gubry-Rangin, C., Kublik, S., Schloter-Hai, B., Rattei, T., Welzl, G., Nicol, G. W., Schrumpf, M., Buscot, F., Prosser, J. I., & Schloter, M. (2015). pH as a Driver for Ammonia-Oxidizing Archaea in Forest Soils. *Microbial Ecology*, *69*(4), 879–883.

Stepnov, A. A., Christensen, I. A., Forsberg, Z., Aachmann, F. L., Courtade, G., & Eijsink, V. G. H. (2022). The impact of reductants on the catalytic efficiency of a lytic polysaccharide monooxygenase and the special role of dehydroascorbic acid. *FEBS Letters*, *596*(1), 53–70.

Stepnov, A. A., Eijsink, V. G. H., & Forsberg, Z. (2022). Enhanced in situ H2O2 production explains synergy between an LPMO with a cellulose-binding domain and a single-domain LPMO. *Scientific Reports*, *12*(1), 6129.

Sterngren, A. E., Hallin, S., & Bengtson, P. (2015). Archaeal Ammonia Oxidizers Dominate in Numbers, but Bacteria Drive Gross Nitrification in N-amended Grassland Soil. *Frontiers in Microbiology*, *6*.

Stoll, S., & Schweiger, A. (2006). EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *Journal of Magnetic Resonance*, *178*(1), 42–55.

Straka, L. L., Meinhardt, K. A., Bollmann, A., Stahl, D. A., & Winkler, M.-K. H. (2019). Affinity informs environmental cooperation between ammonia-oxidizing archaea (AOA) and anaerobic ammonia-oxidizing (Anammox) bacteria. *The ISME Journal*, *13*(8), 1997–2004.

Strandberg, L., & Enfors, S. O. (1991). Factors influencing inclusion body formation in the production of a fused protein in Escherichia coli. *Applied and Environmental Microbiology*, *57*(6), 1669–1674.

Sun, R., Myrold, D. D., Wang, D., Guo, X., & Chu, H. (2019). AOA and AOB communities respond differently to changes of soil pH under long-term fertilization. *Soil Ecology Letters*, *1*(3–4), 126–135.

Svergun, D. I. (1999). Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. *Biophysical Journal*, *76*(6), 2879–2886.

Svistunenko, D. A., Davies, N., Brealey, D., Singer, M., & Cooper, C. E. (2006). Mitochondrial dysfunction in patients with severe sepsis: An EPR interrogation of individual respiratory chain components. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1757*(4), 262–272.

Tago, K., Okubo, T., Shimomura, Y., Kikuchi, Y., Hori, T., Nagayama, A., & Hayatsu, M. (2015). Environmental Factors Shaping the Community Structure of Ammonia-Oxidizing Bacteria and Archaea in Sugarcane Field Soil. *Microbes and Environments*, *30*(1), 21–28.

Takamura, E., Taki, S., Sakamoto, H., Satomura, T., Sakuraba, H., Ohshima, T., & Suye, S. (2021). Site-Directed Mutagenesis of Multicopper Oxidase from Hyperthermophilic Archaea for High-Voltage Biofuel Cells. *Applied Biochemistry and Biotechnology*, *193*(2), 492–501.

Tamburrini, K. C., Terrapon, N., Lombard, V., Bissaro, B., Longhi, S., & Berrin, J.-G. (2021). Bioinformatic Analysis of Lytic Polysaccharide Monooxygenases Reveals the Pan-Families Occurrence of Intrinsically Disordered C-Terminal Extensions. *Biomolecules*, *11*(11), 1632.

Tandrup, T., Frandsen, K. E. H., Johansen, K. S., Berrin, J.-G., & lo Leggio, L. (2018). Recent insights into lytic polysaccharide monooxygenases (LPMOs). *Biochemical Society Transactions*, *46*(6), 1431–1447.

Tang, W., Ward, B. B., Beman, M., Bristow, L., Clark, D., Fawcett, S., Frey, C., Fripiat, F., Herndl, G. J., Mdutyana, M., Paulot, F., Peng, X., Santoro, A. E., Shiozaki, T., Sintes, E., Stock, C., Sun, X., Wan, X. S., Xu, M. N., & Zhang, Y. (2023). Database of nitrification and nitrifiers in the global ocean. *Earth System Science Data*, *15*(11), 5039–5077.

Tanizawa, Y., Fujisawa, T., Kaminuma, E., Nakamura, Y., & Arita, M. (2016). DFAST and DAGA: web-based integrated genome annotation tools and resources. *Bioscience of Microbiota, Food and Health*, *35*(4), 173–184.

Taylor, A. E., Giguere, A. T., Zoebelein, C. M., Myrold, D. D., & Bottomley, P. J. (2017). Modeling of soil nitrification responses to temperature reveals thermodynamic differences between ammonia-oxidizing activity of archaea and bacteria. *The ISME Journal*, *11*(4), 896–908.

Teira, E., Reinthaler, T., Pernthaler, A., Pernthaler, J., & Herndl, G. J. (2004). Combining Catalyzed Reporter Deposition-Fluorescence In Situ Hybridization and Microautoradiography To Detect Substrate Utilization by Bacteria and Archaea in the Deep Ocean. *Applied and Environmental Microbiology*, *70*(7), 4411–4414.

Threatt, S. D., & Rees, D. C. (2023). Biological nitrogen fixation in theory, practice, and reality: a perspective on the molybdenum nitrogenase system. *FEBS Letters*, *597*(1), 45–58.

Tourna, M., Freitag, T. E., Nicol, G. W., & Prosser, J. I. (2008). Growth, activity and temperature responses of ammonia-oxidizing archaea and bacteria in soil microcosms. *Environmental Microbiology*, *10*(5), 1357–1364.

Tourna, M., Stieglmeier, M., Spang, A., Könneke, M., Schintlmeister, A., Urich, T., Engel, M., Schloter, M., Wagner, M., Richter, A., & Schleper, C. (2011a). *Nitrososphaera viennensis* , an ammonia oxidizing archaeon from soil. *Proceedings of the National Academy of Sciences*, *108*(20), 8420–8425.

Tria, G., Mertens, H. D. T., Kachala, M., & Svergun, D. I. (2015). Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ*, *2*(2), 207–217.

Tuveng, T. R., Jensen, M. S., Fredriksen, L., Vaaje-Kolstad, G., Eijsink, V. G. H., & Forsberg, Z. (2020). A thermostable bacterial lytic polysaccharide monooxygenase with high operational stability in a wide temperature range. *Biotechnology for Biofuels*, *13*(1), 194.

Underwood, G. J. C., Dumbrell, A. J., McGenity, T. J., McKew, B. A., & Whitby, C. (2022). *The Microbiome of Coastal Sediments* (pp. 479–534).

Urakawa, H., Garcia, J. C., Nielsen, J. L., Le, V. Q., Kozlowski, J. A., Stein, L. Y., Lim, C. K., Pommerening-Röser, A., Martens-Habbena, W., Stahl, D. A., & Klotz, M. G. (2015). Nitrosospira lacus sp. nov., a psychrotolerant, ammonia-oxidizing bacterium from sandy lake sediment. *International Journal of Systematic and Evolutionary Microbiology*, *65*(Pt_1), 242–250.

Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sørlie, M., & Eijsink, V. G. H. (2010). An Oxidative Enzyme Boosting the Enzymatic Conversion of Recalcitrant Polysaccharides. *Science*, *330*(6001), 219–222.

van Kessel, M. A. H. J., Speth, D. R., Albertsen, M., Nielsen, P. H., Op den Camp, H. J. M., Kartal, B., Jetten, M. S. M., & Lücker, S. (2015). Complete nitrification by a single microorganism. *Nature*, *528*(7583), 555–559.

van Rosmalen, M., Krom, M., & Merkx, M. (2017). Tuning the Flexibility of Glycine-Serine Linkers To Allow Rational Design of Multidomain Proteins. *Biochemistry*, *56*(50), 6565–6574.

van Wezel, G. P., Krabben, P., Traag, B. A., Keijser, B. J. F., Kerste, R., Vijgenboom, E., Heijnen, J. J., & Kraal, B. (2006). Unlocking *Streptomyces* spp. for Use as Sustainable Industrial Production Platforms by Morphological Engineering. *Applied and Environmental Microbiology*, *72*(8), 5283–5288.

van Wolferen, M., Pulschen, A. A., Baum, B., Gribaldo, S., & Albers, S.-V. (2022). The cell biology of archaea. *Nature Microbiology*, *7*(11), 1744–1755.

Vandhana, T. M., Reyre, J., Sushmaa, D., Berrin, J., Bissaro, B., & Madhuprakash, J. (2022). On the expansion of biological functions of lytic polysaccharide monooxygenases. *New Phytologist*, *233*(6), 2380–2396.

Vílchez, R., Pozo, C., Gómez, M. A., Rodelas, B., & González-López, J. (2007). Dominance of sphingomonads in a copper-exposed biofilm community for groundwater treatment. *Microbiology*, *153*(2), 325–337.

Voland, R. W., Wang, H., Abruña, H. D., & Lancaster, K. M. (2025). Nitrous oxide production via enzymatic nitroxyl from the nitrifying archaeon *Nitrosopumilus maritimus*. *Proceedings of the National Academy of Sciences*, *122*(3).

Volkov, V. v., & Svergun, D. I. (2003). Uniqueness of *ab initio* shape determination in small-angle scattering. *Journal of Applied Crystallography*, *36*(3), 860–864.

Votvik, A. K., Røhr, Å. K., Bissaro, B., Stepnov, A. A., Sørlie, M., Eijsink, V. G. H., & Forsberg, Z. (2023). Structural and functional characterization of the catalytic domain of a cell-wall anchored bacterial lytic polysaccharide monooxygenase from Streptomyces coelicolor. *Scientific Reports*, *13*(1), 5345.

Vu, T. T. T., Jeong, B., Yu, J., Koo, B.-K., Jo, S.-H., Robinson, R. C., & Choe, H. (2014). Soluble prokaryotic expression and purification of crotamine using an N-terminal maltose-binding protein tag. *Toxicon*, *92*, 157–165.

Vuong, P., Moreira-Grez, B., Wise, M. J., Whiteley, A. S., Kumaresan, D., & Kaur, P. (2022). From rags to enriched: metagenomic insights into ammonia-oxidizing archaea following ammonia enrichment of a denuded oligotrophic soil ecosystem. *Environmental Microbiology*, *24*(7), 3097–3110.

Walker, C. B., de la Torre, J. R., Klotz, M. G., Urakawa, H., Pinel, N., Arp, D. J., Brochier-Armanet, C., Chain, P. S. G., Chan, P. P., Gollabgir, A., Hemp, J., Hügler, M., Karr, E. A., Könneke, M., Shin, M., Lawton, T. J., Lowe, T., Martens-Habbena, W., Sayavedra-Soto, L. A., … Stahl, D. A. (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proceedings of the National Academy of Sciences*, *107*(19), 8818–8823.

Walker, C. B., de La Torre, J. R., Klotz, M. G., Urakawa, H., Pinel, N., Arp, D. J., Brochier-Armanet, C., Chain, P. S. G., Chan, P. P., Gollabgir, A., Hemp, J., Hügler, M., Karr, E. A., Könneke, M., Shin, M., Lawton, T. J., Lowe, T., Martens-Habbena, W., Sayavedra-Soto, L. A., … Stahl, D. A. (2010). Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(19).

Walsh, G. (2010). Post-translational modifications of protein biopharmaceuticals. *Drug Discovery Today*, *15*(17–18), 773–780.

Walton, P. H., & Davies, G. J. (2016). On the catalytic mechanisms of lytic polysaccharide monooxygenases. *Current Opinion in Chemical Biology*, *31*, 195–207.

Walton, P. H., Davies, G. J., Diaz, D. E., & Franco-Cairo, J. P. (2023). The histidine brace: nature's copper alternative to haem? *FEBS Letters*, *597*(4), 485–494.

Wan, X. S., Hou, L., Kao, S.-J., Zhang, Y., Sheng, H.-X., Shen, H., Tong, S., Qin, W., & Ward, B. B. (2023). Pathways of N_2O production by marine ammonia-oxidizing archaea determined from dual-isotope labeling. *Proceedings of the National Academy of Sciences*, *120*(11).

Wan, X. S., Sheng, H., Shen, H., Zou, W., Tang, J., Qin, W., Dai, M., Kao, S., & Ward, B. B. (2024). Significance of Urea in Sustaining Nitrite Production by Ammonia Oxidizers in the Oligotrophic Ocean. *Global Biogeochemical Cycles*, *38*(10).

Wang, B., Walton, P. H., & Rovira, C. (2019). Molecular Mechanisms of Oxygen Activation and Hydrogen Peroxide Formation in Lytic Polysaccharide Monooxygenases. *ACS Catalysis*, *9*(6), 4958–4969.

Wang, B., Wang, Z., Davies, G. J., Walton, P. H., & Rovira, C. (2020). Activation of O_2 and H_2O_2 by Lytic Polysaccharide Monooxygenases. *ACS Catalysis*, *10*(21), 12760–12769.

Wang, B., Zhao, J., Guo, Z., Ma, J., Xu, H., & Jia, Z. (2015). Differential contributions of ammonia oxidizers and nitrite oxidizers to nitrification in four paddy soils. *The ISME Journal*, *9*(5), 1062–1075.

Wang, J., Kan, J., Zhang, X., Xia, Z., Zhang, X., Qian, G., Miao, Y., Leng, X., & Sun, J. (2017). Archaea Dominate the Ammonia-Oxidizing Community in Deep-Sea Sediments of the Eastern Indian Ocean—from the Equator to the Bay of Bengal. *Frontiers in Microbiology*, *8*.

Wang, M., Wang, S., Long, X., Zhuang, L., Zhao, X., Jia, Z., & Zhu, G. (2019). High contribution of ammonia-oxidizing archaea (AOA) to ammonia oxidation related to a potential active AOA species in various arable land soils. *Journal of Soils and Sediments*, *19*(3), 1077–1087.

