

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/147748>

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

Zinc regulation in an open ocean cyanobacterium

By

Alevtina Mikhaylina

A thesis submitted for the fulfilment of the degree of
Doctor of Philosophy in Chemistry

University of Warwick

Department of Chemistry and School of Life Sciences

September 2019

Table of contents

| | |
|--|-----------|
| Table of contents | II |
| List of Figures | VII |
| List of Tables | XIII |
| Acknowledgments | XVI |
| Declaration | XVII |
| Abstract | XVIII |
| List of abbreviations | XIX |
| Chapter 1. Introduction..... | 1 |
| 1.1 Cyanobacteria | 2 |
| 1.1.1 Marine <i>Synechococcus</i> | 3 |
| 1.2 Metals in cyanobacteria | 6 |
| 1.3 Zinc in cyanobacteria..... | 8 |
| 1.3.1 Zinc sensing in cyanobacteria | 8 |
| 1.3.1.1 Zinc uptake regulator | 10 |
| 1.3.1.1.1 <i>Zur</i> -boxes in cyanobacterial genomes..... | 11 |
| 1.3.1.1.2 <i>zur</i> -mutants | 12 |
| 1.3.2 Dealing with zinc limitation..... | 12 |
| 1.3.2.1 <i>Zur</i> as a transcriptional repressor | 12 |
| 1.3.2.1.1 Outer membrane zinc transport in cyanobacteria..... | 13 |
| 1.3.2.1.2 Inner membrane zinc uptake | 13 |
| 1.3.2.1.3 Putative zinc chaperones (the COG0523 family)..... | 15 |
| 1.3.2.1.4 Other members of the <i>Zur</i> regulon | 16 |
| 1.3.3 Dealing with zinc excess..... | 17 |
| 1.3.3.1 <i>Zur</i> as a transcriptional activator | 17 |
| 1.3.3.1.1 Cation-diffusion facilitator (CDF) | 17 |
| 1.3.3.1.2 <i>P</i> -type ATPases..... | 18 |
| 1.3.3.1.3 <i>RND</i> systems..... | 18 |
| 1.3.3.2 Bacterial metallothioneins..... | 19 |
| 1.4 Zinc metabolism in <i>Synechococcus</i> sp. WH8102..... | 21 |
| 1.4.1 Open questions regarding zinc homeostasis in <i>Synechococcus</i> sp. WH8102..... | 23 |
| 1.5 Aims and objectives..... | 26 |
| Chapter 2. General materials and methods..... | 28 |
| 2.1 <i>Escherichia coli</i> strains and growth conditions | 29 |
| 2.2 General molecular biology techniques | 31 |
| 2.2.1 Polymerase Chain Reaction (PCR)..... | 31 |
| 2.2.2 Agarose gel electrophoresis | 31 |
| 2.2.3 DNA purification following PCR | 32 |
| 2.3 Gene cloning techniques..... | 33 |
| 2.3.1 Digestion of DNA with restriction enzymes..... | 33 |
| 2.3.2 DNA dephosphorylation | 33 |
| 2.3.3 DNA ligation protocol | 33 |
| 2.3.4 Preparation of chemically competent <i>E. coli</i> cells..... | 34 |
| 2.3.5 <i>E. coli</i> transformation..... | 35 |
| 2.3.6 Plasmid isolation from <i>E. coli</i> | 35 |
| 2.3.7 Nanodrop measurements..... | 36 |
| 2.3.8 DNA sequencing | 36 |
| 2.4 Protein monitoring, buffer exchange and analysis | 37 |

| | | |
|--|--|------------|
| 2.4.1 | Sodium Dodecyl Sulfate - Polyacrylamide Gel Electrophoresis (SDS-PAGE)..... | 37 |
| 2.4.2 | Protein concentration and desalting | 37 |
| 2.4.3 | Inductively Coupled Plasma - Optical Emission Spectroscopy (ICP-OES) | 38 |
| 2.4.4 | Electrospray Ionization - Mass Spectrometry (ESI-MS) | 39 |
| Chapter 3. Properties and structure of the <i>Synechococcus</i> sp. WH8102 Zinc Uptake Regulator protein..... | | 40 |
| 3.1 | Introduction..... | 41 |
| 3.2 | Materials and methods | 42 |
| 3.2.1 | Zur purification | 42 |
| 3.2.1.1 | Heterologous over-expression of SynZur in <i>E. coli</i> | 42 |
| 3.2.1.2 | Fast Protein Liquid Chromatography (FPLC)..... | 43 |
| 3.2.1.3 | TEV cleavage of SynZur..... | 44 |
| 3.2.2 | Characterisation of SynZur | 44 |
| 3.2.2.1 | Electrophoretic mobility shift assays | 44 |
| 3.2.2.2 | Metal-responsive DNA binding..... | 46 |
| 3.2.2.3 | Non-denaturing Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis (ND SDS-PAGE) | 47 |
| 3.2.2.4 | Analytical Size Exclusion Chromatography (SEC) of SynZur | 47 |
| 3.2.2.5 | Zinc binding affinity of SynZur | 48 |
| 3.2.3 | Structural characterisation of SynZur | 48 |
| 3.3 | Results and discussion | 52 |
| 3.3.1 | Over-expression and purification of the SynZur protein | 52 |
| 3.3.1.1 | Heterologous recombinant expression of <i>Synechococcus</i> sp. WH8102 Zur | 53 |
| 3.3.1.2 | Protein purification using FPLC | 53 |
| 3.3.2 | Characterisation of SynZur | 57 |
| 3.3.2.1 | Elemental composition and molar absorption coefficient calculation | 57 |
| 3.3.2.2 | Electrospray Ionization–Mass Spectrometry | 60 |
| 3.3.2.3 | Non-denaturing SDS-PAGE..... | 64 |
| 3.3.2.4 | Multimeric state of SynZur as assessed by analytical SEC | 66 |
| 3.3.2.5 | Investigation of SynZur de-metallation..... | 67 |
| 3.3.2.6 | Zinc affinity of SynZur | 72 |
| 3.3.2.7 | Electrophoretic Mobility Shift Assays (EMSA) | 73 |
| 3.3.2.7.1 | <i>Zur-pznuABC stoichiometry determination</i> | 75 |
| 3.3.2.7.2 | <i>Calculation of dissociation constants of SynZur-pznuABC complexes</i> | 79 |
| 3.3.2.8 | Metal-selectivity of SynZur | 82 |
| 3.3.2.8.1 | <i>Metal-responsive DNA-binding</i> | 82 |
| 3.3.2.8.2 | <i>ESI-MS of Zur with different metals</i> | 84 |
| 3.3.3 | Crystal structure | 89 |
| 3.3.3.1 | Comparison of the SynZur structure with Fur-family proteins from other bacteria..... | 94 |
| 3.4 | Conclusions and future work | 106 |
| Chapter 4. Properties of a <i>zur</i> knockout mutant of <i>Synechococcus</i> sp. WH8102 .. | | 108 |
| 4.1 | Introduction..... | 109 |
| 4.2 | Materials and methods | 111 |

| | | |
|-----------|---|-----|
| 4.2.1 | Bacterial strains and growth conditions | 111 |
| 4.2.1.1 | <i>Escherichia coli</i> strains and growth conditions..... | 111 |
| 4.2.1.2 | <i>Synechococcus</i> sp. WH8102 growth conditions | 111 |
| 4.2.1.2.1 | <i>Contamination check</i> | 113 |
| 4.2.1.3 | <i>Ruegeria pomeroyi</i> growth conditions | 113 |
| 4.2.1.4 | Monitoring bacterial growth and cell number calculation | 113 |
| 4.2.1.5 | Plasmids | 114 |
| 4.2.2 | Mutant construction | 115 |
| 4.2.2.1 | DNA extraction from cyanobacteria | 115 |
| 4.2.2.2 | PCR primers | 115 |
| 4.2.2.3 | Plasmid construction and preparation of <i>E. coli</i> for conjugation | 116 |
| 4.2.2.4 | <i>Synechococcus</i> sp. WH8102 conjugation..... | 117 |
| 4.2.2.5 | Purification of the <i>Synechococcus</i> sp. WH8102 <i>zur</i> -mutant from <i>R. pomeroyi</i> | 118 |
| 4.2.3 | Phenotypic comparison of the <i>Synechococcus</i> sp. WH8102 <i>zur</i> -mutant with wild type | 118 |
| 4.2.3.1 | Growth rate comparison | 118 |
| 4.2.3.1.1 | <i>General conditions</i> | 118 |
| 4.2.3.1.2 | <i>Medium preparation</i> | 119 |
| 4.2.3.1.3 | <i>Chelex resin preparation</i> | 119 |
| 4.2.3.2 | UV-vis absorption spectra comparison | 119 |
| 4.2.3.3 | Total cell lysate analysis by SDS-PAGE | 120 |
| 4.2.3.4 | Carbon fixation rate measurements..... | 120 |
| 4.2.3.4.1 | <i>Chlorophyll content measurement</i> | 120 |
| 4.2.3.5 | Inductively Coupled Plasma – Mass Spectrometry (ICP-MS).... | 121 |
| 4.2.4 | Transcriptomics..... | 122 |
| 4.2.4.1 | Sample preparation..... | 122 |
| 4.2.4.2 | RNA isolation..... | 122 |
| 4.2.4.3 | RNA-seq analysis | 123 |
| 4.2.4.3.1 | <i>RNA-seq data analysis</i> | 124 |
| 4.2.4.4 | Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) | 126 |
| 4.3 | Results and discussion | 128 |
| 4.3.1 | Generation of a <i>Synechococcus</i> sp. WH8102 <i>zur</i> interposon mutant . | 128 |
| 4.3.2 | Phenotypic characterization of the <i>Synechococcus</i> sp. WH8102 <i>zur</i> mutant | 132 |
| 4.3.2.1 | Growth rate assessment..... | 132 |
| 4.3.2.2 | Metal accumulation | 135 |
| 4.3.2.3 | Assessment of pigment composition via UV-vis absorption spectra | 139 |
| 4.3.2.4 | Comparison of CO ₂ fixation rates in the <i>Synechococcus</i> sp. WH8102 WT and <i>zur</i> mutant..... | 140 |
| 4.3.2.5 | SDS-PAGE of total cell lysates..... | 141 |
| 4.3.3 | Zur/Zn regulated transcriptomics of <i>Synechococcus</i> sp. WH8102..... | 143 |
| 4.3.3.1 | RNA-seq analysis | 143 |
| 4.3.3.1.1 | <i>Preparation of the libraries for RNA-sequencing</i> | 144 |
| 4.3.3.1.2 | <i>Data processing and analysis</i> | 145 |
| 4.3.3.2 | Identifying the <i>Synechococcus</i> sp. WH8102 Zur-regulon | 145 |
| 4.3.3.2.1 | <i>Genes up-regulated in the mutant compared to the WT</i> | 148 |
| 4.3.3.2.2 | <i>Genes down-regulated in the mutant compared to the WT</i> . 151 | |

| | | |
|---|---|------------|
| 4.3.3.3 | Validating the Zur-regulon..... | 154 |
| 4.3.3.3.1 | <i>Confirmation of RNA-seq data by RT-qPCR</i> | 158 |
| 4.3.3.4 | Functions and properties of newly identified SynZur-regulated genes | 159 |
| 4.3.3.4.1 | <i>Zur-regulon: Activation</i> | 159 |
| 4.3.3.4.2 | <i>Zur-regulon: Repression</i> | 161 |
| 4.3.3.5 | SynZur-independent genes potentially regulated by zinc | 166 |
| 4.3.3.6 | SynZur-independent genes responding to zinc toxicity. | 169 |
| 4.4 | Conclusion and future work..... | 171 |
| <u>Chapter 5. Properties of a Zur-activated metallothionein from <i>Synechococcus</i> sp. WH8102</u> | | 174 |
| 5.1 | Introduction..... | 175 |
| 5.2 | Materials and methods | 177 |
| 5.2.1 | <i>Zur-pbmtA</i> EMSA | 177 |
| 5.2.2 | SynBmtA over-expression and purification..... | 177 |
| 5.2.2.1 | Plasmid construction for SynBmtA over-expression | 177 |
| 5.2.2.2 | BmtA over-expression..... | 179 |
| 5.2.2.2.1 | <i>Over-expression of ¹⁵N labelled SynBmtA</i> | 179 |
| 5.2.2.3 | Purification of SynBmtA..... | 180 |
| 5.2.2.3.1 | <i>Chemical precipitation</i> | 180 |
| 5.2.2.3.2 | <i>Size Exclusion Chromatography of SynBmtA</i> | 181 |
| 5.2.2.3.3 | <i>Silver staining</i> | 181 |
| 5.2.3 | Characterisation of SynBmtA | 182 |
| 5.2.3.1 | ESI-MS and ICP-OES | 182 |
| 5.2.3.2 | pH titration SynBmtA | 182 |
| 5.2.3.3 | Investigation of SynBmtA de-metallation..... | 182 |
| 5.2.3.4 | Titration of apo-BmtA with zinc in the presence of Quin-2 | 183 |
| 5.2.3.5 | SynBmtA/SynZur competition for Zn..... | 183 |
| 5.2.3.6 | Nuclear Magnetic Resonance Spectroscopy | 183 |
| 5.2.3.6.1 | <i>1D ¹H-NMR Spectroscopy</i> | 183 |
| 5.2.3.6.2 | <i>2D Heteronuclear NMR Spectroscopy (HSQC)</i> | 184 |
| 5.2.3.6.3 | <i>3D Heteronuclear NMR Spectroscopy</i> | 184 |
| 5.3 | Results and discussion | 185 |
| 5.3.1 | Binding stoichiometry and affinity of <i>pbmtA</i> to SynZur | 185 |
| 5.3.2 | SynBmtA over-expression and purification..... | 188 |
| 5.3.2.1 | Plasmid construction for SynBmtA purification..... | 188 |
| 5.3.2.2 | SynBmtA over-expression | 190 |
| 5.3.2.2.1 | <i>Chemical precipitation</i> | 190 |
| 5.3.2.2.2 | <i>Size Exclusion Chromatography</i> | 191 |
| 5.3.3 | Characterisation of SynBmtA | 193 |
| 5.3.3.1 | ESI-MS..... | 193 |
| 5.3.3.2 | ICP-OES..... | 196 |
| 5.3.3.3 | pH titration of SynBmtA | 198 |
| 5.3.3.4 | Investigation of SynBmtA de-metallation..... | 200 |
| 5.3.3.5 | Zinc affinity of SynBmtA | 204 |
| 5.3.3.6 | Competition for zinc between SynZur and SynBmtA | 208 |
| 5.3.3.7 | Towards a three-dimensional structure of SynBmtA..... | 209 |
| 5.4 | Conclusions and future work | 219 |
| <u>Chapter 6. General conclusions</u> | | 221 |
| 6.1 | Introduction..... | 222 |

| | | |
|-----------|---|------------|
| 6.2 | Key findings..... | 225 |
| 6.2.1 | Zinc Uptake Regulator from <i>Synechococcus</i> sp. WH8102..... | 225 |
| 6.2.2 | <i>Synechococcus</i> sp. WH8102 metallothionein | 227 |
| 6.2.3 | Discovery of potential new players in zinc homeostasis | 229 |
| 6.2.3.1 | New putative outer membrane zinc uptake transporters | 230 |
| 6.3 | Future work..... | 231 |
| 7. | References..... | 233 |
| 8. | Appendix..... | 251 |

List of Figures

| | |
|--|----|
| Figure 1.1 Pigment types of marine <i>Synechococcus</i> sp. and structures of the associated chromophores. | 2 |
| Figure 1.2 Global distribution and abundance of <i>Prochlorococcus</i> and <i>Synechococcus</i> | 4 |
| Figure 1.3 X-ray crystallography structure of a complex of two Zur dimers with a 31 bp fragment of double-stranded DNA (<i>E. coli</i>). | 11 |
| Figure 1.4 Predicted cyanobacterial Zur binding box taken from the RegPrecise database. | 11 |
| Figure 1.5 The ABC transporter shown is BtuCD-F (4FI3) which is used for cobalamin uptake. | 14 |
| Figure 1.6 The structure of ZIP transporters. | 15 |
| Figure 1.7 NMR solution structure of SmtA from <i>Synechococcus</i> sp. PCC7942..... | 20 |
| Figure 1.8 Predicted mechanism of action of SynZur as a transcriptional repressor in <i>Synechococcus</i> sp. WH8102. | 22 |
| Figure 1.9 Predicted components of zinc metabolism in <i>Synechococcus</i> sp. WH8102. | 23 |
| Figure 3.1 The MRC 96-well plate used for SynZur crystallization..... | 49 |
| Figure 3.2 Scheme of the experimental design of SynZur purification. | 52 |
| Figure 3.3 SDS-PAGE analysis of over-expression of His-tagged SynZur protein in <i>E. coli</i> BL21 (DE3) pLysS. | 53 |
| Figure 3.4 FPLC purification of SynZur. | 54 |
| Figure 3.5 SDS-PAGE analysis of various purification steps during the concentrating and buffer exchange of the SynZur and TEV-protease (TEVp). | 55 |
| Figure 3.6 SDS-PAGE analysis of the second FPLC purification step of the SynZur protein after cleavage of the His-tag with TEV protease. | 56 |
| Figure 3.7 The theoretical sequence of the SynZur protein following purification...56 | |
| Figure 3.8 Amino acid sequence of the SynZur protein with sulfur containing amino acid residues highlighted with bold letters. | 58 |
| Figure 3.9 Sequence of the SynZur protein highlighting residues involved in determination of the extinction coefficient and zinc binding sites. | 60 |
| Figure 3.10 The full ESI-MS spectrum of apo-SynZur at pH 2. | 61 |
| Figure 3.11 A single charge state (+9) of the apo-SynZur ESI-MS spectrum. | 62 |
| Figure 3.12 ESI-MS spectrum of the native SynZur protein. | 63 |
| Figure 3.13 ESI-MS spectrum of a single +13 charge state of the SynZur ₂ Zn ₄ protein in native conditions. | 64 |
| Figure 3.14 Visualisation of SynZur at different concentrations following SDS-PAGE or ND-SDS-PAGE. | 65 |
| Figure 3.15 Analytical SEC of SynZur. | 66 |
| Figure 3.16 Native PAGE of native SynZurZn ₂ at pH 7.9..... | 67 |

| | |
|--|-----|
| Figure 3.17 ESI-MS spectrum of de-metallated Zur..... | 68 |
| Figure 3.18 Comparison of deconvoluted ESI-MS spectra of native SynZur and SynZurZn ₁ | 69 |
| Figure 3.19 Comparison of native SynZur (black) and de-metallated at the sensory site SynZurZn ₁ (grey)..... | 71 |
| Figure 3.20 A representative titration curve of SynZur by zinc in the presence of Quin-2. | 73 |
| Figure 3.21 Agarose gel electrophoresis of the region of <i>pznuABC</i> from <i>Synechococcus</i> sp. WH8102 after PCR. | 74 |
| Figure 3.22 EMSA of <i>pznuABC</i> with SynZur..... | 75 |
| Figure 3.23 EMSA using a 8% gel and explanation of how measurements of relative mobility were made..... | 76 |
| Figure 3.24 Ferguson plots..... | 77 |
| Figure 3.25 Determination of molecular weight of the SynZur–DNA complex based on the calibration curve constructed with MW of standard proteins. | 78 |
| Figure 3.26 Visualisation of the EMSA data and fitted models. | 80 |
| Figure 3.27 EMSA analysis of DNA binding of SynZur in the presence of different metals. | 82 |
| Figure 3.28 Quantification with Image J of the shifted bands within the box for the EMSAs shown in Figure 3.27. | 83 |
| Figure 3.29 Full ESI-MS spectra of SynZur with different metals..... | 85 |
| Figure 3.30 ESI-MS spectrum of the +10 charge state of ZurZn ₁ in native conditions with Cu ²⁺ added..... | 88 |
| Figure 3.31 Microscopy of SynZur crystals obtained during crystallization trials.... | 90 |
| Figure 3.32 The SynZur crystal structure. | 91 |
| Figure 3.33 Structural domains of the SynZur monomer and zinc binding sites..... | 92 |
| Figure 3.34 The His ₄ zinc binding site in SynZur..... | 93 |
| Figure 3.35 Structural alignment of SynZur and ScZur domains. | 96 |
| Figure 3.36 Prediction of secondary structure for SynZur and EcZur (4mtd) using the JPred4 tool (simple result output). | 97 |
| Figure 3.37 Sequential alignment of SynZur with other FFPs with known crystal structures using NCBI BLAST. | 98 |
| Figure 3.38 Superposition of metal binding sites by alignment of the DD in SynZur and 3mwm (<i>S. coelicolor</i> Zur). | 99 |
| Figure 3.39 Superposition of SynZur and CjFur..... | 101 |
| Figure 3.40 Sequence alignment of Zur proteins from cyanobacteria. | 103 |
| Figure 4.1 Haemocytometer grid. | 114 |
| Figure 4.2 Experimental design of the RNA-seq experiment..... | 122 |
| Figure 4.3 The overview of the RNA-seq data analysis. | 124 |
| Figure 4.4 Schematic representation of a single crossover for generation of a <i>zur</i> knockout mutant of <i>Synechococcus</i> sp. WH8102 after conjugation..... | 128 |

| | |
|---|-----|
| Figure 4.5 Steps needed for generation of a WH8102 <i>zur</i> knockout mutant by single crossover. | 129 |
| Figure 4.6 PCR amplification of the internal fragment of the <i>zur</i> gene from <i>Synechococcus</i> sp. WH8102. | 130 |
| Figure 4.7 Construction of plasmid pGP704CmKmZur. | 130 |
| Figure 4.8 Analysis of plasmid integration into the genome of <i>Synechococcus</i> sp. WH8102 by PCR with A and B primers with cells as template. | 131 |
| Figure 4.9 PCR confirmation of the <i>Synechococcus</i> sp. WH8102 mutant construct. | 132 |
| Figure 4.10 Growth curves of the WT and <i>Synechococcus</i> sp. WH8102 <i>zur</i> knockout mutant at different zinc concentrations. | 133 |
| Figure 4.11 Specific growth rates for <i>Synechococcus</i> sp. WH8102 WT and <i>zur</i> -mutant at different zinc concentrations. | 133 |
| Figure 4.12 Normalised elemental content of the <i>zur</i> mutant compared to the <i>Synechococcus</i> sp. WH8102 wild type. | 136 |
| Figure 4.13 Comparison of P-normalised zinc content by ICP-MS. | 138 |
| Figure 4.14 Comparison of UV-vis absorption spectra for WT <i>Synechococcus</i> sp. WH8102 and the <i>zur</i> -mutant. | 140 |
| Figure 4.15 CO ₂ fixation rate measurements at different light intensities. | 141 |
| Figure 4.16 SDS-PAGE of total cell lysates and extracellular proteins. | 142 |
| Figure 4.17 Visualisation of total RNA profiles on an Agilent 2100 Bioanalyzer using a RNA 6000 Pico LabChip. | 144 |
| Figure 4.18 Genes differentially expressed in the <i>Synechococcus</i> sp. WH8102 <i>zur</i> mutant compared to wild type grown under 772 nM zinc. | 146 |
| Figure 4.19 High expression of SYNW2401 in the <i>zur</i> knockout mutant compared to the wild type. | 149 |
| Figure 4.20 Location of SYNW0968, SYNW0972 and SYNW0973 on the genome. | 150 |
| Figure 4.21 Location of <i>zur</i> -boxes in <i>pbmtA</i> and their position towards RNA-polymerase binding site (-10 and -35 elements). | 152 |
| Figure 4.22 Venn diagrams of three different comparisons: Mutant/Wild Type at 0 zinc, Mutant/Wild Type at 772 nM zinc and Wild Type at 0/Wild Type at 772 nM zinc. | 157 |
| Figure 4.23 RT-qPCR analysis of <i>znuA</i> and <i>bmtA</i> gene expression. | 159 |
| Figure 4.24 Sequences of SYNW0968, SYNW0972 and SYNW0973. | 162 |
| Figure 4.25 Predicted secondary structure of SYNW0968 (A), SYNW0972 (B) and SYNW0973 (C) proteins based on superposition of a selection of hits using Phyre2 analysis. | 163 |
| Figure 4.26 Cyanorak genome context analysis of SYNW0972. | 165 |
| Figure 4.27 Venn diagrams for comparison of M0/M772 with WT0/WT772. | 168 |
| Figure 4.28 Venn diagrams for comparison of WT772/WT2500 nM zinc with Mut0/Mut772. | 170 |

| | |
|---|-----|
| Figure 4.29 A plot of normalised gene expression for <i>znuA</i> and <i>bmtA</i> measured by RT-qPCR against normalised zinc content. | 172 |
| Figure 5.1 EMSA of <i>pbmtA</i> with SynZur. | 185 |
| Figure 5.2 Determination of the molecular weight of the SynZur- <i>pbmtA</i> complex. | 186 |
| Figure 5.3 Comparison of <i>pznuA</i> -Zur and <i>pbmtA</i> -Zur complexes by EMSA. | 187 |
| Figure 5.4 Visualisation of the EMSA data and fitted models. | 188 |
| Figure 5.5 Agarose gel electrophoresis of PCR products obtained with primers BmtA_F and BmtA_Re using <i>Synechococcus</i> sp. WH8102 genomic DNA as template. | 189 |
| Figure 5.6 Agarose gel electrophoresis of pET24a(+) <i>BmtA</i> plasmids after PCR with T7 primers. | 189 |
| Figure 5.7 SDS-PAGE of <i>BmtA</i> over-expression. | 190 |
| Figure 5.8 SDS-PAGE of <i>BmtA</i> fractions following chemical precipitation visualized by silver-staining. | 191 |
| Figure 5.9 Size exclusion chromatography of Syn <i>BmtA</i> | 192 |
| Figure 5.10 ESI-MS spectrum of Syn <i>BmtA</i> at pH 2. | 194 |
| Figure 5.11 ESI-MS spectrum of native Syn <i>BmtA</i> at pH 7.9. | 195 |
| Figure 5.12 Sulfur containing residues of Syn <i>BmtA</i> without the first methionine residue. | 197 |
| Figure 5.13 pH titration of Syn <i>BmtA</i> | 199 |
| Figure 5.14 Investigation of Syn <i>BmtA</i> de-metallation. | 200 |
| Figure 5.15 ESI-MS spectrum of apo-Syn <i>BmtA</i> at pH 7.8. | 203 |
| Figure 5.16 Reconstitution of apo-Syn <i>BmtA</i> with zinc, pH 7.8. | 204 |
| Figure 5.17 A representative titration curve of Syn <i>BmtA</i> by zinc in the presence of Quin-2. | 205 |
| Figure 5.18 Absorbance equilibration of Syn <i>BmtA</i> upon titration with zinc in the presence of Quin-2. | 207 |
| Figure 5.19 ESI-MS spectra of native SynZur, apo-Syn <i>BmtA</i> and their mixture. | 208 |
| Figure 5.20 ¹ H- ¹⁵ N HSQC spectrum of Syn <i>BmtA</i> Zn ₄ | 210 |
| Figure 5.21 Sequential alignment of Syn <i>BmtA</i> with SmtA from <i>Synechococcus elongatus</i> strain PCC 7942. | 210 |
| Figure 5.22 Strips of 3D NMR corresponding to selected characteristic residues from Syn <i>BmtA</i> | 211 |
| Figure 5.23 Secondary structure prediction for Syn <i>BmtA</i> based on the CSIs of H α | 215 |
| Figure 5.24 Prediction of the secondary structure based on CSI and its comparison with Phyre2 and Jpred4 predictions and with the known SmtA structure. | 215 |
| Figure 5.25 Comparison of CSI values for H α of Syn <i>BmtA</i> and SmtA. | 216 |
| Figure 5.26 Structure of Syn <i>BmtA</i> predicted by the Phyre2 server. | 217 |
| Figure 5.27 Predicted Zn ₄ cluster in Syn <i>BmtA</i> based on the model of SmtA and sequential alignment shown in Figure 5.21. | 218 |

| | |
|--|-----|
| Figure 6.1 Proposed mode of zinc regulation in <i>Synechococcus</i> sp. WH8102..... | 227 |
| Figure 6.2 Zinc transfer between SynZur and BmtA..... | 229 |
| Figure 6.3 Components of the Zur (green) regulon in <i>Synechococcus</i> sp. WH8102 identified in Chapter 4 with their suggested locations. | 230 |
| Figure 8.1 Plasmid used for Zur over-expression and a codon-optimized synthetic SynZur sequence. | 251 |
| Figure 8.2 Over-expression and purification of TEV protease by FPLC. | 252 |
| Figure 8.3 Dynafit script used for calculation of SynZurZn ₁ K _D | 253 |
| Figure 8.4 Scripts used in DynaFit software for fitting the K _D to the experimental EMSA data. | 254 |
| Figure 8.5 Dynafit script for Model 3 and the fitted curve. | 255 |
| Figure 8.6 Superposition of SynZur (dark grey) and PaFur (light grey) performed by the Swiss PDBD viewer. | 257 |
| Figure 8.7 Site 3 in <i>Rhizobium leguminosarum</i> Mur. | 258 |
| Figure 8.8 Bash-script used for RNA-seq data analysis. | 258 |
| Figure 8.9 Example of the DESeq2 script in R software used for comparison of one condition containing three replicates with another condition containing three replicates (mutant at 0 zinc with WT at 0 zinc). | 260 |
| Figure 8.10 Agilent Bioanalyzer profiles of the RNA samples after Qiagen RNeasy columns. | 261 |
| Figure 8.11 Agilent Bioanalyzer profiles of the RNA samples after using the RiboZero kit. | 261 |
| Figure 8.12 Agilent Bioanalyzer profiles of the RNAseq libraries..... | 262 |
| Figure 8.13 Post-processing statistics of RNA-sequencing data. | 262 |
| Figure 8.14 Identification of the predicted zur-boxes..... | 264 |
| Figure 8.15 Transcript coverage of SYNW0971 and SYNW0972 in the <i>zur</i> mutant (top) and WT (bottom) checked by IGV software. | 265 |
| Figure 8.16 Location of Zur boxes and their position in the promoter elements of <i>Anabaena</i> sp. PCC 7120 <i>znuA</i> (<i>all0833</i>). | 265 |
| Figure 8.17 Location of a possible zur-box and RNA-polymerase binding sites in the operator-promoter region of SYNW0360. | 266 |
| Figure 8.18 Transcript coverage of SYNW0359 and SYNW0360 in the mutant (top) and WT (bottom) checked by IGV software. | 266 |
| Figure 8.19 Genes differentially expressed in the <i>zur</i> knockout mutant of <i>Synechococcus</i> sp. WH8102 compared to the wild type at 0 added zinc. | 267 |
| Figure 8.20 Genes differentially expressed in <i>Synechococcus</i> sp. WH8102 at 0 added zinc compared to 772 nM zinc. | 268 |
| Figure 8.21 qPCR amplification using a known concentration of genomic DNA as a template and corresponding primers. | 270 |
| Figure 8.22 Location of SYNW0360 on the genome towards the 3' end of the tmRNA. | 271 |

| | |
|--|-----|
| Figure 8.23 Analysis of WP_042503314.1 (top) and SYNW0973 (bottom) using the SignalP server. | 271 |
| Figure 8.24 Genes differentially expressed in the <i>Synechococcus</i> sp. WH8102 <i>zur</i> -mutant at 0 added zinc compared to 772 nM zinc. | 272 |
| Figure 8.25 Genes differentially expressed in the <i>Synechococcus</i> sp. WH8102 <i>zur</i> -mutant at 772 nM zinc compared to 2.5 μ M zinc. | 273 |
| Figure 8.26 pET-24a(+) vector map. | 274 |
| Figure 8.27 Sequence of <i>pbmtA</i> used in this study. | 274 |
| Figure 8.28 Sequence of the SYNW0359 PCR product. <i>bmtA</i> is underlined. | 275 |
| Figure 8.29 DynaFit script used for calculation of SynBmtA- _{Zn} K_D | 277 |

List of Tables

| | |
|---|-----|
| Table 1.1 Summary of marine <i>Synechococcus</i> clades and their ecological distribution. | 5 |
| Table 1.2 Common metalloproteins present within marine phytoplankton. | 7 |
| Table 1.3 Genes predicted to be regulated by Zur in <i>Synechococcus</i> sp. WH8102...22 | 22 |
| Table 2.1 <i>Escherichia coli</i> strains used in this work..... | 29 |
| Table 2.2 List of the plasmids used in this work. | 30 |
| Table 2.3 PCR components..... | 31 |
| Table 2.4 Default PCR conditions. | 31 |
| Table 2.5 Determination of elements of interest by ICP-OES by characteristic emission lines..... | 38 |
| Table 3.1 Primers used to amplify the operator-promoter region of <i>Synechococcus</i> sp. WH8102 <i>znuABC</i> | 44 |
| Table 3.2 PAGE gel composition. Values are given for 10 ml or 2 gels. | 45 |
| Table 3.3 ICP-OES analysis of SynZur: the content of various trace metals in the analysed SynZur samples..... | 58 |
| Table 3.4 Comparison of the calculated molecular mass of apo-SynZur with that determined by ESI-MS..... | 61 |
| Table 3.5 ESI-MS results and comparison with the calculated molecular mass of SynZur under native conditions. | 63 |
| Table 3.6 ESI-MS of de-metallated SynZur, showing the experimentally determined mass compared with the predicted molecular mass. | 69 |
| Table 3.7 Mobility and Rf's of different bands using 8 and 12% acrylamide gels used to build Ferguson plots..... | 76 |
| Table 3.8 Determination of molecular weights of the SynZur-DNA complex (C) based on the calibration curve..... | 78 |
| Table 3.9 Species observed by ESI-MS for SynZur supplemented with different metals. | 86 |
| Table 3.10 List of Zur-species oxidized by Cu ²⁺ | 88 |
| Table 3.11 Superposition of SynZur monomers within its asymmetric unit. | 92 |
| Table 3.12 Metal distances in the crystal structure of SynZur at the two zinc sites. .93 | 93 |
| Table 3.13 Structural alignment of SynZur with other FFP members..... | 95 |
| Table 3.14 Summary of sequential alignment of FFPs using NCBI BLAST..... | 98 |
| Table 3.15 Comparison of metal-binding sites in different FFP members with SynZur..... | 100 |
| Table 3.16 Summary of sequential alignment of annotated Zur proteins from cyanobacteria using NCBI BLAST..... | 102 |
| Table 4.1 Composition of ASW _{-zn} medium..... | 112 |
| Table 4.2 List of plasmids used in this work. | 114 |
| Table 4.3 PCR primers used in this study. Restriction sites are highlighted. | 116 |

| | |
|--|-----|
| Table 4.4 16S rRNA gene primers used for assessing DNA contamination. | 123 |
| Table 4.5 qPCR primers used in this study. | 127 |
| Table 4.6 ICP-MS (moles/kg) assessment of trace metal accumulation..... | 135 |
| Table 4.7 Cellular zinc content and calculation of zinc concentrating factors. | 138 |
| Table 4.8 Genes differentially expressed in the <i>zur</i> knockout mutant of <i>Synechococcus</i> sp. WH8102 compared to the wild type under 772 nM zinc. | 147 |
| Table 4.9 Genes differentially expressed in the <i>zur</i> knockout mutant of <i>Synechococcus</i> sp. WH8102 compared to the wild type grown under 0 zinc..... | 155 |
| Table 4.10 Genes differentially expressed in WT <i>Synechococcus</i> sp. WH8102 at 0 added zinc compared to 772 nM zinc. | 156 |
| Table 4.11 Genes differentially expressed in the <i>Synechococcus</i> sp. WH8102 <i>zur</i> - mutant at 0 added zinc compared to 772 nM zinc. | 167 |
| Table 4.12 Differentially expressed genes in common for comparisons M0/M772 and W0/W772. | 168 |
| Table 4.13 Genes differentially expressed in the wild type <i>Synechococcus</i> sp. WH8102 at 772 nM zinc compared to 2.5 μ M. | 169 |
| Table 4.14 Differentially expressed genes in common for comparisons M0/M772 and W772/W2500 which are probably involved in zinc toxicity. | 170 |
| Table 5.1 Primers used for amplification of a 143 bp region of <i>pbmtA</i> | 177 |
| Table 5.2 Primers used for amplification of <i>synbmtA</i> for subsequent cloning into pET24a(+). | 178 |
| Table 5.3 T7 primers used for checking the successful ligation of the <i>synbmtA</i> insert. | 178 |
| Table 5.4 M9 Minimal Medium recipe. | 180 |
| Table 5.5 Values of theoretical and found molecular masses for apo-SynBmtA without the initiator methionine. | 193 |
| Table 5.6 Found and calculated molecular mass values for SynBmtA at pH 7.9.... | 196 |
| Table 5.7 Calculation of the ϵ_{230} for SynBmtA. | 197 |
| Table 5.8 ICP-OES analysis of SynBmtA: | 198 |
| Table 5.9 pH values of half displacement for metallothioneins from diferent organisms. | 199 |
| Table 5.10 Preliminary 3D [^1H , ^{15}N , ^1H] TOCSY-HSQC and NOESY-HSQC resonance assignments of SynBmtAZn ₄ at 298 K and pH 7.9 recorded at 700 MHz. | 212 |
| Table 5.11 Correlational peaks from SynBmtA confirming formation of hydrogen bonds based on the SmtA structure. | 213 |
| Table 8.1 SynZur structural information..... | 256 |
| Table 8.2 RNA sample concentrations and purity assessment by NanoDrop..... | 260 |
| Table 8.3 RNA-Seq analysis statistics. | 263 |
| Table 8.4 List of genes common to the different comparisons found from Venn diagrams. | 269 |

| | |
|--|-----|
| Table 8.5 DynaFit scripts used for calculation of SynZur- <i>pbmtA</i> K_D | 276 |
| Table 8.6 Side chain signals of Asn, Gln and un-assigned residues of SynBmtA... | 278 |

Acknowledgments

First of all I would like to thank my PhD supervisors Dr. Claudia Blindauer and Prof. David Scanlan for giving me the opportunity to undertake this project, for all the help and support and most importantly for introducing me to the amazing worlds of bioinorganic chemistry and marine microbiology.

I would like to express my gratitude my family: my husband Dr. Alexey Korepanov and my mom Valentina Belosludtseva for their care, love and support.

I am greatly thankful for all the help and collaboration to Dr. Rachael Wilkinson, Prof. Vilmos Fülöp and Dr. Joseph Christie-Oleza who significantly influenced my project. I want to thank my colleagues from Dr. Blindauer and Prof. Scanlan groups for their enormous help and friendly attitude: Dr. Agnieszka Mierek-Adamska, Dr. Fran Pitt, Dr. James Coverdale, Dr. Rich Puxty and Dr. Sabine Bowman-Grahl. A special acknowledgement goes to Dr. Amira Ksibe for her initial discoveries and assistance. I am also very grateful to my students Ben Ashley, Eleanor Marks and Luke Scott for their responsible and thoughtful work and ideas. In addition, I would like to credit my friends and colleagues Gabriel Erni Cassola Barata, Audam Chhun, Marco Pinto Corujo, Eric Otto and Phoenix Scott for their help, friendship and support. A particular recognition goes to Vinko Zadjelovic Varas and Robyn Wright for maintaining my physical fitness level and motivation during this project.

I am deeply grateful to Dr. Ivan Prokes, Dr. Lijiang Song and Mr. Phil Aston for the training and help with equipment. I really appreciate members of the Centre for Genomic Research, Institute of Integrative Biology at the University of Liverpool for the RNA library preparation, sequencing and help with troubleshooting.

I would like to also thank my former supervisors: Prof. Stepan Kalmykov, Dr. Elena Zakharova and Dr. Andrey Frolov for all the initial work and assistance that allowed me to start this project.

And finally, I am very grateful to the University of Warwick and Warwick Chancellor's International Scholarship for providing an amazing research environment and funding for my project.

Declaration

I, Alevtina Mikhaylina, declare that the work presented in this thesis is original and has not been submitted for any other degree. Experimental work was conducted by me, unless stated otherwise, under the supervision of Dr. Claudia A. Blindauer and Prof. David J. Scanlan, University of Warwick. All materials and work from other sources have been fully acknowledged.

Crystallisation experiments and structure resolution were performed by Dr. Rachael Wilkinson and Prof. Vilmos Fülöp. The NMR assignment was performed with the assistance of Dr. Claudia A. Blindauer. The recombinant plasmid for overexpression of SynZur was constructed by Dr. Amira Ksibe. Carbon fixation experiment was performed with the help of Dr. Rich Puxty. ICP-MS measurements were performed by Dr. James Coverdale. RNA libraries preparation and sequencing were performed by the Centre for Genomic Research, Institute of Integrative Biology, at the University of Liverpool.

Alevtina Mikhaylina

Abstract

Cyanobacteria of the genera *Synechococcus* and *Prochlorococcus* are key players in marine CO₂ fixation contributing ~25% of the global total. Since oceanic zinc concentrations may directly affect the CO₂ fixation process via the role of carbonic anhydrase in cyanobacterial photosynthesis, understanding cyanobacterial zinc metabolism is of critical importance.

The current study used the model open ocean cyanobacterium *Synechococcus* sp. WH8102. In this species zinc uptake is thought to be regulated by a predicted zinc uptake regulator (SynZur) which controls expression of at least four genes thought to be involved in zinc homeostasis: *znuA*, *znuB*, *znuC* and *bmtA* (encoding a predicted high affinity zinc uptake transporter ZnuABC and a metallothionein SynBmtA). These components and Zur are largely uncharacterized.

SynZur was over-expressed recombinantly in *E. coli*, purified, crystallized and its X-ray crystal structure resolved. This is the first member of the Fur family of proteins from cyanobacteria to have a known structure. Two zinc binding sites were identified: a conserved structural site 1 (Cys₄), and a novel sensory site (His₂AspCys). Both zinc binding sites reside within the dimerization domain. ESI-MS detected a decrease in the abundance of the SynZur dimer in the absence of zinc at the sensory site, hinting at a potentially novel zinc sensing mechanism that might be typical of cyanobacterial Zur proteins. The affinity of the sensory site for zinc was $K_D = 8.27 \times 10^{-13}$ M. SynZur with two zinc ions per monomer bound also showed selective binding to *pznuABC* with a $K_D \sim 5$ nM per dimer or ~ 10 nM per monomer. ESI-MS revealed that Cd²⁺ can replace Zn²⁺ in the structural site suggesting that it is not kinetically inert.

A *Synechococcus* sp. WH8102 *zur* mutant was also constructed and characterised. Growth rate comparisons showed that the mutant has a lower tolerance to high zinc levels compared to the WT, whilst ICP-MS analysis revealed different accumulation behaviour for trace metals and most significantly for zinc. This is consistent with SynZur repressing a predicted *znuABC* system. RNA-seq analysis of the WT and *zur* mutant facilitated experimental identification of new members of the *Synechococcus* sp. WH8102 Zur regulon. Phyre2 structural prediction suggested that two of the newly identified genes (SYNW0972 and SYNW0973) encode components of a novel outer membrane zinc uptake system. Significantly, these two genes are conserved in other marine cyanobacteria. Furthermore, the RNA-seq data (confirmed using RT-qPCR) showed that SynZur can activate a metallothionein (SYNW0359, SynBmtA) in *Synechococcus* sp. WH8102 previously thought to be repressed by SynZur.

SynZur binds to the operator-promoter region of SynBmtA (*pbmtA*) in a 4:1 stoichiometry, with an affinity of 26.7 nM per monomer. The SynBmtA protein was over-expressed and purified with four zinc ions bound. SynBmtA was de-metallated and its affinity to Zn²⁺ was determined ($K_D = 1.13 \times 10^{-14}$ M). ESI-MS showed that apo-SynBmtA is able to remove zinc from metallated SynZurZn₂ in agreement with the order of zinc affinities for SynZur and SynBmtA. Finally, SynBmtA was analysed by ¹H, ¹⁵N-NOESY-HSQC and ¹H, ¹⁵N-TOCSY-HSQC NMR spectroscopy. The spectra were partially assigned and the assignment indicated that the characteristic zinc-finger fold of bacterial MTs is indeed present in SynBmtA, but that the C-terminus is likely to differ significantly from that of SmtA from the *Synechococcus elongatus* strain PCC 7942.

List of abbreviations

| | |
|----------------|--|
| aa | Aminoacid(s) |
| ABC | Adenosine triphosphate-binding cassette |
| AdcR | Adhesin competence repressor |
| Amp | Ampicillin |
| ASW | Artificial seawater |
| ATP | Adenosine tri-phosphate |
| A _x | Absorbance at x nm |
| Å | Angstrom |
| b | Base(s) |
| BLAST | Basic Local Alignment Search Tool |
| bmtA | Bacterial metallothionein |
| bp | Base pair |
| Bs | <i>Bacillus subtilis</i> |
| c | Centi |
| °C | Degrees Celsius |
| CDF | Cation diffusion facilitator |
| cDNA | Complementary DNA |
| CIAP | Calf intestinal alkaline phosphatase |
| Cj | <i>Campylobacter jejuni</i> |
| CL | Crude lysate |
| Cm | Chloramphenicol |
| CTAB | Cetyl trimethyl ammonium bromide |
| COG | Cluster of Orthologous Groups of proteins |
| CSI | Chemical Shift Indexes |
| CV | Coefficient of variation / Relative standard deviation |
| D | Dimencional |
| Da | Dalton |
| DBD | DNA-binding domain |
| DD | Dimerization domain |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxyribonucleotide triphosphate |
| DTT | Dithiothreitol |
| E | Einstein / mole of photons |
| <i>E. coli</i> | <i>Escherichia coli</i> |
| Ec | <i>Escherichia coli</i> |
| e. g. | For example |
| EDDS | Ethylenediamine-Disuccinate |
| EDTA | Ethylenediaminetetraacetic acid |
| EMSA | Electrophoretic mobility shift assay |
| ESI-MS | Electrospray ionisation - mass spectrometry |
| eV | Electronvolt |
| ε _x | Molar attenuation coefficient at x nm / extinction coefficient at x nm |
| E-value | Expect value |
| EXAFS | Extended X-ray Absorption Fine Structure |
| f | Femto |
| FFP | Fur-family protein |
| FPLC | Fast protein liquid chromatography |
| Fur | Ferric uptake regulator |

| | |
|-----------------------|--|
| g | Gram(s) |
| His-tag | Polyhistidine tag |
| HPLC | High performance liquid chromatography |
| HSQC | Heteronuclear single quantum coherence |
| Hz | Hertz |
| ICP-MS | Inductively coupled plasma - mass spectrometry |
| ICP-OES | Inductively coupled plasma - optical emission spectroscopy |
| i.e. | In other words |
| IMAC | Immobilized Metal ion Affinity Chromatography |
| <i>in vitro</i> | In a controlled environment outside of a living organism |
| <i>in vivo</i> | Within a living organism |
| ITC | Isothermal titration calorimetry |
| IPTG | Isopropyl-beta-D-thiogalactopyranoside |
| k | Kilo |
| K_D | Dissociation constant |
| Km | Kanamycin |
| Kr | Slope of a Ferguson plot |
| L | Litre(s) |
| L | Protein ladder |
| LB | Luria-Bertani medium or liquid broth |
| LL | Low light |
| log ₂ fold | Log ₂ fold |
| M | Molar |
| M | Mega |
| m | Meter(s) |
| m | Milli |
| m/z | Mass over charge |
| MarR | Multiple antibiotic resistance repressor |
| MB | Marine broth |
| MES | 2-(N-morpholino)ethanesulfonic acid |
| min | Minute(s) |
| MM | Minimal medium |
| MM | Molecular weight / Molecular mass |
| MOPS | 3-(N-morpholino)propanesulfonic acid |
| mRNA | Messenger RNA |
| MT | Metallothionein |
| Mt | <i>Mycobacterium tuberculosis</i> |
| Mur | Manganese uptake regulator |
| Mut | Mutant |
| MW | Molecular weight / Molecular mass |
| μ | Micro |
| n | Nano |
| NCBI | National Center for Biotechnology Information |
| ND | Non-denaturing |
| NMR | Nuclear magnetic resonance |
| NOESY | Nuclear Overhauser effect spectroscopy |
| nt | Nucleotide(s) |
| Nur | Nickel uptake regulator |
| OD | Optical density |
| p | Pico |

| | |
|--------------------------|---|
| Pa | <i>Pseudomonas aeruginosa</i> |
| PAGE | Polyacrylamide gel electrophoresis |
| PCR | Polymerase chain reaction |
| PDB | Protein data bank |
| PE | phycoerythrin |
| PEG | Polyethylene glycol |
| PerR | Hydrogen peroxide transcriptional regulator of Fur family |
| pH _{1/2} | pH value of half displacement of a metal |
| PMSF | Phenylmethanesulfonyl fluoride |
| ppb | Parts per billion |
| ppm | Parts per million |
| p-value | Probability value |
| Q | Ohm |
| Quin-2 | 2-[(2-Amino-5-methylphenoxy)methyl]-6-methoxy-8-aminoquinoline-N,N,N',N'-tetraacetic acid tetrapotassium salt |
| R ² | Coefficient of determination / Correlation coefficient |
| Rf | Relative mobility |
| Rl | <i>Rhizobium leguminosarum</i> <i>bv. viciae</i> |
| RMSD | Root-mean-square deviation |
| RNA | Ribonucleic acid |
| RND | Resistance nodulation division |
| rpm | Revolutions per minute |
| rRNA | Ribosomal RNA |
| RT-qPCR | Reverse transcription - quantitative polymerase chain reaction |
| Sc | <i>Streptomyces coelicolor</i> |
| sec | Seconds |
| SEC | Size exclusion chromatography |
| SDS | Sodium dodecyl sulfate |
| smtA | Metallothionein from <i>Synechococcus</i> sp. PCC7942 |
| SBP | Substrate-binding protein |
| Syn | <i>Synechococcus</i> sp. WH8102 |
| TBDRs | TonB-dependent receptors |
| TCEP | 3,3',3''-Phosphanetriyltriopropanoic acid |
| TEV | Tobacco etch virus |
| TEVp | TEV-protease |
| tmRNA | Transfer-messenger RNA |
| TOCSY | Total Correlation Spectroscopy |
| TPEN | N,N,N',N'-Tetrakis(2-pyridinylmethyl)-1,2-ethanediamine |
| Tris | 2-Amino-2-(hydroxymethyl)propane-1,3-diol |
| UV-vis | Ultra violet - visible |
| V | Volts |
| v | Volume |
| w | Weight |
| WT | Wild type |
| ZIP | <u>Z</u> RT/ <u>I</u> RT-like proteins |
| ZnuABC | High affinity inner membrane zinc transporter of ABC family |
| <u>Z</u> RT/ <u>I</u> RT | <u>z</u> inc or <u>i</u> ron <u>r</u> egulated <u>t</u> ransporters |
| Zur | Zinc uptake regulator |

Chapter 1.
Introduction

1.1 CYANOBACTERIA

Cyanobacteria are photosynthetic Gram-negative prokaryotes found in a variety of habitats from polar regions to hot springs¹. They were previously known as blue-green algae due to the presence of the blue pigment phycocyanin and the green pigment chlorophyll *a*². However, different cyanobacterial species can be of different colours depending on their ecological niche and light harvesting chromophores (see Figure 1.1)².

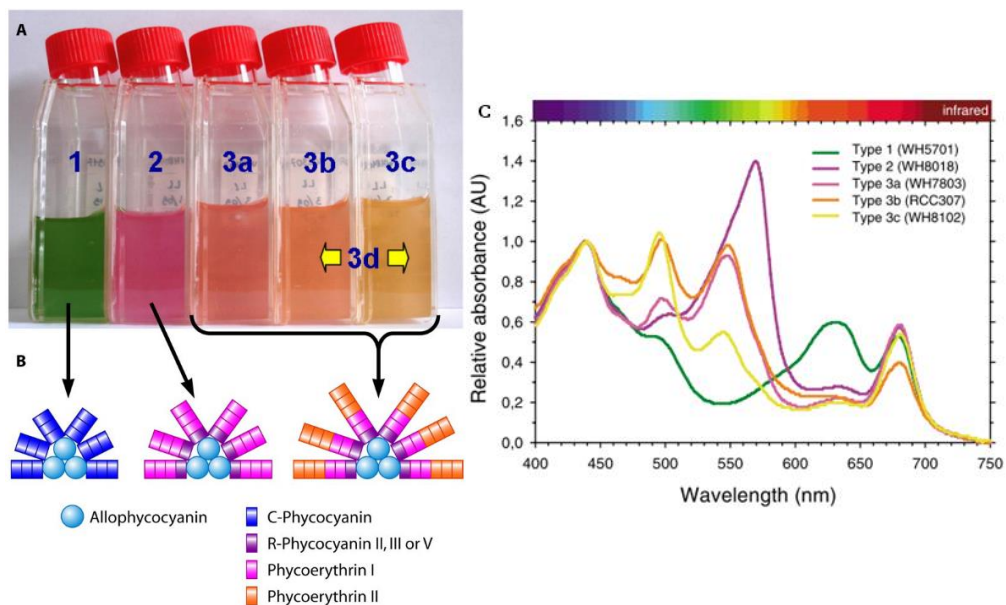


Figure 1.1 Pigment types of marine *Synechococcus* sp. and structures of the associated chromophores.

A. Strains representative of the three main pigment types. B. Proposed models of phycobilisome structures corresponding to the three main types of chromophores. Figures are taken from^{2,3}.

Cyanobacteria evolved about 2.75-3.2 billion years ago⁴. They were the first oxygenic photosynthetic organisms, ultimately altering the reducing atmosphere during the Archaean and Proterozoic Eras into the current oxidizing one^{4,5}. This in turn shaped the direction of the evolution of life throughout Earth's history.

Another significant contribution of cyanobacteria is the development of plant life on our planet. The genesis of chloroplasts occurred via endosymbiosis of cyanobacteria⁶. Apparently in the late Proterozoic and/or early Cambrian, a cyanobacterium was engulfed by a eukaryotic cell, fixing atmospheric carbon for their host in a structure now known as the chloroplast^{6,7}.

In the present day, cyanobacteria are still globally important for the biogeochemical cycling of various elements, especially carbon. Marine phytoplankton including cyanobacteria are responsible for about half of the carbon fixed globally⁸ with cyanobacteria contributing about a quarter of this⁹. This contribution is of great significance due to the increasing levels of carbon dioxide in the atmosphere and associated climate change, which is predicted to massively impact life on Earth.

Some cyanobacteria are also able to fix atmospheric nitrogen being unique photosynthetic nitrogen fixers¹⁰. Amongst these are several genera such as *Anabaena* which are extremely important for agriculture by providing nitrogen in a form accessible to globally essential crops such as rice¹.

1.1.1 Marine *Synechococcus*

The most abundant photosynthetic organisms on Earth, contributing around 70% of the total fixed carbon through photosynthesis in some oceanic regions, are marine cyanobacteria of the genera *Prochlorococcus* and *Synechococcus*¹¹. These two genera contribute 25% of the carbon dioxide fixed globally in marine ecosystems while only comprising <1% of the biomass compared to other photosynthetic organisms^{9,11,12}. Although cyanobacteria of the genera *Synechococcus* are less numerous than those of *Prochlorococcus*, *Synechococcus* has a wider global distribution, occupying a broader range of habitats, including the Arctic and Southern Oceans as shown in Figure 1.2⁹. *Synechococcus* also has higher cell specific CO₂ fixation rates compared to *Prochlorococcus* and thus contributes around 17% of ocean net primary production compared to 8.5% by *Prochlorococcus*⁹.

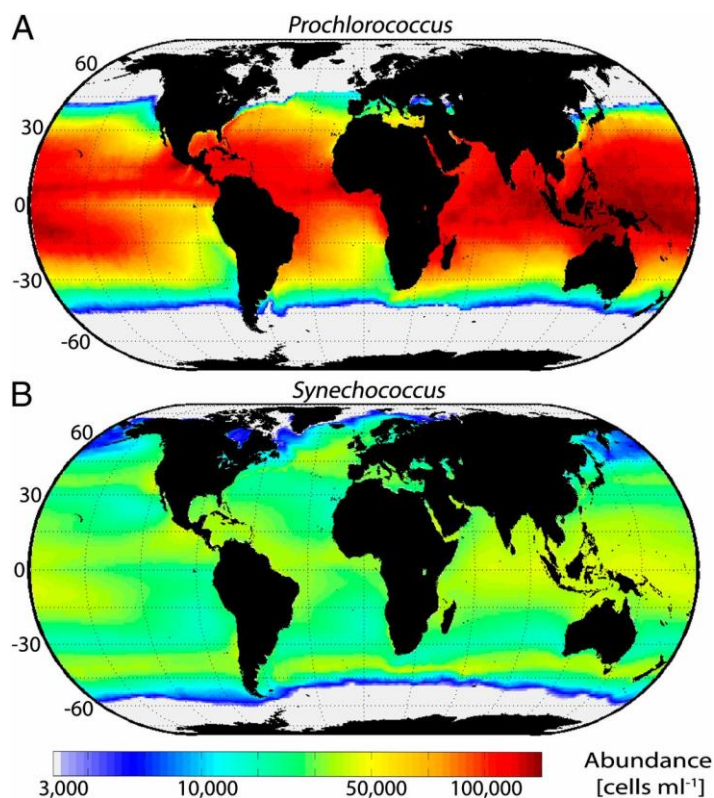


Figure 1.2 Global distribution and abundance of *Prochlorococcus* and *Synechococcus*, taken from⁹.

Marine *Synechococcus* strains are small unicellular coccoid cyanobacteria approximately 1 μm in size¹³. They are a more ancient and genetically more variable group than *Prochlorococcus* and are adapted to a variety of conditions meaning they occupy environments with wide variations in temperature, nutrients and light³.

The genome sizes of marine *Synechococcus* vary from 2.2 to approximately 3 Mb, making them on average much smaller than those of other bacteria, which could be an adaptation to their nutrient-poor environment³. Phylogenetic analysis using 16S rRNA gene sequences of marine *Synechococcus* strains divides them into three sub-clusters, 5.1, 5.2 and 5.3. The most ubiquitous marine sub-cluster, 5.1, is further divided into several clades occupying a variety of habitats. The ecological distribution of different sub-clusters and clades of marine *Synechococcus* is shown in Table 1.1.

Table 1.1 Summary of marine *Synechococcus* clades and their ecological distribution^{3,13,14,15,16}.

| Sub-cluster | Clade | Ecological conditions |
|--------------------|--|---|
| 5.1 | I | Coastal and/or temperate mesotrophic open ocean |
| | II | Offshore, continent shelf, oligotrophic tropical or subtropical waters |
| | III | Open-ocean waters |
| | IV | Coastal and/or temperate mesotrophic open ocean waters |
| | V | Wide distribution, low abundance in mesotrophic upwelling regions |
| | VI | Wide distribution, low abundance in mesotrophic upwelling regions |
| | VII | Wide distribution, low abundance in mesotrophic upwelling regions |
| | VIII | Hypersaline pools |
| | IX | East China Sea, Red Sea and upwelling off of Africa |
| | X | Isolated from Mediterranean, Central North Sea and Long Island Sound |
| | XVI | Atlantic, Mediterranean Sea and Pacific |
| | XV | Open-ocean and East China Sea |
| | XX | Phosphate-depleted area of the transition zones (30–35° S and 30–35° N) |
| | EnvA | Cold, intermediate, and warm waters with various degrees of Fe limitation |
| | EnvB | |
| | CRD1 | Tropical and sub-tropical oceans, Costa Rica upwelling |
| UC-A | South and North Pacific Subtropical Gyres | |
| WPC1 | Observed in open ocean and nearshore waters | |
| CB1 | Estuarine and coastal waters, polar and subpolar | |
| 5.2 | | Temperate estuarine/coastal waters, some are prevalent in polar/subpolar waters |
| 5.3 | | Sargasso Sea in spring |

1.2 METALS IN CYANOBACTERIA

When cyanobacteria evolved oxygenic photosynthesis, they eventually enriched the Earth's atmosphere with oxygen¹⁷. This in turn affected the availability of metals in the aquatic environment⁵. Previously available Fe^{2+} became oxidised to significantly less soluble Fe^{3+} which precipitates as ferric hydroxide at neutral pH. Concentrations of Co^{2+} , Ni^{2+} and Mn^{2+} dramatically decreased as well¹⁸. In contrast, insoluble sulfides of Cu^{2+} , Cd^{2+} and Zn^{2+} were oxidised to the corresponding soluble sulfates. This radical change in the chemical composition of the environment caused development of effective mechanisms for metal sensing and homeostasis in cyanobacteria¹⁹.

Interestingly, in comparison to other prokaryotes, cyanobacterial demands for metals are higher, probably due to the presence of photosynthetic machinery¹⁹. Both cyanobacteria and plants require iron for photosystems I and II and iron-rich photosynthetic electron transport^{20,21}. Manganese is essential for the enzyme oxidising water and releasing electrons required for photosynthesis²². The copper-containing electron carrier plastocyanin participates in the transfer of electrons between cytochrome and photosystem I²³. Additional examples of metal co-factors utilised by phytoplankton including cyanobacteria are given in Table 1.2.

Fe, Co, Cu, Mo, and Zn show nutrient-like vertical distribution profiles in seawater columns. Their concentrations are reduced in the euphotic zone due to assimilation by plankton, and increased with depth due to remineralisation as a result of cellular degradation²⁴. While the important role that iron plays in cyanobacterial metabolism and global biogeochemical cycles has been recognized for some time²⁵, the importance of other trace metals is less clear despite nutrient-like profiles and hypothesized demands for picophytoplankton.

Table 1.2 Common metalloproteins present within marine phytoplankton.

The Table is taken from Twining and Baines, 2012¹⁸ with minor adjustments.

| Metal | Protein | Function |
|--------------|---------------------------------|--|
| Fe | Cytochromes | Electron transport in photosynthesis and respiration |
| | Ferredoxin | Electron transport in photosynthesis and N fixation |
| | Other Fe-S proteins | Electron transport in photosynthesis and respiration |
| | Nitrate and nitrite reductase | Conversion of nitrate to ammonia |
| | Chelatase | Porphyrin and phycobiliprotein synthesis |
| | Nitrogenase | N fixation |
| | Catalase | Conversion of hydrogen peroxide to water |
| | Peroxidase | Reduction of reactive oxygen species |
| | Superoxide dismutase | Disproportionation of superoxide to H ₂ O ₂ and O ₂ |
| Zn | Carbonic anhydrase | Hydration and dehydration of carbon dioxide |
| | Alkaline phosphatase | Hydrolysis of phosphate esters |
| | RNA polymerase | Nucleic acid replication and transcription |
| | tRNA synthetase | Synthesis of tRNA |
| | Carboxypeptidase | Hydrolysis of peptide bonds |
| | Cu,Zn-Superoxide dismutase | Disproportionation of superoxide to H ₂ O ₂ and O ₂ |
| Mn | O ₂ -evolving enzyme | Oxidation of water during photosynthesis |
| | Superoxide dismutase | Disproportionation of superoxide to H ₂ O ₂ and O ₂ |
| | Arginase | Hydrolysis of arginine to ornithine and urea |
| | Phosphotransferases | Phosphorylation reactions |
| Ni | Urease | Hydrolysis of urea |
| | Superoxide dismutase | Disproportionation of superoxide to H ₂ O ₂ and O ₂ |
| Cu | Plastocyanin | Photosynthesis electron transport |
| | Cytochrome oxidase | Mitochondrial electron transport |
| | Ascorbate oxidase | Ascorbic acid oxidation and reduction |
| | Cu,Zn-Superoxide dismutase | Disproportionation of superoxide to H ₂ O ₂ and O ₂ |
| | Multicopper ferroxidase | High-affinity transmembrane Fe transport |
| Co | Vitamin B12-dependent enzymes | A number of enzymes, C and H transfer reactions |
| Cd | Carbonic anhydrase | Hydration and dehydration of CO ₂ in diatoms |
| Mo | Nitrate reductase | Conversion of nitrate to ammonia |
| | Nitrogenase | N fixation |

1.3 ZINC IN CYANOBACTERIA

Zinc, typically the second most abundant cellular trace metal after iron, has been predicted to be required for between 5 and 9% of the predicted proteomes of most organisms²⁶. This is in most cases more than either the predicted iron or copper sub-proteomes²⁶. Indeed, for autotrophic picoplankton zinc cellular contents were of a similar order of magnitude as for iron¹⁸. In particular, zinc is thought to be required in the biochemical process of carbon fixation²⁷, arguably one of the most important biogeochemical processes. Carbonic anhydrase, an enzyme facilitating the hydration/de-hydration of carbon dioxide usually utilises zinc as a co-factor²⁸. Indeed, a previously formulated “zinc hypothesis” has linked global atmospheric CO₂ levels and oceanic zinc concentrations, with higher zinc levels promoting a reduction in atmospheric CO₂²⁹.

In addition, zinc might be required for phosphorus utilisation via alkaline phosphatases that hydrolyze organic phosphates³⁰. It has previously been shown that upon addition of zinc to seawater samples with low concentrations of inorganic phosphate but replete organic phosphates, rates of chlorophyll *a*-normalized alkaline phosphatase activity increased significantly³¹. The effect was less noticeable for samples with elevated concentrations of inorganic phosphate.

Interestingly, marine cyanobacteria show little or no measurable requirement for zinc⁴, possibly due to adaptation to very zinc-limited conditions in the oceans. However, these organisms have a remarkable ability to concentrate zinc 10⁵-10⁷ fold from their environment³². Certainly, most marine cyanobacterial genomes checked contain putative members of zinc-sensing and uptake-related protein families such as the “zinc uptake regulator” (Zur) and a high affinity inner membrane zinc transporter (ZnuABC; see sections 1.3.2.1 and 1.3.2.1.2 respectively)³³. Moreover, bioinformatic prediction of Zur-binding sites (see section 1.3.1.1.1) in a variety of sequenced cyanobacterial genomes suggests that zinc could be an essential element in cyanobacterial metabolism³³.

1.3.1 Zinc sensing in cyanobacteria

In bacteria, expression of zinc uptake/efflux components is often controlled by zinc sensors or zinc-responsive transcriptional regulators, which are DNA-binding

proteins. They bind to operator-promoter regions upstream of certain genes in a zinc-dependent manner, regulating their transcription. Depending on their role, regulators can be separated into two types. One type (SmtB, ZiaR, AztR, BxmR and ZntR, see below) deals with zinc toxicity when zinc levels are high, by de-repressing or activating the expression of genes encoding efflux systems (e.g. ATPases (see section 1.3.3.1.2), CDF (cation-diffusion facilitators, see section 1.3.3.1.1) and RND (resistance-nodulation-division, see section 1.3.3.1.3) proteins) or sequestration proteins (metallothioneins, see section 1.3.3.2). The other type of sensors (Zur, see section 1.3.1.1 and AdcR, see below) function during zinc deficiency by de-repressing or activating the expression of proteins engaged in the uptake (ABC systems, porins, ZIP proteins), acting as chaperones, or proteins reducing the necessity for zinc or substituting zinc-containing proteins^{34,35}.

SmtB is a zinc excess sensor found in cyanobacteria that usually acts to prevent high zinc levels accumulating in the cytosol. SmtB in its apo-form is a transcriptional repressor of genes encoding zinc efflux or sequestration proteins. Upon binding zinc SmtB de-represses their transcription by changing into a non-DNA binding conformation³⁶. SmtB is common in freshwater cyanobacteria – but apparently not present in marine picocyanobacteria^{37,38}. *smtB* was identified as a divergently transcribed gene to a zinc metallothionein *smtA* (see section 1.3.3.2) in the genome of the freshwater *Synechococcus* sp. PCC7942³⁹. SmtB is a part of the ArsR family of repressors involved in resistance to heavy metals and *smtB* has been shown to belong to the *smt* locus together with *smtA* in that species.

SmtB-like zinc sensors are also found in cyanobacteria: ZiaR in *Synechocystis* sp. PCC6803⁴⁰, BxmR in *Oscillatoria brevis*⁴¹ and AztR in *Anabaena* PCC 7120⁴². All apparently repress transcription of genes in the absence of zinc: *ziaA*, *bxa1* and *aztA* respectively, encoding metal efflux P-type ATPases (see section 1.3.3.1.2). *ziaR* and *aztR* share a common operator-promoter region with the corresponding genes negatively regulated by themselves. BxmR regulates a divergently transcribed operon encoding *bxmR* itself and *bmtA* (see section 1.3.3.2) while *bxa1* is located in a discrete operon. *ziaA* mutants of *Synechocystis* sp. PCC6803 are hypersensitive to zinc. De-repression of BxmR regulated genes in *Oscillatoria brevis* is mediated by Ag¹⁺, Cu¹⁺, Zn²⁺ and Cd²⁺. Finally, the AztR sensor of *Anabaena* PCC 7120 is Zn²⁺/Cd²⁺/Pb²⁺ specific.

In contrast, another zinc excess sensor, ZntR, is an activator in its metallated state and belongs to the MerR-family of activators⁴³. It is only known from some proteobacteria³⁸. In *E. coli* it regulates expression of *zntA* encoding a zinc exporting ATPase (ZntA, see section 1.3.3.1.2 for ATPases)⁴⁴.

Of all the prokaryotic zinc deficiency sensors, Zur and AdcR are the most well-known. AdcR belongs to the multiple antibiotic resistance repressor (MarR) family⁴⁵. Under conditions of zinc sufficiency AdcR represses expression of a high-affinity zinc importer AdcABC, which is an ABC type high affinity zinc uptake system in some Gram-positive bacteria analogous to ZnuABC in Gram-negative bacteria (see section 1.3.2.1.2). No homologues of AdcR have been identified in cyanobacteria.

The genome of *Synechococcus* sp. WH8102, the marine cyanobacterium under study in this thesis (see section 1.4), only contains one putative zinc sensor (Zur). Properties of Zur proteins are reviewed next (see section 1.3.1.1).

1.3.1.1 Zinc uptake regulator

One of the most prevalent genes for zinc homeostasis identified in a majority of cyanobacterial genomes is a gene likely to encode the zinc sensor Zur³³. It is a member of the Fur (ferric uptake regulator) family of transcriptional repressors. Other members of this family are specific sensors for iron (Fur), manganese (Mur), nickel (Nur) and hydrogen peroxide (PerR)⁴⁶.

In zinc-replete conditions, Zur with zinc bound allosterically switches into a DNA-binding configuration and binds to a specific sequence in the operator-promoter regions upstream of Zur-regulated genes (see section 1.3.1.1.1), repressing their transcription³³. Like for other sensors of this family, when zinc is below a certain threshold, no zinc is bound in the sensory site of the protein, which leads to lower DNA-binding affinity, allowing RNA polymerase to bind to the promoter elements. This leads to the transcription of genes involved in e.g. zinc uptake to be de-repressed. Typically, the zinc binding affinity of Zur proteins usually lies in the pico- to femtomolar range; this is the case for Zur proteins from a variety of phyla³⁵.

No structures of any of the cyanobacterial Zurs or Fur-family proteins are available. However, three resolved structures of bacterial Zurs are known: *Streptomyces coelicolor* 3mwm⁴⁷, *E. coli* 4mtd⁴⁸ and *Mycobacterium tuberculosis*

2o03⁴⁹. The structure of a complex of two Zur dimers bound to the zur-box of an *E. coli* DNA fragment is shown in Figure 1.3.

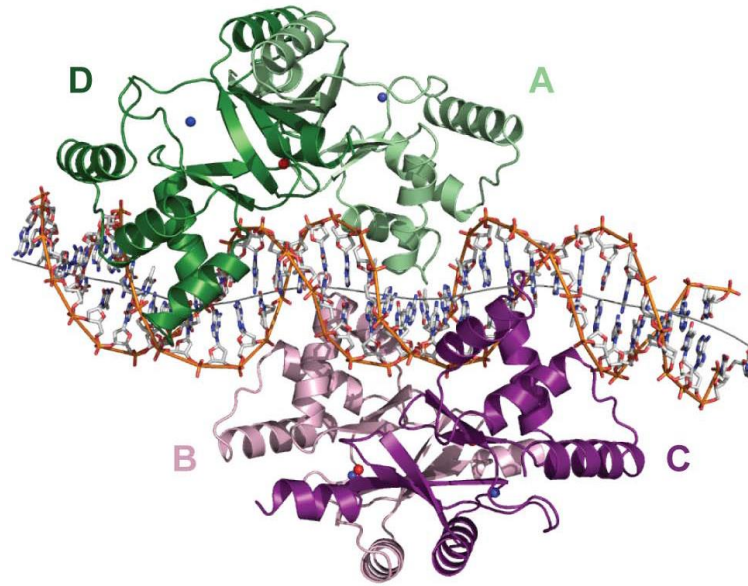


Figure 1.3 X-ray crystallographic structure of a complex of two Zur dimers with a 31 bp fragment of double-stranded DNA (*E. coli*), taken from⁴⁸.

The four Zur subunits are labelled A–D. Zinc ions at structural sites are shown as red balls; zinc ions in sensory sites are shown as blue balls.

1.3.1.1.1 *Zur*-boxes in cyanobacterial genomes

Like many transcriptional repressors, Zur is able to bind to certain palindromic sequences of DNA called “zur-boxes” located upstream of Zur-regulated genes in the operator-promoter regions. The consensus sequence for the predicted cyanobacterial zur-box based on 84 different genes from 14 cyanobacterial species is shown in Figure 1.4³⁸. Since for different bacterial phyla consensus sequences differ, this implies that the proteins that bind to these may also differ accordingly.



Figure 1.4 Predicted cyanobacterial Zur binding box taken from the RegPrecise database³⁸.

Larger letters indicate highly conserved nucleotides.

DNA binding K_{DS} of Zur proteins to zur-boxes are typically in the nanomolar range³⁵.

1.3.1.1.2 *zur*-mutants

In order to identify target genes of a transcriptional regulator as well as to understand its role *in vivo*, it is common to use an approach involving gene deletion or disruption. For *zur* this approach has been used for bacterial and cyanobacterial species^{50–52}. Since expression of Zur-regulated genes is expected to be de-repressed at any level of zinc in *zur*-mutants, this potentially allows the detection of these *zur*-regulated genes which would otherwise be difficult to trace. For example, this approach together with bioinformatics allowed to identify a new set of genes for *Amycolatopsis japonicum* involved in the synthesis of a potential zincophore ethylenediamine-disuccinate (EDDS)⁵³. Other examples of the components of zinc uptake potentially regulated by Zur and found either in *zur* mutants or by bioinformatics approaches are described below (see section 1.3.2).

Specifically for marine *Synechococcus*, a *zur* deletion mutant was constructed in the euryhaline strain PCC7002 and used to investigate zinc homeostasis mechanisms⁵⁴. Open ocean picocyanobacteria are predicted to have different zinc bio-accumulation mechanisms due to a dissimilar and smaller genome as a consequence of significant differences in the environmental zinc concentration: in open ocean waters zinc is significantly less abundant compared to coastal waters where zinc has higher bio-availability⁵⁵.

1.3.2 Dealing with zinc limitation

1.3.2.1 Zur as a transcriptional repressor

Probably the most well-known Zur-regulated genes are those involved in zinc uptake. This includes outer membrane (see section 1.3.2.1.1) and inner membrane components (see section 1.3.2.1.2). In addition, the Zur regulons of cyanobacteria could include zinc chaperones (see section 1.3.2.1.3), ribosomal proteins, Zur itself and a number of enzymes (see section 1.3.2.1.4). Members of bacterial Zur regulons have been reviewed in detail previously³⁵.

Since only a fraction of the zinc metabolism components regulated by Zur in cyanobacteria have been studied experimentally, I include in the sections below some members of the Zur regulon from other phyla where applicable.

1.3.2.1.1 Outer membrane zinc transport in cyanobacteria

Transport of zinc across the cyanobacterial outer membrane is poorly understood. However, Barnett *et al.* (2014)³² suggested that active mechanisms should operate because total zinc concentrations in marine surface waters are only in the low nanomolar range, with bioavailable free zinc in the picomolar range^{56,57}. Thus, passive diffusion is unlikely.

TonB-dependent receptors (TBDRs) are outer membrane proteins found in Gram-negative bacteria. They are responsible for the active uptake of extracellular molecules such as siderophores, vitamin B12, nickel complexes, and carbohydrates⁵⁸. Napolitano and colleagues identified two genes encoding TonB-dependent receptors which are regulated by Zur in the freshwater cyanobacterium *Anabaena* sp. PCC7120⁵⁹. They found that *alr3242* and *alr4028-4029*, which encode TBDRs, are induced and reach high levels of expression in an *Anabaena zur* deletion mutant, suggesting that these are probably TBDRs specific for the import of zinc or zinc complexes across the outer membrane.