Wang, Q., Li, S., Li, J., & Huang, D. (2024). The Utilization and Roles of Nitrogen in Plants. *Forests*, *15*(7), 1191.

Ward, L. M., Johnston, D. T., & Shih, P. M. (2021). Phanerozoic radiation of ammonia oxidizing bacteria. *Scientific Reports*, *11*(1), 2070.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, *25*(9), 1189–1191.

Weik, M., & Colletier, J.-P. (2010). Temperature-dependent macromolecular X-ray crystallography. *Acta Crystallographica Section D Biological Crystallography*, *66*(4), 437–446.

Wilkinson, D. L., & Harrison, R. G. (1991). Predicting the Solubility of Recombinant Proteins in Escherichia coli. *Nature Biotechnology*, *9*(5), 443–448.

Wilson, M. A. (2022). Mapping Enzyme Landscapes by Time-Resolved Crystallography with Synchrotron and X-Ray Free Electron Laser Light. *Annual Review of Biophysics*, *51*(1), 79–98.

Winkler, J., Seybert, A., König, L., Pruggnaller, S., Haselmann, U., Sourjik, V., Weiss, M., Frangakis, A. S., Mogk, A., & Bukau, B. (2010). Quantitative and spatio-temporal features of protein aggregation in Escherichia coli and consequences on protein quality control and cellular ageing. *The EMBO Journal*, *29*(5), 910–923.

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences*, *87*(12), 4576–4579.

Wolff, A. M., Nango, E., Young, I. D., Brewster, A. S., Kubo, M., Nomura, T., Sugahara, M., Owada, S., Barad, B. A., Ito, K., Bhowmick, A., Carbajo, S., Hino, T., Holton, J. M., Im, D., O'Riordan, L. J., Tanaka, T., Tanaka, R., Sierra, R. G., … Thompson, M. C. (2023). Mapping protein dynamics at high spatial resolution with temperature-jump X-ray crystallography. *Nature Chemistry*, *15*(11), 1549–1558.

Woo, Y., Cruz, M. C., & Wuertz, S. (2022). Selective Enrichment of Nitrososphaera viennensis-Like Ammonia-Oxidizing Archaea over Ammonia-Oxidizing Bacteria from Drinking Water Biofilms. *Microbiology Spectrum*, *10*(6).

Wright, C. L., & Lehtovirta-Morley, L. E. (2023). Nitrification and beyond: metabolic versatility of ammonia oxidising archaea. *The ISME Journal*, *17*(9), 1358–1368.

Wright, C. L., Schatteman, A., Crombie, A. T., Murrell, J. C., & Lehtovirta-Morley, L. E. (2020). Inhibition of Ammonia Monooxygenase from Ammonia-Oxidizing Archaea by Linear and Aromatic Alkynes. *Applied and Environmental Microbiology*, *86*(9).

Wu, L., Chen, X., Wei, W., Liu, Y., Wang, D., & Ni, B.-J. (2020). A Critical Review on Nitrous Oxide Production by Ammonia-Oxidizing Archaea. *Environmental Science & Technology*, *54*(15), 9175–9190.

Xi, R., Long, X.-E., Huang, S., & Yao, H. (2017). pH rather than nitrification and urease inhibitors determines the community of ammonia oxidizers in a vegetable soil. *AMB Express*, *7*(1), 129.

Xiao, R., Qiu, Y., Tao, J., Zhang, X., Chen, H., Reberg-Horton, S. C., Shi, W., Shew, H. D., Zhang, Y., & Hu, S. (2020). Biological controls over the abundances of terrestrial ammonia oxidizers. *Global Ecology and Biogeography*, *29*(2), 384–399.

Xu, A., Li, L., Xie, J., Gopalakrishnan, S., Zhang, R., Luo, Z., Cai, L., Liu, C., Wang, L., Anwar, S., & Jiang, Y. (2022). Changes in Ammonia-Oxidizing Archaea and Bacterial Communities and Soil Nitrogen Dynamics in Response to Long-Term Nitrogen Fertilization. *International Journal of Environmental Research and Public Health*, *19*(5), 2732.

Xu, S.-Q., Qian, X.-X., Jiang, Y.-H., Qin, Y.-L., Zhang, F.-Y., Zhang, K.-Y., Hong, Q., He, J., Miao, L.-L., Liu, Z.-P., Li, D.-F., Liu, S.-J., & Qiu, J.-G. (2022). Genetic Foundations of Direct Ammonia Oxidation (Dirammox) to N_2 and MocR-Like Transcriptional Regulator DnfR in Alcaligenes faecalis Strain JQ135. *Applied and Environmental Microbiology*, *88*(6).

Yamamoto, N., Otawa, K., & Nakai, Y. (2010). Diversity and Abundance of Ammonia-Oxidizing Bacteria and Ammonia-Oxidizing Archaea During Cattle Manure Composting. *Microbial Ecology*, *60*(4), 807–815.

Yan, L., Li, Z., Bao, J., Wang, G., Wang, C., & Wang, W. (2015). Diversity of ammonia-oxidizing bacteria and ammonia-oxidizing archaea during composting of municipal sludge. *Annals of Microbiology*, *65*(3), 1729–1739.

Yang, Y., Liu, H., Zhang, Y., Fang, X., Zhong, X., & Lv, J. (2023). Contribution of ammonia-oxidizing archaea and bacteria to nitrogen transformation in a soil fertilized with urea and organic amendments. *Scientific Reports*, *13*(1), 20722

Yasuda, S., Palomo, A., Smets, B. F., & Terada, A. (2026). Potential survival strategies of novel comammox and nitrite-oxidizing Nitrospira synthesizing osmoprotectants in a wastewater microbiome treating high-ammonia brackish landfill leachate. *Microbiome*, *14*(1), 98.

Yin, C., Tan, C., Chen, H., Ye, M., Fan, X., Zheng, W., Gao, Z., Peng, H., & Liang, Y. (2022). Insight into the role of competition in niche differentiation between ammonia-oxidizing archaea and bacteria in ammonium-rich alkaline soil: A network-based study. *Soil Biology and Biochemistry*, *168*, 108638.

Yin, Z., Bi, X., & Xu, C. (2018a). Ammonia-Oxidizing Archaea (AOA) Play with Ammonia-Oxidizing Bacteria (AOB) in Nitrogen Removal from Wastewater. *Archaea*, *2018*, 1–9.

Yin, Z., Bi, X., & Xu, C. (2018b). Ammonia-Oxidizing Archaea (AOA) Play with Ammonia-Oxidizing Bacteria (AOB) in Nitrogen Removal from Wastewater. *Archaea*, *2018*, 1–9.

Yokamo, S., Irfan, M., Huan, W., Wang, B., Wang, Y., Ishfaq, M., Lu, D., Chen, X., Cai, Q., & Wang, H. (2023). Global evaluation of key factors influencing nitrogen fertilization efficiency in wheat: a recent meta-analysis (2000-2022). *Frontiers in Plant Science*, *14*.