Various other TBDRs which are up-regulated in zinc-depleted conditions and regulated by Zur have been found in other bacteria: *Neisseria meningitides*⁶⁰, *Acinetobacter baumannii*⁶¹ and *Pseudomonas protegens*⁶². These zinc uptake components were called ZnuD. Homologues of ZnuD were also identified by a BLAST search in several genera of pathogenic bacteria: *Moraxella*, *Haemophilus*, *Mannheimia*, *Acinetobacter*, *Pasteurella*, *Bordetella*, *Actinobacillus* and *Neisseria*⁶⁰.

1.3.2.1.2 Inner membrane zinc uptake

There are two families of inner-membrane transporters involved in zinc acquisition: high affinity ZnuABC systems and lower affinity ZIP-transporters (ZRT/IRT-like proteins i.e. zinc or iron regulated transporters).

ATP-binding cassette (ABC) transporters

Of all the known bacterial zinc uptake systems ZnuABC is the most common³⁴. It is a member of the ATP-binding cassette (ABC) transporter family and may play a role in the virulence of some pathogens⁶³. The ZnuABC system is the only zinc-specific ABC-type transporter described for cyanobacteria, in particular in *Synechocystis*⁵. Putative ZnuABC systems have been found in the majority of cyanobacterial species^{3,37}.

ZnuABC usually consists of three proteins: a cytosolic dimeric ATPase (ZnuC), a membrane-spanning dimeric permease (ZnuB), and a monomer of a substrate-binding protein (SBP, ZnuA). In Gram-positive bacteria, the analogue of the latter protein is located at the cell surface (e.g. AdcA from *Streptococcus pneumoniae*), whilst in Gram-negative bacteria (Znu systems), the ZnuA SPBs are periplasmic and often anchored to the inner membrane³⁴. The structure of an ABC system is shown in Figure 1.5.

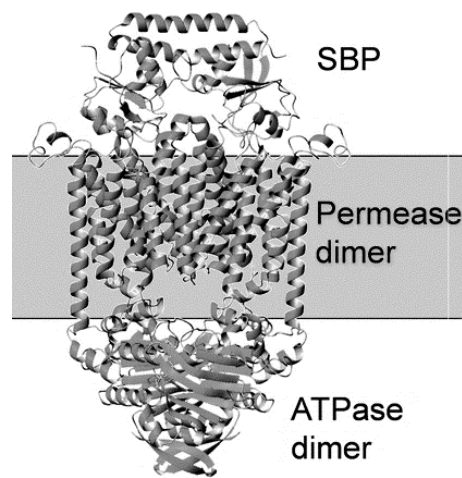


Figure 1.5 The ABC transporter shown is BtuCD-F (4FI3) which is used for cobalamin uptake, taken from³⁴.

Expression of ZnuABC is often regulated by Zur in bacteria³⁵, including cyanobacteria^{54,59}. Moreover, the genes encoding ZnuABC are up-regulated in *zur* deletion mutants of *Anabaena* sp. PCC7120⁵⁹ and *Synechococcus* sp. PCC7002⁵⁴. In *Synechococcus* sp. WH8102 (see section 1.4) the periplasmic binding protein component (ZnuA, SYNW2481) of a predicted ABC system has been shown to bind to a Zn²⁺-charged IMAC column³² adding support to the idea that this protein is involved in zinc uptake in this organism.

ZIP transporters

ZIP transporters are lower affinity inner membrane transporters compared to ZnuABC. An example is *zupT*, encoding the *E. coli* ZIP protein⁶⁴. Usually, ZIP proteins from different phyla have similar structures: they contain eight transmembrane helices, with their N- and C-termini being in the periplasm facing

outwards from the inner membrane (see Figure 1.6). In addition, some of them have a long histidine-rich loop located between transmembrane helices 3 and 4⁶⁵.

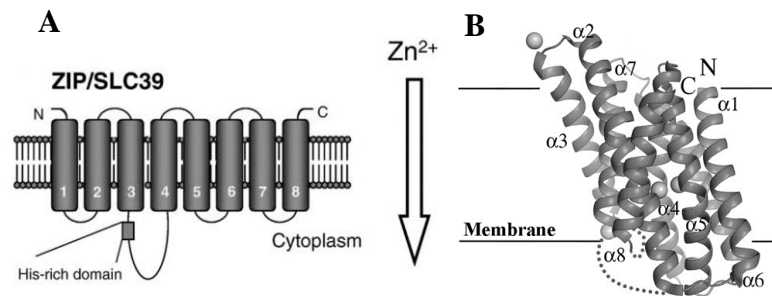


Figure 1.6 The structure of ZIP transporters.

A. Topology with numbered transmembrane helices, taken from⁶⁵. B. ZIP from *Bordetella bronchiseptica* (pdb 5TSA), a Gram-negative pathogen, taken from⁶⁶.

Two ZIP transporters Zip11 (*Npun_F3111*) and Zip63 (*Npun_F2202*) were identified in the cyanobacterium *Anabaena punctiforme* sp. PCC73120⁶⁷. Mutants lacking one or both of these transporters showed reduced zinc uptake. Furthermore, heterologous overexpression of both of these ZIP proteins in an *E. coli* *zupT736::kan* mutant strain i.e. lacking *zupT*, restored zinc uptake. These findings provide evidence that ZIP transporters might also mediate zinc uptake in some cyanobacteria.

Bioinformatic analysis identified a zur-box upstream of a putative ZIP-transporter (*sync_2443*) in the genome of *Synechococcus* sp. CC9311³³. This indicates that ZIP-transporters might be regulated by Zur. However, further work is required to clarify this hypothesis. The only other recognizable homologs of *sync_2443* are found in *Synechococcus* sp. WH8020 (another clade I marine *Synechococcus*) and *Synechococcus* sp. BIOS-E4-1 (a CRD1b clade strain) but no other marine *Prochlorococcus* or *Synechococcus* including *Synechococcus* sp. WH8102.

1.3.2.1.3 Putative zinc chaperones (the COG0523 family)

Putative genes for *zur* and *znuABC* are often co-localised with a member of the COG0523 family in a significant number of cyanobacterial genomes^{33,37}. Indeed, the cyanobacterium *Anabaena* sp. PCC7120 has two COG0523 genes regulated by Zur⁵⁹. Members of this family were found in most organisms sequenced and could be separated into several subgroups depending on their function⁶⁸ e.g. proteins facilitating incorporation of Ni into urease (UreG) or hydrogenase (HypB) or an iron-chaperone

participating in maturation of nitrile hydratase (Nha3). COG0523 proteins apparently play a role in the assembly of metalloenzymes and could be metallochaperones.

It has been suggested that a significant number of COG0523 family members are also involved in zinc homeostasis⁶⁸⁻⁷⁰. For *Bacillus subtilis*, the COG0523 member YciC is Zur-regulated, leading to the suggestion that it may function as a metallochaperone to deal with zinc limitation⁶⁹. The *yciC* gene was up-regulated under low zinc conditions and in a *zur* knockout mutant. The target enzyme(s) of these zinc chaperones are still unclear. It has been suggested that it is unlikely that there is a specific target zinc enzyme, because there are too many destinations for this metal³³.

1.3.2.1.4 Other members of the Zur regulon

Other potential members of bacterial Zur regulons dealing with zinc limitation are ribosomal proteins, enzymes and Zur itself.

De-repressing zinc-free ribosomal proteins under zinc limiting conditions apparently facilitates the release of Zn²⁺ from ribosomes which can then be utilized for other target processes³⁵. This strategy is very common for prokaryotes. However, regarding cyanobacteria, so far only the genome of *Prochlorococcus* sp. CCMP1986 was predicted to have zur-boxes upstream of a few ribosomal proteins (S7, S12, and S14p/S29e)³³.

Zur-regulated enzymes are common in bacteria³⁵. Similar to ribosomal proteins, these are apparently zinc-free alternatives of enzymes that normally require zinc as a co-factor. For example, Zur regulons of the two cyanobacterial species *Anabaena* sp. PCC7120 and *Synechococcus* sp. PCC7002 contain *hemB2* genes encoding porphobilinogen synthase^{52,54}.

Autoregulation of Zur has also been commonly observed for a variety of prokaryotic phyla. Several species of cyanobacteria are predicted to have zur-boxes upstream of *zur* genes: *Synechococcus* sp. PCC7002, *Synechocystis* sp. PCC6803, *Cyanothece* sp. ATCC 51142 and *Cyanothece* sp. PCC7425³⁸. Since usually Δzur mutants are constructed to identify Zur regulons, expression of *zur* itself might be completely abolished due to the absence of this gene in the genomes of the mutants. Experimentally, autoregulation could be confirmed by electrophoretic mobility shift assays (EMSAs) showing binding of a Zur protein to its own upstream region and by

differential expression of *zur* mRNA levels as a response to changing zinc concentrations for the wild type organism.

1.3.3 Dealing with zinc excess

Despite zinc being an essential element for all kingdoms of life including cyanobacteria, its excess could be toxic. Recently it was shown that Zur, despite being called a “deficiency sensor” and a transcriptional repressor, can act to protect cells from zinc excess by activating expression of genes involved in zinc efflux (see Mikhaylina *et al.* (2018)³⁵ for a review). This aspect is briefly described in section 1.3.3.1. Other well-known transcriptional regulators dealing with zinc excess in prokaryotes are the SmtB/SmtB-family and ZntR proteins (see section 1.3.1).

1.3.3.1 Zur as a transcriptional activator

Currently, there is no evidence for Zur acting as a transcriptional activator in cyanobacteria. However, in other bacterial phyla this has already been observed (see below). Typically, Zur activates genes encoding components dealing with zinc toxicity such as zinc efflux pumps. Most marine cyanobacterial strains do not possess these pumps, probably due to the very low concentrations of zinc in open-ocean seawaters³³.

1.3.3.1.1 Cation-diffusion facilitator (CDF)

Cation-diffusion facilitator proteins are transmembrane proteins that efflux metal ions such as Cd²⁺, Zn²⁺ and Co²⁺ out of the cytosol when they are in excess^{34,71}. CDFs are found in both prokaryotes and eukaryotes and are secondary active transporters that do not work by diffusion as might be assumed from their name³⁴. Zur-activated expression of two CDF efflux systems has been shown for *czcD* from *Xanthomonas campestris*⁷² and *zitB* from *Streptomyces coelicolor*⁷³.

Interestingly, DNaseI footprinting experiments showed that *X. campestris* Zur bound to a 59-bp GC-rich sequence upstream of *czcD* which was dissimilar to the consensus zur-box from this species. In contrast, the operator-promoter region of the *zitB* gene from *S. coelicolor* contained a zur-box typical for this species. In the latter case, activation apparently takes place through Zur oligomerisation at higher zinc concentrations. Therefore, the activation mechanism apparently varies in different phyla, an aspect that requires further investigation³⁵.

Two zinc efflux transporters of the CDF family were identified in the cyanobacterium *Anabaena punctiforme* (Npun_F0707 and Npun_F1794)⁷⁴. However, no regulation mechanisms were suggested for these, and it is unclear whether they are regulated by Zur.

1.3.3.1.2 P-type ATPases

Two further examples of the activation of zinc-efflux systems by Zur were shown in *Corynebacterium glutamicum* (a non-pathogenic soil γ -proteobacterium)⁷⁵ and *Caulobacter crescentus* (an oligotrophic aquatic α -proteobacterium)⁷⁶. Deletion of *zur* in these bacteria caused repression of *zra* and *zntA*, respectively, which both encode P-type ATPases likely used for zinc efflux. P-type ATPases are found in prokaryotic and eukaryotic cells and function as active transmembrane transporters for a variety of specific cations³⁴.

Several cyanobacteria are predicted to have zinc-transporting P-type ATPases: *Anabaena* sp. PCC7120, *Anabaena variabilis* sp. ATCC29413, *Lyngbya* sp. PCC8106, *Synechocystis* sp. PCC6803 and *Oscillatoria brevis*³⁷. However, none was predicted to be regulated by Zur³⁸. Interestingly, the zinc ATPases of *Synechocystis* sp. PCC6803 (encoded by *ziaA*⁴⁰), *Oscillatoria brevis* (encoded by *bxaI*⁴¹) and *Anabaena* PCC7120 (encoded by *aztA*⁴²) were found to be regulated by SmtB-like zinc excess sensors ZiaR, BxmR and AztR, respectively (see section 1.3.1).

1.3.3.1.3 RND systems

In addition to *zntA*, expression of another putative zinc efflux system, *czr*, was repressed in the *C. crescentus* Δ zur mutant⁷⁶. This gene encodes a Resistance-Nodulation-Division (RND) system. Metal transporting RND systems are thought to be involved in exporting metals to the cell exterior or medium and usually comprise three proteins: an inner membrane trimeric A protein, an outer membrane trimeric C protein and six periplasmic B proteins linking A and C³⁴.

Two putative RND metal transporters (all7631 and all7618) have been found in *Anabaena* sp. PCC7120. However, it is unclear whether these are zinc exporters⁷⁷. In addition, no predicted zur-boxes were identified upstream of these genes³⁸.

1.3.3.2 Bacterial metallothioneins

Numerous eukaryotes and some prokaryotes including cyanobacteria also possess metallothioneins (MTs) to deal with zinc excess. MTs are low-molecular weight proteins with a high cysteine content and significant metal ion binding capacity (usually for Zn^{2+} , Cu^+ and Cd^{2+})⁷⁸. It is thought that bacterial MTs, and some cyanobacterial MTs, are expressed in zinc (and cadmium) excess conditions in order to sequester a surplus of these metals within the cell. Using such a mechanism could de-toxify excess zinc within the cytosol without the requirement to efflux this potentially useful resource. Thus, MTs could also be utilised as internal zinc storage proteins⁷⁹. This could be beneficial for cyanobacteria in oligotrophic environments in comparison to those expressing efflux pumps. Another proposed role for MTs is as a metallochaperone which may potentially transfer metal ions to other subcellular locations⁷⁹.

A cyanobacterial metallothionein identified in the freshwater cyanobacterium *Synechococcus* sp. PCC7942 (encoded by *smtA*) was shown to be transcriptionally regulated by the zinc sensor SmtB and was de-repressed under high zinc³⁹. Knockout mutants lacking *smtA* showed hypersensitivity to zinc and cadmium compared to the WT, while no difference was observed for copper⁸⁰.

Bacterial MTs, unlike most eukaryotic metallothioneins, contain secondary structure elements that form a zinc finger fold^{81,82}. This facilitated determination of the NMR solution structure of SmtA from *Synechococcus* sp. PCC7942 (see Figure 1.7)⁸¹.

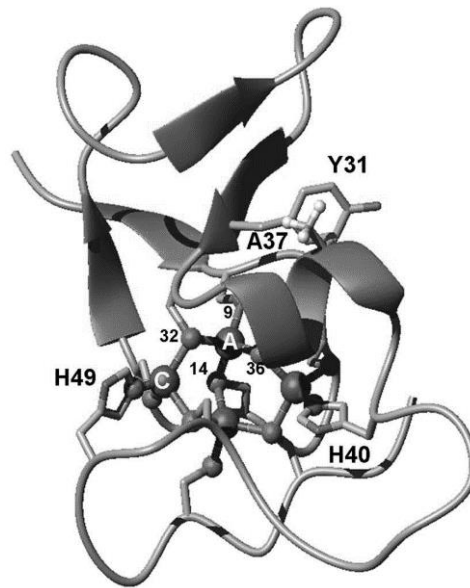


Figure 1.7 NMR solution structure of SmtA from *Synechococcus* sp. PCC7942⁸¹, taken from³³.

MTs from non-vertebrate organisms including bacteria, unlike those from vertebrates, may contain aromatic residues especially histidine residues, that participate in metal-binding³⁷.

Although absent in *Prochlorococcus* genomes, MTs are prevalent in marine *Synechococcus* – including *Synechococcus* sp. WH8102 (SYNW0359), despite not having identified homologs of SmtB-type regulators³³. So far, it has been unclear which transcriptional regulators control their expression.

1.4 ZINC METABOLISM IN *SYNECHOCOCCUS* SP. WH8102

Synechococcus sp. WH8102 is a representative of clade III from sub-cluster 5.1, a clade that occupies oligotrophic open ocean environments. *Synechococcus* sp. WH8102 was isolated from the tropical Atlantic Ocean (22°30'N, 65°36'W)⁸³ and possesses a genome of ~2.43 Mb with 2,519 protein-coding genes⁸⁴.

Synechococcus sp. WH8102 is unusual amongst all other marine *Synechococcus* in being able to carry out swimming motility, albeit via an unknown mechanism, presumably an adaption to the oligotrophic marine environment it inhabits⁸⁵. Two unique genes were found to be required for motility: *swmA*⁸⁶ and *swmB*⁸⁷. *swmB* is 32375 bp and constitutes >1% of the total genome length⁸⁴. The genome of *Synechococcus* sp. WH8102 also contains sixteen putative phage integrases indicative of horizontal gene transfer. In addition, about 5-6% of the predicted open reading frames encode transporters, predominantly of the ABC type. Only one putative copper P-type ATPase has been identified⁸⁴.

Putative genes involved in zinc metabolism in this species were identified by bioinformatics and include *zur*, two *znuABC* systems, *bmtA*, and a member of the COG0523 family³³. Interestingly, as predicted by the RegPrecise database, one of the *znuABC* systems and *bmtA* contain predicted cyanobacterial *zur*-boxes in their operator-promoter regions³⁸. These putative Zur regulated components are shown in Table 1.3. The ability of the *zur*-boxes of this species to bind SynZur have been shown recently⁸⁸. A schematic representation of the suggested Zur regulation mechanism is shown in Figure 1.8.

Table 1.3 Genes predicted to be regulated by Zur in *Synechococcus* sp. WH8102³⁸.

Genes for *znuB* and *znuC* are divergently transcribed with *znuA* and share the same predicted zur-box.

| Zur-box location | Gene name | Annotation | Predicted function |
|------------------|-----------|------------|--|
| | SYNW0970 | cce_1485 | Mn/Zn ABC transporter, ATP-binding protein (ZnuC) |
| | SYNW0969 | cce_1486 | Mn/Zn ABC transporter, permease protein (ZnuB) |
| | SYNW0971 | cce_1484 | Mn/Zn ABC transporter, substrate-binding protein (ZnuA) |
| | SYNW0359 | SmtA | Zn metallothionein (BmtA) |

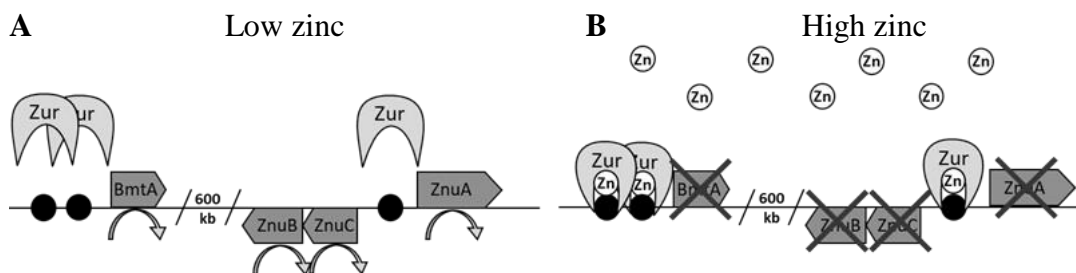


Figure 1.8 Predicted mechanism of action of SynZur as a transcriptional repressor in *Synechococcus* sp. WH8102.

A. At low zinc concentrations SynZur is unable to bind to the operator-promoter region of SynZur-regulated genes and transcription of genes is de-repressed. **B.** When the free intracellular zinc concentration is high, zinc binds to SynZur and changes its conformation so that SynZur is able to bind to zur-boxes and stop transcription of genes by preventing RNA-polymerase from binding.

In addition, a number of zinc-binding proteins were identified in *Synechococcus* sp. WH8102 using a zinc-charged IMAC column³². Amongst those identified were several enzymes including alkaline phosphatase, several periplasmic binding proteins including a predicted ZnuA (SYNW2481) and two outer membrane porins (SYNW2224 and SYNW2227). Intriguingly, the predicted ZnuA (SYNW2481) captured on the IMAC column is encoded by a different gene to the ZnuA (SYNW0971) predicted to be regulated by SynZur. No binding of SynZur to the operator-promoter region of SYNW2481 was observed by EMSA⁸⁸. Therefore, it is possible that *Synechococcus* sp. WH8102 possesses two independent zinc uptake systems.

Zinc transport across the outer membrane (and metal transport in general) of this cyanobacterium and related species is unclear. No homologs of TonB-dependent transporters have been identified in *Synechococcus* sp. WH8102. However, one of the identified outer membrane porins mentioned above (SYNW2224) was significantly more abundant under zinc-limiting conditions³². Similarly, expression of porins in *Pseudomonas protegens*⁶², and in *Pseudomonas aeruginosa*⁸⁹ was shown to be zinc-dependent. However, further experimental work is required in order to demonstrate the role of SYNW2224 in zinc bio-accumulation in *Synechococcus* sp. WH8102.

Predicted components of zinc metabolism in *Synechococcus* sp. WH8102 are summarised in Figure 1.9.

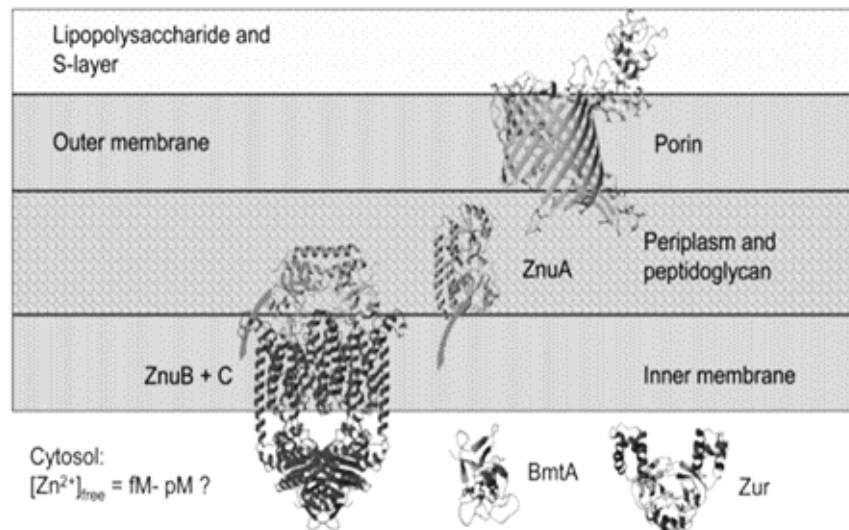


Figure 1.9 Predicted components of zinc metabolism in *Synechococcus* sp. WH8102, taken from³².

1.4.1 Open questions regarding zinc homeostasis in *Synechococcus* sp. WH8102

SynZur is the only predicted zinc uptake regulator in *Synechococcus* sp. WH8102. However, direct experimental evidence is required to assess its metal-binding specificity and function and thus confirm whether SynZur is indeed a true zinc-specific uptake regulator or another member of the Fur family of proteins. Assessment of the capacity of the predicted SynZur protein to bind DNA selectively in the presence of zinc using EMSAs should allow determination of whether this uptake regulator is a true Zur³⁵. It should be remembered though that DNA binding by SynZur does not have to be zinc specific since *in vivo* not only does DNA-binding

ability need to be considered, but also the zinc binding affinity of the regulator which should be in the range of expected intracellular free Zn^{2+} concentrations⁹⁰. Considering that affinities (K_{DS}) of Zur proteins for zinc are below picomolar in a variety of bacteria³⁵, a similar value is expected for SynZur since intracellular levels of free zinc in a cyanobacterium occupying oligotrophic waters are expected to be in a similar range or lower than for other bacterial phyla.

As well as no direct experimental evidence for the function of Zur in *Synechococcus* sp. WH8102 there are also no structures of any cyanobacterial Zur proteins, or indeed any cyanobacterial Fur-family proteins, yet. Sequential alignment of Fur-family proteins from different bacterial species, together with selected cyanobacterial sequences including that of the predicted Zur from *Synechococcus* sp. WH8102, showed that the sensory site in cyanobacterial Zur homologs has to be significantly different from those found in published structures³³. Several metal-binding residues identified in the few published structures are not conserved in cyanobacterial Zurs, and the cyanobacterial zur-box has some distinct features in comparison to zur-boxes from other phyla. Therefore, it is assumed that DNA-protein interactions for cyanobacterial Zur proteins could also differ significantly. Determining the structure of at least one of the cyanobacterial Zur proteins could clarify these differences.

Furthermore, the predicted Zur regulon of *Synechococcus* sp. WH8102 contains only four genes (three components of *znuABC* and *bmtA*). Typically, the oligotrophic marine environment is very nutrient poor and an active mechanism for zinc uptake across the outer membrane may be expected. Since no other zinc sensors have been identified, it is possible that the Zur regulon of *Synechococcus* sp. WH8102 contains some outer membrane components for zinc uptake as well as other genes. Identification of the *Synechococcus* sp. WH8102 Zur regulon could be achieved by differential gene expression analysis of a *zur*-knockout mutant in comparison to the WT. This would clarify whether the porins SYNW2224 and SYNW2227 and other zinc homeostasis components described in section 1.4 belong to the Zur regulon of *Synechococcus* sp. WH8102. Besides, phenotypic characterisation of a *zur*-mutant would clarify what role this gene plays in *Synechococcus* sp. WH8102 and help to determine whether it is a true zinc uptake regulator.

Finally, the putative MT (BmtA, SYNW0359) also deserves attention. As was mentioned in section 1.3.3.2, in freshwater cyanobacteria transcription of MTs is regulated by SmtB, a protein not found in *Synechococcus* sp. WH8102. However, in *Synechococcus* sp. WH8102 transcription of *bmtA* appears to be under the control of Zur as judged by the presence of two zur-boxes upstream of the gene. Zur is usually considered as a deficiency sensor and represses transcription of its regulon in zinc-rich conditions. This would contradict roles in dealing with zinc excess that metallothioneins are known to play and raises the question whether BmtA in *Synechococcus* sp. WH8102 acts as a classical MT, responsible for intracellular zinc sequestration in excess conditions. It was hypothesised that SynZur might work as a transcriptional activator for *synBmtA*⁸⁸. However, this has not been confirmed experimentally. An alternative hypothesis is that SynBmtA is expressed by cells as a response to zinc limitation³³. The presence of SynBmtA in the cytosol might support zinc uptake by lowering the free intracellular concentration, or it could act as a zinc chaperone providing zinc as a co-factor to enzymes such as alkaline phosphatase and carbonic anhydrase, the latter located in the carboxysome.

To understand what particular role BmtA plays in *Synechococcus* sp. WH8102, it is essential to characterise the isolated protein; this should include determination of its zinc affinity and a comparison with SynZur. Whether, and in which direction, zinc transfer between these two proteins occurs is also worthy of investigation. Based on the outcomes of these experiments, it may become possible to propose a hypothesis about the potential role and mode of action of BmtA in zinc metabolism. NMR studies could also clarify if SynBmtA has a structure similar to the known SmtA from freshwater *Synechococcus* sp. PCC7942. This species utilises its MT to deal with zinc excess; if the structure of SynBmtA is conserved it could mean that its function is conserved as well.

1.5 AIMS AND OBJECTIVES

In order to answer questions defined in section 1.4.1 and to confirm previous findings, the specific objectives of this project are outlined below:

1. Structural and functional characterization of the putative Zur protein from *Synechococcus* sp. WH8102 in order to confirm its role and unique properties (Chapter 3):
 - ❖ Optimisation of the over-expression and purification of SynZur using a previously constructed plasmid and method⁸⁸;
 - ❖ Biophysical characterization to verify metal stoichiometry and oligomeric state of SynZur;
 - ❖ Optimisation of DNA binding conditions and assessment of metal binding selectivity for DNA binding in order to verify SynZur function;
 - ❖ Determination of the zinc binding affinity of SynZur in order to compare it to other Zur proteins and to assess if the value lies in the physiologically acceptable range;
 - ❖ X-ray crystallography to structurally characterise SynZur in order to compare it to other Zur proteins with known structures and to verify the suggested distinctive properties of cyanobacterial Zurs.

2. Identification of the Zur regulon in *Synechococcus* sp. WH8102 and establishing the role of SynZur (Chapter 4):
 - ❖ Generation of a *Synechococcus* sp. WH8102 *zur* interposon mutant;
 - ❖ Phenotypic characterisation of the mutant to confirm involvement of SynZur in zinc metabolism;
 - ❖ Transcriptomic analysis by RNA-sequencing in order to identify the Zur regulon, to confirm regulation of predicted Zur-regulated genes and to verify the role of SynZur as a zinc sensor.

3. Structural and functional characterisation of the BmtA protein from *Synechococcus* sp. WH8102 in order to understand its role in zinc metabolism (Chapter 5):
 - ❖ Construction of an over-expressing plasmid, overexpression and purification of SynBmtA;

- ❖ Biophysical characterization of the protein to identify it after purification and to find its metal stoichiometry;
- ❖ Assessment of zinc binding affinity of SynBmtA and comparison with SynZur in order to understand the physiological role of the MT and to suggest a regulation mechanism;
- ❖ Structural studies of the SynBmtA protein in order to compare its fold with known metallothioneins expressed in response to zinc excess to see if SynBmtA possesses any unique properties due to potentially being regulated by the putative deficiency sensor SynZur.

Chapter 2.
General materials and methods

2.1 *ESCHERICHIA COLI* STRAINS AND GROWTH CONDITIONS

All *E. coli* strains were grown in liquid Luria-Bertani (LB) medium or streaked onto solid LB agar plates with either kanamycin (Km, 50 $\mu\text{g mL}^{-1}$), ampicillin (Amp, 100 $\mu\text{g mL}^{-1}$) and chloramphenicol (Cm, 30 $\mu\text{g mL}^{-1}$) added whenever appropriate. *E. coli* cells were grown overnight at 37 °C in sterile LB medium unless otherwise stated. Agar plates were prepared by adding Bacto-Agar (Difco; 10 g L⁻¹) to the appropriate amount of LB medium, autoclaving the mixture, cooling it down to 37 °C and adding the appropriate antibiotics followed by pouring the mixture into disposable polystyrene Petri-dishes (Fisher Scientific). Plates were kept at 4 °C and were warmed in the incubator for 1 hour preliminary to streaking.

For longer storage, to 1 mL overnight grown cells 1 mL sterile 80% (v/v) glycerol in LB medium was added. Cells were then snap-frozen in liquid nitrogen and kept at -80 °C. In order to recover *E. coli*, stock cultures were thawed on ice, streaked on plates and left overnight at 37 °C.

Table 2.1 *Escherichia coli* strains used in this work.

| Strain | Description | Supplier |
|-----------------------------------|---|---------------------------|
| S17-1 λ Pir ⁹¹ | TpR SmR <i>recA</i> , <i>thi</i> , <i>pro</i> , <i>hsdR</i> -M+RP4: 2-Tc:Mu: Km Tn7 λ pir | Dr. Joseph Christie-Oleza |
| Rosetta2 (DE3) pLysS | F ⁻ <i>ompT</i> <i>hsdSB</i> (rB ⁻ mB ⁻) <i>gal dcm</i> (DE3) pLysSpRARE23(Cm ^R) | Novagen |
| BL21 (DE3) pLysS | F ⁻ <i>ompT</i> <i>hsdSB</i> (rB ⁻ , mB ⁻) <i>gal dcm rne131</i> (DE3) pLysS (Cm ^R) | Novagen |
| DH5- α | F ⁻ Φ 80lacZ Δ M15 Δ (lacZYA-argF) U169 <i>recA1 endA1 hsdR17</i> (rK ⁻ , mK ⁺) <i>phoA supE44</i> λ - <i>thi-1 gyrA96 relA1</i> | Invitrogen |
| TEV system ⁹² | <i>E. coli</i> BL21 (DE3) CodonPlus-RIL containing the TEV protease expression vector pRK793 | Dr. Goran M. Rashid |

Table 2.2 List of the plasmids used in this work.

| Plasmid | Description | Supplier |
|---------------------------------|---|--------------------|
| pGP704CmKm ⁹¹ | Ap ^r Km ^r Cm ^r ; pGP704-derivative containing the Km ^r and Cm ^r gene at the <i>EcoRI</i> site of its MCS | Dr. Christie-Oleza |
| pGP704CmKm Zur | pGP704CmKm-derivative containing the <i>zur</i> (21-314) insert at the <i>XbaI</i> and <i>SalI</i> sites of its MCS | This study |
| pET155/D-topo Zur ⁸⁸ | pET155/D-topo derivative containing synthetic <i>zur</i> insert cloned into the topo-cloning site | Dr. Amira Ksibe |
| pET-24a(+) | T7 promoter-driven high-efficiency protein expression and sequencing vector; encodes Km ^r | Novagen |
| pET-24a BmtA | pET-24a(+)- derivative containing <i>bmtA</i> insert at the <i>NdeI</i> and <i>SacI</i> sites of its MCS | This study |

2.2 GENERAL MOLECULAR BIOLOGY TECHNIQUES

2.2.1 Polymerase Chain Reaction (PCR)

In order to amplify target genes or to perform a preliminary check of successful ligations or a crossover during conjugation, polymerase chain reactions (PCR) were carried out with the T100™ Thermal Cycler, Bio-Rad laboratories. If this reaction was carried out for amplifying the gene for further usage, the default volume of the reaction was 50 µL. If the PCR was used for checking, the default volume was 20 µL (amounts of all the components were proportionally decreased). MyTaq™ Red DNA Polymerase kit from Bioline (BIO-21108) was used. 5x MyTaq Reaction Buffer which was included in the kit already contained dNTPs, reaction buffer, MgCl₂ and a red loading dye at the optimal concentrations and allowed loading the reaction components directly onto the gel following the PCR. PCR components are shown in Table 2.3 and were used as recommended in the PCR kit instructions. Default PCR conditions were modified and are shown in Table 2.4.

Table 2.3 PCR components.

| Components | Amount |
|--------------------------|-------------|
| 5x MyTaq Reaction Buffer | 10 µL |
| Template DNA | 5-500 ng |
| Primers, 20 µM | 1 µL each |
| MyTaq Red DNA Polymerase | 0.25 – 1 µL |
| MilliQ water | Up to 50 µL |

Table 2.4 Default PCR conditions.

| PCR steps | Temperature | Time | Cycles |
|----------------------|-------------|-----------------------|--------|
| Initial denaturation | 95 °C | 3 min | 1 |
| Denaturation | 95 °C | 30 sec | 35 |
| Annealing | 55 °C | 30 sec | |
| Extension | 72 °C | 30 sec for every 1 kb | |
| Final Extension* | 95 °C | 5 min | 1 |

2.2.2 Agarose gel electrophoresis

Gel preparation: 0.7% (w/v) agarose in TBE buffer (Tris 10.8 g L⁻¹, pH 8.3; 5.5 g L⁻¹ boric acid and 10 mM EDTA) was heated to boiling point until all the agarose

was dissolved. 0.01% (v/v) ethidium bromide was added to the hot agarose solution and the solution poured into a plastic mould and left to harden for about 1 hour.

If the loading dye was required, 20% (volume) of a 6x loading dye (#B7024S, NEB) was added to each sample containing DNA. 10 μ L of the resulting mixture was loaded into each well. Two Quick-Load purple ladders from NEB were used to estimate the band sizes: 1 kb ladder (#N0552, NEB) and 100bp ladder (#N0551, NEB). TBE buffer was used to run gels. A potential of 100 V was applied to the electrodes for 40 minutes. The electrolyser system from Bio-Rad was used (PowerPac Basic™ Power Supply and Mini-Sub® Cell GT Cell, Bio-Rad).

2.2.3 DNA purification following PCR

To clean up DNA after the PCR step, the Promega Wizard® SV Gel and PCR Clean-Up System was used. The protocol included in the kit was used without modification.

An equal volume of Membrane Binding Solution was added to the PCR amplification product and transferred onto the SV Minicolumn where it was incubated for 1 minute. The SV Minicolumn was installed into the collection tube and spun down for 1 minute at 13000 rpm. Then 700 μ L of Membrane Wash Solution with ethanol was added on top of the column and centrifuged for 1 minute. This step was repeated with 500 μ L of Membrane Wash Solution and centrifuged for 5 minutes. Following this step, the column was left for 1 minute to allow evaporation of any residual ethanol. After drying, the column was transferred into a clean 1.5 mL microcentrifuge tube and 50 μ L nuclease-free water added on top of the minicolumn. The assembly was incubated at room temperature for 1 minute and centrifuged for 1 minute at 13000 rpm, resulting in the DNA solution being collected in the microcentrifuge tube. DNA concentration and purity was measured using a NanoDrop (see section 2.3.7).

2.3 GENE CLONING TECHNIQUES

2.3.1 Digestion of DNA with restriction enzymes

Typically, to amplify inserts PCR primers with designed restriction sites were used. The digestion buffer was tested for each restriction enzyme separately and for the vector prior to the main double digestion reaction, followed by agarose gel electrophoresis in order to assess digestion yield. Simultaneous double digestion of the vector and the insert, performed separately, was used for the cloning experiments. Further components of the double digestions were:

- 1 μ g vector DNA (pGP704CmKm or pET-24a) or the insert DNA (PCR amplified (see section 2.2.1) and purified (see section 2.2.3) *zur*₂₁₋₃₁₅ or *bmtA*);
- 1 μ L (10 units) of each of the two restriction enzymes (all from Promega): *XbaI* and *SalI* for *zur*₂₁₋₃₁₅ or *NdeI* and *SacI* for *bmtA*;
- 5 μ L 10x MultiCore Buffer (Promega);
- nuclease-free water was added up to 50 μ L total volume.

The components were mixed gently by pipetting in a PCR tube and incubated for 1 hour at 37 °C in a PCR machine.

2.3.2 DNA dephosphorylation

Dephosphorylation of the vector was performed using Calf Intestinal Alkaline Phosphatase (CIAP) from Invitrogen. A simplified protocol was used: 1 μ L (1 unit) of CIAP was added to the digestion mixture after the digestion reaction was completed. Then the mixture was left for 5 minutes at 37 °C; the resulting cut and dephosphorylated vector was purified as described in section 2.2.3 prior to ligation.

2.3.3 DNA ligation protocol

For all ligation reactions 200 ng vector DNA was used. The amount of the insert in ng was calculated from Equation 2.1 given in the protocol supplied with the ligation kit (Promega LigaFast™ Rapid DNA Ligation System, # M8221).

$$x = \frac{200 \text{ ng}}{v} * i * n \quad \text{Equation 2.1}$$

Where:

x – is the amount of the insert needed for the ligation, ng;

v – the vector size after double digestion, kb;

i – the insert size after double digestion, kb;

n – the insert/vector ratio, 2 was used by default, unless stated otherwise.

For the ligation reaction further components included 10 μL 2x ligation buffer; 2 μL T4 ligase; water was added up to 20 μL .

Reaction components were mixed gently by pipetting and left for 5 minutes at room temperature since the ligase used was a rapid ligase. This ligation mixture was used directly, i.e. without a DNA purification step, to transform competent *E. coli* as described in section 2.3.5. Competent *E. coli* cells were obtained as described in section 2.3.4.

2.3.4 Preparation of chemically competent *E. coli* cells

E. coli cells (S17-1 λPir , Rosetta2 (DE3) pLysS, BL21 (DE3) pLysS, DH5- α) were grown overnight at 37 $^{\circ}\text{C}$ in 5 mL LB with 35 $\mu\text{g mL}^{-1}$ chloramphenicol where appropriate. 100 μL overnight culture were inoculated into 100 mL sterile LB with chloramphenicol if required and incubated for 3-4 hours until OD_{600} reached a value of 0.4-0.6. Following this, cells were incubated on ice for 5 minutes and centrifuged for 5 minutes at 2500 rpm at 4 $^{\circ}\text{C}$.

Then, cells were gently re-suspended in 50 mL cold sterile CaCl_2 (0.1 M) and incubated on ice for 30 minutes. After centrifuging the mixture for 15 minutes at 2500 rpm at 4 $^{\circ}\text{C}$, the cell pellet was re-suspended in 1 mL cold sterile solution of 70% (w/v) 0.1 M CaCl_2 in MilliQ water and 30% (v/v) glycerol. Subsequently, 30 μL aliquots were pipetted into 1.5 mL micro centrifuge tubes using cold tips and snap-frozen in liquid nitrogen. Competent cells were kept at -80 $^{\circ}\text{C}$.

2.3.5 *E. coli* transformation

First, a 30 μ L aliquot of competent cells (see section 2.3.4.) was thawed on ice and 200 μ L cold sterile solution of 0.1 M CaCl₂ was added to competent cells together with the plasmid aliquot (1 μ L) or a ligation product (see section 2.3.3). The mixture was incubated on ice for 30 minutes followed by heat shock, which involved heating the tube in a water bath for 45 sec at 42 °C. Following that, the tube was placed on ice for 5 more minutes followed by addition of 1 mL warm LB without antibiotics. The mixture was incubated at 37 °C for 1 hour. Subsequently, if a purified plasmid was used for transformation, then 100 μ L of the mixture were spread onto an LB plate with the appropriate antibiotics. If a ligation product was used, then the tube was centrifuged for 3 minutes at 3000 rpm and LB medium was discarded by inverting the tube gently so that approximately 100 μ L of LB medium remained in the tube. The cell pellet was mixed with the remaining LB medium and spread onto an LB plate with antibiotics. Plates were incubated overnight at 30 °C. If colonies had appeared they were checked for the correct insert by PCR (see section 2.2.1) followed by agarose gel electrophoresis (see section 2.2.2). If the sizes of the PCR products were as predicted, the colony was inoculated into 5 mL liquid LB medium, grown overnight and plasmids were purified by Miniprep (see section 2.3.6) and sent for sequencing (see section 2.3.8).

2.3.6 Plasmid isolation from *E. coli*

The QIAprep® Spin Miniprep Kit was used for plasmid extractions. *E. coli* cells from a single colony were transferred into 5 mL sterile liquid LB medium with the appropriate antibiotics and grown overnight at 37 °C. 5 mL overnight culture was centrifuged at 4000 rpm (Heraeus Biofuge Pico, Kendro Laboratory Products) for 10 minutes. The cell pellet then was re-suspended in 250 μ L buffer P1 and transferred into 1.5 mL Eppendorf tubes. Then the standard manufacturer's protocol included in the kit was applied to the mixture.

The flow-through of the last step contained clean plasmids. Their presence was monitored by DNA gel electrophoresis (see section 2.2.2). Concentrations of plasmids were measured using a NanoDrop (see section 2.3.7).

2.3.7 Nanodrop measurements

NanoDrop (NanoDrop 2000, Thermo Fisher Scientific) was used to measure DNA concentration at a wavelength of 260 nm. In addition, NanoDrop automatically measures A_{230} and A_{280} . The 260/280 absorbance ratio was used to assess the purity of the DNA. A ratio of 1.8 and higher was considered pure. Low 260/280 ratios can be caused by residual phenol or other reagents associated with the extraction protocol or by very low concentrations of nucleic acid. The 260/230 absorbance ratio was also used to assess contamination of DNA by other components. Values in the range of 2.0-2.2 were considered acceptable. Low 260/230 ratios could be the result of carbohydrate carryover, residual phenol, or guanidine and glycogen, used for precipitation.

2.3.8 DNA sequencing

DNA sequencing was performed by GATC (Eurofins-GATC, Germany). The sample for sequencing was prepared by mixing of 5 μL DNA template and 5 μL primer. DNA concentration was 80-100 $\text{ng } \mu\text{L}^{-1}$ for plasmids and 20-80 $\text{ng } \mu\text{L}^{-1}$ for PCR products. Primer DNA concentration was 5 μM . All the primers used for sequencing had to satisfy the following criteria:

- G or C at the 3' end, but not more than 3 Gs or Cs.
- The primer sequence was a good mix of all 4 nucleotides with no more than 4 identical bases in a row.

Sequencing results were monitored and aligned using the BioEdit software.

2.4 PROTEIN MONITORING, BUFFER EXCHANGE AND ANALYSIS

2.4.1 Sodium Dodecyl Sulfate - Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis was used to monitor the presence, size and purity of proteins (SynZur and SynBmtA). Further buffers were used for SDS-PAGE (all in MilliQ water):

- Buffer 1: 6.9 g L⁻¹ NaH₂PO₄, 17.53 g L⁻¹ NaCl, 1.36 g L⁻¹ imidazole, pH 8.0.
- 5x Laemmli Buffer: 36.3 g L⁻¹ Tris-HCl (pH 6.8), 605 g L⁻¹ glycerol, 100 g L⁻¹ SDS, 277.5 g L⁻¹ β-mercaptoethanol, 0.5 g L⁻¹ bromophenol blue.
- Running buffer: 3 g L⁻¹ Tris base, 14.4 g L⁻¹ glycine, 1 g L⁻¹ SDS

Frozen cell pellets were re-suspended in Buffer 1 (10 mL Buffer 1 per 1 g cell pellet). Mixtures, while being kept on ice, were sonicated twice for 15 seconds using an ultrasonic disintegrator (PG1284 12/76.MK2, MSE) to disrupt the cells. 100 μL lysate was centrifuged at 13000 rpm using a table top centrifuge (Centrifuge 5415 R, Eppendorf) at 4 °C. For the soluble fraction 40 μL clear supernatant was mixed with 10 μL 5x Laemmli buffer in a PCR tube and heated to 95 °C for 5 minutes. Then, 5 μL protein ladder (#SM1811, PageRuler™ Plus, Fermentas) and the samples were loaded into each well of a pre-cast polyacrylamide gel (14% Tris-glycine gel, Novex) using the Invitrogen Novex XCell SureLock™ system. SDS-PAGE was performed at 200 V for 45 minutes using a PowerPac 300 Power Supply, Bio-Rad. Subsequently, gels were removed from the plastic cassettes and washed 3 times with water for 5 minutes, followed by incubating for 1 hour with 20 mL SimplyBlue™ SafeStain (LC6060, Invitrogen) followed by 1 hour wash with 50 mL water. Following the last step, gels were transferred into plastic bags and scanned using a Ricoh Aficio MFD located in the School of Life Sciences, University of Warwick.

2.4.2 Protein concentration and desalting

Protein solutions were concentrated using Amicon filters (Amicon Ultra-15 centrifugal filters 10000 (for SynZur) and 3000 (for SynBmtA), Merck Millipore) to a volume of 2.5 mL. Amicon tubes were centrifuged at 4000 rpm at 4 °C until the desired volume was reached. Protein presence was also monitored in the flow-through fraction by SDS-PAGE (see section 2.4.1). When the protein solution volumes reached

2.5 mL, PD-10 desalting columns (Sephadex G-25, 2.5 mL, GE Healthcare) were used for buffer exchange. 2.5 mL concentrated protein solution was loaded onto the pre-equilibrated column with the desired buffer as described in the manufacturer's protocol. Then 3.5 mL of the desired buffer was loaded onto the column and the flow-through solution containing the protein in the desired buffer was collected. The column was then washed with buffer, water and 20% (v/v) ethanol, and kept at 4 °C for future use.

The desired buffer can in principle be any required buffer. For example, for cleavage of His-tagged Zur with TEV-protease, protein was exchanged into "Cleavage buffer", and for crystallization trials "Crystallization buffer" was used etc. The composition of each buffer is shown in the corresponding sections.

2.4.3 Inductively Coupled Plasma - Optical Emission Spectroscopy (ICP-OES)

ICP-OES was employed in order to assess the elemental composition of proteins and determine UV-vis extinction coefficients for subsequently faster determination of protein concentration. Protein solution was gravimetrically diluted with 0.01 M HNO₃. An ICP-OES spectrometer (PerkinElmer 5300DV ICP-OES) was used for the analysis. 1000 ppm standards for calibration of S, Zn, Cu, Co, Ni and Cd (Fisher Scientific) were diluted gravimetrically to make a 50 ppm stock solution in 0.01 M HNO₃ and then this stock was gravimetrically diluted to prepare standards of 0, 0.2, 0.4, 0.6, 1, 2 and 5 ppm in 0.01 M HNO₃. Wavelengths at which emission was measured are shown in Table 2.5.

Table 2.5 Determination of elements of interest by ICP-OES by characteristic emission lines.

Wavelengths (2 per element) at which emission was measured are shown.

| Element | Wavelength, nm |
|---------|------------------|
| Zn | 206.199, 213.858 |
| S | 180.672, 181.976 |
| Ni | 221.645, 231.605 |
| Cd | 214.445, 228.801 |
| Cu | 324.754, 327.393 |
| Mn | 257.606, 259.37 |

Parameters for ICP-OES measurements: argon flow rate 13.0 L min⁻¹, auxiliary gas flow rate 0.2 L min⁻¹, nebuliser flow rate 0.8 L min⁻¹ and radio frequency power at 1300 W. Washing time between the samples was 60 sec and each sample was measured in triplicate. ICP-OES data was analysed using Microsoft Excel.

2.4.4 Electrospray Ionization - Mass Spectrometry (ESI-MS)

ESI-MS was used to identify the target proteins and estimate protein to zinc stoichiometry. Sample concentrations were 5-10 µM in 10 or 20 mM ammonium bicarbonate solution. Samples were filtered through 0.22 µm pore size syringe filters (Minisart® NML Syringe Filters 1653K, Sartorius) prior to analysis. For acidic samples, the pH was adjusted to 2 using either acetic or formic acid solutions before filtering. A Bruker Compact ESI-MS spectrometer (Bruker) was used for the analysis (except for some spectra, where a MaXis Plus, Bruker was used). The spectrometer was operated in positive mode with the following parameters: flow rate 240 µL h⁻¹, range (500 - 4500 m/z), end plate offset 500 V, capillary 3900 V, nebulizer 0.8 Bar, drying gas 4 L min⁻¹ and drying gas temperature 190 °C. Spectra were recorded 1-2 minutes after stabilisation. For data analysis Bruker Daltonics DataAnalysis 4.3 software was used. The deconvolution method used was maximum entropy deconvolution within the range 14–65 kDa for Zur and 1-80 kDa for BmtA.

Chapter 3.

Properties and structure of the *Synechococcus* sp. WH8102 Zinc Uptake Regulator protein

3.1 INTRODUCTION

Several members of the putative zinc metabolic network were predicted by bioinformatics for the vast majority of examined cyanobacteria, namely the transcription factor Zinc Uptake Regulator (Zur) and the ZnuABC uptake system^{33,37}. However, it is difficult to distinguish Zur from its close homologues, especially other members of the Fur family, based on amino acid sequence alone. In order to decipher if a protein is a true Zur it has to satisfy certain criteria. These properties have been summarised in our recent review on bacterial Zinc Uptake Regulator proteins³⁵. Thus, a true Zur should have i) a conserved fold of two domains (as does any Fur-family protein (FFP)), ii) tetrahedral coordination of a sensory site which is usually composed of 2N, O and S, iii) an ability to bind DNA at the operator-promoter region of Zur-regulated genes within the nanomolar range and iv) have a Zur-Zn dissociation constant in the pico- to femtomolar range.

In this chapter, a predicted Zinc Uptake Regulator from the model marine cyanobacterium *Synechococcus* sp. WH8102 (SynZur) was over-expressed in *E. coli*, purified and functionally and structurally characterised. The properties of SynZur investigated include: elemental composition, oligomerisation state, DNA-binding properties, metal selectivity for DNA binding, and zinc affinity. In addition, together with Dr. Rachael Wilkinson and Prof. Vilmos Fülöp, it was possible to crystallize and solve the X-ray crystal structure. Thus, SynZur was structurally characterized and compared with all known structures of other bacterial Zurs (*Streptomyces coelicolor* 3mwm⁴⁷, *E. coli* 4mtd⁴⁸, *Mycobacterium tuberculosis* 2o03⁴⁹), and a selection of other members of the FFP: Furs (*Pseudomonas aeruginosa* 1mzb⁹³, *Campylobacter jejuni* 4ets⁹⁴), PerR (*Bacillus subtilis* 3f8n⁹⁵), Nur (*Streptomyces coelicolor* 3eyy⁹⁶) and Mur (*Rhizobium leguminosarum* 5fd6⁹⁷).

3.2 MATERIALS AND METHODS

3.2.1 Zur purification

3.2.1.1 Heterologous over-expression of SynZur in *E. coli*

SynZur over-expression was performed using a previously designed construct made by Dr. Amira Ksibe⁸⁸: the codon-optimised synthetic SynZur gene was cloned into the pET151/D-TOPO vector (see Appendix, Figure 8.1A) containing a His-tag and Tobacco Etch Virus (TEV) protease cleavage site at the N-terminus. The codon-optimised sequence and the vector map are shown in the Appendix, Figure 8.1B.

The *E. coli* cell line used for over-expression was BL21 (DE3) pLysS. BL21 (DE3) pLysS cells (originally from Novagen) were cultured on LB agar plates containing chloramphenicol ($35 \mu\text{g ml}^{-1}$) (see section 2.1). For SynZur over-expression, *E. coli* BL21 (DE3) pLysS cells were made competent and later transformed with pET151/D-TOPO-Zur as described in sections 2.3.4 and 2.3.5.

For over-expression of His-tagged SynZur only freshly transformed cells were used. After transformation several single colonies were picked simultaneously from the agar plates and transferred into 5 mL LB medium with the appropriate antibiotics and incubated shaking (180 rpm) overnight at 37 °C.

For over-expression of His-tagged TEV protease, which was used to cleave the His-tag from the His-tagged SynZur protein, overnight cultures were prepared from frozen TEV system *E. coli* cells⁹² by streaking them onto LB agar plates, incubating overnight at 37 °C followed by inoculating several single colonies into 5 mL LB containing chloramphenicol ($35 \mu\text{g ml}^{-1}$) and ampicillin ($100 \mu\text{g ml}^{-1}$) and left shaking (180 rpm) overnight at 37 °C.

Next day the overnight cultures of *E. coli* cells were added into 1 L LB medium with the appropriate antibiotics at a 1:100 ratio and incubated at 23 °C for 7-8 hours with shaking at 180 rpm. The optical density at 600 nm was monitored and when in the range of 0.6-0.8 isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to the cultures to induce protein over-expression. The final concentration of IPTG used was 0.5 mM.

The following day cultures were centrifuged at 5000 rpm (Eppendorf centrifuge 5810R) for 15 minutes and the cell pellet transferred into 50 mL Falcon tubes and frozen at -20 °C.

The presence of the TEV-protease and SynZur proteins was monitored by SDS-PAGE as described in section 2.4.1.

3.2.1.2 Fast Protein Liquid Chromatography (FPLC)

Two steps of FPLC with concentrating and desalting were used in order to purify both SynZur and the TEV protease. Frozen cell pellets (see section 3.2.1.1.) were re-suspended in Loading Buffer 1 (50 mM NaH₂PO₄, 300 mM NaCl, 20 mM imidazole, pH 8.0). For each gram of cell pellet 10 mL Loading Buffer 1 was used. Cell mixtures were then sonicated at 4 °C for 5 minutes with amplitude 60 and pulse 4 (Ultrasonic Processor SONICS) and centrifuged at 5000 rpm for 15 min (Eppendorf centrifuge 5810R).

Cell lysates at this stage were called crude lysates (only the soluble fraction was used). The crude lysates were then filtered using 0.22 µm pore size syringe filters (Minisart® NML Syringe Filters 1653K, Sartorius) and loaded onto a 5 mL Ni-Sepharose His-Trap column (GE Healthcare), prewashed with Loading Buffer 1 using an AKTA purifier with UPC-900 monitor, P-900 sample pump, and fraction collector Frac-950 (Amersham Biosciences). The flow rate was 1 mL min⁻¹. The presence of protein was monitored at 280 nm.

After the high absorbance level caused by the crude lysate washing through began to decrease, the column was washed with 50 mL of Loading Buffer 1 until a stable background level was achieved. After that, the concentration of elution Buffer 2 (50 mM NaH₂PO₄, 300 mM NaCl, 250 mM imidazole, pH 8) was gradually increased over 10 column volumes to 100% in order to elute the SynZur protein from the column. Then the eluted fractions, mainly containing Buffer 2, were checked for the presence of the SynZur and TEV protease proteins by SDS-PAGE as described in section 2.4.1.

Fractions containing SynZur were concentrated using Amicon Filter and the buffer was exchanged into the buffer required with PD-10 column (see section 2.4.2). Then, TEV cleavage was performed in order to remove the His-tag (see section 3.2.1.3.). After that step a second FPLC purification step was used to separate Zur from the TEV-protease and the removed His-Tag. During the second FPLC, filtered protein mixture was mixed with Buffer 1 and loaded onto the same 5ml Ni-Sepharose His-Trap column. Then the column was washed with Buffer 1 until the background

level became stable. The cleaved protein was in the flow through fractions which were then monitored by SDS-PAGE (see section 2.4.1). The column was then washed with Buffer 2 until the background was stable to remove bound proteins, washed with water and kept under 20% (v/v) ethanol.

3.2.1.3 TEV cleavage of SynZur

After the desalting step both the His-tagged SynZur and TEV-protease were resuspended in cleavage buffer (50 mM Tris-HCl (pH 8.0), 1 mM DTT). TEV-protease was added to SynZur at a 1:100 concentration as determined by OD₂₈₀ as suggested in⁹² and left overnight on a shaker at 4 °C. Cleavage rate was monitored by SDS-PAGE. Remaining TEV-protease was kept at -80 °C for future use.

3.2.2 Characterisation of SynZur

3.2.2.1 Electrophoretic mobility shift assays

For Electrophoretic Mobility Shift Assays (EMSAs) an upstream region of genes SYNW0971, SYNW0970 and SYNW0969 (which are thought to have the same promoter) (*pznuABC*) was amplified by PCR as described in section 2.2.1 (see Table 3.1 for the primers used), purified using a Promega PCR purification kit (see section 2.2.3) and its concentration measured using a NanoDrop as described in section 2.3.7.

Table 3.1 Primers used to amplify the 256 bp operator-promoter region of *Synechococcus* sp. WH8102 *znuABC*, taken from⁸⁸.

| Name in this study | Sequence |
|--------------------|-----------------------|
| ZnuABC_F | GAGATTCCTGGCTGCTAGATG |
| ZnuABC_Re | GTTCCATCCACAGCCACAAC |

The purified *pznuABC* PCR product was diluted in EDTA-free EMSA binding buffer (10 mM Tris-HCl, pH 8, 50 mM KCl, 2 mM MgCl₂•6H₂O, 5% (v/v) glycerol, 0.05 mg mL⁻¹ BSA, 1 mM dithiothreitol (DTT), 3 mM spermidine) to the required concentration. The specificity of SynZur binding to *pznuABC* was shown previously using *pntcA* as a control⁸⁸.

SynZur protein (after purification by FPLC as described in section 3.2.1.2.) was diluted to the required concentration in EMSA binding buffer. Aliquots containing 5 ng DNA were mixed with different amounts of Zur (0-100 μ M) and diluted with EMSA binding buffer so that the total reaction volume was 10 μ L and left at room temperature for 30 minutes. After that, reaction mixtures were loaded onto 10% polyacrylamide gels (0.1M Tris, pH 8.3) and then PAGE was performed in EDTA-free Tris-glycine running buffer (0.025 M Tris, 0.187 M glycine) at 100V for ~90 minutes at 4 °C.

Gels for EMSA were cast in advance using a Mini-PROTEAN® Tetra Cell Casting Module (Bio-Rad) by mixing all the components in the order they are listed in Table 3.2 and pouring the mixture inside the Mini-PROTEAN® glass cassettes (Bio-Rad). Gels were left to solidify for at least 40 minutes and pre-run for at least 10 minutes under running conditions before loading the samples.

Table 3.2 PAGE gel composition. Values are given for 10 mL or 2 gels.

| Acrylamide percentage | 6% | 8% | 10% | 12% | 15% |
|--|-------------|-------------|-------------|-------------|-------------|
| Acrylamide/Bis-acrylamide (30%/0.8% w/v) | 2 mL | 2.6 mL | 3.4 mL | 4 mL | 5 mL |
| 1 M Tris (pH 8.3) | 1 mL | 1 mL | 1 mL | 1 mL | 1 mL |
| Water | 6.89 mL | 6.29 mL | 5.49 mL | 4.89 mL | 3.89 mL |
| 10% (w/v) ammonium persulfate | 100 μ L | 100 μ L | 100 μ L | 100 μ L | 100 μ L |
| TEMED | 10 μ L | 10 μ L | 10 μ L | 10 μ L | 10 μ L |

After PAGE, DNA was visualized with SYBR-Green I. SYBR-Green I stock in DMSO (S9430, Sigma-Aldrich, 10 mg mL⁻¹) was diluted 1:10000 in Tris-glycine running buffer (0.025 M Tris, 0.187 M glycine) and gels were stained for 30 minutes in plastic containers wrapped in foil to protect them from light. Visualisation was done using a luminescent image analyser ImageQuant LAS 4000 (GE Healthcare Bio-Sciences AB).

For Ferguson plots, protein standards (P77125 New England Biolabs Color Prestained Protein Standard, Broad Range) were run together with DNA-Zur complexes using 8% and 12% gels (see Table 3.2) under the standard EMSA conditions described above. In addition, a 5 μ L of MyTaq red buffer containing an

inert dye of low molecular weight was loaded into a separate well as a low mass control.

Gels were scanned to visualize the protein ladder and low MW size control prior to staining with SYBR-Green I for DNA visualisation as described above. Gel images were combined using Gimp software (GNU Image Manipulation Program, 2.8.10). Measurement of the mobility of DNA bands in pixels was performed using “Measure tool” in Gimp.

3.2.2.2 Metal-responsive DNA binding

All the work described here was conducted in a glove-box under an atmosphere of nitrogen containing <1% oxygen. All solutions and buffers were purged with argon prior to loading them into the glove-box.

For the metal specificity experiment SynZur (20-40 μM) containing 2 zinc ions per monomer was de-metallated by mixing with 1 mM EDTA pH 8, 1 mM DTT overnight at room temperature. The next day a PD-10 column was used to separate the de-metallated SynZur (SynZurZn₁) from EDTA and zinc as described in section 2.4.2. 10 mM Tris buffer (pH 8) was used for elution. To check that SynZur was de-metallated, similar experiments were conducted, except that the elution buffer was 20 mM NH₄HCO₃ and de-metallation was checked by ESI-MS (see section 2.4.4).

Eluted SynZurZn₁ was aliquoted into several separate tubes containing different metals so that the protein concentration was 320 nM. The concentration of all the metals was 3.2 μM dissolved in 10 mM Tris buffer (pH 8). Solutions of Zn²⁺, Cd²⁺, Cu²⁺, Co²⁺, Fe²⁺, Mn²⁺ and Ni²⁺ were prepared in advance.

After de-metallated SynZur was premixed with metals, 5 μL of each SynZur+metal containing solution were mixed with 5 ng *pznuABC* in 10 mM Tris buffer (pH 7.5) with glycerol (5% (v/v)) so that the resulting SynZur concentration was 160 nM. Reaction mixtures were left for 30 minutes at room temperature and directly loaded onto 10% native acrylamide gels containing 10 mM Tris buffer and run and visualized under the same conditions as described in section 3.2.2.1.

To assess binding of Zur with different metals by ESI-MS, Zur was de-metallated with EDTA overnight. After double buffer exchange into 20 mM NH₄HCO₃ using a PD-10 column, the resulting SynZurZn₁ was mixed with 2 times

molar excess of metal salts (Zn^{2+} , Cd^{2+} , Cu^{2+} , Co^{2+} , Fe^{2+} , Mn^{2+} , Ni^{2+}). The mixtures were then filtered through 0.22 μm pore size syringe filters (Minisart® NML Syringe Filters 1653K, Sartorius) and subjected to native ESI-MS (see section 2.4.4).

3.2.2.3 Non-denaturing Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis (ND SDS-PAGE)

The protocol for ND SDS-PAGE used was as described in⁹⁸. Protein sample was mixed with 4x NSDS sample buffer (100 mM Tris HCl, 150 mM Tris base, 10% (v/v) glycerol, 0.0185% (w/v) Coomassie G-250, 0.00625% (w/v) Phenol Red, pH 8.5). Native gels were cast using the same recipe as for EMSAs (see section 3.2.2.1.). Running buffer contained 50 mM Tris (pH 7.3), 50 mM MOPS, 0.0375% (w/v) SDS. ND SDS-PAGE was performed at 4 °C at 100 V until the dye front reached the bottom of the gel (usually about 90 minutes). Gels were visualised using (SimplyBlue™ SafeStain, Life Technologies) and scanned.

In addition, SynZur samples at a concentration of 5 and 10 μM were run with Tris-glycine buffer (0.025 M Tris, 0.187 M glycine).

3.2.2.4 Analytical Size Exclusion Chromatography (SEC) of SynZur

Analytical SEC of SynZur was used for an assessment of oligomerisation state. SEC was run for proteins with known molecular weight using an AKTA purifier with the same settings as described in section 3.2.1.2 at 1 mL min^{-1} . The column was Superdex™ 75, 10/300 (GE Healthcare) and the buffer used was crystallization buffer (50 mM Tris (pH 8.0), 100 mM NaCl).

200 μL standards with a concentration of 1 mg mL^{-1} were separately loaded using a 500 μL loop and run. Standards were: Blue Dextran, 2000 kDa; Bovine Serum Albumin, 66,5 kDa, Carbonic Anhydrase, 29.2 kDa, Cytochrome c, 12.2 kDa and Insulin chain B oxidized, 3.5 kDa. Then, 200 μL SynZur sample (at a 1 mg mL^{-1} concentration) was also loaded and run.

Values of elution volumes of the standards divided by the elution volume of Blue Dextran were plotted against molecular weights of the standards using Microsoft Excel.

3.2.2.5 Zinc binding affinity of SynZur

SynZur at a concentration of 32 μM in 20 mM ammonium bicarbonate, pH 7.9 was mixed with 1 mM EDTA, 1 mM DTT and left overnight at 4 °C. After zinc was removed from the sensory site, a PD-10 column was used to purify SynZurZn₁ inside the anaerobic polyethylene glove bag (Alfa Aesar). 20 mM ammonium bicarbonate, pH 7.9 was used twice to elute the protein and an aliquot was used for ESI-MS to assess SynZur de-metallation.

Upon successful de-metallation, ~ 10 μM SynZurZn₁ with 0.1 mM TCEP in 20 mM ammonium bicarbonate, pH 7.9 was mixed with ~ 15 μM Quin-2 and titrated with 710 μM ZnSO₄ in triplicate. Quin-2 (2-[(2-Amino-5-methylphenoxy)methyl]-6-methoxy-8-aminoquinoline-N,N,N',N'-tetraacetic acid tetrapotassium salt) was measured spectrophotometrically at 261 nm as well as quantified at this wavelength using an extinction coefficient of 37500 $\text{cm}^{-1} \text{M}^{-1}$ ⁹⁹. Protein concentration was estimated by absorbance at 280 nm (A_{280}) and using an extinction coefficient of 3485 $\text{cm}^{-1} \text{M}^{-1}$ (as determined in section 3.3.2.1.). Zinc concentration was estimated using ICP-OES (see section 2.4.3). A UV-vis spectrum was measured after each addition of ZnSO₄ several times until the spectrum stopped changing (up to 15 minutes per addition of Zn²⁺ aliquot).

The K_D was calculated using DynaFit¹⁰⁰ software based on Quin-2 $K_{DZn} = 3.7 \times 10^{-12}$ M.

3.2.3 Structural characterisation of SynZur

Crystallization experiments were performed with the help of Dr. Rachael Wilkinson and Prof. Vilmos Fülöp, School of Life Sciences, University of Warwick.

After the second IMAC purification step (see section 3.2.1.2.) purified SynZur was concentrated and buffer was exchanged to crystallization buffer (50 mM Tris–HCl (pH 8.0), 100 mM NaCl) on a PD-10 column (see section 2.4.2). In addition, before setting up crystallization trials, SynZur was purified by SEC using a Sephacryl S-200 column (HiPrep 26/60, GE Healthcare) pre-equilibrated in 50 mM Tris-HCl pH 8, 150 mM NaCl at 0.5 mL min^{-1} with collection of 5 mL fractions. Fractions containing SynZur were pooled and concentrated to 10 mg mL^{-1} .

Initial trials were set up with the commercial crystallization screening buffers (Molecular Dimensions): Structure Screen 1 & 2 HT-96, ProPlex HT-96, MIDAS™ HT-96, and JCSG-plus™ HT-96.

Crystallization trials were performed using the hanging drop method in MRC 96-well plates (see Figure 3.1). Firstly, 50 μ L of the appropriate crystallization screening solution was loaded into larger chambers of each well (marked with I in Figure 3.1B). A Mosquito LCP crystallization robot (Warwick Macromolecular Crystallisation Facility, School of Life Sciences) was used to dispense 200 nL of the screening buffers into chambers II and III (see Figure 3.1B). Then, the robot added 200 nL SynZur solution into chambers II and 200 nL of the crystallization buffer into chambers III, the latter to control crystallization of the buffer components without SynZur.

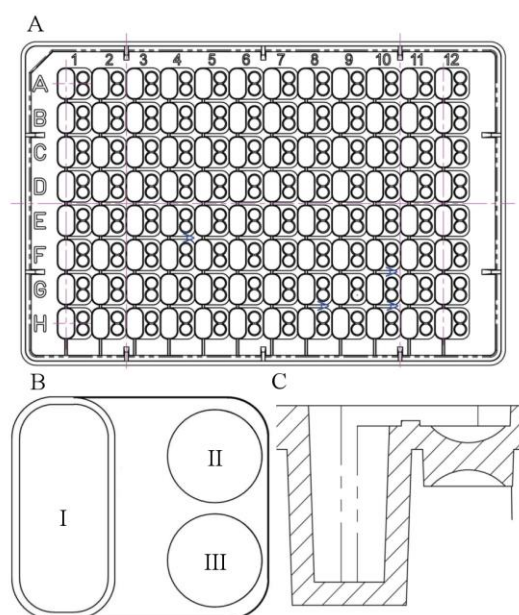


Figure 3.1 The MRC 96-well plate used for SynZur crystallization.

A. Overall top view. B: A top view of one of the 96 wells. Each well contains three chambers: I, a larger chamber, and two smaller chambers, II and III. Chamber I is for the screening buffer; chamber II is for protein crystallization and chamber III is to control for crystallization of the buffer components. C: Front view projection of each well.

After setting up all the components for crystallization, the plates were sealed with MicroAmp adhesive film (Applied Biosystems) and left for crystals to grow at 18 °C. During the first week crystals were monitored daily. During weeks 2 – 4 plates were monitored weekly and after that they were monitored once a month.

After crystals were formed they were picked using loops into cryoprotectant solutions and snap frozen in liquid nitrogen. Each cryoprotectant had the same composition as the screening buffer in the appropriate well together with 20% (v/v) glycerol. X-ray diffraction data were collected at the X-ray Diamond Light Source, Macromolecular Crystallography Facility, Harwell Science and Innovation Campus, Oxfordshire.

Initial hits were observed on F3 of the Proplex screen (Molecular Dimensions). The composition of this solution is 100 mM magnesium acetate, 100 mM MES, 10% (w/v) PEG 10000, pH 6.5. However, further optimisation was required for well-diffracting crystals. Crystals were grown using the hanging drop method in 24 well plates (ThermoFisher Scientific) covered by round CrystalClene cover slips 22 mm (Molecular Dimensions). 1 μ L of SynZur protein was mixed with 1 μ L of crystallisation solution and placed on a cover slip. Then, 1 μ L 50 mM Tris pH 8, 150 mM NaCl was mixed together with crystallisation solution separately on the same cover slip. Subsequently, 0.5 mL crystallisation solution was added into each well and cover slips were sealed with white grease so that the droplets on the cover slip were upside down and water evaporation was prevented by the grease.

Crystallisation optimisation was performed by Dr. Rachael Wilkinson. This included slightly varying concentration of each component separately and checking the best conditions. Final crystals were produced using the following conditions: 100 mM magnesium acetate, 100 mM MES pH 6, 16% (w/v) PEG 10000 at 4 °C. Crystals were harvested using a 0.08 mm mounted Litholoop (Molecular Dimensions) and flash-frozen in liquid nitrogen in crystallisation solution containing 20% (w/v) ethylene glycol.

X-ray diffraction data collection and structural assignment were conducted by Dr. Rachael Wilkinson and Prof. Vilmos Fülöp. Data were collected at the zinc absorption edge (9666 eV) using beamline I03 and a Pilatus 6M detector at the Diamond Light Source, Didcot, Oxfordshire, UK. All data were indexed, integrated and scaled using the XDS package¹⁰¹. Further data handling was carried out using the CCP4 software package¹⁰². The structure was solved by single wavelength anomalous diffraction using SHELX¹⁰³. Refinement of the structure was carried out by alternate cycles of manual refitting using Coot¹⁰⁴ and Refmac¹⁰⁵. Water molecules were added to the atomic model automatically using ARP¹⁰⁶, at the positions of large positive

peaks in the difference electron density, only at places where the resulting water molecule fell into an appropriate hydrogen bonding environment. Restrained isotropic temperature factor refinements were carried out for each individual atom. The polypeptide chain was traced continuously through electron density maps ($2F_o - F_c$ and $F_o - F_c$) from residue 13 to 108 and from residue 115 to 132.

3.3 RESULTS AND DISCUSSION

3.3.1 Over-expression and purification of the SynZur protein

In order to investigate the properties of SynZur and subsequently characterise it, the protein was obtained by heterologous recombinant expression in *E. coli* and purified as described in section 3.2.1. SynZur over-expression was initially performed using a previously designed construct (pET151/D-TOPO-Zur) and previously optimised over-expression and purification conditions⁸⁸.

However, the above protocol was lengthy and required the use of large concentrations of TEV protease. As a result a new purification approach was designed using IPTG induction at room temperature (~23 °C) overnight, instead of 6 hours at 30 °C. Furthermore, an additional buffer exchange step was incorporated using a PD-10 column in the purification process in order to remove imidazole contained within the loading and elution buffers. This step could potentially greatly decrease the TEV-protease requirement⁹² and would allow to over-express and purify TEV-protease only once which would be enough for the whole project if it was kept frozen at -80 °C.

Plasmid pET151/D-TOPO-Zur was sequenced with T7 primers and matched the theoretical prediction by 100% (data not shown). The experimental design of SynZur purification is shown in Figure 3.2.

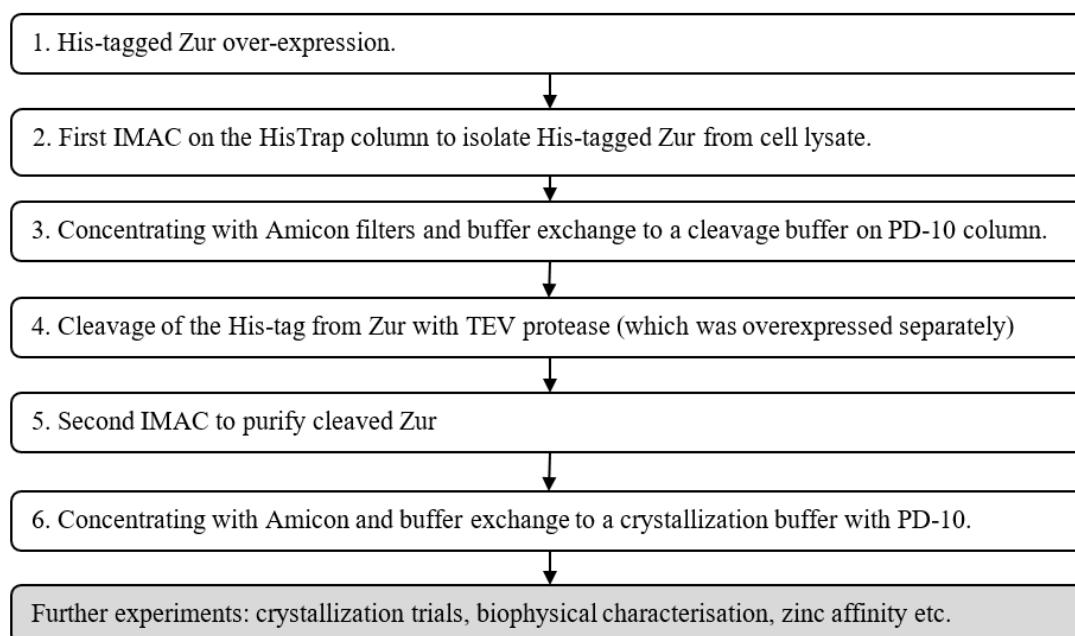


Figure 3.2 Scheme of the experimental design of SynZur purification.

All steps were assessed by SDS-PAGE analysis.

3.3.1.1 Heterologous recombinant expression of *Synechococcus* sp. WH8102 Zur

SynZur protein over-expression levels were assessed under various conditions: 23 °C, 30 °C and 36 °C; 6h induction with IPTG and overnight induction. It was found that there was almost no visible effect of the different over-expression conditions on SynZur expression levels when normalised to cell biomass (see Figure 3.3). Thus, the most convenient conditions were chosen: inoculation of the overnight culture at a ratio of 1:100 into LB containing ampicillin and chloramphenicol, incubation for 7-8 hours at 23 °C until the OD₆₀₀ reached 0.6-0.8 and induction with 0.5 mM IPTG overnight as described in section 3.2.1.1. Cells were then harvested and crude lysates prepared as also described in section 3.2.1.1. Crude lysates were used for further purification.