You, X.-Y., Liu, C., Wang, S.-Y., Jiang, C.-Y., Shah, S. A., Prangishvili, D., She, Q., Liu, S.-J., & Garrett, R. A. (2011). Genomic analysis of Acidianus hospitalis W1 a host for studying crenarchaeal virus and plasmid life cycles. *Extremophiles*, *15*(4), 487–497.

Zayed, O., Hewedy, O. A., Abdelmoteleb, A., Ali, M., Youssef, M. S., Roumia, A. F., Seymour, D., & Yuan, Z.-C. (2023). Nitrogen Journey in Plants: From Uptake to Metabolism, Stress Response, and Microbe Interaction. *Biomolecules*, *13*(10), 1443.

Zeng, H., & Yang, A. (2019). Quantification of proteomic and metabolic burdens predicts growth retardation and overflow metabolism in recombinant *Escherichia coli*. *Biotechnology and Bioengineering*, *116*(6), 1484–1495.

Zhalnina, K., de Quadros, P. D., Camargo, F. A. O., & Triplett, E. W. (2012). Drivers of archaeal ammonia-oxidizing communities in soil. *Frontiers in Microbiology*, *3*.

Zhalnina, K. v., Dias, R., Leonard, M. T., Dorr de Quadros, P., Camargo, F. A. O., Drew, J. C., Farmerie, W. G., Daroub, S. H., & Triplett, E. W. (2014). 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*(7), e101648.

Zhang, B., Gu, H., Yang, Y., Bai, H., Zhao, C., Si, M., Su, T., & Shen, X. (2019). Molecular Mechanisms of AhpC in Resistance to Oxidative Stress in Burkholderia thailandensis. *Frontiers in Microbiology*, *10*.

Zhang, H., Zhou, H., Zhao, Y., Li, T., & Yin, H. (2023). Comparative studies of two AA10 family lytic polysaccharide monooxygenases from *Bacillus thuringiensis*. *PeerJ*, *11*, e14670.

Zhang, I. H., Sun, X., Jayakumar, A., Fortin, S. G., Ward, B. B., & Babbin, A. R. (2023). Partitioning of the denitrification pathway and other nitrite metabolisms within global oxygen deficient zones. *ISME Communications*, *3*(1).

Zhang, L., Koay, M., Maher, M. J., Xiao, Z., & Wedd, A. G. (2006). Intermolecular Transfer of Copper Ions from the CopC Protein of *Pseudomonas syringae* . Crystal Structures of Fully Loaded Cu^{I} Cu^{II} Forms. *Journal of the American Chemical Society*, *128*(17), 5834–5850.

Zhang, L.-M., Hu, H.-W., Shen, J.-P., & He, J.-Z. (2012). Ammonia-oxidizing archaea have more important role than ammonia-oxidizing bacteria in ammonia oxidation of strongly acidic soils. *The ISME Journal*, *6*(5), 1032–1045.

Zheng, Y., Wang, B., Gao, P., Yang, Y., Xu, B., Su, X., Ning, D., Tao, Q., Li, Q., Zhao, F., Wang, D., Zhang, Y., Li, M., Winkler, M.-K. H., Ingalls, A. E., Zhou, J., Zhang, C., Stahl, D. A., Jiang, J., … Qin, W. (2024b). Novel order-level lineage of ammonia-oxidizing archaea widespread in marine and terrestrial environments. *The ISME Journal*, *18*(1).

Zhong, X., Zhang, L., van Wezel, G. P., Vijgenboom, E., & Claessen, D. (2022). Role for a Lytic Polysaccharide Monooxygenase in Cell Wall Remodeling in Streptomyces coelicolor. *mBio*, *13*(2).

Zhong, Y., Yan, W., Canisares, L. P., Wang, S., & Brodie, E. L. (2023). Alterations in soil pH emerge as a key driver of the impact of global change on soil microbial nitrogen cycling: Evidence from a global meta-analysis. *Global Ecology and Biogeography*, *32*(1), 145–165.

Zhou, X., & Zhu, H. (2020). Current understanding of substrate specificity and regioselectivity of LPMOs. *Bioresources and Bioprocessing*, *7*(1), 11.

Appendix A List of Tables

Table A.1. FWM components for AOA enrichment of soil derived AOA

Salts	**Final Concentration (mol/L)**
NaCl	0.017
$MgCl_2$ $6H_2O$	0.0019
$CaCl_2$ $2H_2O$	6.8×10^{-4}
KCl	0.0067
$NaHCO_3$	0.002
Nutrients	
NH_4Cl	0.002
$KHPO_4$	0.0014
Buffer	
HEPES (pH 7.5)	0.01
FeNaEDTA solution	7.5×10^{-6}
Non-Chelated Trace Elements Mixture	
HCl	1×10^{-3}
H_3BO_3	5×10^{-7}
$MnCl_2$ $4H_2O$	5×10^{-7}
$CoCl_2$ $6H_2O$	8×10^{-7}
$NiCl_2$ $6H_2O$	1×10^{-7}
$ZnSO_4$ $7H_2O$	5×10^{-7}
Na_2MoO_4 $2H_2O$	1.5×10^{-7}
$CuCl_2$ H_2O[a]	1.0×10^{-8}
Vitamins	
Biotin	8×10^{-8}
Folic acid	5×10^{-8}
Pyrodoxine HCl	4.8×10^{-7}
Thiamine HCl	1.4×10^{-7}
Riboflavin	1.3×10^{-7}
Nicotinic acid	4.0×10^{-7}
DL Panthothenic acid	2.2×10^{-7}
p-aminobenzoic acid	3.6×10^{-7}
Choline chloride	1.4×10^{-5}
Vitamin B12	7.3×10^{-9}
Antibiotics	
Ampicillin Sodium	1.35×10^{-4} (50 µg/mL)
Streptomycin Sulfate	6.86×10^{-5} (50 µg/mL)

Table A.2. Amino Acid identity of amo protein matches identified in C2 and C4-2 MAGs using stand-alone Blast analysis.

MAG	Target Gene	Reference Accession	Reference Taxon	% Identity	Reference
C2	amoB	VFJ12993	*Ca. Nitrosocosmicus franklandus* (C13)	100	(Lehtovirta-Morley, *et al*., 2016)
	amoC	VFJ14124	*Ca. Nitrosocosmicus franklandus* (C13)	100	(Lehtovirta-Morley, *et al*., 2016)
	amoX	VFJ12998	*Ca. Nitrosocosmicus franklandus* (C13)	100	(Lehtovirta-Morley, *et al*., 2016)
C4-2	amoB	KAF0869013	*Ca. Nitrosocosmicus agrestis* (SS)	100	(Liu, *et al*., 2021)
	amoC	GKS61403	*Ca. Nitrosocosmicus sp*.	100	(Yasuda, *et al*., 2026)
	amoX	VFJ12998	*Ca. Nitrosocosmicus franklandus* (C13)	94.6	(Lehtovirta-Morley, *et al*., 2016)

Table A.3. Taxonomic classification and genome quality assessment of C2 MAGs using GTDB-Tk and CheckM2. ANI = Average nucleotide identity and AF = Alignment fraction