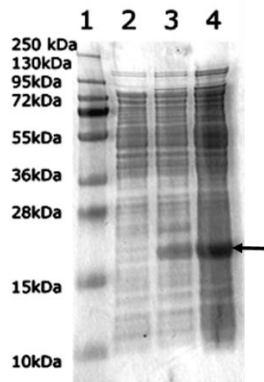


Figure 3.3 SDS-PAGE analysis of over-expression of His-tagged SynZur protein in *E. coli* BL21 (DE3) pLysS.

Lane 1 – protein ladder. Lanes 2-4 are crude cell lysates as follows: Lane 2 un-induced cells; Lane 3 Cell culture induced with IPTG under the conditions described in section 3.2.1.1; Lane 4 Cell culture induced with IPTG at 30 °C for 6 hours. The protein band marked with a black arrow approximately corresponds to the predicted length of the SynZur protein containing a His-tag, V5 epitope and TEV cleavage site (18.6 kDa). Lanes 3 and 4 have different intensities due to the different total protein concentrations loaded.

3.3.1.2 Protein purification using FPLC

As a first step of purification, crude lysates were purified via FPLC using a Ni-Sepharose IMAC (Immobilized Metal ion Affinity Chromatography) column (see section 3.2.1.2.). The result is shown in Figure 3.4.

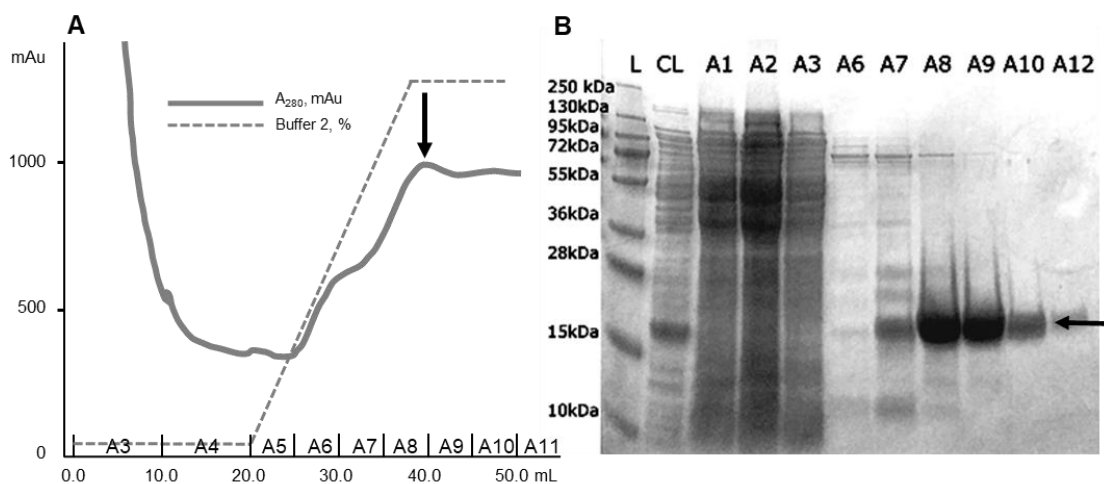


Figure 3.4 FPLC purification of SynZur.

A: The elution profile of the SynZur protein during the first IMAC step. The absorbance is shown with a grey line and the concentration of elution Buffer 2 is shown with a dashed line (100% at the top). The fractions collected are shown at the bottom. **B:** SDS-PAGE analysis of fractions obtained during FPLC. L: protein ladder; CL: crude lysate. The SynZur protein with a theoretical size of 18.6 kDa is indicated with a black arrow.

The peak for SynZur looked very low on the chromatogram due to the low extinction coefficient which was calculated in section 3.3.2.1.

After this step purified His-tagged SynZur was concentrated using an Amicon filter and the buffer was exchanged to a TEV cleavage buffer (see section 2.4.2). Cleaving off the His-tag for metal sensor proteins is crucial since it would contribute towards their metal binding ability. The His-tagged TEV protease used for cleavage was over-expressed and purified using conditions similar to those used for SynZur as shown in the Appendix (see Figure 8.2). Concentration and buffer exchange steps for SynZur and the TEV protease, and the cleavage rate of His-Zur, were analysed by SDS-PAGE (see Figure 3.5).

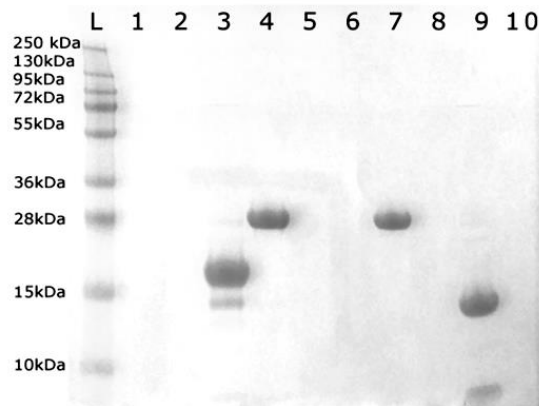


Figure 3.5 SDS-PAGE analysis of various purification steps during the concentrating and buffer exchange of the SynZur and TEV-protease (TEVp).

L: protein ladder. Lanes 1 and 2: Zur and TEVp flow through during Amicon 10000 filtration. Lanes 3 and 4: Concentrated SynZur and TEVp in imidazole buffer. Lanes 5 and 6: SynZur and TEVp flow through from a PD-10 column (first 2.5 mL). Lanes 7 and 8: TEVp elution from a PD-10 column after buffer exchange to buffer-1. Lane 7: first 3.5 mL; Lane 8: elution after 4 mL more (TEVp is 30 kDa). Lanes 9 and 10: SynZur elution from a PD-10 column after cleavage with TEVp overnight at room temperature. Lane 9: first 3.5 mL. The band at the bottom is the cleaved His-tagged tail. Cleaved SynZur is 15.5 kDa. Lane 10: elution after 4 mL more.

After overnight cleavage with TEV-protease, the cleavage buffer in which SynZur, the TEV-protease and the cleaved His-tag was contained was exchanged with Loading Buffer 1 using a PD-10 column (see Figure 3.5, Lanes 9-10) (in later purifications this step was omitted). Then, a second IMAC step was used to separate the cleaved SynZur protein from the cleaved His-tag and TEV-protease (see Figure 3.6).

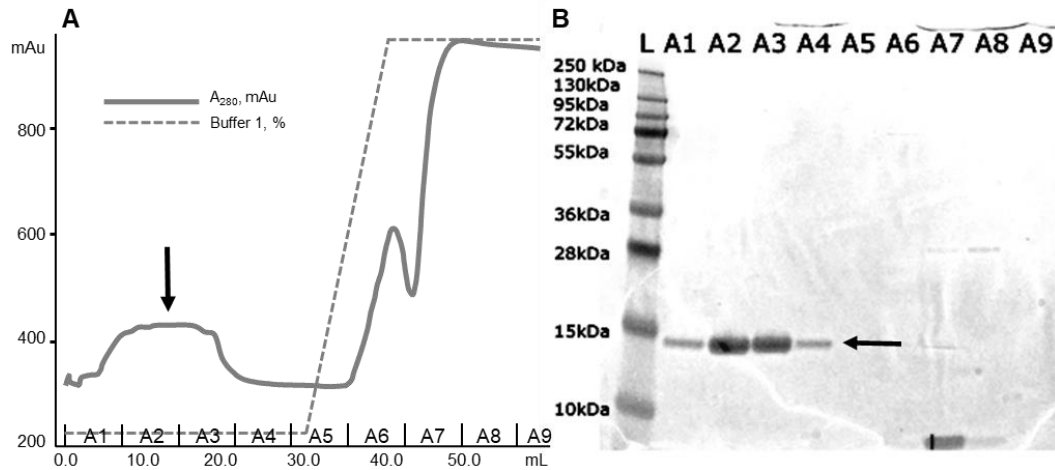


Figure 3.6 SDS-PAGE analysis of the second IMAC purification step of the SynZur protein after cleavage of the His-tag with TEV protease.

A: The elution profile of the SynZur protein during the second IMAC step. Absorbance is shown with a grey line and the concentration of elution Buffer 2 is shown with a dashed line (100% at the top, see section 3.2.1.2). The fractions collected are shown at the bottom. **B:** SDS-PAGE analysis of the fractions obtained during the second IMAC step. The purified SynZur protein with a theoretical size of 15.5 kDa is indicated by a black arrow. The cleaved His-tag and TEV-protease are also visible in fractions A7 and A8.

Since the His-tag had been removed by cleavage, the Zur protein was unable to bind to the Ni-sepharose column in the presence of imidazole and it was subsequently collected in the flow-through fractions with high purity. The theoretical size of cleaved Zur is 15466.43 Da (calculated by ExPASy Compute pI/Mw tool webserver¹⁰⁷) and in addition to the expected amino acids it has 6 more residues due to the expression construct used. The theoretical sequence of the resulting Zur protein is shown in Figure 3.7.

```

-6                                     GIDPFT      0
  1 MTGSSPALNARQQALLTALNACGDEMSGQQLHRSLLDDEASMGLATVYRNLRLQLQQRGLVR  60
 61 CRHLPTGEALYAPVDRDRHHLTCVDCGTTQVLHDHCPHIGIDVPADSRGDFELLFHTLEFF 120
121 GFCSSCRPQRSSKP                                                         134

```

Figure 3.7 The theoretical sequence of the SynZur protein following purification.

The six additional amino acid residues are highlighted with bold letters. Amino acid numbering corresponds to the original SynZur sequence.

After the second FPLC purification step the SynZur protein was considered pure enough for further experiments. FPLC flow-through fractions containing SynZur were pooled, concentrated using an Amicon filter and buffer exchanged to a

crystallization buffer with a PD-10 desalting column as described in section 2.4.2. The yield of purified SynZur protein was about 7–8 mg per litre of *E. coli* culture.

A deconvoluted ESI-MS spectrum of the SynZur protein acidified to pH 2 using formic acid showed a single predominant (>95%) peak with MW 15466.52 Da. This is in excellent agreement with the predicted mass of apo-SynZur based on the sequence (see Figure 3.7, 15466.43 Da calculated by ExPASy¹⁰⁷). The ESI-MS experiments are discussed in detail in section 3.3.2.2. Thus, pure SynZur protein was over-expressed and purified and ready for characterisation.

3.3.2 Characterisation of SynZur

The SynZur protein was characterized using a range of different methods including ICP-OES, ESI-MS, SDS-PAGE, ND SDS-PAGE and analytical SEC. In addition, SynZur de-metallation with EDTA was investigated, its zinc affinity was determined and EMSAs were used for studying the interactions of SynZur with DNA.

3.3.2.1 Elemental composition and molar absorption coefficient calculation

Inductively Coupled Plasma - Optical Emission Spectroscopy (ICP-OES) is one of the most precise methods for simultaneous determination of metal and sulfur content of a protein sample which then allows calculation of the metal/protein ratio⁷⁸. Knowledge of the sulfur content allows calculation of the protein concentration as well as its extinction coefficient (ϵ_{280} ; correctly referred to as molar absorption coefficient): if the absorbance at 280 nm (A_{280}) of a protein sample is measured prior to ICP-OES, knowing the protein concentration would allow calculation of its molar absorption coefficient. This in turn would enable fast determination of the protein concentration by spectrophotometry which is faster and less resource consuming than ICP-OES.

Previously, it was found that traditional methods for calculating the SynZur protein concentration, such as the Bradford or cysteine assays, did not work well giving significantly lower concentrations, unlike ICP-OES⁸⁸. However, the absorption coefficient determined by this method (ϵ_{280} of SynZur) was $4354 \text{ M}^{-1} \text{ cm}^{-1}$, which was significantly higher than the predicted value ($3400 \text{ M}^{-1} \text{ cm}^{-1}$)⁸⁸. The reason for this was unclear so it was decided to repeat this experiment.

SynZur samples and standards of different metals for ICP-OES were prepared as described in section 2.4.3. A_{280} of the SynZur stock was also measured prior to ICP-OES sample preparation. However, upon dissolving the SynZur protein in 2% (v/v) HNO_3 (a concentration initially recommended during ICP-OES training), the protein formed a precipitate. This could be a possible reason for the higher ϵ_{280} value previously found: precipitation would lead to a reduced protein concentration in the sample solution which for a sample with the same absorbance would give a higher absorption coefficient. Thus, it was decided to use a lower concentration of nitric acid (0.01 M or 0.063%) to avoid precipitation. With this change the sample remained transparent.

The concentration of SynZur in the samples was calculated based on sulfur content which was found to be 2.44 ± 0.01 ppm based on ICP-OES. Since each SynZur protein molecule contains 10 sulfur atoms as shown below (see Figure 3.8), the protein concentration was calculated as $7.61 \pm 0.04 \times 10^{-6}$ M.

GIDPFT**M**TGSSPALNARQQALLTALNA**C**GDE**M**SGQQLHRSLDDEAS**M**GLATVYRNL
 RQLQQRGLVLR**C**RHLPTGEALYAPVDRDRHHLT**C**VD**C**GTTQVLDH**C**PIHGIDVPADS
 RGDFELLFHTLEFFGF**C**SS**C**RPQRSSKP

Figure 3.8 Amino acid sequence of the SynZur protein with sulfur containing amino acid residues highlighted with bold letters.

Cysteine residues: 7; Methionine residues: 3.

In the same way, concentrations of trace metals were calculated and based on sulfur content, ratios of metals to SynZur were calculated. Results of metal content are shown in Table 3.3.

Table 3.3 ICP-OES analysis of SynZur: the content of various trace metals in the analysed SynZur samples.

| Metal, M | M to protein ratio | St. error |
|-----------------|---------------------------|------------------|
| Zn | 2.0529 | 0.0049 |
| Ni | 0.0165 | 0.0054 |
| Cd | 0.0036 | 0.0028 |
| Cu | 0.0083 | 0.0012 |
| Mn | 0.0098 | 0.0021 |

It can be seen that there are ~2 zinc ions per molecule of SynZur present, as was predicted³³ and confirmed in⁸⁸. It should be pointed out that the LB medium used contained no additional zinc but nonetheless *E. coli* produced SynZur with zinc bound, which may hint at a high zinc selectivity of the *Synechococcus* sp. WH8102 Zur protein. In addition, a slightly elevated nickel content could be explained by using Ni affinity chromatography and leakage of Ni into the sample.

Since samples of known A_{280} were analysed and the protein concentration determined based on sulfur content, the molar absorption coefficient could be calculated using the following formula:

$$\epsilon_{280} = A_{280}/(C \times l) = 3485 \pm 14 \text{ M}^{-1} \text{ cm}^{-1} \quad \text{Equation 3.1.}$$

Where: ϵ_{280} – molar absorption coefficient at 280 nm, $\text{M}^{-1} \text{ cm}^{-1}$;

A_{280} – absorbance at 280 nm;

C – protein concentration, M;

l – path length, cm. It is usually 1 cm.

This value is in good agreement with a range of theoretically predicted extinction coefficient (280 nm) values based on the formula¹⁰⁸.

$$\epsilon_{280}(\text{M}^{-1} \text{ cm}^{-1}) = \# \text{Trp} \times 5,500 + \# \text{Tyr} \times 1,490 + \# \text{cys} \times 125 = 2980 - 3855 \text{ M}^{-1} \text{ cm}^{-1} \quad \text{Equation 3.2.}$$

Where #Trp is the number of tryptophan residues in the protein, #Tyr is the number of tyrosine residues and #cys is the number of cysteine residues forming disulfide bonds (see Figure 3.9 for numbers).

The variation here is due to the unknown number of disulfide bonds in SynZur. Despite SynZur containing 7 cysteine residues (see Figure 3.9), some of them are likely to participate in binding zinc and will not form disulfide bonds. Based on a homology model produced using other Zurs and FFP, it was suggested that Zur from *Synechococcus* sp. WH8102 has two zinc binding sites: 4 cysteine residues form a structural zinc site and 1 Cys together with 2 His and 1 Asp form a sensory zinc site³³.

This would narrow the predicted extinction coefficient to within the range of 2980-3230 M⁻¹ cm⁻¹.

The extinction coefficient value obtained was 3485 M⁻¹ cm⁻¹, which is higher than the highest predicted value by 7.9%. The difference could be due to the presence of some impurities with high extinction coefficients but undetectable by SDS-PAGE (thus with negligible concentration). In addition, Phe and His could also slightly contribute to A₂₈₀. Over-estimating the molar absorption coefficient could lead to underestimation of the protein concentration in solution and might affect the values of K_D calculated below. However, if the error appeared due to impurities in the SynZur sample prepared by this purification protocol, then they would possibly be present every time SynZur was purified this way. Thus, it was decided to use this experimentally calculated absorption coefficient.

GIDPFTMTGSSPALNARQQALLTALNACGDEMSGQQLHRSLLDDEASMGLAT
VYRNLRQLQQRGLVRCRHLPTGEALYAPVDRDRHHLTCVDCGTTQVLDHC
PIHGIDVPADSRGDFELLFHTLEFFGFCSSCRPQRSSKP

Figure 3.9 Sequence of the SynZur protein highlighting residues involved in determination of the absorption coefficient and zinc binding sites³³.

7 cysteine residues are highlighted in yellow, 2 tyrosine residues are highlighted in red. SynZur contains no Tryptophan residues. Residues predicted to form the sensory zinc binding site are shown in bold blue letters. Cysteine residues predicted to form the structural zinc binding site are shown in bold green letters.

3.3.2.2 Electrospray Ionization–Mass Spectrometry

The next step was to analyse the molecular mass of the SynZur protein using ESI-MS. This method is based on ionization of molecules in the sample when they acquire positive or negative charges in the gas phase. Ions are then detected by a mass analyser according to their mass/charge ratio¹⁰⁹. Usually several ions of different charges are formed and it is possible to find the molecular mass by deconvolution or matching integer values of charges to peaks so that the resulting molecular mass is the same for all of them. Native ESI-MS allows precise assessment of molecular mass, purity of the sample and sometimes oligomeric state and co-factors including metal ions^{109,110}.

In order to further characterise SynZur and corroborate the ICP-OES results, ESI-MS analysis was performed. Proteins in acidic conditions (pH 2) usually appear in their apo-form because amino acid residues are protonated and do not bind metals leading to the proteins being unfolded. Thus, formic acid was added to a SynZur sample to give pH 2 prior to ESI-MS analysis to determine its apo-molecular mass. The full MS spectrum for apo-Zur is shown in Figure 3.10. After deconvolution, the neutral protein mass was 15466.52 Da which is in excellent agreement with the predicted mass based on the sequence. The predicted mass calculated using the ExPASy¹⁰⁷ web service based on the actual SynZur sequence (see Figure 3.9) was found to be 15466.43 Da. Comparison of theoretical and experimentally determined neutral masses of SynZur is presented in Table 3.4.

Table 3.4 Comparison of the calculated molecular mass of apo-SynZur with that determined by ESI-MS.

| | Found | Theoretical | Difference |
|----------------|----------|-------------|------------|
| Apo-Zur (pH 2) | 15466.52 | 15466.43 | -0.09 |

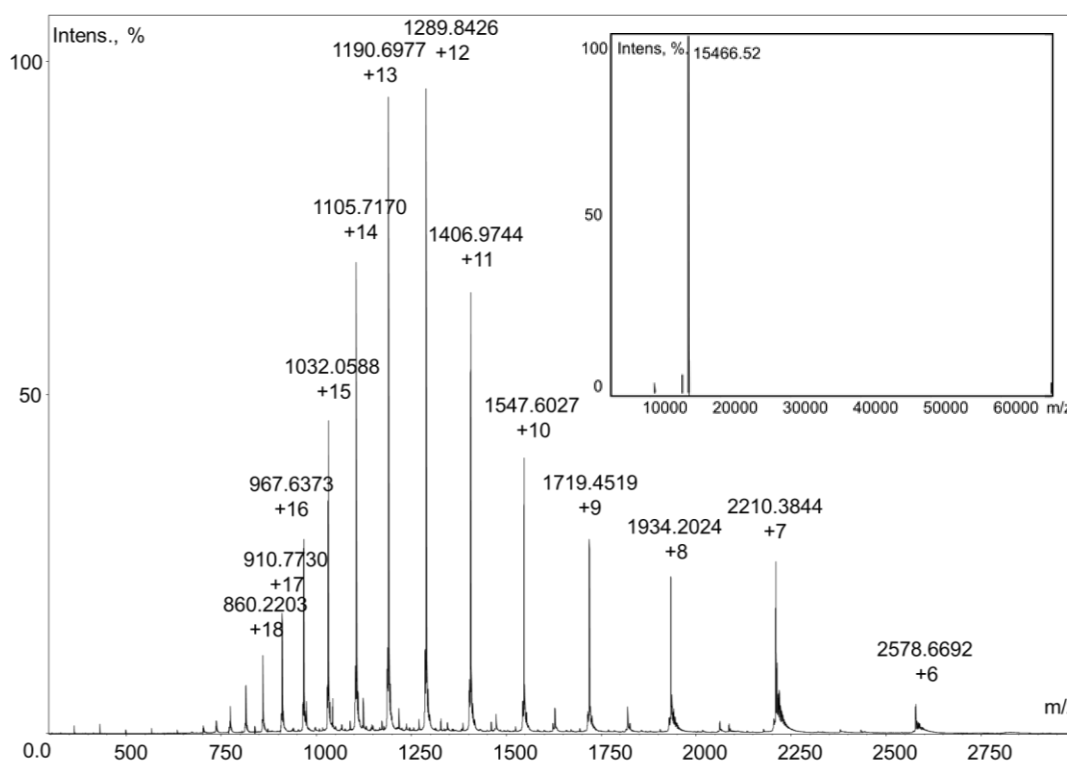


Figure 3.10 The full ESI-MS spectrum of apo-SynZur at pH 2.

The deconvoluted spectrum is shown in the top right hand corner. The pH of 5 μ M SynZur in 20 mM NH_4HCO_3 was adjusted to 2.0 with formic acid.

If zoomed in on one charge state (+9), it can be seen that in addition to the main SynZur peak of MW 15466.52 Da, minor peaks of molecular masses of 15431.44, 15487.51, 15529.48 and 15552.44 are present (see Figure 3.11). The mass difference with the main peak correspond to -34.08, 21.99, 63.96 and 86.92. The first peak was not identified, (though it could be H₂S eliminated from the sample due to an unknown reason) and the remaining 3 peaks likely correspond to the minor adducts SynZurNa, SynZurZn and SynZurZnNa, respectively.

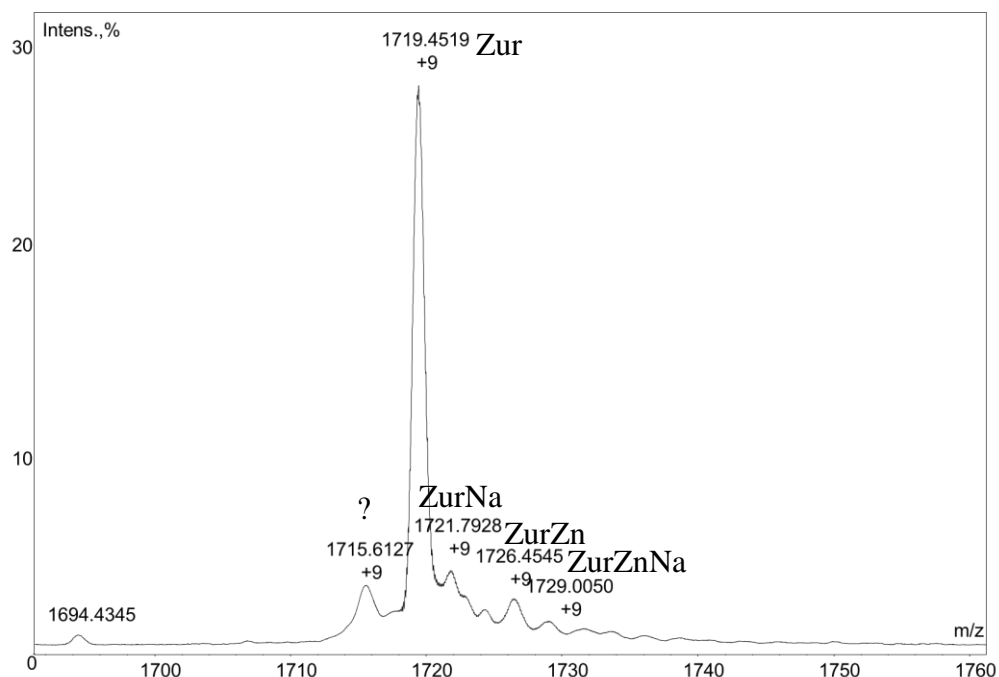


Figure 3.11 A single charge state (+9) of the apo-SynZur ESI-MS spectrum. The full range spectrum is shown in Figure 3.10.

Under neutral pH conditions (pH 7.9) proteins are expected to be in their native form and SynZur is supposed to keep Zn²⁺ within its structure. The total ESI-MS spectrum of native SynZur is shown in Figure 3.12.

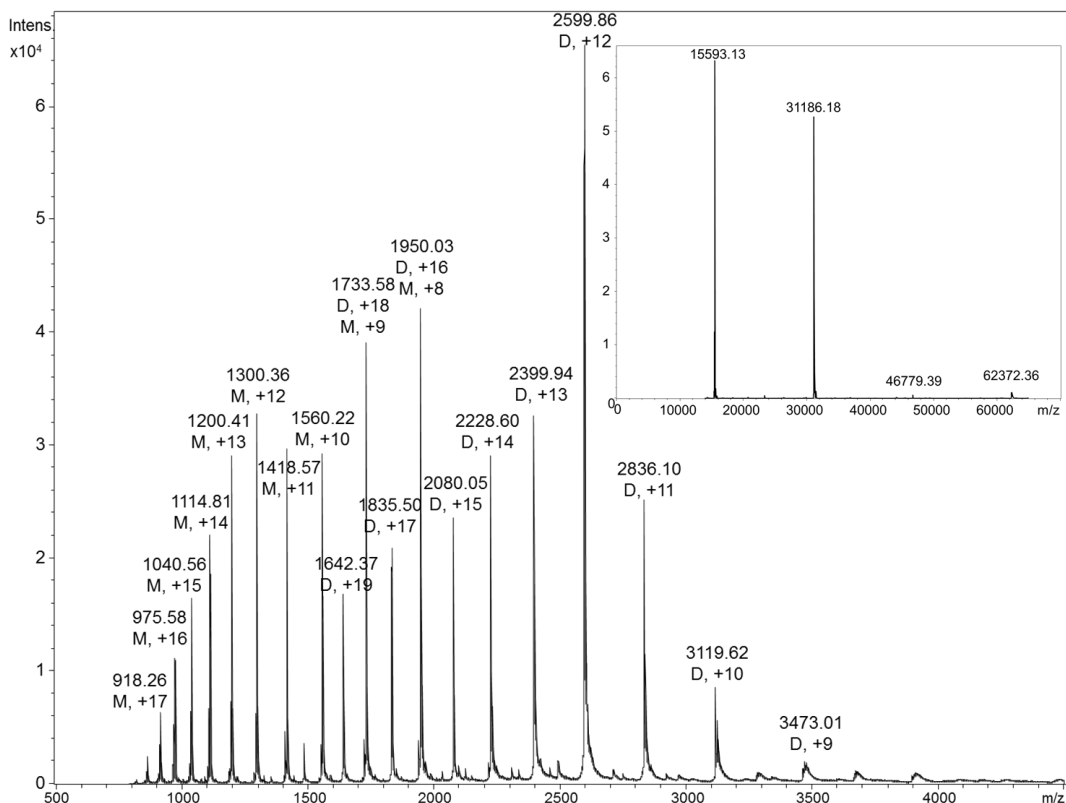


Figure 3.12 ESI-MS spectrum of the native SynZur protein.

SynZur is observed overwhelmingly in monomeric and dimeric forms. Peaks corresponding to the monomer are marked with the letter M, whilst dimer peaks are marked with the letter D. The deconvoluted spectrum shows a monomer, a dimer and small peaks corresponding to a trimer and a tetramer. The sample was 5 μM SynZur in 20 mM NH_4HCO_3 , pH 7.9.

Molecular weight values after deconvolution for a monomer and a dimer are shown in Table 3.5. Values found correspond to either SynZurZn₂ or SynZur₂Zn₄, which is in agreement with the ICP-OES data (see section 3.3.2.1.).

Table 3.5 ESI-MS results and comparison with the calculated molecular mass of SynZur under native conditions.

| | Found | Theoretical | Difference |
|----------------------------------|----------|-------------|------------|
| ZurZn ₂ | 15593.13 | 15593.19 | 0.06 |
| Zur ₂ Zn ₄ | 31186.18 | 31186.36 | 0.18 |

To calculate the molecular weight of SynZurZn₂ the following formula was used:

$$\text{MM (ZurZn}_2\text{)} = \text{MM (apo-Zur)} + 2 \times (\text{MM(Zn)} - 4 \times \text{MM(H)}) = 15593.19 \text{ Equation 3.3.}$$

A single charge state (+13) of a dimer is shown in Figure 3.13. The main peak with mass 31186.22 Da corresponds to SynZur₂Zn₄. Also, several minor peaks corresponding to 31224.18, 31249.27, 31261.10 and 31286.19 Da are present. The differences in mass with SynZur₂Zn₄ are 37.96, 63.05, 74.88 and 99.97, and likely correspond to SynZur₂Zn₄K, SynZur₂Zn₅, SynZur₂Zn₄K₂ and SynZur₂Zn₅K, respectively. Potassium adducts potentially appear due to pH adjustment of the buffer using a pH electrode containing KCl.

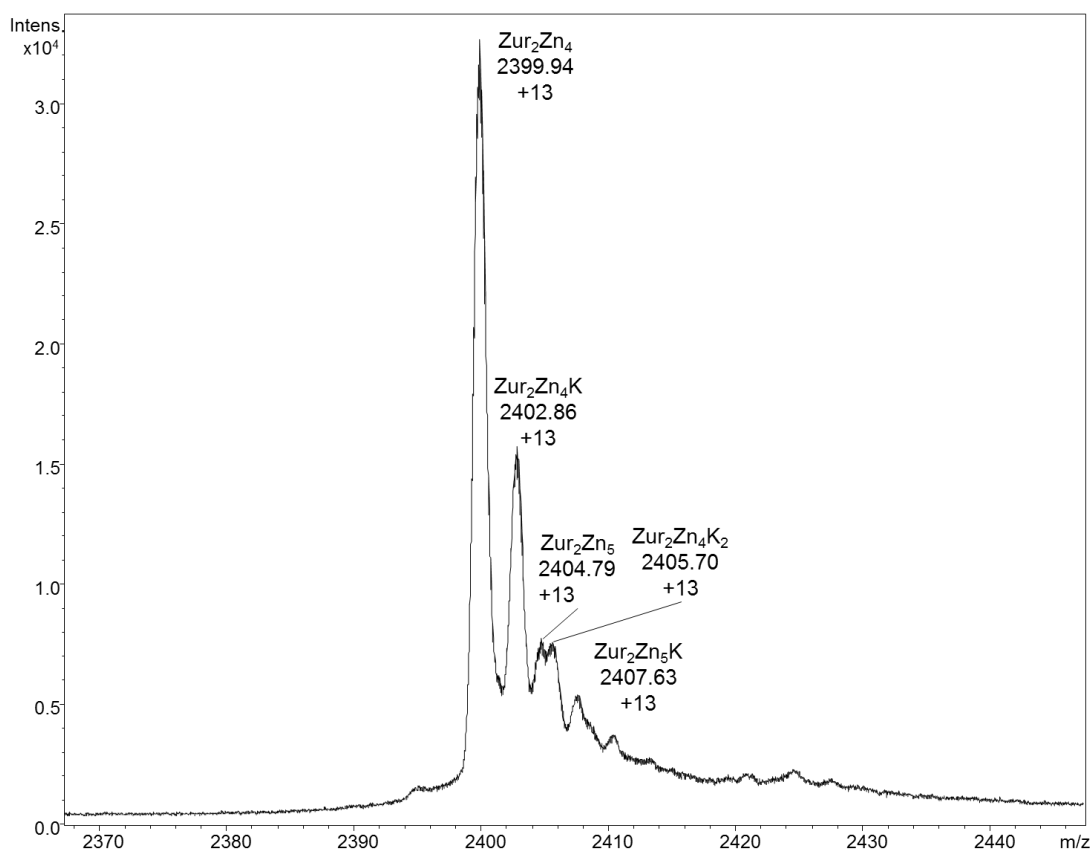


Figure 3.13 ESI-MS spectrum of a single +13 charge state of the SynZur₂Zn₄ protein in native conditions.

The full range spectrum is given in Figure 3.12.

Results of ICP-OES and ESI-MS analysis are overall in good agreement with the theoretical prediction³³ and as previously observed⁸⁸.

3.3.2.3 Non-denaturing SDS-PAGE

In order to assess oligomeric states of the SynZur protein in addition to using ESI-MS, ND SDS-PAGE was performed, as described in section 3.2.2.3⁹⁸. Unlike

conventional SDS-PAGE, this method does not denature proteins, allowing to observe them in their native state and having a better resolution than conventional blue native PAGE. Results of ND SDS-PAGE for SynZur at different concentrations, in comparison with SDS-PAGE, are presented in Figure 3.14.

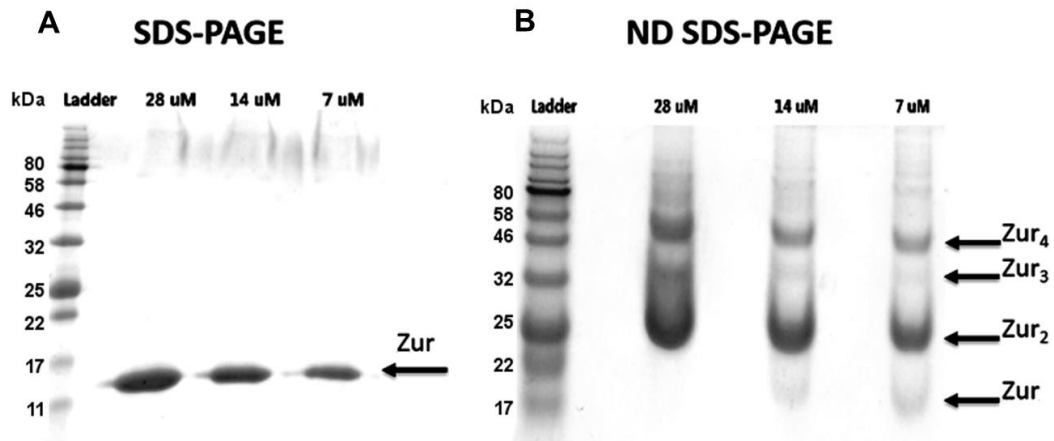


Figure 3.14 Visualisation of SynZur at different concentrations following SDS-PAGE or ND-SDS-PAGE.

A. A single band of the SynZur monomer is observed following SDS-PAGE. B. Several bands comprising different oligomeric states of SynZur are visible following ND SDS-PAGE.

Non-denaturing SDS-PAGE showed several oligomeric states of the SynZur protein, but still with the dimer form the most intense. A similar picture with the presence of several oligomers was previously observed for *Anabaena* PCC 7119 Fur¹¹¹. The presence of small amounts of trimer and tetramer were also observed by ESI-MS (see section 3.3.2.2). However, the amount of higher oligomers observed was lower than on the ND SDS-PAGE gel, probably due to the lower stability or ionization efficiency of these oligomers in ESI-MS. Interestingly, the sizes of the bands on this ND gel for dimer, trimer and tetramer show up lower than expected: for the dimer it is about 23-24 kDa, for the trimer it is ~35 kDa and for the tetramer it is about 48 kDa. Only the monomer shows up at the expected size of about 16 kDa.

To further clarify SynZur oligomeric state, it was decided to run analytical SEC described in the next part (see section 3.3.2.4.).

3.3.2.4 Multimeric state of SynZur as assessed by analytical SEC

The oligomeric state of SynZur was assessed by analytical size exclusion chromatography (SEC) in 100 mM NaCl, 50 mM Tris-HCl (pH 8) as described in section 3.2.2.4. Analytical SEC showed that SynZur elutes as a species with a calculated molecular weight of 45 kDa (see Figure 3.15).

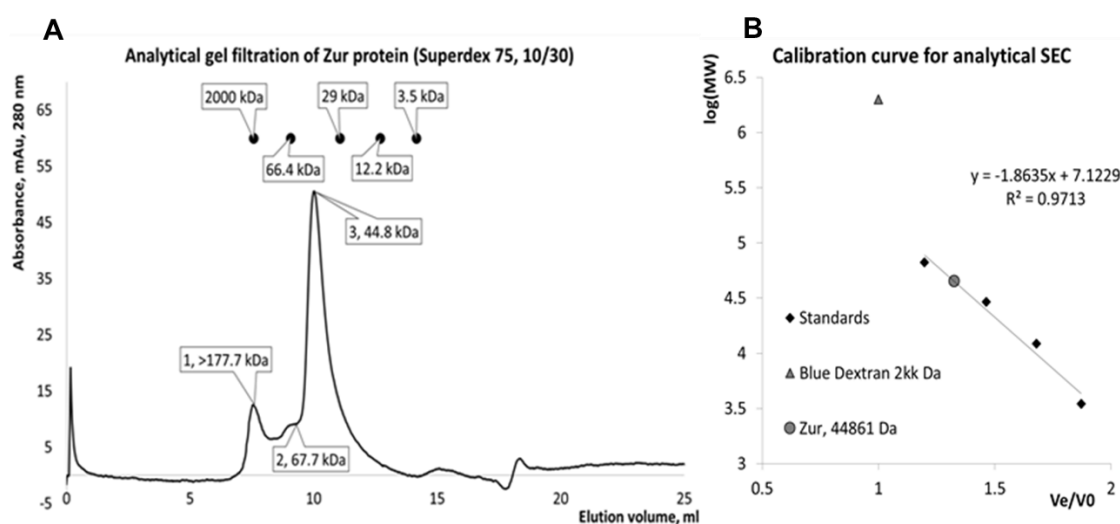


Figure 3.15 Analytical SEC of SynZur.

Standards used were: Blue Dextran, 2000 kDa; BSA, 66.4 kDa; Carbonic anhydrase, 29 kDa; Cytochrome c, 12.2 kDa; Insulin chain B oxidized, 3.5 Da. A. The SynZur SEC chromatogram. B. The calibration curve for analytical SEC. The 3 peaks of SynZur are marked with numbers.

This molecular weight approximately corresponds to that of a trimer. However, previously none of the Ferric Uptake Regulator family proteins was reported to occur as a trimer under native conditions, and the trimer was also at very low abundance in ND SDS-PAGE. In most studies Zur proteins were reported to be dimers^{49,112–114}. The apparent MW of a protein determined by SEC could be different from its real MW and thus, it was still unclear what oligomeric state SynZur has. Therefore, it was decided to run SynZur PAGE under completely native conditions using only Tris-Glycine as the running buffer. The result is shown in Figure 3.16.

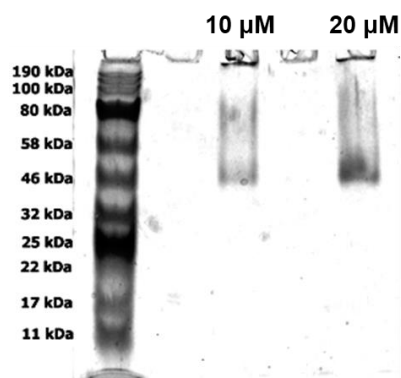


Figure 3.16 Native PAGE of SynZurZn₂ at pH 7.9.

Without addition of SDS in the running buffer, the observed main band of SynZur is about 47 kDa in size. The profile is very similar to that determined by analytical size exclusion chromatogram as well as being a similar size (45 kDa). In addition, the determined MW is also similar to the ND SDS-PAGE tetramer (~48 kDa). It will subsequently be shown (see section 3.3.3) that SynZur also crystallised as a tetramer. Thus, while it is still unclear what oligomeric state SynZur has under the conditions tested, it is likely that it is predominantly a tetramer at higher concentrations.

However, the species observed at 45-47 kDa could still be a dimer as well which might show up at higher MW on SEC due to a shape that significantly deviates from globular. Crystals might not represent its real solution state while PAGE profiles may not be reliable enough for MW determination. In addition, the oligomeric state of SynZur *in vivo* might be different from *in vitro* conditions due to various factors such as its concentration, ionic strength and redox potential and the presence of zinc, DNA and other proteins.

3.3.2.5 Investigation of SynZur de-metallation

To understand how SynZur would behave upon removing zinc from the sensory site, it was de-metallated as described in section 3.2.2.5. Using EDTA allows removing zinc only from the sensory site, but retaining it within the SynZur structural site as was found previously⁸⁸. Correspondingly, it was shown by Zn-EXAFS (Extended X-ray Absorption Fine Structure) that treatment of *E. coli* Zur with EDTA (leaving ZurZn₁) produced a predominant sulfur environment for the remaining zinc which corresponds to a structural Cys₄ site¹¹⁵. It is unclear if this site binds zinc with significantly higher affinity than EDTA ($K_{D \text{ Zn-EDTA}} = 2.3 \times 10^{-14} \text{ M}$ at pH 7.4¹¹⁶) which

is applied in large excess, or if there are some structural/electrostatic/kinetic restrictions³⁵.

After de-metallation overnight with 1 mM EDTA SynZur (SynZurZn₁) was purified from EDTA by twice passing through a PD-10 column and checked by ESI-MS under native conditions. The ESI-MS spectrum of SynZurZn₁ is shown in Figure 3.17.

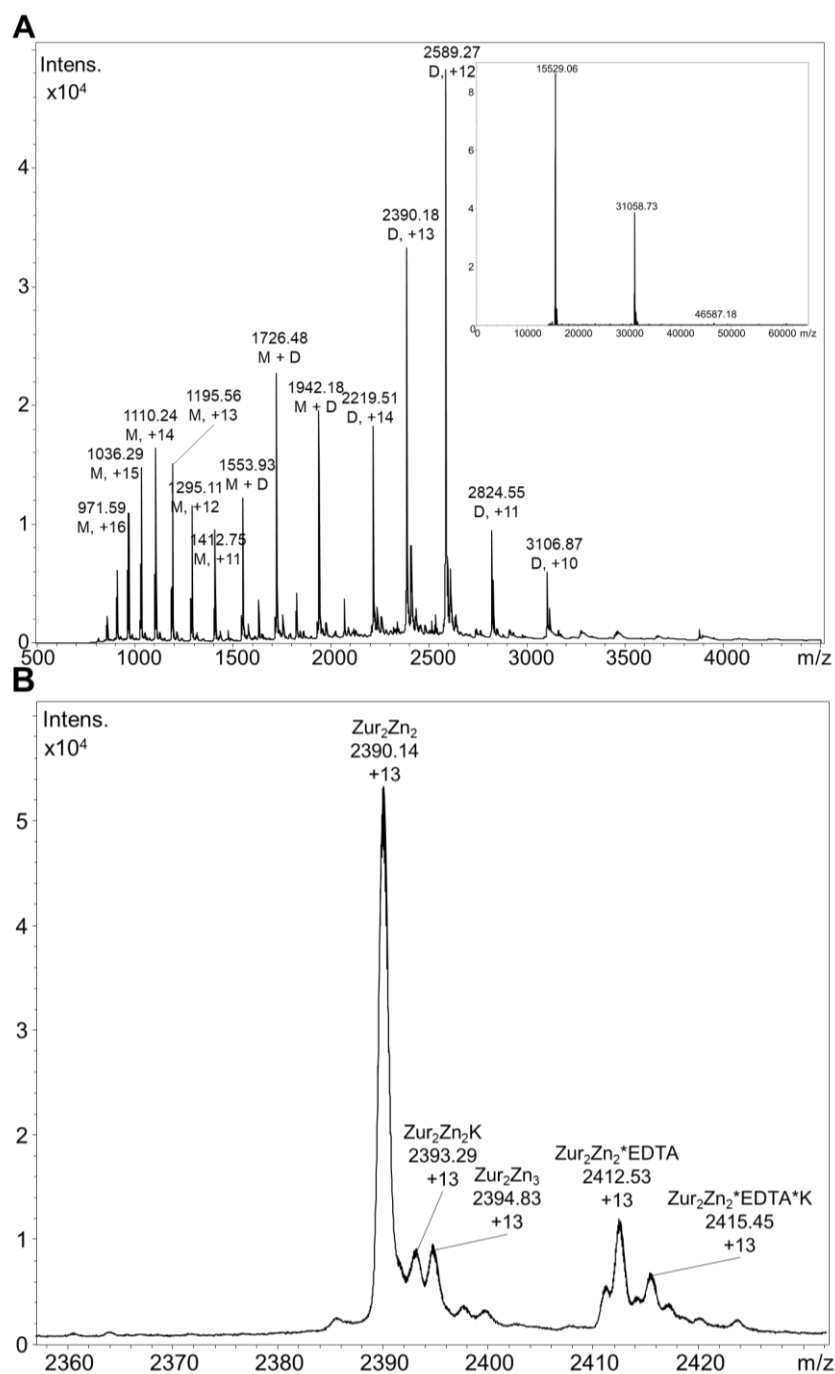


Figure 3.17 ESI-MS spectrum of de-metallated Zur.

A. Full range spectrum. **B.** A single charge state (+13) of a dimer of SynZurZn₁. 5 μM SynZurZn₁ was used, in 10 mM NH₄HCO₃, pH 7.9. Smoothing resulted in a slight change of the peak values.

The molecular mass of the de-metallated SynZur based on the spectrum is in very good agreement with the predicted mass, shown in Table 3.6. The predicted molecular mass was found using ExPASy¹⁰⁷ and using Equation 3.3 with the zinc coefficient equal to 1 and H coefficient equal to 2 for the monomer.

Table 3.6 ESI-MS of de-metallated SynZur, showing the experimentally determined mass compared with the predicted molecular mass.

| Species | Found by ESI-MS | Calculated by ExPASy | Difference |
|---------|-----------------|----------------------|------------|
| Monomer | 15529.06 | 15529.79 | 0.73 |
| Dimer | 31058.73 | 31059.59 | 0.86 |

Comparison of deconvoluted spectra of native SynZur and SynZurZn₁ revealed that the amount of dimer in SynZurZn₁ is almost half that in the SynZur species (see Figure 3.18). In addition, the tiny peak for the tetramer has completely disappeared in SynZurZn₁. Since FFP are typically able to bind DNA as dimers¹¹⁷, a decreased amount of dimer is expected to lead to lower DNA binding affinity of SynZurZn₁ as was shown previously⁸⁸.

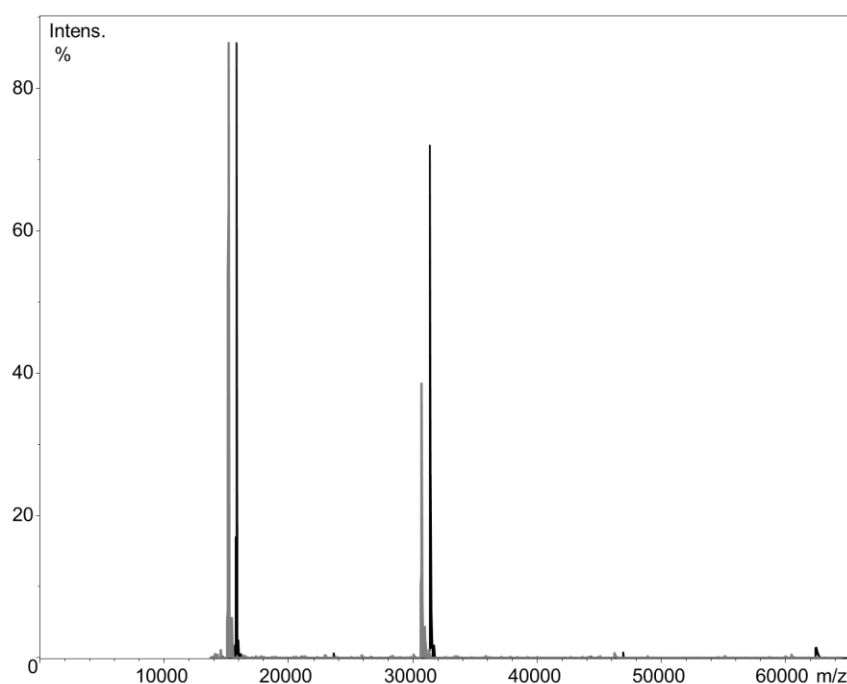


Figure 3.18 Comparison of deconvoluted ESI-MS spectra of native SynZur and SynZurZn₁.

Native SynZur is shown in black, SynZurZn₁ is shown in grey. Spectrum of SynZurZn₁ is slightly moved along the x-axis to better visualise the difference in dimer peak height. Also, the spectrum height of SynZurZn₁ was normalized on SynZur so that the monomers have visually the same size.

Ksibe (2016) showed that under native conditions SynZur runs as a mixture of a monomer, a dimer and another dimer with a different conformation (dubbed “well-folded dimer” based on it having low charge states)⁸⁸. However, the ESI-MS spectrum of native SynZur, acquired on a different instrument and shown here, looks different (see Figure 3.12). There is a much higher proportion of the charge states previously attributed to the well-folded dimer than observed in previous work, and from the native spectrum of SynZur₂Zn₄ it is unclear whether more than one dimeric charge-state is present. These differences could be explained by different ionisation conditions; in the present case, these were specifically optimised for proteins by Dr. Lijiang on a Bruker Compact instrument used in this study while Dr. Ksibe used a Bruker Daltonics MicroTOF.

Ksibe (2016) also showed that upon removal of zinc from the sensory site, peaks corresponding to the well-folded dimer almost disappeared after reacting 32 μ M Zur with 64 μ M EDTA⁸⁸. However, in the current study, using the ESI-MS instrument with higher sensitivity and apparently softer ionisation conditions, treatment of 32 μ M SynZur with 10 mM EDTA overnight caused no disappearance of the well-folded dimer peak (see Figure 3.19). However, a decrease in the intensity of the lowest charge states – corresponding to a higher degree of folding - is evident. It has to be mentioned that in the current study the SynZurZn₁ sample was purified through a PD-10 column twice before running the ESI-MS experiment. The presence of residual EDTA in less rigorously purified samples might cause decreased sensitivity; in addition the formation of EDTA adducts can lower the intensity of main peaks which might have caused the apparent disappearance of the peaks corresponding to the well-folded dimer in previous work.

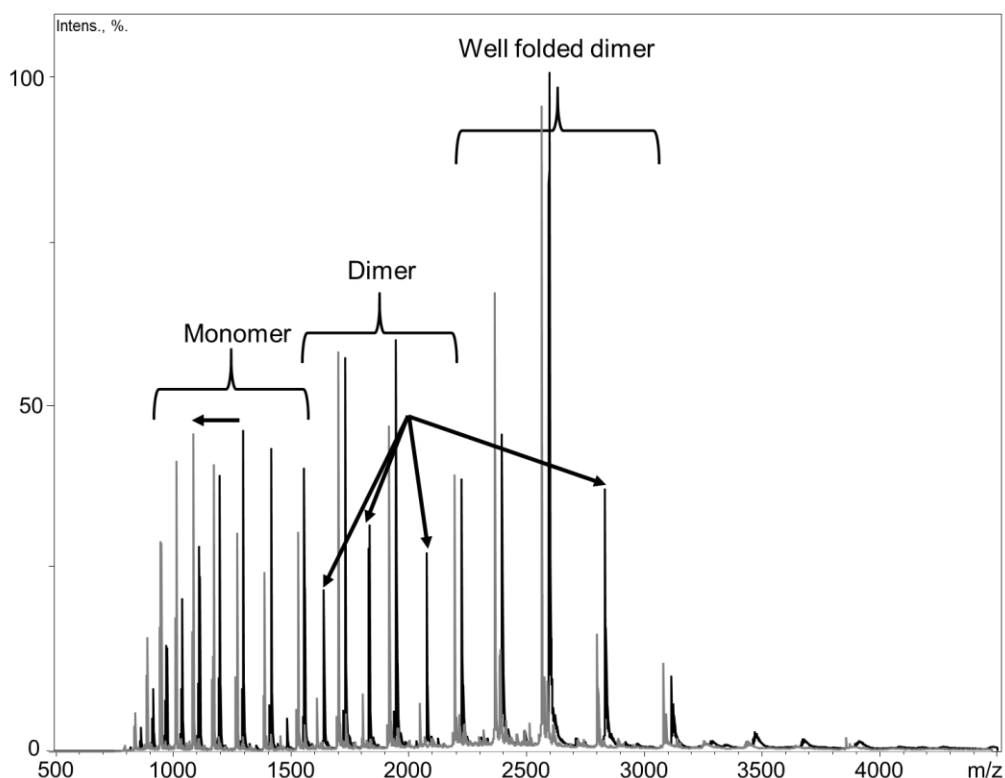


Figure 3.19 Comparison of native SynZur (black) and de-metallated at the sensory site SynZurZn₁ (grey).

The SynZurZn₁ spectrum was scaled to SynZur so that monomer envelopes have similar intensity. The speciation is taken from Ksibe (2016)⁸⁸. The most significant difference in the spectra are marked with black arrows. The SynZurZn₁ spectrum is moved slightly along the x-axis to make peaks not overlapping at this scale.

From the overlay of the SynZur and SynZurZn₁ native ESI-MS spectra, it can be seen that the distribution of the monomer peaks also moved towards lower m/z values (i.e. higher charge states) and some peaks corresponding to dimers significantly decreased. An overall decrease in dimers is confirmed by the deconvolution (see Figure 3.18). However, it is hard to tell which fraction of the dimer is more decreased without knowing the contribution of each species into each peak.

Ksibe (2016) showed previously that removal of the sensory zinc caused a significant decrease in the fraction corresponding to the “well-folded dimer” which led to the conclusion that binding sensory zinc is responsible for the allosteric switch towards the well folded dimer, which was hypothesised to correspond to the DNA-binding competent “closed” form of Zur⁸⁸. For the spectra acquired in the present study, it is unclear whether that hypothesis was correct.

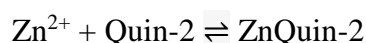
3.3.2.6 Zinc affinity of SynZur

Foster *et al.* (2014) showed that affinities of metal sensors correlate with intracellular free metal concentrations⁹⁰. If the K_D value of a metal sensor is far removed from the expected intracellular metal concentration, it is unlikely that this metal is a cognate species of this sensor. Thus, estimating SynZur affinity for zinc and its comparison with affinities of Zurs from other species could give additional confirmation to the idea of zinc being the sensed metal species for SynZur.

There are various methods for determining zinc affinity constants: titration of the apo-protein in the presence of metallochromic dyes such as Quin-2 or Mag-Fura2, Isothermal Titration Calorimetry (ITC), *in vitro* assays by EMSA or *in vivo* transcription assays. Importantly, if a dye is used, its affinity has to be similar to that of the protein, otherwise there is no effective competition between the protein and the dye. In addition, some methods can only detect K_D in a limited range. For example, ITC (direct titration) would not allow calculation of a K_D value $<10^{-8}$ M for reactions with a net heat of 10 kcal/mol with the concentration of the macromolecule equal to 10 nM¹¹⁸. A lower protein concentration would lead to heat changes below the detection limit.

To measure the affinity of SynZur for zinc, a basic SynZur property, titration of SynZurZn₁ with zinc in the presence of the metallochromic dye Quin-2 was chosen. Based on data from a variety of other species including cyanobacteria, Zurs should have K_{DS} in the picomolar to femtomolar range, a range suitable for use of this metallochromic dye³⁵.

After de-metallation of SynZur with EDTA, leading to an empty sensory site (see section 3.3.2.5.), it was titrated with ZnSO₄ in the presence of Quin-2. A representative Quin-2 competition titration is shown in Figure 3.20, with parameter values obtained from a simple step-wise binding model:



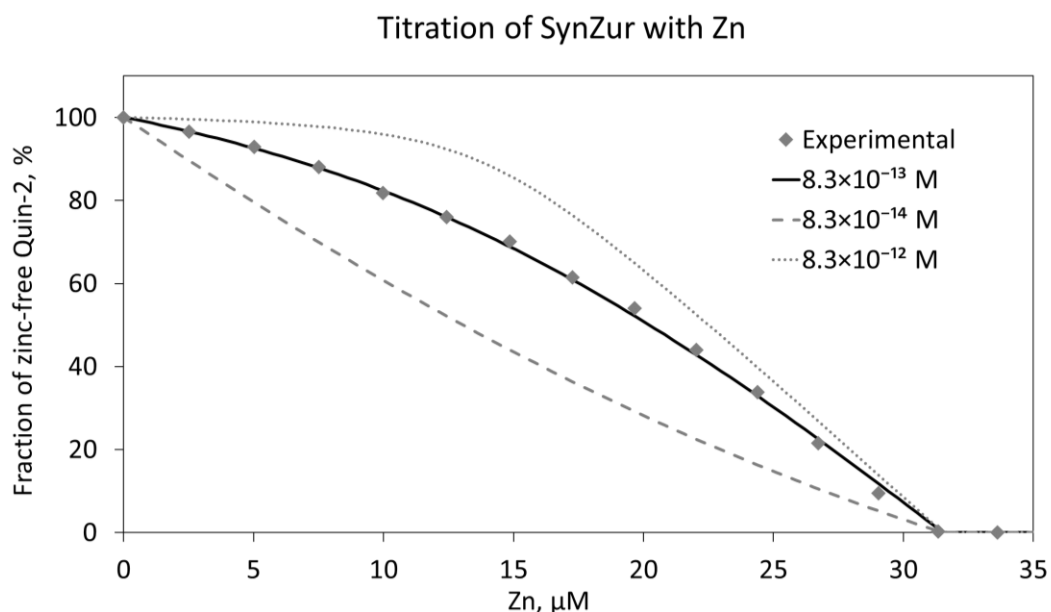


Figure 3.20 A representative titration curve of SynZur by zinc in the presence of Quin-2.

Zn^{2+} was titrated into 17.4 μM Quin-2 and 13.07 μM SynZur (as monomer). The solid curve represents the best fit using DynaFit¹⁰⁰ software. Dashed lines are simulated curves describing K_D of one order higher and lower values. The K_D found based on 3 replicates was $8.27 \pm 2.50 \times 10^{-13} \text{ M}$.

The script for calculation of the SynZurZn₁ dissociation constant used in DynaFit¹⁰⁰ software is shown in Figure 8.3 in the Appendix.

It should be mentioned that since de-metallation was not always complete, some variation in the titratable SynZur concentration was observed and for this reason, the Zur concentration was used as a variable. Three replicates were used to calculate K_D . It was found that the amount of zinc binding per each SynZurZn₁ molecule was 1.06. This approximately corresponds to binding of 1 zinc ion per monomer of SynZurZn₁ for which the K_D value was detectable by Quin-2. The dissociation constant of SynZur fitted by Dynafit¹⁰⁰ software was found to be $K_D = 8.27 \pm 2.50 \times 10^{-13} \text{ M}$ based on three replicates. This value is in very good agreement with other zinc affinity values found for Zur sensors from different microorganisms, especially those calculated using competitive titration by Zur in the presence of Quin-2: *B. subtilis* Zur: $5.5 \times 10^{-14} \text{ M}$ ¹¹³, *S. enterica* Zur: $6.36 \pm 0.41 \times 10^{-13} \text{ M}$ ¹¹⁹, *Synechocystis* sp. PCC 6803 Zur: $2.3 \pm 1.9 \times 10^{-13} \text{ M}$ ¹²⁰.

3.3.2.7 Electrophoretic Mobility Shift Assays (EMSA)

ZnuABC is a high affinity zinc uptake system and a member of the ATP-binding cassette (ABC) transporter protein family. It is often regulated by Zur in

bacteria. Ksibe (2016) showed binding of SynZur to *pznuABC* but not to the *ntcA* promoter which does not have a zur-box⁸⁸. However, in these experiments EDTA was used in the running buffer for EMSAs. Since EDTA removes zinc from the SynZur sensory site (see section 3.3.2.5) this may affect the DNA-binding ability of SynZur. Although Ksibe (2016) counter-balanced the presence of EDTA by supplying a high concentration of Zn²⁺ (10 mM) in the samples, it was desirable to conduct EMSA experiments under better controlled conditions, where the DNA-binding species was clearly defined. Hence, EMSAs were performed under conditions without EDTA.

The K_D value for DNA binding of Zur from other organisms varies dependent on the target gene, but ranges from $0.025 \times 10^{-18} \text{ M}^2$ (*ppl31p*) for *E. coli* for 2 Zur dimers (which roughly corresponds to $\sim 0.16 \text{ nM } K_D$ per dimer)⁴⁸, to 2.5 nM (*ppall4725*) in *Anabaena* sp. PCC 7120⁵⁹ and 74.9 nM (*pznuA*) in *S. coelicolor*⁴⁷.

In order to assess SynZur binding to the operator-promoter region of the *Synechococcus* sp. WH8102 *znuABC* genes by EMSA, a region containing a predicted zur-box, *pznuABC* was amplified by PCR (Figure 3.21).

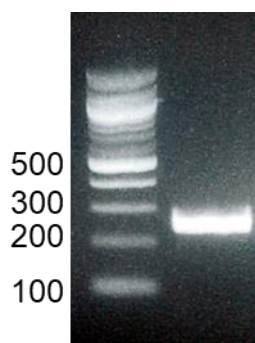


Figure 3.21 Agarose gel electrophoresis of the region of *pznuABC* from *Synechococcus* sp. WH8102 after PCR.

A product of 256 bp is observed as predicted.

Preliminary EMSA experiments were performed as described in section 3.2.2.1 across a wide range of SynZur concentration (0-100 μM) (data not shown) and optimal conditions were identified for observing both free and complexed DNA (see Figure 3.22).

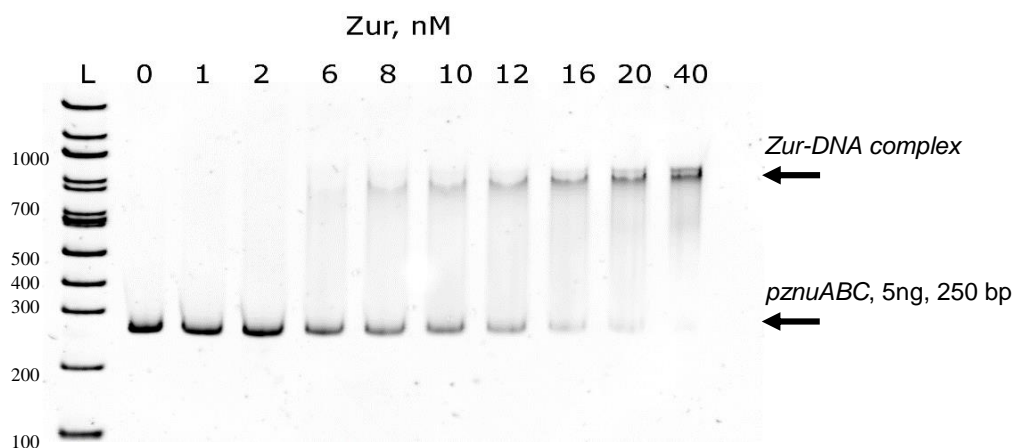


Figure 3.22 EMSA of *pznABC* with SynZur.

Free DNA and the shifted SynZur-DNA complex are indicated with black arrows. EMSA conditions are given in section 3.2.2.1. Total DNA concentration is 3.16 nM. Zur concentration is given for the monomer.

A shifted band is evidence of formation of a SynZur-*pznABC* complex. A shifted band, with the same intensity as the free DNA band, was observed at a SynZur concentration of 10–12 nM (see Figure 3.22), a value much lower than the 200 nM concentration observed by Ksibe (2016)⁸⁸, suggesting that under the improved EDTA-free conditions used here, a lower K_D is apparent. The free SynZur concentration when the SynZur-DNA complex is half dissociated represents the dissociation constant, assuming 1:1 stoichiometry of the complex. However, to calculate a meaningful dissociation constant, it is important to know the stoichiometry of the protein-DNA complex. This was determined using Ferguson Plots (see section 3.3.2.7.1).

3.3.2.7.1 *Zur-pznABC stoichiometry determination*

Because it is not possible to determine the molecular weight of a protein-DNA complex based on the molecular weight of DNA or protein ladders, Ferguson analysis was used to determine the stoichiometry of these complexes. It is known that the logarithm of relative mobility (Rf) of protein-DNA complexes depends linearly on the gel percentage¹²¹. Hence, it was decided to use a simplified protocol with gels of just two concentrations: 8 and 12% (see section 3.2.2.1). The EMSAs performed and the measurements taken are shown in Figure 3.23.

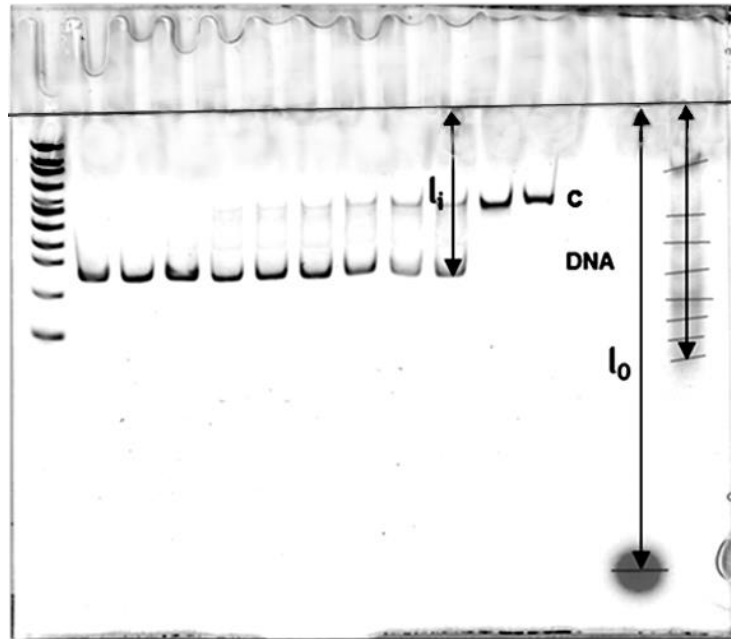


Figure 3.23 EMSA using a 8% gel and explanation of how measurements of relative mobility were made.

Rf values were calculated by dividing l_i/l_0 . C: The SynZur-DNA complex.

Values for measured mobility and relative mobility (Rf) are presented in Table 3.7.

Table 3.7 Mobility and Rf values of different bands using 8 and 12% acrylamide gels used to build Ferguson plots.

Distances are given in pixels and measured using Gimp software.

| | $l_{i,12}$ | Rf_{12} | $l_{i,8}$ | Rf_8 |
|--------------|------------|-----------|-----------|--------|
| Dye | 976 | 1 | 1250 | 1 |
| 11 kD | 364 | 0.373 | 683 | 0.5464 |
| 17 kD | 313 | 0.3207 | 628 | 0.5024 |
| 22 kD | 282 | 0.2889 | 577 | 0.4616 |
| 25 kD | 237 | 0.2428 | 523 | 0.4184 |
| 32 kD | 183 | 0.1875 | 449 | 0.3592 |
| 46 kD | 122 | 0.1250 | 371 | 0.2968 |
| 58 kD | 80 | 0.0820 | 296 | 0.2368 |
| 80 kD | 33 | 0.0338 | 164 | 0.1312 |
| C | 85 | 0.0871 | 248 | 0.1984 |
| DNA | 246 | 0.2520 | 444 | 0.3552 |

Based on the data obtained, plots of $\log R_f$ ($100 \times \log(R_f \times 100)$) vs gel percentage were built (Ferguson plots) for the unknown complex, free DNA and ladder proteins as described in Gilston *et al.*⁴⁸. Ferguson plots built using calculated R_f coefficients are shown in Figure 3.24, with the slopes (K_r) of each line also given.

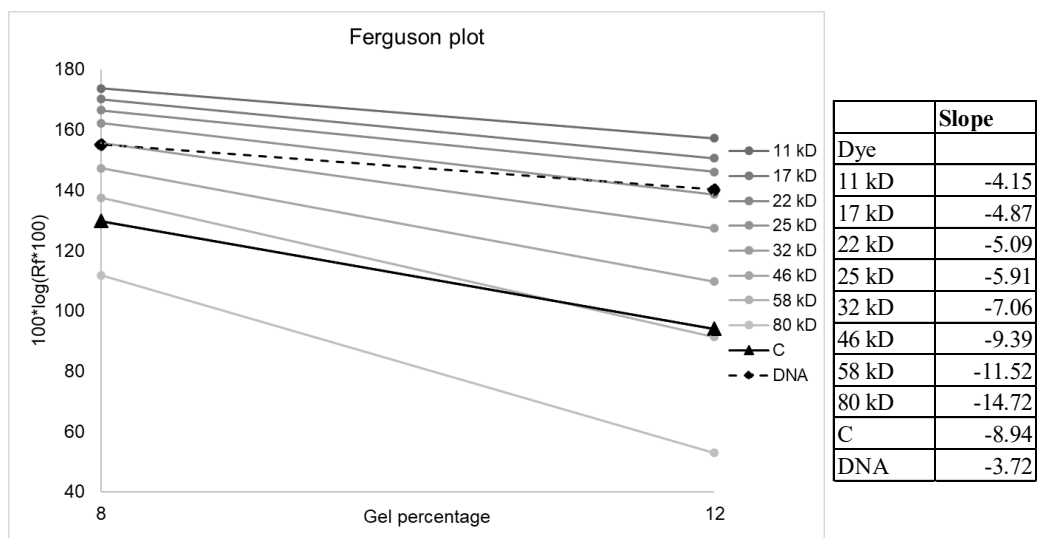


Figure 3.24 Ferguson plots. Values of slopes (K_r) are given in the Table.

Subsequently, a plot of K_r versus the molecular weight of the SynZur–DNA complex, free DNA and protein standards was made (see Figure 3.25). It is almost perfectly linear ($R^2=0.9956$) allowing “molecular weight” values for the complex and free DNA to be determined (see Table 3.8).

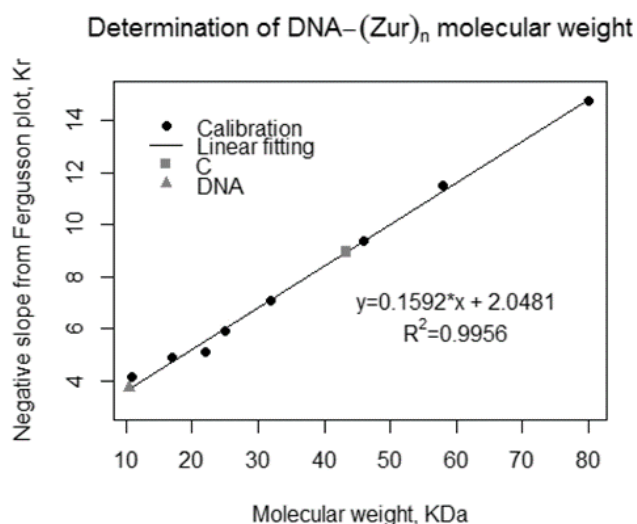


Figure 3.25 Determination of molecular weight of the SynZur-DNA complex based on the calibration curve constructed with MW of standard proteins.

Table 3.8 Determination of molecular weights of the SynZur-DNA complex (C) based on the calibration curve.

| | "MW" | "MW"-DNA | Predicted bound species | Theoretical MW | Error, % |
|-----|--------|----------|----------------------------------|----------------|----------|
| C | 43.286 | 32.755 | Zur ₂ Zn ₄ | 31.186 | 5.03 |
| DNA | 10.531 | | | | |

Since a protein ladder was used for calibration of molecular weights, and DNA has a much higher negative charge value per Da compared to proteins, those values cannot be used directly as molecular weights of the complexes. However, after subtracting the DNA contribution, the correct molecular weight of the protein in the complex can be calculated. The results are presented in Table 3.8.

The molecular weight of the SynZur-DNA complex after subtraction of the DNA component (32.755 kDa) corresponds closely to the size of the SynZur dimer (31.19 kDa) bound to DNA (see Table 3.8). Using a similar approach, the stoichiometry of *E. coli* Zur-DNA complex was determined and was found to be Zur₄ – DNA⁴⁸.

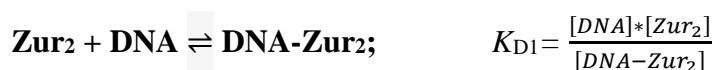
Now that the stoichiometry of the SynZur-*pznuABC* complex is known, it is possible to calculate a dissociation constant, as performed in section 3.3.2.7.2.

3.3.2.7.2 Calculation of dissociation constants of SynZur-pznuABC complexes

This was carried out by fitting models based on the equilibria described below to the experimental values. Three replicates of EMSA-gels were used and band densities were quantified using ImageJ software.

Since the stoichiometry of the SynZur₂-pznuABC complex is now known, a mathematical model of the interaction between SynZur and DNA can be proposed. The most straightforward is a simple 1:1 interaction of a SynZur dimer with DNA (Model 1):

Model 1: Reaction with Zur dimer only:



Fitting was done using Dynafit¹⁰⁰ software and visualized in Microsoft Excel (see Figure 3.26, grey curve). The script for model 1 is shown in the Appendix (see Figure 8.4A). The fitted value for K_{D1} was found to be $4.78 \pm 0.82 \times 10^{-9}$ M in the absence of EDTA. Previously, in the presence of EDTA in the running buffer, the same model gave a value of 58.29 nM⁸⁸. If compared to another cyanobacterium *Anabaena* sp. PCC7120, Zur-DNA K_D defined by EMSA were in a similar range: 2.5 nM for *pall4725* and 7 nM for *pall4723*⁵⁹.

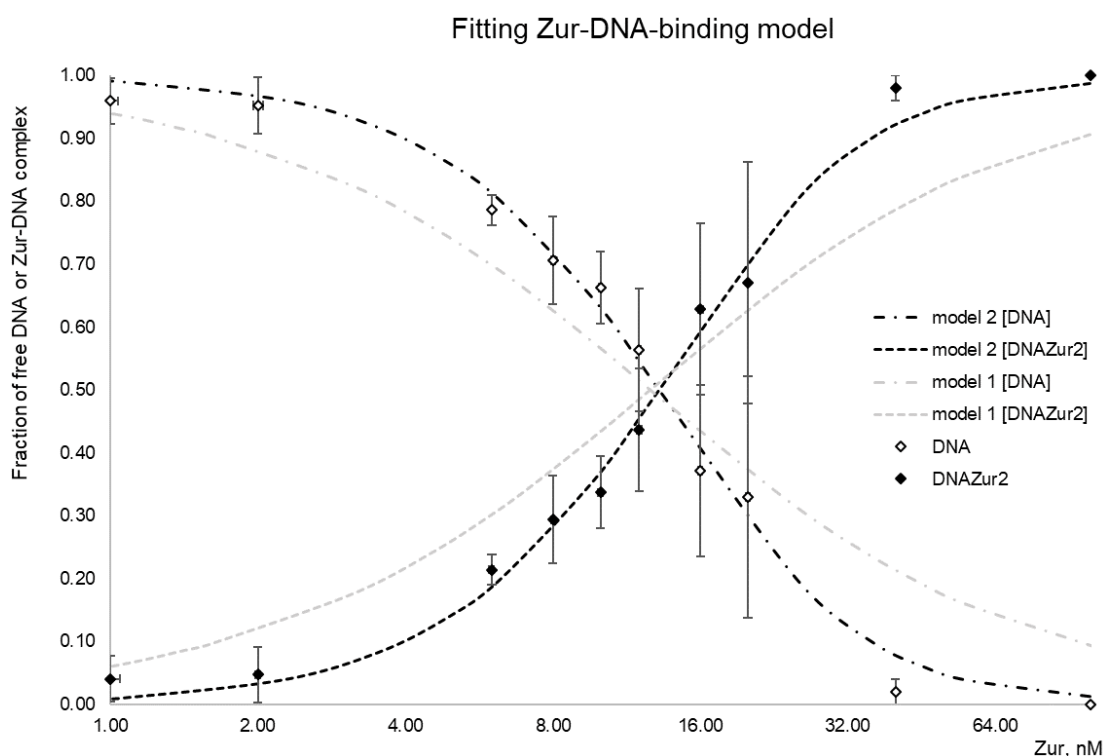


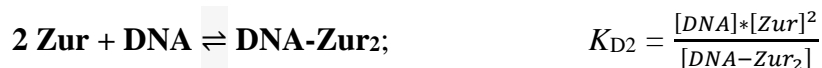
Figure 3.26 Visualisation of the EMSA data and fitted models.

Fitting was performed to the average values at each SynZur concentration. Error bars indicate errors between three replicates.

However, it was observed that Model 1 does not fit the experimental data very well: the fitted curves are much flatter than the experimental data points and the coefficient of variation (CV) was equal to 17.2%. However, it is possible that at the highest and lowest Zur concentrations weaker intensity bands were not quantified correctly due to the detection limit and gave values close to 0 leading to the steeper behaviour of the experimental data.