MAG/Bin	GTDB-Tk Classification	ANI (%)	AF	Completeness (%)	Contamination (%)	Genome Size (Mb)	GC (%)	Total Contigs	N50 (bp)
C2 (Bin.9)	d_Archaea; s_Nitrosocosmicus franklandus	99.98	0.9758	99.35	0.03	2.78	34	117	37,242
Bin12	d_Bacteria; s_Bosea thiooxidans; sp011764485	99.07	0.9286	90.70	1.15	5.17	67	129	58,275
Bin13	d_Bacteria; s_Reyranella massiliensis; sp020160995	98.80	0.8397	25.24	1.80	1.24	65	3	38,863
Bin5	d_Bacteria; s_Reyranella massiliensis; sp020160995	98.45	0.9463	72.69	0.40	4.56	65	29	255,971
Bin7	d_Bacteria; s_Mesorhizobium terrae	98.91	0.8658	90.94	3.20	4.40	65	65	116,116
Bin1	d_Bacteria; f_Iamiaceae; g_Iamia	N/A	N/A	87.22	0.88	4.02	74	585	7,992
Bin10	d_Bacteria; f_Nocardioidaceae; g_Nocardioides	N/A	N/A	61.55	50.34	3.01	71	628	4,945
Bin11	d_Bacteria; f_Sphingomonadaceae; g_Sphingomicrobium	N/A	N/A	99.85	0.04	2.22	64	8	342,845
Bin2	d_Bacteria; f_Solirubrobacteraceae; g_Conexibacter	N/A	N/A	75.57	2.50	3.57	72	666	5,726
Bin3	d_Bacteria; f_Caulobacteraceae; g_Phenylobacterium	N/A	N/A	98.86	1.72	4.34	70	134	48,840
Bin4	d_Bacteria; f_Rhizobiaceae; g_Shinella	N/A	N/A	56.53	0.79	3.10	66	662	4,936
Bin6	d_Bacteria;; f_Nocardioidaceae; g_Nocardioides	N/A	N/A	99.69	1.98	4.39	72	129	55,115
Bin8	d_Bacteria; c_Phycisphaerae; o_Phycisphaerales	N/A	N/A	25.08	0.00	0.65	72	186	3,506

Table A.4. Taxonomic classification and genome quality assessment of C2 MAGs using GTDB-Tk and CheckM2. ANI = Average nucleotide identity and AF = Alignment fraction

MAG/Bin	GTDB-Tk Classification	Closest Reference Genome	ANI (%)	AF	Completeness (%)	Contamination (%)	Genome Size (Mb)	GC (%)	Total Contigs	N50 (bp)
C4-2 (Bin6)	d_Archaea; s_Nitrosocosmicus sp008389435	GCF_008389435.1	97.06	0.8039	98.99	0.33	3.19	33	106	49,422
Bin7	d_Bacteria; o_Phycisphaerales; sp023429345	GCA_023429345.1	96.88	0.5803	35.98	0.15	1.64	69	411	3,979
Bin8	d_Bacteria; f_Hyphomicrobiaceae; sp036380915	GCA_036380915.1	99.22	0.8102	42.58	0.37	1.78	63	479	3,727
Bin1	d_Bacteria; c_Limnocylindria; o_Limnocylindrales	N/A	N/A	N/A	92.55	1.80	3.57	73	233	21,197
Bin2	d_Bacteria; o_Planctomycetales; f_Planctomycetaceae	N/A	N/A	N/A	99.67	0.82	5.51	65	81	162,416
Bin3	d_Bacteria; f_Lacipirellulaceae; g_Bythopirellula	N/A	N/A	N/A	34.60	0.05	2.29	58	593	3,924
Bin4	d_Bacteria; c_Gemmatimonadetes; o_Longimicrobiales	N/A	N/A	N/A	12.10	0.03	0.52	68	163	3,087
Bin5	Unclassified	N/A	N/A	N/A	6.48	0.06	0.39	45	94	4,405

Table A.5. ICP-MS measurements of Cu^{2+} concentrations in FWM, washed AOA cells, trace metals with no Cu^{2+} addition and Cu^{2+} stock solution. Volume of Cu^{2+} stock was adjusted to ensure experimental cultures were at desired Cu^{2+} concentration.

Sample	Purpose	Cu^{2+} Concentration (µM)
FWM without added Cu^{2+}	Asses background Cu^{2+} contamination in media	<0.000
Trace metal mix with no added Cu^{2+}	Confirm trace metal contains no Cu^{2+}	<0.000
AOA inoculum (*N. franklandus* and C4-2 enrichment) cells after PBS wash	Confirm AOA cells do not carry over Cu^{2+}	<0.000
Cu^{2+} stock solution	To verify Cu^{2+} concentration of stock used for experiment	47.84

Table A.6. Primers for each protein target used in thesis.

Target	Forward Primer	Reverse Primer
Nmar_1131 pET26b	5′-ATTACCATGGCAGCAAGAGGAAAACACCTTC-3′	5′-TAATAAGCTTTCACGTACCGTAACCTTCCGA-3′
Nmar_1354 pET28a	5′-ATTACATATGCAATCTCCCGAAATCCAGACTACC-3′	5′- ATTAAAGCTTTCATCCACGGGTGCAGGG-3′
Nmar_1667 pET28a	5′-ATTACATATGCAGAGCCTGGATATGCAG-3′	5′-TAATAAGCTTTTAACTGGTCAAAATTTCGGCG-3′
NFRAN_2427 pET28a	5′-ATTACATATGGGCAAACAAGAAAGCAGCTTTAACG-3′	5′-TAATAAGCTTTTAATCCTGCGGTGTAAACGGATC-3′
***Sli*LPMO10F**	5′- ATTACCATGGCGCACGGCTGGGTGGTTCTCG -3′	5′-TAATAAGCTTTTAGTTGACGTCGATGCAGGCG-3′

Table A.7. The hyperfine interaction constants, assume to be isotropic for each of the three nitrogen's, used in the simulations of the four EPR spectra as shown in **Figure 6.7** and **Figure 6.8**.

	$A_{N1(iso)}$ MHz	$A_{N2(iso)}$ MHz	$A_{N3(iso)}$ MHz
*Sli*LPMO10C	35	35	35
*Sli*LPMO10F Component I	35	35	35
*Sli*LPMO10F Component II (identical to the two-domain *Sli*LPMO10G spectrum)	35	20.5	29.9
*Sli*LPMO10G (LPMO domain only)	35	35	35

Table A.8. *Sli*LPMO10E reduced in ascorbic acid at room temperature structure (XFEL). Crystallographic processing and refinement statistics for *Sli*LPMO10E. Values in parenthesis are for the highest resolution shell.