Since it was unclear what oligomerisation state SynZur has (see sections 3.3.2.3.-3.3.2.5), it is also possible that SynZur is not acting as a pre-formed dimer in solution. In addition, ND SDS-PAGE (see section 3.3.2.3) showed that a decrease in Zur concentration leads to the formation of some monomer at 7 μM total SynZur. Therefore, it is possible that at nanomolar concentrations which were used for EMSAs, SynZur could contain a significant amount of monomer or even be 100% monomer. Thus, it was decided to test a model for two monomers binding to DNA (Model 2):

Model 2: Reaction of DNA with SynZur monomers:

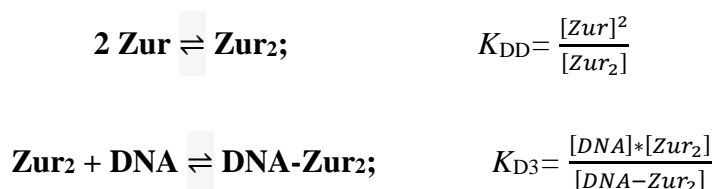


The resulting curves fitted the experimental data significantly better (CV = 7.4 %) and are shown in Figure 3.26 in black. The script for model 2 is shown in the Appendix (see Figure 8.4B). The fitted value for K_{D2} was found to be $9.99 \pm 0.74 \times 10^{-17} \text{ M}^2$. This apparent lower value of K_{D2} represents two SynZur protein molecules binding DNA. In order to compare affinities with model 1 or other values from the literature based on dimeric Zur, a square root has to be taken giving $K_{D2}' \sim 10 \text{ nM}$.

It should be mentioned that the *in vivo* SynZur concentration might be different to that used here. In this project it was not analysed, but for example in *Streptomyces coelicolor* it was reported that the Zur concentration within cells was $3.7 \mu\text{M}^{73}$ which is significantly higher than the concentration used for K_D determination by EMSA in this study. Whilst the second model proposes the interaction of two SynZur monomers with DNA to fit the experimental data, this may not necessarily represent the actual DNA-SynZur binding mechanism within *Synechococcus* cells.

It is also possible that if monomers are present in solution at low SynZur concentrations, they could form a dimer first and then the dimer could bind to DNA giving Model 3:

Model 3: Reaction of DNA with a SynZur dimer formed from monomers:



However, the dissociation constant of SynZur dimers to monomers (K_{DD}) is unknown and it was not possible to use this model since fitting two K_D s gave ambiguous results. For example, whilst it was possible to fit the curve manually in Dynafit¹⁰⁰ with set up $K_{DD} = 50 \text{ nM}$ this gave a $K_{D3} = 1.13 \pm 0.11 \times 10^{-9} \text{ M}$ with a CV=10%. This K_{D3} would be on the low end of values reported in the literature. Setting up higher values of K_{DD} gave even lower values of K_{D3} which is unlikely. Moreover, lower values of K_{DD} resulted in a less well fitted curve (CV>10%). The script and the fitted plot with the given values are shown in the Appendix, Figure 8.5.

Overall, the value for DNA-Zur₂ K_D is estimated to be in the range of 1–5 nM based on models 1 and 3 or $3 \pm 2 \times 10^{-9}$ M for the dimer (Models 1 and 3) or 10.0 ± 0.3 nM based on the monomer (Model 2).

3.3.2.8 Metal-selectivity of SynZur

3.3.2.8.1 Metal-responsive DNA-binding

In order to assess whether zinc is the cognate species for SynZur in *Synechococcus* sp. WH8102, EMSAs were used to assess SynZur–DNA binding in the presence of different divalent metals such as Zn, Cd, Co, Cu, Fe, Mn and Ni. If SynZur binds to DNA selectively in the presence of zinc, this would be an additional confirmation of it being a zinc sensor.

First, purified SynZur protein was de-metallated with EDTA as described in section 3.2.2.5. After removal of free EDTA and Zn–EDTA using a PD-10 column inside an anaerobic glove-box, SynZurZn₁ was mixed with different metals and 5 ng *pznuABC*. Then EMSAs were performed (see Figure 3.27).

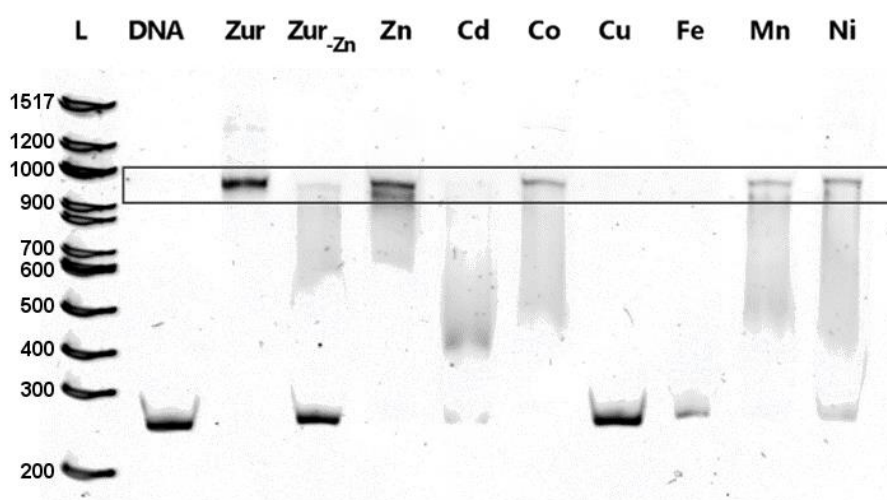


Figure 3.27 EMSA analysis of DNA binding of SynZur in the presence of different metals.

The box shows the position of the SynZur₂-DNA complex and the area used for ImageJ quantification per lane (see Figure 3.28). EMSA conditions are described in section 3.2.2.2. Total DNA concentration is 3.16 nM. Zur concentration is 160 nM as monomer and concentration of metals is 1.6 μ M.

De-metallated SynZur showed an intense band corresponding to free DNA and a weak smeared band corresponding to the Zur–DNA complex (see Figure 3.27, lane “Zur-zn”). This might be explained by some residual weak binding of de-metallated Zur to DNA. Alternatively, it is likely that Zur was not completely de-metallated and

some ZurZn₂ remained in the reaction mixture allowing complex to be formed. In addition, zinc re-distribution from the structural site to the sensory site of Zur could also take place which would result in a small amount of ZurZn₀ and DNA-binding ZurZn₂. Finally, all the reagents might contain zinc or other metal impurities which would bind to SynZurZn₁ resulting in the observation of a SynZur-DNA complex.

The strongest bands, corresponding to the Zur₂-DNA complex, are present for the original SynZur and SynZurZn₁ following zinc reintroduction (see Figure 3.27, lanes “Zur” and “Zn”, respectively). This indicates that SynZur has the highest DNA-binding propensity in the presence of zinc under these experimental conditions.

Quantification of the shifted band intensities with ImageJ is shown in Figure 3.28.

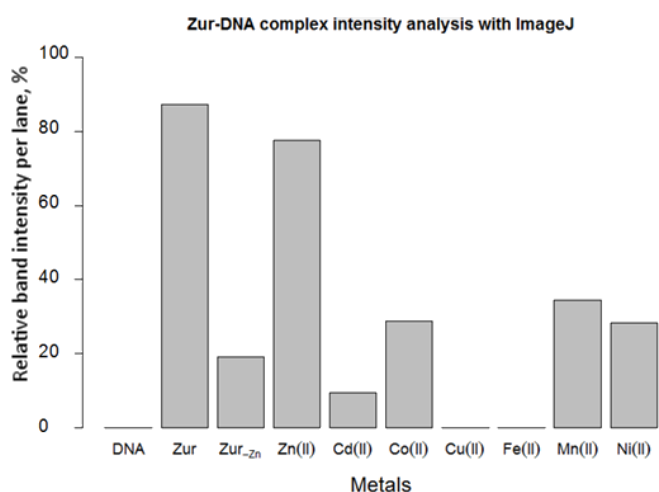


Figure 3.28 Quantification with Image J of the shifted bands within the box for the EMSAs shown in Figure 3.27.

The relative intensities of the bands inside the box compared to the total lane intensity are shown.

Despite some complex formation observed in the presence of Co, Mn and Ni, the intensities of the bands were only slightly higher compared to de-metallated SynZur (see Figure 3.28). This is probably due to the lower DNA-binding affinity of SynZur with the listed metals bound (see section 3.3.2.8.2) compared to zinc. Intensities of the bands with Cd, Cu and Fe added are lower than de-metallated Zur, hence showing no binding.

It was shown previously that some metals such as Cd²⁺ and Co²⁺ could substitute Zn²⁺ in metalloproteins whilst maintaining or even improving the activity of the protein^{122,123}. In the case of SynZur, the lack of response to cadmium was unexpected due to the fact that the tetrahedral coordination geometry and electronic

configuration of cadmium is similar to zinc. Some response to Co^{2+} can be expected, as Co^{2+} is well known to substitute for Zn^{2+} in protein sites, including those with tetrahedral geometries^{124,125}. This is unlikely to be an issue *in vivo*, as Zn^{2+} generally outcompetes Co^{2+} based on the Irving-Williams series¹²⁶.

Interestingly, there was no response to Cu^{2+} which is important, as copper is usually able to outcompete Zn^{2+} if present at the same levels¹²⁶. In addition, both Cu^+ and Cu^{2+} can in principle adopt tetrahedral coordination geometry in proteins¹²⁷, the same as Zn^{2+} . It is suggested that Cu^{2+} may react with free cysteines, including those in the sensory site, leading to their oxidation. This could lead to formation of intermolecular and intramolecular disulfide bridges¹²⁸. This might cause unfolding and the complete inability of Zur to bind DNA in the presence of Cu^{2+} .

3.3.2.8.2 ESI-MS of Zur with different metals

To clarify whether the binding of different metals to SynZurZn₁ affects its oligomerisation state and conformation, or whether the binding is possible at all, SynZur, de-metallated at the sensory site, was prepared in 20 mM NH_4HCO_3 mixed with different metals and analysed by ESI-MS as described in section 3.2.2.2. Full range ESI-MS spectra are shown in Figure 3.29. For this experiment, metals were added to SynZurZn₁ in slight excess to compensate for any remaining EDTA since small amounts of EDTA adduct had been observed in the SynZurZn₁ ESI-MS spectrum (Figure 3.17B).

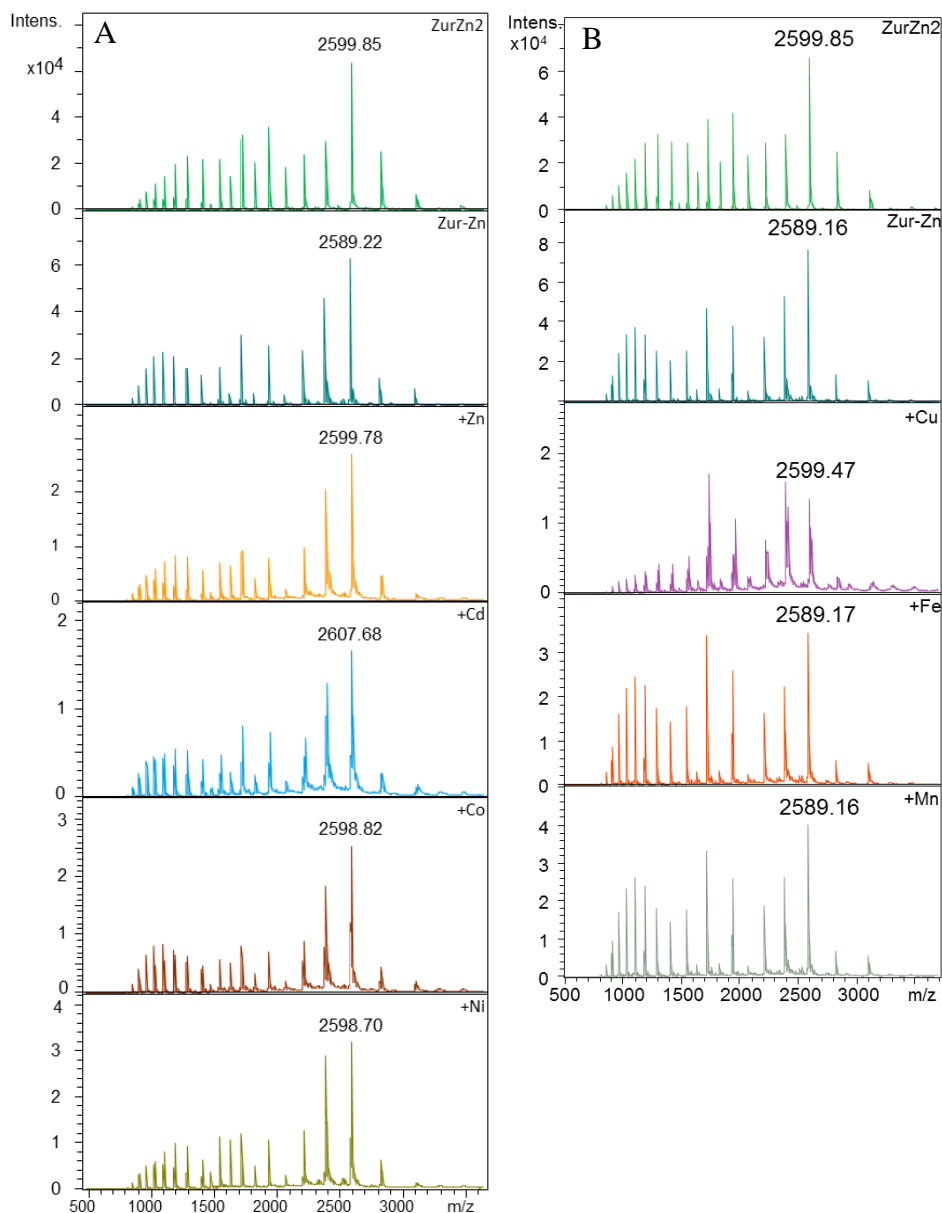


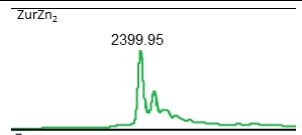


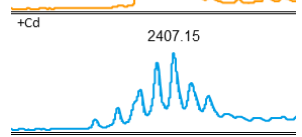
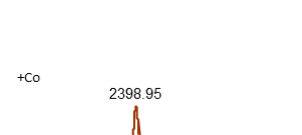
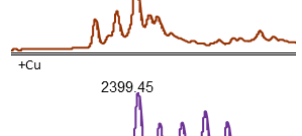
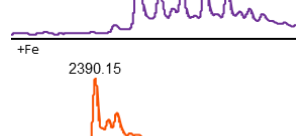
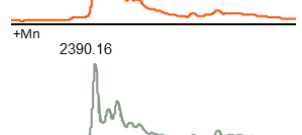
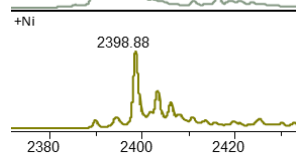
Figure 3.29 Full ESI-MS spectra of SynZur with different metals.

The top panels in both A and B show spectra of native SynZur and SynZurZn₁ for comparison. A. Spectra of SynZurZn₁ with different metals which showed binding, and where the overall peak distribution was similar to the original SynZur. B. Spectra where either no binding with metals was observed (Fe, Mn) or where the peak distribution changed (Cu). Two fold excess of metals was added to $\sim 5 \mu\text{M}$ SynZurZn₁ in 10 mM NH_4HCO_3 , pH 7.9.

Single charge states (Dimer, +13) with peak assignments are presented in Table 3.9 for all preparations.

Table 3.9 Species observed by ESI-MS for SynZur supplemented with different metals.

The main peaks are highlighted in different colours. Species where cadmium replaced zinc in the structural site are shown in bold.

| Spectrum name | Molecular mass of the peaks after deconvolution, Da | Species | ESI-MS spectra of a single charge state (+13, Dimer) |
|--------------------|--|--|---|
| ZurZn ₂ | 31186.4; 31224.1 | Zur ₂ Zn ₄ ; Zur ₂ Zn ₄ K |  |
| ZurZn ₁ | 31058.8; 31350.9 | Zur ₂ Zn ₂ ; Zur ₂ Zn ₂ ×EDTA |  |
| +Zn ²⁺ | 31184.9, 31224.8, 31247.2, 31284.8 | Zur ₂ Zn ₄ , Zur ₂ Zn ₄ K, Zur ₂ Zn ₅ , Zur ₂ Zn ₅ K |  |
| +Cd ²⁺ | 31058.6, 31121.2, 31170.71, 31184.5, 31231.4, 31279.5 , 31327.5, 31377.4 | Zur ₂ Zn ₂ , Zur ₂ Zn ₃ , Zur ₂ Zn ₂ Cd, Zur ₂ Zn ₄ , Zur ₂ Zn ₃ Cd, Zur₂Zn₂Cd₂ , Zur₂ZnCd₃ , Zur₂Cd₄ |  |
| +Co ²⁺ | 31058.6, 31115.7, 31172.8 , 31210.2, 31231.3 | Zur ₂ Zn ₂ , Zur ₂ Zn ₂ Co, Zur₂Zn₂Co₂ , Zur ₂ Zn ₂ Co ₂ K, Zur ₂ Zn ₂ Co ₃ |  |
| +Cu ²⁺ | 31179.6 , 31240.1, 31302.7, 31368.2, 31429.3 | Zur ₂ Zn ₂ Cu ₂ , Zur ₂ Zn ₂ Cu ₃ , Zur ₂ Zn ₂ Cu ₄ , Zur ₂ Zn ₂ Cu ₅ , Zur ₂ Zn ₂ Cu ₆ |  |
| +Fe ²⁺ | 31058.6 , 31095.9, 31120.7 | Zur₂Zn₂ , Zur ₂ Zn ₂ K, Zur ₂ Zn ₃ |  |
| +Mn ²⁺ | 31058.5, 31096.2, 31120.5 | Zur ₂ Zn ₂ , Zur ₂ Zn ₂ K, Zur ₂ Zn ₃ |  |
| +Ni ²⁺ | 31059.2, 31115.9, 31172.6 , 31229.8, 31268.8 | Zur ₂ Zn ₂ , Zur ₂ Zn ₂ Ni, Zur₂Zn₂Ni₂ , Zur ₂ Zn ₂ Ni ₃ , Zur ₂ Zn ₂ Ni ₃ K |  |

Firstly, no binding of SynZurZn₁ with either Fe²⁺ or Mn²⁺ was observed by ESI-MS. This is in agreement with the metal-selective EMSA experiments where either no DNA binding was observed (Fe) or only a weak band was observed (Mn) which was of comparable intensity to the SynZurZn₁ band (see Figure 3.27). It is

unclear why no free DNA was observed for Mn by EMSA. Possibly, SynZur-Mn binding is weak and not observed under ESI-MS ionization conditions while in solution SynZurZnMn might still be produced forming weak SynZurMn-DNA complexes. They might also show up smeary on the EMSA gel due to a constant association-dissociation process.

For Co and Ni, ESI-MS spectra showed the formation of SynZurZn₁Me₁ species (and the corresponding dimers) with the peak distribution similar to SynZurZn₂ suggesting a similar oligomerisation state and protein conformation. For both metals weak bands were observed during EMSA (see Figure 3.27, Co and Ni). However, both were weaker than SynZurZn₂. This leads to the conclusion that despite SynZur being able to bind Co and Ni and despite the overall oligomeric distribution in the ESI-MS spectra looking similar, SynZur with Co or Ni at the sensory site has a reduced ability to bind to *pznuABC*. It is unknown whether this is due to subtle changes in protein conformation or something else and would need further investigation.

A different picture was observed for copper. Despite it being added at the same double excess ratio, additional peaks of over-metallated SynZurCu species were observed in the ESI-MS spectrum (see Table 3.9). Moreover, the overall distribution of the peaks looked different: visually, the ratio of monomer to dimer decreased. It is possible that a small amount of SynZur is oxidized by Cu²⁺ causing the metal to be in larger excess. However, no intense peaks for the oxidized species were observed for the dimer. Nevertheless, upon investigating the monomer peaks (see Figure 3.30), two peaks were found corresponding to MW 15458.95 Da and 15523.90 Da. Upon subtraction of the MW of apo-SynZur, this gave a difference of -7.57 and 57.38 which could correspond to Zur_{-8H/-7H} and ZurZn_{-6H} (or ZurCu¹⁺_{-5H}/ ZurCu²⁺_{-4H}), respectively. Comparison of the found species with calculated MW is given in Table 3.10.

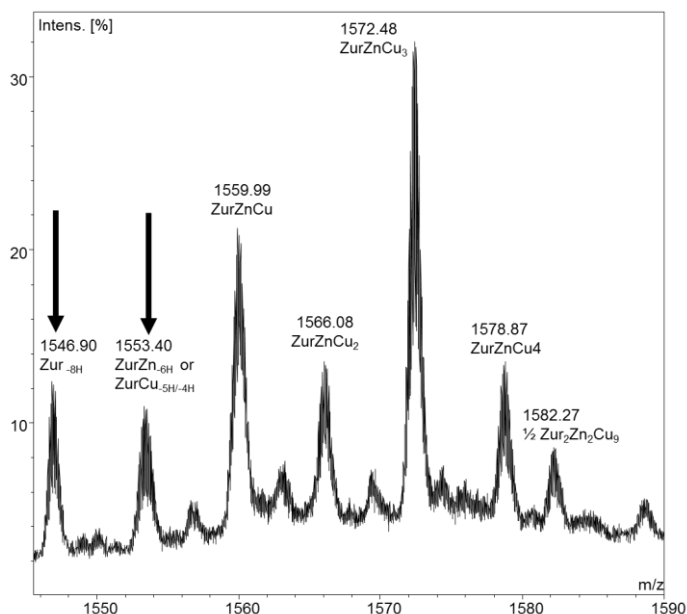


Figure 3.30 ESI-MS spectrum of the +10 charge state of ZurZn₁ in native conditions with Cu²⁺ added.

Two peaks corresponding to oxidized species are marked with black arrows. The full range spectrum is shown in Figure 3.29.

Table 3.10 List of Zur-species oxidized by Cu²⁺.

| MW of a peak with z = +10 | Found MW | Corresponding species | Calculated MW | Difference |
|---------------------------|----------|---|---------------|---------------|
| 1546.90 | 15458.95 | Zur-. ₈ H (Zur-. ₇ H) | 15458.37(+1) | -0.43 (+0.57) |
| 1553.40 | 15523.90 | ZurZn-. ₆ H | 15523.76 | -0.14 |
| | | ZurCu ¹⁺ -. ₅ H | 15523.93 | 0.03 |
| | | ZurCu ²⁺ -. ₄ H | 15523.93 | 0.03 |

The apparent decrease in the intensities of the monomer charge state envelope was due to the distribution of the monomer between non-oxidized and oxidized peaks in comparison to the dimer, where no intense oxidized peaks were observed (see Table 3.10). Cu²⁺ apparently oxidized some or all of the cysteines in the SynZur protein leading to the formation of disulfide bonds and apparently misfolding since no dimer was observed for the oxidized species. Interestingly, SynZur contains only 7 cysteines and it is unclear how Zur-.₈H species could be formed. Alternatively for Zur-.₇H, even if all the cysteines are oxidized, it is unclear how an odd number of oxidized Cys residues

can form disulfide bridges within the monomer. Possibly, other residues are involved. Although these observations are interesting from a chemical perspective, it should be acknowledged that *in vivo* where reducing conditions exist inside cells, copper would be Cu^+ which would behave differently upon SynZur binding. Thus, this experiment would need to be repeated in the presence of a reducing agent to reduce Cu^{2+} to Cu^+ while being compatible with the ESI-MS machine. This was not done in the current study. Since for the metal-selective EMSA experiments a 10-fold excess of Cu^{2+} was used, SynZur was likely oxidized to a greater extent and no binding was observed, possibly due to misfolding.

Finally, upon addition of cadmium to Zur_2Zn_2 , a surprising range of peaks was observed corresponding to Zur_2Zn_2 , Zur_2Zn_3 , $\text{Zur}_2\text{Zn}_2\text{Cd}$, Zur_2Zn_4 , $\text{Zur}_2\text{Zn}_3\text{Cd}$, $\text{Zur}_2\text{Zn}_2\text{Cd}_2$, $\text{Zur}_2\text{ZnCd}_3$, and importantly, Zur_2Cd_4 (see Table 3.9). It was noticed that a re-distribution of cadmium from the sensory site took place and led to the formation of some ZurCd_2 and ZurZn_2 suggesting that cadmium can replace zinc in the structural Cys_4 site. Thus, during EMSAs, upon addition of 10-fold excess cadmium, the main species formed was probably Zur_2Cd_4 (and its over-metallated species). Cd^{2+} has a larger ionic radius compared to Zn^{2+} and such replacement at the structural site might affect the conformation of SynZur leading to a much weaker DNA-binding which was indeed seen by EMSAs where only a smeary band was observed.

3.3.3 Crystal structure

To assess whether the structure of SynZur is similar to other FFPs, such as *E. coli* Zur, crystallization trials were set up using the hanging drop method as described in section 3.2.3. Similar to SynZur, EcZur also binds 2 zinc ions per monomer⁴⁸ (see sections 3.3.2.1 and 3.3.2.2.). However, the sensory zinc site was predicted to be different to that of EcZur³³. Thus, it was important to characterize SynZur structurally.

Two to three weeks after setting up the plates (see section 3.2.3.), the first defined crystals appeared. Large enough crystals for X-ray diffraction analysis were formed in four of the conditions used. Images of these four crystals, and the conditions used, are shown in Figure 3.31. They all were checked at the Diamond Light Source synchrotron for their X-ray diffraction patterns.

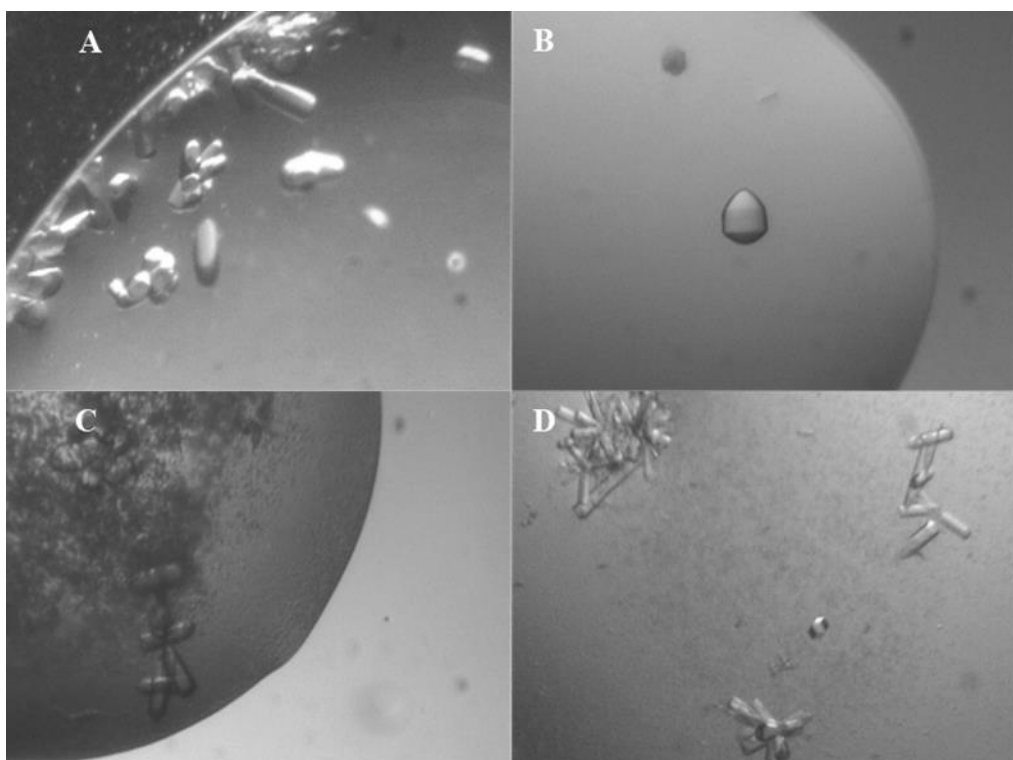


Figure 3.31 Microscopy of SynZur crystals obtained during crystallization trials.

A: The MIDAS-HT96 screen, C3 (0.1M Sodium malonate dibasic monohydrate, 0.1 M MES, 25% Jeffamine® SD-2001, pH 5.5). **B:** The ProPlex-HT96 D5 screen (0.1M MgCl₂, 0.1 M MES, 8% (w/v) PEG6000). **C:** The ProPlex-HT96 F3 screen (0.1M Mg(CH₃COO)₂, 0.1 M MES, 10% (w/v) PEG10000). **D:** The ProPlex-HT96 H8 screen (0.05M Ca(CH₃COO)₂, 0.1 M Sodium cacodylate, 25% (v/v) 2-Methyl-2,4-pentanediol).

All four crystals checked appeared to be protein crystals and the best resolution at around 3 Å was for the ProPlex-HT96 F3 screen (0.1 M Mg(CH₃COO)₂, 0.1 M MES, 10% (w/v) PEG10000). It was attempted to solve the structure by molecular replacement, using Zur structural data from other bacterial species found in the Protein Databank (Zur proteins from *E. coli* (4MTD), *Mycobacterium tuberculosis* (2O03) and *Streptomyces coelicolor* (3MWM)), as well as using a previously generated homology model for *Prochlorococcus marinus* Zur³³. Unfortunately, based on those data it was not possible to solve the structure.

After optimisation of the ProPlex-HT96 F3 conditions, we were able to obtain crystals which allowed structural determination. The conditions used were: 0.1 M Mg(CH₃COO)₂, 0.1 M MES pH 6.0, 16% (w/v) PEG 10000 at 4 °C.

The crystal structure of *Synechococcus* sp. WH8102 Zur was subsequently resolved by Dr. Rachael Wilkinson and Prof. Vilmos Fülöp. The resolution of the structure is 2.24 Å. Although no zinc was added during culture growth, purification or

crystallization stages, SynZur was found to contain zinc as shown previously by ICP-OES. SynZur structural information was provided by Dr. Rachael Wilkinson and is shown in the Appendix (see Table 8.1).

The asymmetric unit of the crystal with the space group $P6_5$ contains four protein molecules (see Figure 3.32A). SynZur crystallized as a zinc-bridged dimer of dimers (see Figure 3.34) which was in agreement with the conclusion made in section 3.3.2.4. Whether this dimer of dimers is relevant *in vivo* is unknown. However, it was shown in section 3.3.2.7.1 that SynZur binds *pznuABC* as a single dimer while for *pbmtA* two dimers are bound per DNA molecule (see section 5.3.1.).

A homodimer, together with its zinc binding sites, is shown in Figure 3.32B.

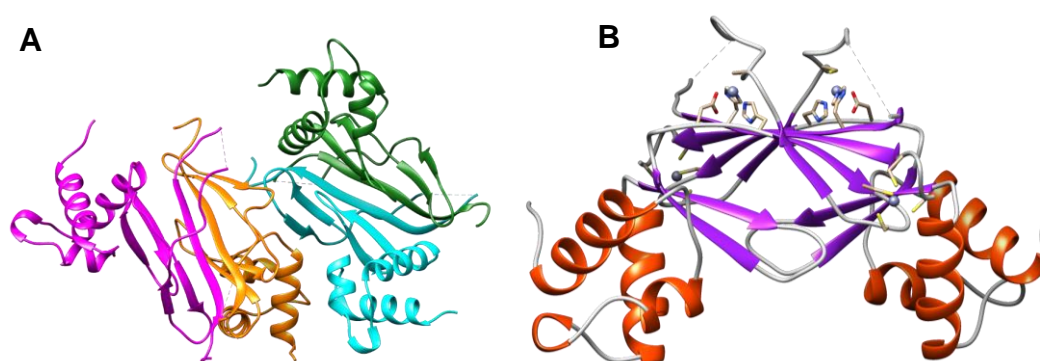


Figure 3.32 The SynZur crystal structure.

A. The asymmetric unit of SynZur with two dimers. Different chains are presented with different colours: A – orange, B – blue, C – pink, D – green. A+C and B+D constitute the two dimers. **B.** A side view image of the SynZur homodimer coloured by secondary structure features: β -sheets are coloured with purple and α -helices are red. Zinc ions are shown as grey balls.

Dimers within the asymmetric unit differ slightly. Structural alignment of the dimers over each other (chains AC over BD) using Swiss PDB Viewer (Magic fit) gives an RMSD value of 1.51 Å over 904 backbone atoms. Values for comparison of different chains separately are shown in Table 3.11. It is unclear if this variation in conformation is relevant for DNA-binding or an artefact of crystallisation. It does however indicate that structural dynamics may be of relevance in the protein's activity.

Table 3.11 Superposition of SynZur monomers within its asymmetric unit.

Since A and C units are completely identical, chains of B-D are compared only with chain A and each other. All calculations were done using Swiss PDB viewer software, Magic fit.

| Reference chain | Superposed chain | Number of backbone atoms | RMSD value |
|-----------------|------------------|--------------------------|------------|
| A | B | 448 | 1.22 Å |
| A | C | 448 | 0 Å |
| A | D | 448 | 0.89 Å |
| B | D | 452 | 0.97 Å |

Each monomer is divided into two domains: a DNA-binding domain (DBD) (~residues 7-74) and a dimerization domain (DD) (~residues 75-126) as shown in Figure 3.33.

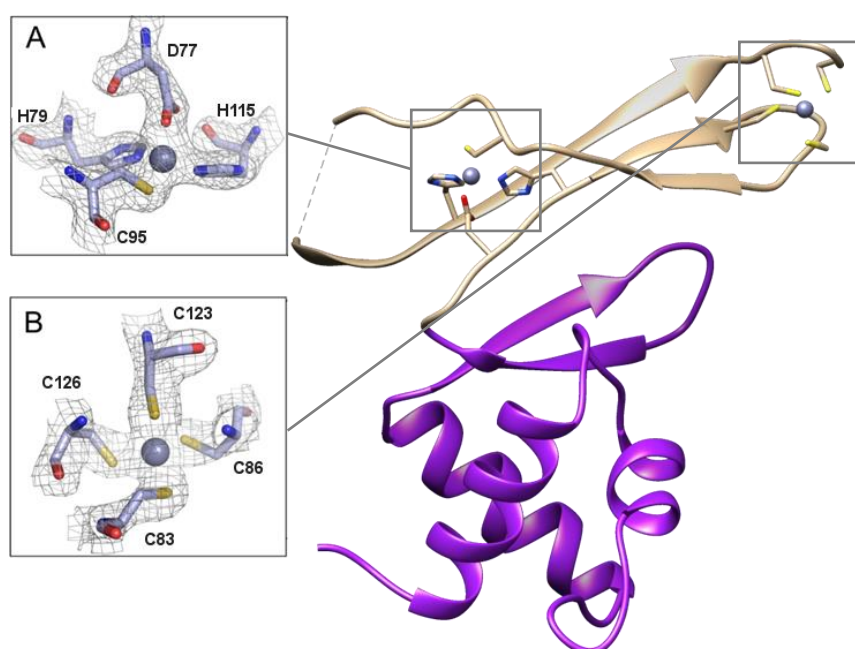


Figure 3.33 Structural domains of the SynZur monomer and zinc binding sites.

Domains are shown in different colours: DBD - purple, DD - tan. Electron density maps were provided by Dr. Rachael Wilkinson. Site A is the sensory site and site B is the structural site. Both binding sites were previously predicted³³. Zinc ions are shown as grey balls.

The DBD contains three α -helices and 2 β -sheets and the DD has 3 β -sheets (see Figure 3.33). Each monomer contains two zinc binding sites within the DD. The sensory zinc ion is formed by His79, His115, Cys95 and Asp77. The structural zinc is bound by four cysteine residues (Cys83, Cys86, Cys123 and Cys126) stemming from

two CXXC motifs. All the metal-to-ligand distances are within the expected range for zinc ions¹²⁹ (see Table 3.12).

Table 3.12 Metal distances in the crystal structure of SynZur at the two zinc sites.

| Structural Site | (Å) | Sensory Site | (Å) |
|-----------------|-------------|--------------|-------------|
| Zn-S Cys-83 | 2.35 ± 0.06 | Zn-O Asp-77 | 1.98 ± 0.05 |
| Zn-S Cys-86 | 2.30 ± 0.00 | Zn-N His-79 | 2.10 ± 0.00 |
| Zn-S Cys-123 | 2.35 ± 0.06 | Zn-S Cys-95 | 2.33 ± 0.05 |
| Zn-S Cys-126 | 2.33 ± 0.05 | Zn-N His-115 | 2.08 ± 0.05 |

In addition to 8 zinc ions in the asymmetric unit, one more symmetry related zinc ion was found between chains C and D, suggested to be important for packing during crystallization. This site comprises His94 and His98 on each of the respective chains and is shown in Figure 3.34. Since no hydrogen bonds were observed by PISA analysis between the dimers (chains A and B are adjusted, see Table 8.1 in the Appendix), it is possible that this zinc site is responsible for tetramer formation. To judge whether this His₄ site is important *in vivo*, it would be interesting to assess its zinc affinity to understand if it is biologically relevant.

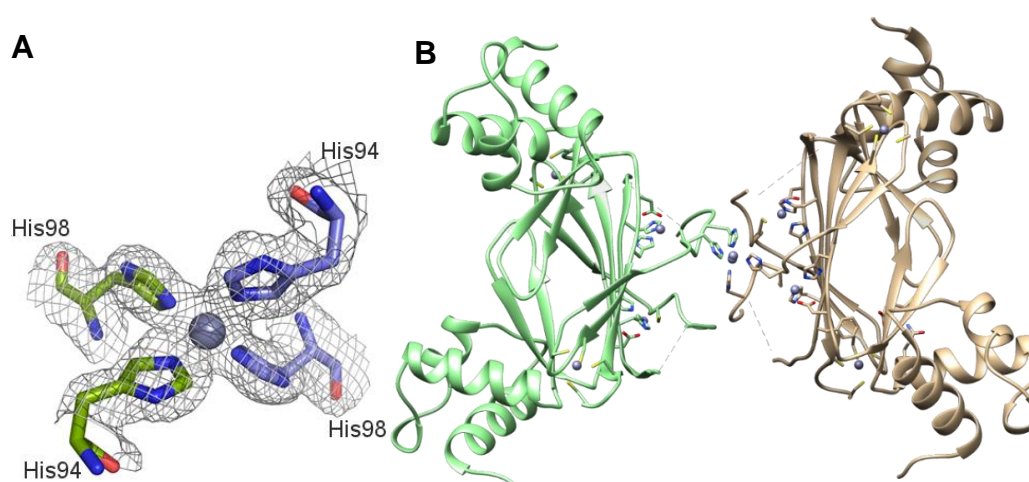


Figure 3.34 The His₄ zinc binding site in SynZur.