Space group	P1
Unit Cell (Å)	a = 33.6, b = 36.4, c = 62.2, 82.6° 78° 65°
Resolution (Å)	32.92 - 1.42 (1.4 – 1.4)
Unique Reflections	49771 (2487)
Completeness (%)	96.9 (100)
Mn(I/sd)	7.9 (2)
Wilson B-factor (Å^2)	0.614
$CC_{1/2}$	0.901 (0.631)
R_{work}	0.189
R_{free}	0.212
RMSD bond lengths (Å)	0.0130
RMSD bond angles (°)	1.955
Ramachandran most favoured (%)	97.61
PDB accession code	**TBC**

Table A.9. *Sli*LPMO10E reduced in cysteine at room temperature structure (XFEL). Crystallographic processing and refinement statistics for *Sli*LPMO10E. Values in parenthesis are for the highest resolution. shell.

Space group	P1
Unit Cell (Å)	a = 33.7, b = 36.4, c = 62.2, 82.6° 78° 65°
Resolution (Å)	32.93 - 1.42 (1.4 – 1.4)
Unique Reflections	49803 (2478)
Completeness (%)	96.8 (100)
Mn(I/sd)	8.5 (2.3)
Wilson B-factor (Å^2)	0.978
$CC_{1/2}$	0.923 (0.685)
R_{work}	0.185
R_{free}	0.220
RMSD bond lengths (Å)	0.0126
RMSD bond angles (°)	1.937
Ramachandran most favoured (%)	97.01
PDB accession code	**TBC**

Table A.10. *Sli*LPMO10E Soaked in HPBC photocage with addition of cysteine, room temperature (XFEL). Crystallographic processing and refinement statistics for *Sli*LPMO10E. Values in parenthesis are for the highest resolution shell.

Space group	P1
Unit Cell (Å)	a = 33.6, b = 36.4, c = 61.2, 82.6° 78.2° 65°
Resolution (Å)	32.97 - 1.42 (1.4 – 1.4)
Unique Reflections	48774 (2571)
Completeness (%)	96.7 (100)
Mn(I/sd)	10.5 (3.2)
Wilson B-factor (Å^2)	0.930
$CC_{1/2}$	0.938 (0.785)
R_{work}	0.182
R_{free}	0.208
RMSD bond lengths (Å)	0.0124
RMSD bond angles (°)	1.900
Ramachandran most favoured (%)	96.72
PDB accession code	**TBC**

Appendix B List of Figures

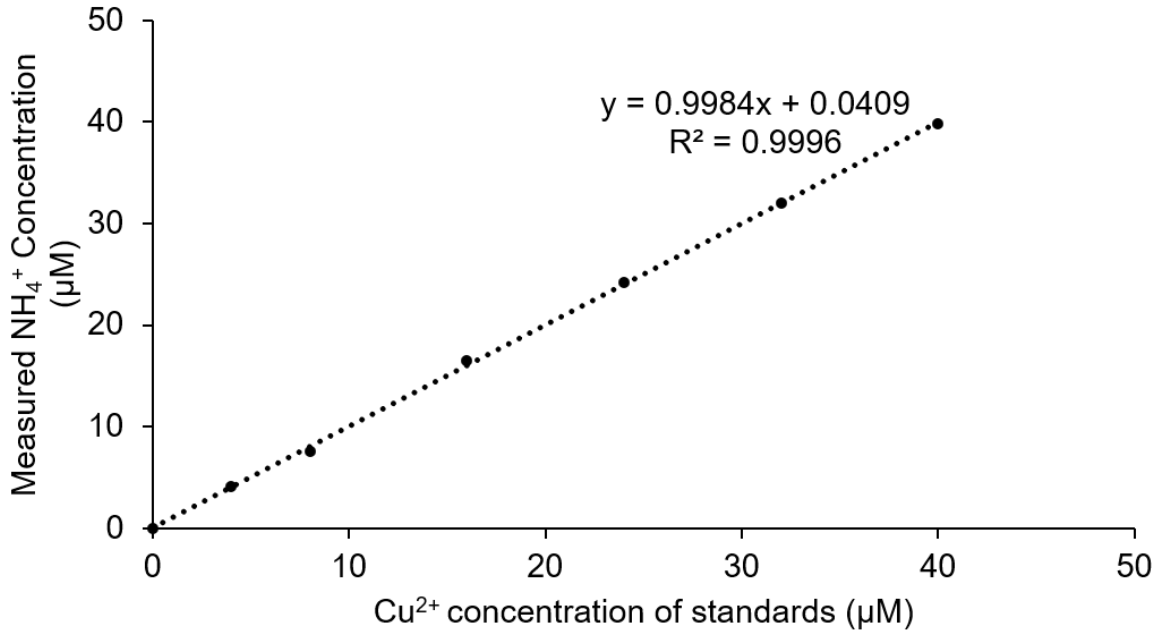


Figure B.1. Standard curve used for quantification of NH_4^+ during AOA physiology experiments. Known NH_4^+ standards were prepared across concentration ranges of 0 – 40 µM to generate standard curve. Samples over 40 µM were diluted to ensure measured values were within the range of the standard curve.

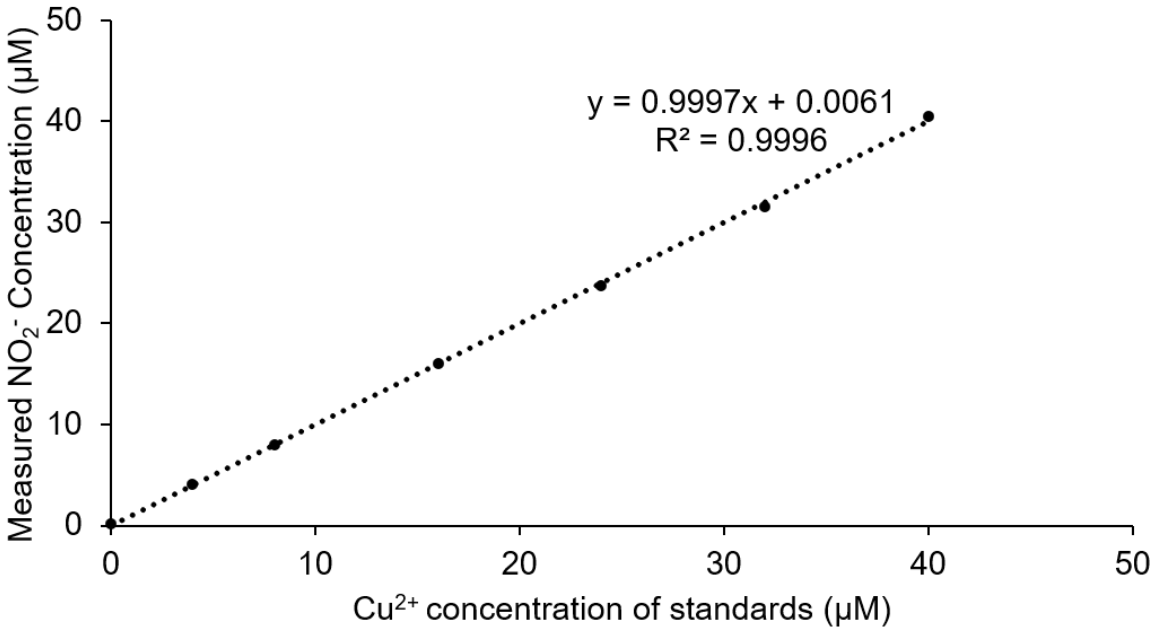


Figure B.2. Standard curve used for quantification of NO_2^- during AOA physiology experiments. Known NO_2^- standards were prepared across concentration ranges of 0 – 40 µM to generate standard curve. Samples over 40 µM were diluted to ensure measured values were within the range of the standard curve.

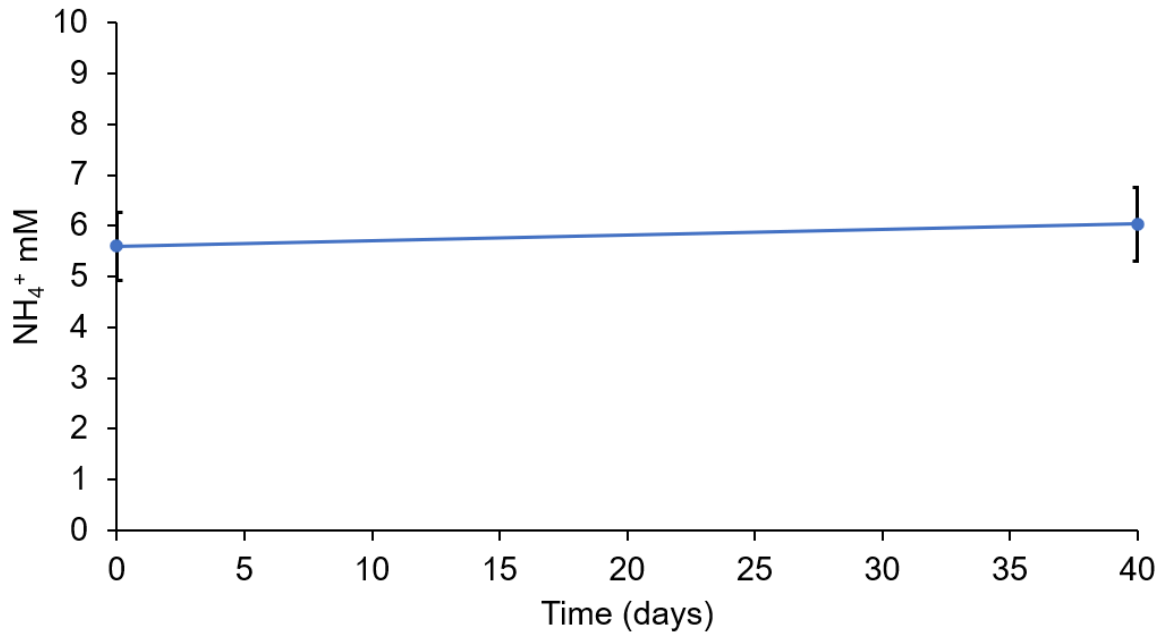


Figure B.3. Negative abiotic control for C4-2 physiology experiments. Media was supplemented with 5 mM NH_4^+ with no added AOA cells to determine whether NH_4^+ depleted and NO_2^- increased during incubation at 37°C.

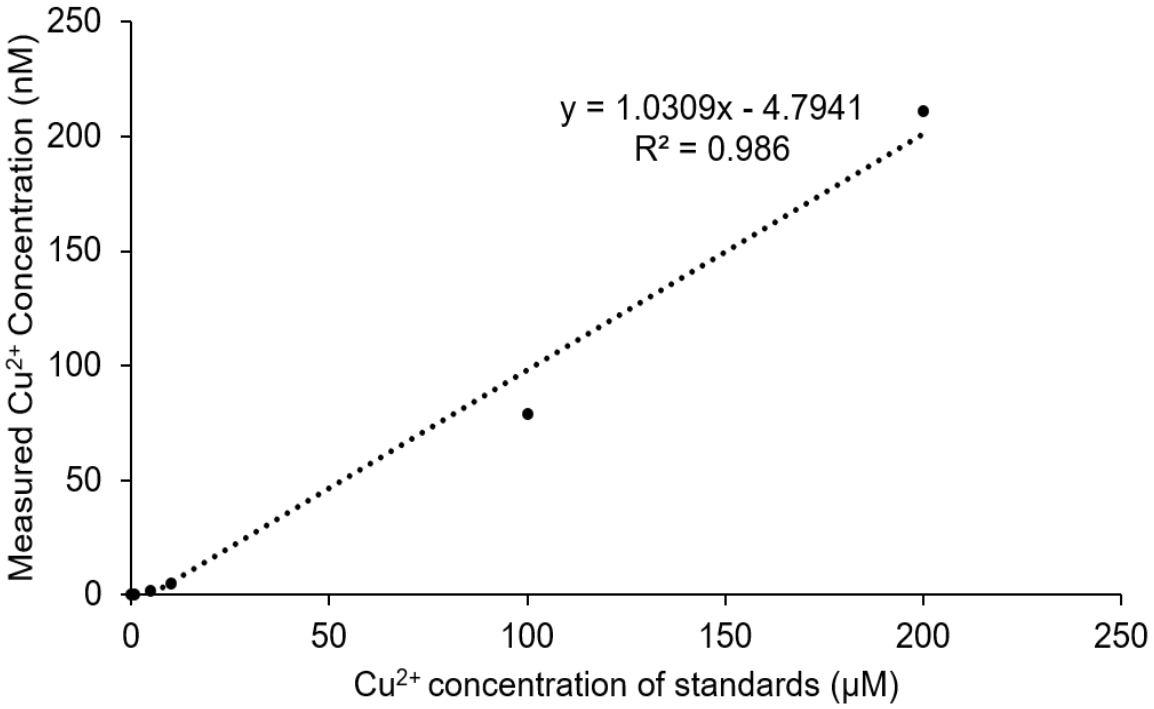


Figure B.4. ICP-MS copper standard curve. A series of Cu standard concentrations were plotted against measured Cu concentrations. Linear regression reveals a strong positive relationship between expected and measured Cu concentration, with $R^2 = 0.986$.

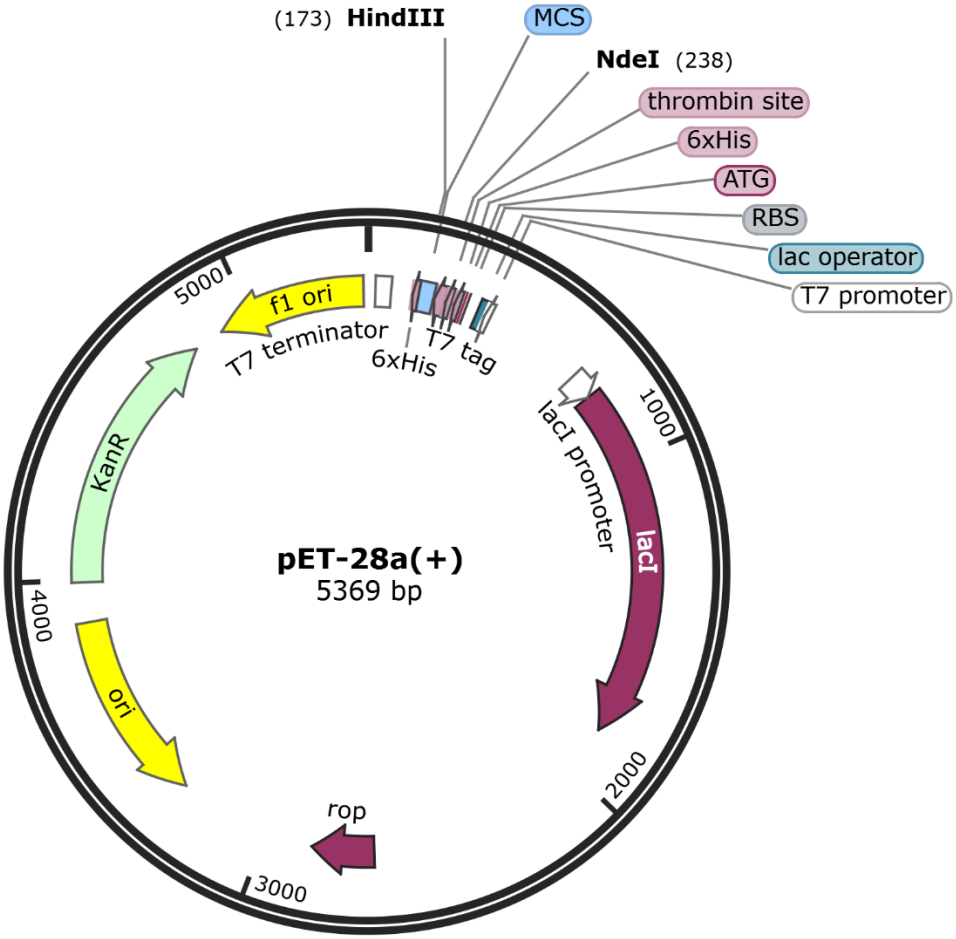


Figure B.5. Plasmid map of pET28a(+) (5369 bp) showing NdeI and HindIII restriction sites used for cloning. Map highlights vector features, including the T7 promoter site, lac operator site, ribosome-binding site (RBS), N-terminal 6xhis tag, thrombin cleavage site, multiple cloning site (MCS), kanamycin resistance gene, and origin of replication (ori).

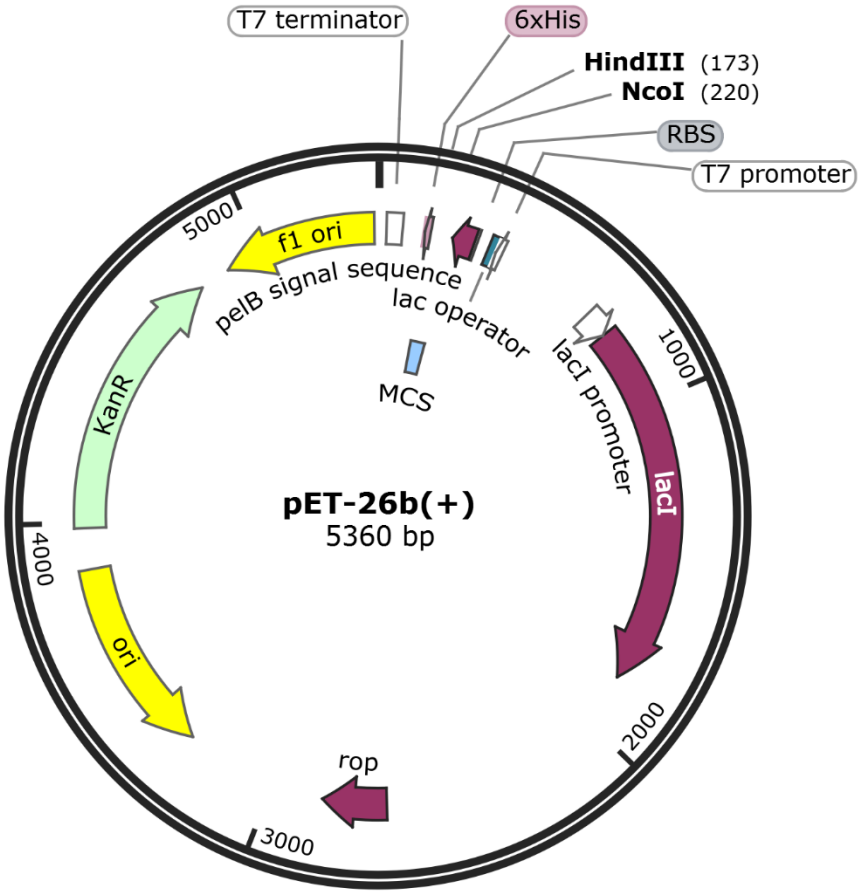


Figure B.6. Plasmid map of pET26b(+) (5360 bp) showing NcoI and HindIII restriction sites used for cloning. Map highlights vector features, including the T7 promoter site, lac operator site, ribosome-binding site (RBS), N-terminal pelB signal sequence, multiple cloning site (MCS), kanamycin resistance gene, and origin of replication (ori).

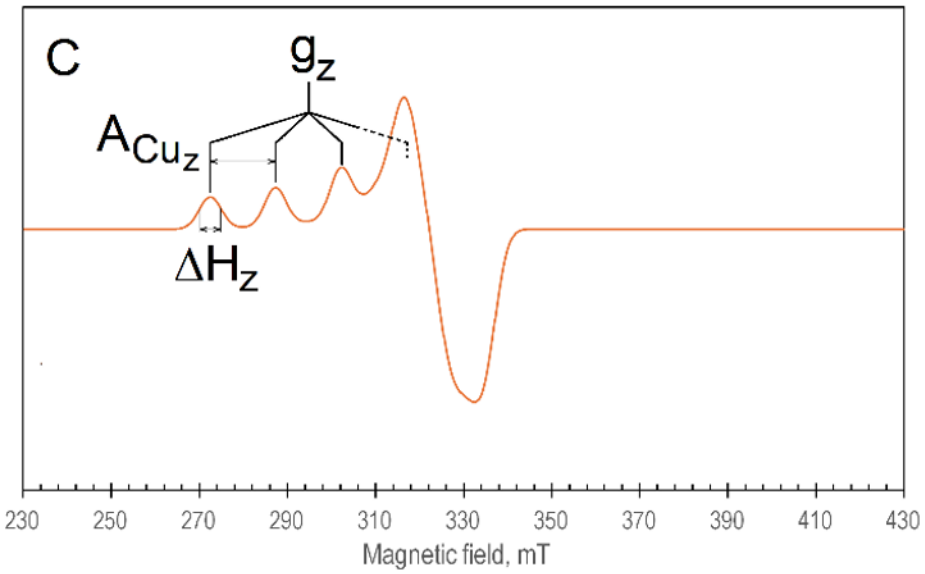


Figure B.7. A Cu^{2+} EPR spectrum example showing how the three basic parameters might be measured directly form the recorded line shape: g_z, A_{Cuz} and DH_z.