A. Symmetry-related zinc site residues with the electron density map (σ level 1.5) displayed as a mesh. Residues from chain D are depicted as light blue sticks and residues from symmetry mate chain C are depicted as light green sticks. The image was provided by Dr. Rachael Wilkinson. **B.** An alternative arrangement of the tetramer if the His₄ site is centred. Zinc ions are shown as grey balls.

A similar observation has been made previously in the structure of *Pseudomonas aeruginosa* Fur with symmetry-related zinc ions observed between subunits A-B or C-D co-ordinated by H76 and D73 from one subunit and E36 of the other subunit⁹³. However, those residues were located in the DBD of the protein unlike SynZur, where the site is between two DD domains.

If the tetramer is formed as a consequence of the presence of additional zinc, this could explain why under native conditions a tetramer was observed (see section 3.3.2.4.), although the origin of zinc impurities is unclear in either case. The presence of small amounts of SDS during ND SDS-PAGE might have changed SynZur folding and affected zinc binding in potentially the weakest His₄ site.

Potentially this additional zinc site between chains C and D could play a role *in vivo* being a 3rd binding site in SynZur. Previously, Choi *et al.*, (2017) showed that ScZur might bind DNA as larger oligomers under higher zinc concentrations causing transcriptional activation of a zinc exporter⁷³. This would be in agreement with our data: under higher zinc concentrations two Zur dimers could bind an additional zinc forming a tetramer which then might act as an activator (see sections 4.3.3.2.2 and 5.3.1). However, experimental evidence would be required to confirm this.

3.3.3.1 Comparison of the SynZur structure with Fur-family proteins from other bacteria

Simple structural alignment (Swiss PDB viewer, Magic Fit) of a dimer of SynZur with known structures of Zurs and other FFPs suggests very little similarity (Table 3.13). However, superposition of single domains gives a much higher similarity. Superposition of full-length monomers was impossible since Swiss PDB viewer was aligning them by single domains if Magic Fit was used.

Table 3.13 Structural alignment of SynZur with other FFP members.

The first number gives the number of atoms involved, the second the RMSD.

| PDB code | Organism | Protein name | Full-length Dimer | DBD | DD |
|-----------------|--|---------------------|--------------------------|-------------|-------------|
| 3mwm | <i>Streptomyces coelicolor</i> | Zur | 892, 20.98 Å | 252, 1.20 Å | 148, 1.43 Å |
| 4mtd | <i>E. coli</i> | Zur | 188, 13.33 Å | 236, 1.22 Å | 140, 1.33 Å |
| 3f8n | <i>E. coli</i> | PerR | 904, 8.34 Å | 244, 1.29 Å | 296, 1.28 Å |
| 2o03 | <i>Mycobacterium tuberculosis</i> | Zur | (monomer) | 248, 1.37 Å | 132, 1.16 Å |
| 1mzb | <i>Pseudomonas aeruginosa</i> | Fur | (monomer) | 300, 1.31 Å | 124, 1.21 Å |
| 4ets | <i>Campylobacter jejuni</i> | Fur | 440, 14.19 Å | 200, 1.26 Å | 144, 1.27 Å |
| 3eyy | <i>Streptomyces coelicolor</i> | Nur | 768, 5.74 Å | 268, 1.22 Å | 152, 1.82 Å |
| 5fd6 | <i>Rhizobium leguminosarum</i> <i>bv. viciae</i> | Mur | 828, 9.00 Å | 252, 1.27 Å | 127, 1.38 Å |

It was noticed that individual domains (DBD or DD) of SynZur and other proteins shown in Table 3.13 could be aligned very well, but that the orientation of those two domains towards each other in SynZur differs considerably from that of other Zur proteins. For example, superposition of different domains of SynZur with 3mwm (*Streptomyces coelicolor* Zur) is shown in Figure 3.35.

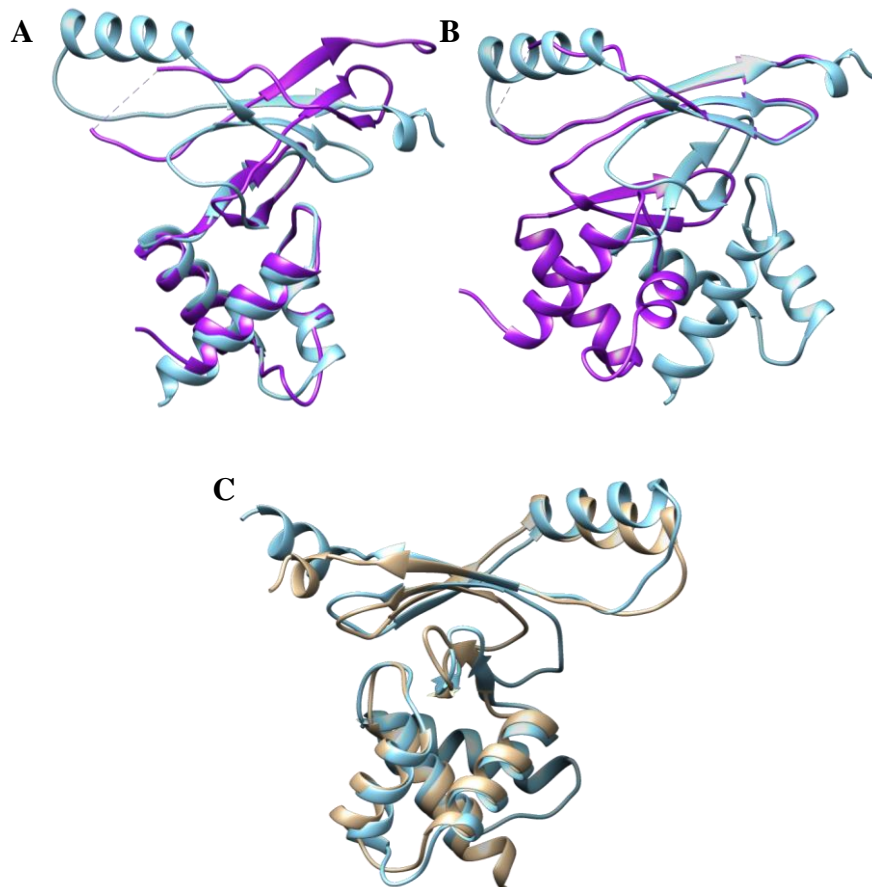


Figure 3.35 Structural alignment of SynZur and ScZur domains.

The alignment was performed using Swiss PDB viewer, Domain Alternate Fit and visualised with Chimera software. RMSD values are given in Table 3.13. A. Superposition of SynZur (purple) and ScZur (light-blue) by DBD. B. Superposition of SynZur (purple) and ScZur (light-blue) by DD. C. Example of superposition of ScZur (tan) and EcZur (light-blue) (RMSD is 1.55 Å over 380 backbone atoms). The orientation of domains in those two is similar, but unlike that in SynZur.

It is possible that the shorter hinge (Asp75-Arg79) between β -sheets of the DBD and DD in SynZur compared to other FFP might affect the orientation of SynZur domains towards each other. For comparison, in EcZur it is 9 amino acids long (His 89-Ser98) and in ScZur it is 8 residues long (Arg78-His85). However, in MtZur and PaFur the hinge is only 5 amino acids (Ser77-His81 and Asp83-His87, respectively) long. Nonetheless, the orientation of the DD and DBD is different between SynZur and MtZur or SynZur and PaFur. However, the PaFur monomer showed the most similarity to SynZur from the point of view of orientation of the domains (see the Appendix, Figure 8.6). Thus, it is possible that hinge length affects the orientation of the DD to the DBD.

Many other Fur family regulators appear to have an additional α -helix at the beginning of the DBD (EcZur, PaFur, CjFur, BsPerR, ScNur, RlMur). In the SynZur


```

SynZur 1 -----MTGSSPALNARQALLTALNACGD-EMSGQLHRSLD--DE-ASMGLATVYRNLRLQLQQRGL 58
MtZur 1 MSAAGV-----RSTRQRAAISLLETLDLDD-FRSAQELHDELRR--RRGENIGLTTVYRTLQSMASSSL 59
ScZur 1 MTTAGPPVK-----GRATRQRAAVSAAALQVEVEE-FRSAQELHDMK--HKGDVGLTTVYRTLQSLADAGE 63
BsPerR 1 MAAHEL-----KEALETLKETGVRI TPQRHAILEYLVNSMA-PTADDIYKALE--GKFPNMSVATVYNNLRVFRSGL 71
PaFur 1 -----GSMVENS--ELR-KAGLKVTLPRVKILQMLDSAEQFMSAEDVYKALM--EAGEDVGLATVYRVLTFQFEAAGL 68
CjFur 1 ATYAMM[ 8] VLLERFKKILRQGGKLYTKQREVLLKTLYHSDT-HYTPESLYMEIKqaEPDLNVGIATVYRTLNLLEEAEM 84
RlMur 1 -----MTDVAKTLEELATERGMRMTEQRRVIARILEDSED--PDVEELYRRSV--KVDAKISISTVYRVTVKLFEDAGI 70
ScNur 1 -----MVSTDWKS DLRQRGYRLTPQRQLVLEAVDTLE--RATPDDILGEVR--KTASGINISTVYRTELLEELGL 67
EcZur 1 MEKTTTQELLAQAEEKICAQRNVRLTPQRLEVLRLMSLQD-GAISAYDLDLLR--EAEPQAKPPPTVYRALDFLLEQGF 75

SynZur 59 VRCRHLPTGEALYAPVDR--DRHHL-TCVDCGTQVLLPPIIGIDVVPADSRGDFELLEHTLEFFFGCSSRQPRSSkp- 134
MtZur 60 VTLHTDTGESVYRRSE-HHRL-VCRSCGSTIEVGDHEVAWAAEVATKHGFSVDSHTIEIFGTCSDCRS----- 130
ScZur 64 VVLRTAEGESVYRRSTGDHHL-VCRACGKAVEVEGPAVKWAEAIAAEHGYVNVAVTVEIFGTCADQAGASGG--- 139
BsPerR 72 VKELTYGDASSRFVTS--DLYSA-ICENCCKIVPHYPGLDEVEQLAAHVTGFKVSHHRLIYGVQCEQSKKENH--- 145
PaFur 69 VVRHNFDDGGHAFELADS-GHLLM-VCVDTGCVIEMDAEIKRQKEIVRERGFELVDNLVLY-VRK---KK----- 136
CjFur 85 VTSISFGSAGKKYELANK-PHHHM-ICKNCGKIIIEFENPIIRQQALIAKEHGFKLTGLMLQLYGVCGDNNQKAKvki 162
RlMur 71 IARHDFRDGRSRYTVPE-EHML-IDLKTGTVIFRSPEIALQERIAREHGFRVLDHRLLELYGVPL---KKEDL--- 142
ScNur 68 VSAALGHGAPTYHLADR-EHML-VCRDCTNVIADLSVAADF TAKLREQFGFDTMCKFAIFGRCESCSLKGSTtds 145
EcZur 76 VKVESTNSYVLEHFLDQPTFSAMFICDRGAVKCEAEGVEDIMHTLAKMGFALRHNVIEAHGLCAAC-VEVEACR 153

```

Figure 3.37 Sequential alignment of SynZur with other FFPs with known crystal structures using NCBI BLAST¹³¹.

Regions of higher similarity are shown in red font and regions of lower similarity are shown in blue. The software does not allow multiple alignment of proteins with e-values of less than 0.001; therefore EcZur was automatically excluded and aligned manually. Residues responsible for zinc binding are highlighted with grey (Site 1; structural Cys₄ site), red (Site 2; major sensory site) and yellow (Site 3; secondary sensory site). Sites in Fur/Nur/PerR proteins are highlighted in dark and light green (except Site 1 which is grey). Sites are named according to Barnett *et al.*, (2012)³³. Residues for unexpected zinc binding sites in SynZur are highlighted in magenta. The number in brackets represents additional un-align residues.

Table 3.14 Summary of sequential alignment of FFPs using NCBI BLAST¹³¹.

| Organism | Protein name | PDB code | Total score | Query cover | E value | Ident. |
|-----------------------------------|--------------|----------|-------------|-------------|-------------------|--------|
| <i>Mycobacterium tuberculosis</i> | Zur | 2o03 | 69.7 | 89% | 3e ⁻²⁰ | 37% |
| <i>Streptomyces coelicolor</i> | Zur | 3mwm | 52.8 | 94% | 9e ⁻¹⁴ | 29% |
| <i>Bacillus subtilis</i> | PerR | 3f8n | 48.5 | 92% | 5e ⁻¹² | 23% |
| <i>Pseudomonas aeruginosa</i> | Fur | 1mzb | 43.9 | 70% | 2e ⁻¹⁰ | 31% |
| <i>Campylobacter jejuni</i> | Fur | 4ets | 43.9 | 69% | 3e ⁻¹⁰ | 24% |
| <i>Rhizobium leguminosarum</i> | Mur | 5fd6 | 36.2 | 85% | 2e ⁻⁰⁷ | 22% |
| <i>Streptomyces coelicolor</i> | Nur | 3eyy | 33.9 | 93% | 1e ⁻⁰⁶ | 30% |
| <i>E. coli</i> | Zur | 4mtd | 22.7 | 61% | 0.009 | 26% |

From the sequential alignment it was noticed that ScZur (3mwm) and MtZur (2o03) have histidine-rich stretches containing 4 and 5 histidine residues in a row

which are parts of sites 2 and 3 as shown previously in ³³. However, SynZur has only two histidine residues there.

Also, based on the sequential alignment (see Figure 3.37), some metal binding sites are clearly conserved in FFPs. If their structures are aligned by the dimerization domain, as shown in Figure 3.38, positions of the metal binding sites are also conserved in these structures.

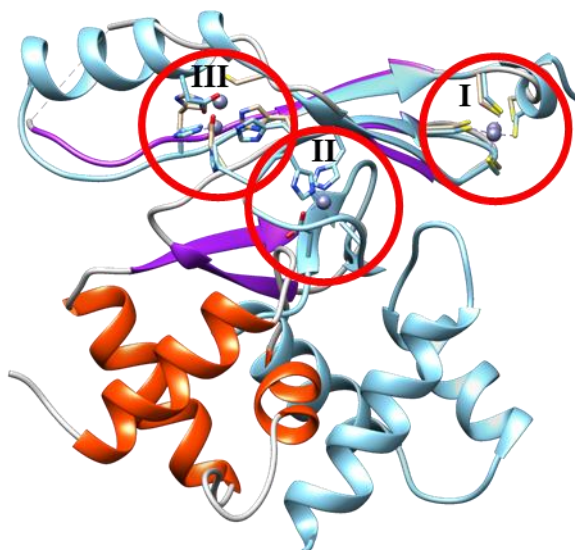


Figure 3.38 Superposition of metal binding sites by alignment of the DD in SynZur and 3mwm (*S. coelicolor* Zur).

SynZur is shown with its secondary structure (α -helices are red, β -sheets are purple, and random coils are shown in grey) and 3mwm is shown in light blue. Sites are named according to³³. Zinc ions are shown as grey balls.

Thus, the structure of *S. coelicolor* Zur showed that binding sites could be organized into three groups: structural site 1 in the DD, main sensory interdomain site 2 and a tuning site 3, located also in the DD⁴⁷. As can be seen from Figure 3.38, the location and composition of site 1 in SynZur are matching those for ScZur. The location of site 3 (the sensory site in SynZur) is matching that in ScZur, but one of the His residues in ScZur is replaced with a Cys in SynZur. Site 2 in ScZur – thought to be the main sensory site in *S. coelicolor* Zur – is located between the DBD and the DD, but is absent in SynZur.

In addition to the inter-domain hinge length, it is possible that the lack of site 2 between the domains in SynZur causes a different orientation of the domains in SynZur compared to ScZur and other FFPs making it not align properly (Figure 3.35).

Thus, it was decided to compare the presence of different metal binding sites in FFPs. This data is summarized in Table 3.15.

Table 3.15 Comparison of metal-binding sites in different FFP members with SynZur.

SynZur is highlighted in grey. *Campylobacter jejuni* Fur is highlighted in bold font because its structure contains zinc ions at the same sites as SynZur.

| PDB code | Protein name | Site 1 and coordination | Site 2 and coordination | Site 3 and coordination |
|-------------|--------------|-------------------------|-------------------------|---|
| SynZur | Zur | 4 Cys (4) | - | 2 His, 1 Cys, 1 Asp (4) |
| 3mwm | Zur | 4 Cys (4) | 2 His, 1 Cys, 1 Asp (4) | 3 His, 1 Glu (4) |
| 4mtd | Zur | 4 Cys (4) | 2 His, 1 Cys, 1 Gln (4) | - |
| 2o03 | Zur | 4 Cys (4) | 2 His, 1 Cys, 1 Asp (4) | 3 His, 1 Glu (4) |
| 1mzb | Fur | - | 2 His, 2 Gln (4) | 2 his, 1 Glu, 1 Asp, water (6) |
| 4ets | Fur | 4 Cys (4) | - | 1 His, 1 Asp, 1 Glu, 2 water (6) |
| 3f8n | perR | 4 Cys (4) | Mn, 3 His, 2 Asp, O (6) | - |
| 3eyy | Nur | - | 4 His (4) | Ni, 3 His, 3 O (6) (different location) |
| 5fd6 | Mur | - | 3 His, 1 Gln (4) | 2 His, 1 Glu, 1 Asp (4 or 6, see Figure 8.7.) |

It can be seen from Table 3.15 that EcZur despite possessing only two zinc binding sites like SynZur, has its sensory site located at site 2 unlike SynZur which has site 3 as the sensory site. A few other FFPs also have only 2 metal binding sites (PaFur, ScNur and RlMur). However, they lack the otherwise highly conserved structural Cys₄ site 1. Finally, Fur from *Campylobacter jejuni* (4ets) has two zinc ions bound at the same metal binding sites as SynZur in the structure. However, it was crystallised without a metal ion in sensory site 2, even though its residues are present⁹⁴. In addition, the DBD of CjFur is rotated by 180° in comparison to other FFPs, and the orientation of domains in this protein is completely different (Figure 3.39).

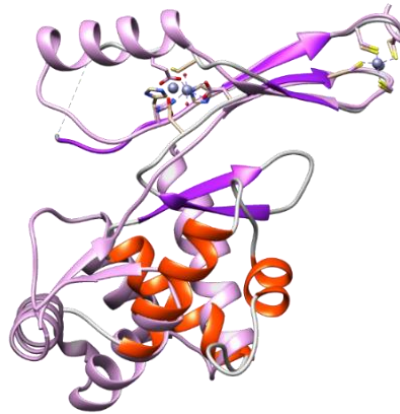


Figure 3.39 Superposition of SynZur and CjFur.

SynZur is shown with its secondary structure (α -helices are red, β -sheets are purple, and random coils are shown in grey) and CjFur is in pink. Zinc ions are shown as grey balls.

Thus, it is impossible to conclude if the lack of a metal binding site 2 affects domain orientation. Since no other FFP has currently been found without site 2 apart from SynZur making it very unique.

Based on sequential and structural alignments, it can be concluded that:

1. The structural Cys₄ site 1 is conserved in most FFPs, including SynZur, but not all.
2. Only the sensory site of Zur proteins (SynZur, EcZur, MtZur and ScZur) have Cys residues in them. All have the same tetrahedral coordination: 2N, O, S and usually include 2 histidine residues, one cysteine and one aspartate or glutamate residue.
3. ScZur (3mwm) and MtZur (2o03) have histidine-rich stretches containing 4 and 5 histidine residues in a row which are part of sites 2 and 3. However, in SynZur this region is shorter and has only two His residues.
4. Apart from SynZur, no other FFPs have a cysteine residue at site 3, and this site comprises residues located differently in the sequence compared to site 3 residues from other FFPs.
5. Other FFPs do not have conserved histidine residues 94 and 98 and formation of a tetramer through them could be a coincidence for SynZur.

Based on the information above, it was decided to compare Zur proteins from other cyanobacteria to understand if the unique properties of SynZur are exclusive to *Synechococcus* sp. WH8102 or whether they are common amongst cyanobacteria.

Since no other cyanobacterial Zur protein crystal structure is available only sequential alignment is possible (Figure 3.40). Annotated Zur genes for cyanobacteria were taken from the RegPrecise³⁸ database. The NCBI BLAST¹³¹ protein alignment summary is given in Table 3.16.

Table 3.16 Summary of sequential alignment of annotated Zur proteins from cyanobacteria using NCBI BLAST¹³¹.

The genes are taken from the RegPrecise³⁸ database.

| Name | Total score | Query cover | E- value | Ident. |
|--|-------------|-------------|-------------------|--------|
| <i>Prochlorococcus marinus</i> str. MIT 9313 | 172 | 92% | 1e ⁻⁶⁰ | 66% |
| <i>Nostoc</i> sp. PCC 7120 | 108 | 91% | 3e ⁻³⁵ | 41% |
| <i>Microcystis aeruginosa</i> NIES-843 | 107 | 87% | 1e ⁻³⁴ | 42% |
| <i>Cyanothece</i> sp. PCC 7425 | 105 | 86% | 3e ⁻³⁴ | 44% |
| <i>Synechococcus elongatus</i> PCC 7942 | 103 | 89% | 2e ⁻³³ | 40% |
| <i>Cyanothece</i> sp. ATCC 51142 | 100 | 85% | 4e ⁻³² | 41% |
| <i>Cyanothece</i> sp. PCC 8801 | 100 | 86% | 5e ⁻³² | 40% |
| <i>Thermosynechococcus elongatus</i> BP-1 | 99.8 | 91% | 7e ⁻³² | 41% |
| <i>Synechococcus</i> sp. PCC 7002 | 90.9 | 93% | 3e ⁻²⁸ | 37% |
| <i>Trichodesmium erythraeum</i> IMS101 | 89.7 | 85% | 7e ⁻²⁸ | 34% |
| <i>Synechocystis</i> sp. PCC 6803 | 88.2 | 89% | 3e ⁻²⁷ | 36% |
| <i>Synechococcus</i> sp. JA-3-3Ab | 71.6 | 91% | 1e ⁻²⁰ | 32% |
| <i>Gloeobacter violaceus</i> PCC 7421 | 68.6 | 89% | 1e ⁻¹⁹ | 32% |

| | | | |
|--|----|--------|--|
| Synechococcus sp. WH 8102 | 1 | MT [2] | -SPALNARQQAALLTALNACGDDEMSGQQLHRSIDDEA-SMGLAATVYRNLRQLQQRGLVRCRHLPTGEALYAPVD--75 |
| Prochlorococcus marinus str. MIT 9313 | 1 | MP [2] | -QSQPTLRQNQLLKELRCCCEDEMSGQQLYRALQDSPYAMGLATVYRNLRQLQQRGLVRCRHLPTGEALYAPVE--76 |
| Nostoc sp. PCC 7120 | 1 | -- | MRALFTRSSQERILNKLQTTIKQGGISAQDLYVELLRNRNDSMGGLATVYRSLEALKLEGLVQVRLTPEGEALYSLAQ--73 |
| Microcystis aeruginosa NIES-843 | 1 | MK | -SQR-TRSSQERIIHLKTLNRAYSAQDLYVELLRQDPMGLATVYRSLEALKLQGAQVQVRLTAGESLYSSVQ--73 |
| Cyanothece sp. PCC 7425 | 1 | -- | MKAQRTSSQERIKLQVLLQTLNRAISAQDLYELLRQDQPMGLATVYRSLEALKLGGVQVQARLIPNGEALYSVQ--73 |
| Synechococcus elongatus PCC 7942 | 1 | -- | MARRTONQKILLEVLLQGLKQETSAQDLYELLRQDQPMGLATVYRSLEALKLGGVQVRLTSGESLYSVHR--73 |
| Cyanothece sp. ATCC 51142 | 1 | -- | MK--RTRSQANIVHVLKSLSHPISAQDLYIEMKNRDLGGLATVYRSLEALKLIDGEGVQVRLPENGAVVSVLH--71 |
| Cyanothece sp. PCC 8801 | 1 | -- | MGQRTSSQERIVRLKSLNHAVSAQDLYLEIKTRDQALGLATVYRSLEALKLEGTQVQARTLNGEAVVSVLVA--73 |
| Thermosynechococcus elongatus BP-1 | 1 | MP | -PRR-TRSSQAILDALRASGRSLSAQDLYELLRQRTPLGLATVYRTLDRIRHGQVQARPLPSGELVYSLVQ--73 |
| Synechococcus sp. PCC 7002 | 1 | M- | -SKLTKNQRTILQLIQAQETESAQNLMHFEMQQTGLKIGLATVYRAKLLKVEGKIQRVYDEGESFYSKIGmt75 |
| Trichodesmium erythraeum IMS101 | 1 | -- | MKTRTHSQERILKLNKQKRSLSAQDLYIKLIRYDRPLGLATVYRSLEALKLEKGGVIVRLTVNGESVYS [LQ--73 |
| Synechocystis sp. PCC 6803 | 1 | MS [9] | LESLTVNQRL-VLQALQRETEPLSAQALFAKIRE-TKKIGLAATVYRALDALKLAGFIQHQAATMNGELLLYQTLE--82 |
| Synechococcus sp. JA-3-3Ab | 1 | MS | -SSRQTRSSQKLLILEVLKQSLRPLSAQEIFLELRNANHSVGLATVYRALLEALIRNEGKLAQAVNLGDQQTYYQVLPsn76 |
| Gloeobacter violaceus PCC 7421 | 1 | MP | -ARKLTKGQQAVALLEALASSKRPLCAQDDIYLLQLRGTGEQEVGLATVYRSLEALLADDDRIQLIDVYRDNQAHY-LMgra75 |
| Synechococcus sp. WH 8102 | 76 | -RDR | HLTGVDCGTTQVLDHCPTEFGIDYPAADSRGDFELLFHTLEFFFGFCSSCRPQRSSK [1] 134 |
| Prochlorococcus marinus str. MIT 9313 | 77 | -RDL | HLTGVDCGKTLPIEHCPLFNLNIPQEQSHNFQLIFHTLEFFFGHCESCQKDKQTLT 134 |
| Nostoc sp. PCC 7120 | 74 | -QDK | HLTGVDCGVSIPIHQCPVFNLEBQLQTAHKFKIFHTLEFFFGCGKQMNHASE [1] 132 |
| Microcystis aeruginosa NIES-843 | 74 | -QDQ | HLTGVDCGRSIVFNECPVFELEBKELEKSHQFKVYHTLEFFFGCDRCQPS--- 127 |
| Cyanothece sp. PCC 7425 | 74 | -EDR | HLTGVDCGQSAIDCPVFELESQLRQSHFEIYHTLEFFFGISPCQAGLPSG [1] 132 |
| Synechococcus elongatus PCC 7942 | 74 | -EDR | HLTGVDCRNIISAIDCPVFALEQDLQQA YFGIFYHTLEFFFGICQTCQAS--KA [1] 130 |
| Cyanothece sp. ATCC 51142 | 72 | -QDQ | HLTGVDCGQSEFIDCPVFDLEKKEIISHQFKVYHTLEFFFGICHLCHGGE--- 126 |
| Cyanothece sp. PCC 8801 | 74 | -KDD | HLTGVDCGQSEFIDECPVFDLEQRLKESHRFQVYHTLEFFFGICYDQLSAERN [1] 132 |
| Thermosynechococcus elongatus BP-1 | 74 | -QDQ | HYLTLHCGTSIEIEQCPVQPLETELQECYHFKVYHTLEFFFGICAKQVMAQR- 130 |
| Synechococcus sp. PCC 7002 | 76 | eHHH | HLNLCVNGGQSYPLDSCPISQKLTDMCQTQKFKVYHTLEFFFGICADQQTQADN- [7] 140 |
| Trichodesmium erythraeum IMS101 | 74 | -QDQ | YVNLGVKCGTSSIPINKCPVYNLEDKLQOSYKFKIFHTLEFFFGICERCITLTEDIS 131 |
| Synechocystis sp. PCC 6803 | 83 | -QDQ | HLTGVDCGQSEVPIQCPVQSLLENLQANYSFRIYHTLEFFFGICQLCAKGS- 138 |
| Synechococcus sp. JA-3-3Ab | 77 | qHNR | HLTGVDCRQVVPPLPDCPVGDLEEQLASRYQFVIEYHVLDFYGTCAECRGESSSN [14] 149 |
| Gloeobacter violaceus PCC 7421 | 76 | nHSQ | HLTGVDCRQVVPPLDHCQVPSALEQHLSDHEFQIAYHVLDFYGTGGEQRQAVVGA- 133 |

Figure 3.40 Sequence alignment of Zur proteins from cyanobacteria.

All available genes from the RegPrecise³⁸ database are compared. Site 1 residues are highlighted in grey. Site 3 residues are highlighted in yellow. Residues for unexpected His₄ Zn binding sites potentially responsible for the formation of a tetramer are highlighted in blue. Species inhabiting oligotrophic oceanic waters are shown with bold letters. The unique Cys71 residue in *Trichodesmium erythraeum* IMS101 is highlighted in green. Numbers in brackets represent additional un-align residues.

It can be seen from Figure 3.40 and Table 3.16 that Zur proteins in cyanobacteria are more conserved and show higher overall identity compared to Zur and FFPs from other phyla as might be expected. The highest degree of identity is observed for *Prochlorococcus marinus* str. MIT9313, a sister taxon to marine *Synechococcus* which inhabits a very similar habitat to *Synechococcus* sp. WH8102, namely extremely oligotrophic open ocean conditions³.

Site 1 seems to be completely conserved in all cyanobacterial species compared and sensory site 3 containing Cys is very highly conserved amongst almost all of the Zur proteins. No histidine-rich hinge was observed for any species except estuarine *Synechococcus* sp. PCC7002. However, it is unclear if this is due to this protein having three zinc binding sites, similar to ScZur and MtZur.

Conservation of site 3 as a sensory site for the majority of cyanobacteria differs from Zurs and FFPs from other phyla where the sensory site is located at site 2. Zurs from other phyla also have the same sensory site 2 environment (2N, O, S) as site 3 in cyanobacteria. Previously, it was suggested that site 3 in Zurs of other phyla was a tuning site in *S. coelicolor*⁴⁷ or a second structural site in *M. tuberculosis* due to its location in the DD⁴⁹. In both cases, those sites did not contain a cysteine residue in them.

Crucially, site 2 (absent in Zurs of cyanobacteria) is located between the DBD and DD and upon binding zinc would probably allow the FFP dimer to change the orientation of its DBD leading to a more favourable DNA-binding configuration. However, since site 3 in SynZur is just located in the DD, and the overall orientation of domains is different from other FFPs, it is likely that cyanobacterial Zurs utilise a completely new sensing mechanism that is different from all other FFPs.

Interestingly, *Trichodesmium erythraeum* IMS101, *Synechococcus* sp. JA-3-3Ab and *Gloeobacter violaceus* PCC 7421 do not have a fully conserved sensory site 3 (Figure 3.40). While in the last two species there are His residues nearby which could be involved, it is unclear what this site comprises in *Trichodesmium erythraeum* IMS101. Perhaps a unique cysteine residue (Cys71), which is absent in all other species, could be the 4th zinc binding residue potentially giving that sensory site with 2 Cys residues a higher zinc affinity. This has not been reported before for any other Zur proteins.

Finally, two additional zinc binding histidine residues (highlighted in blue in Figure 3.40), potentially responsible for formation of a tetramer, were found to be present in some of the compared sequences, if replacement of one histidine with glutamate or aspartate is considered - but are not fully conserved. It remains unclear whether this plays an actual physiological role. Potentially, replacement of nitrogen by oxygen could lead to weaker binding of zinc causing formation of a tetramer-activator at higher zinc concentrations. Figure 3.40 shows both oligotrophic species (*Synechococcus* sp. WH8102 and *Prochlorococcus marinus* MIT9313) have 2 histidine residues in this location whilst freshwater or estuarine cyanobacteria either do not have that site conserved or have one histidine replaced with with glutamate or aspartate. Due to the fact that freshwater or estuarine cyanobacteria are adjusted to living under potentially higher trace metal concentrations, they might require activation of a zinc exporter at higher intracellular zinc levels. Alternatively, the majority of them have a zinc excess sensor SmtB or other ArsR family regulators, namely *Nostoc* sp. PCC 7120, *Microcystis aeruginosa* NIES-843, *Cyanothece* sp. PCC 7425, *Synechococcus elongatus* PCC 7942, *Cyanothece* sp. PCC 8801, *Thermosynechococcus elongatus* BP-1, *Synechococcus* sp. PCC 7002, *Synechocystis* sp. PCC 6803³⁸ and thus potentially not requiring Zur to be an activator of zinc efflux genes unlike *Synechococcus* sp. WH8102 where no such sensor can be identified.

3.4 CONCLUSIONS AND FUTURE WORK

In this chapter the Zur protein from *Synechococcus* sp. WH8102 was over-expressed in *E. coli*, purified and characterised. Based on the results of this chapter, the following major conclusions can be drawn:

1. SynZur was purified without any addition of zinc at any step of *E. coli* growth, over-expression and purification, yet it contained two molar equivalents of zinc per monomer. This was demonstrated by ICP-OES, ESI-MS and X-ray crystallography. This might be an indication that zinc is highly competitive for binding to SynZur.
2. SynZur was crystallized as the first member of a FFP from cyanobacteria. Two zinc binding sites were identified: a conserved site 1, which is a structural site composed of four cysteine residues, and a novel sensory site at a location equivalent to that of site 3 in *S. coelicolor* Zur⁴⁷.
3. The location of the sensory site in SynZur is unique compared to other known Zur or FFP structures. Whilst the major sensory sites in other FFPs are inter-domain sites (site 2), the sensory site in cyanobacterial Zurs is completely located within the DD (site 3). The site 2 inter-domain location probably leads to an allosteric switch upon zinc binding while the mode of action of site 3 in SynZur is unknown.
4. The sensory site for SynZur adopts tetrahedral geometry (2 histidine residues, 1 aspartate and 1 cysteine) and is well-conserved in cyanobacterial Zurs. This site possesses the same environment as Zur sensory sites from other phyla (2N, O, S) despite having a different location.
5. The sensory site in SynZur has a high affinity for zinc ($K_D = 8.27 \pm 2.50 \times 10^{-13}$ M) which is comparable with Zurs from other organisms.
6. Superposition of SynZur with other FFPs showed high similarity of its DBD and DD but not for the orientation of domains towards each other. Probably, the shorter inter-domain hinge of SynZur might cause this and/or the absence of a sensory site between the domains.
7. SynZur was found to bind to *pznuABC* with a $K_D = 3 \pm 2$ nM per dimer or ~ 10 nM per monomer.

8. SynZur seems to bind *pznuABC* selectively in the presence of zinc despite being able to bind other divalent metals. Furthermore, ESI-MS analysis suggests that a similar oligomeric state is adopted with most divalent metals tested.
9. The structural site does not react with EDTA and it is possible that this site has a higher affinity for zinc. It also might be sterically or kinetically inert as well as restricted electrostatically. ESI-MS revealed that Cd^{2+} can replace Zn^{2+} in the structural site.

All this information suggests that SynZur is a ‘real’ zinc sensor. It possesses unique properties compared to Zur proteins from other phyla, hinting at a potentially novel zinc sensing mechanism that might be characteristic for cyanobacteria. However, it is unclear how the sensing mechanism works in SynZur or, indeed, other cyanobacterial Zurs. However, a decrease in dimer abundance for SynZur without zinc at the sensory site was detected by ESI-MS which might be important for DNA affinity.

Future work:

Crystallisation studies of SynZur with an appropriately designed section of the *pznuABC* promoter might help to understand the DNA-binding mechanism. So far only the *E. coli* Zur structure is available with DNA bound⁴⁸.

Clarification of the sensing mechanism is required. This could be achieved with single site mutagenesis of the residues involved in zinc binding and further investigation of the mutant protein. Also, additional experiments and/or computational analysis on SynZurZn₁ might reveal whether a change in conformation or oligomeric state is at play: analytical SEC, ND SDS-PAGE, X-ray crystal structure analysis and Molecular Dynamics simulations could help to solve this question.

Finally, SynZur was crystallized as a dimer of dimers. It is possible that the tetramer observed by various techniques is formed via a zinc ion bridging two dimers. The histidine residues forming this site are conserved in some cyanobacterial Zurs and it might be a third unique zinc binding site. This requires further investigation in order to understand if this site is physiologically relevant. Finding its zinc affinity might answer this question. Site-directed mutagenesis of one or both of those two histidine residues might clarify if the tetramer is formed in their absence and whether it is physiologically important.

Chapter 4.
Properties of a *zur* knockout mutant of
***Synechococcus* sp. WH8102**

4.1 INTRODUCTION

In order to confirm whether the predicted SynZur protein from *Synechococcus* sp. WH8102 investigated in Chapter 3 is a true Zinc Uptake Regulator, I have constructed a *zur* knockout mutant. Phenotypic characterisation of the mutant and its comparison to wild type (WT) under different zinc levels helped to clarify and prove the role of SynZur in this cyanobacterium as a zinc sensor.

Previously, it has been shown that Zur knock-out mutants (including cyanobacteria) show an elevated sensitivity towards higher zinc concentrations: this was the case for *Xanthomonas campestris*⁷², *Streptococcus suis*⁴⁶, *Anabaena* sp. strain PCC 7120⁵⁹ and *Synechococcus* sp. IU 625¹³². The absence of a zinc sensor, which normally transcriptionally represses expression of zinc uptake components under sufficient zinc, could potentially lead to accumulation of zinc within cells, possibly to toxic levels. It is possible though that some bacteria possess a Zur-independent zinc efflux system, in which case a *zur*-mutant would be able to compensate for this elevated zinc uptake. For example, Δ *zur*-mutants of *Staphylococcus aureus*¹³³ and *Corynebacterium diphtheriae*¹³⁴ did not show an increased sensitivity to zinc compared to the WT. Furthermore, *Staphylococcus aureus* possesses an independent zinc resistance sensor, CzrA, of the ArsR regulator family³⁸ while *Corynebacterium diphtheria* contains Znr, another transcription factor of the ArsR regulator family which itself is predicted to regulate transcription of *zur*, but not the opposite³⁸. This could also be a potential reason why *zur*-mutants of these species do not show sensitivity towards elevated zinc concentrations. In contrast, no genes relating to any zinc efflux mechanism have been identified in *Synechococcus* sp. WH8102, and given the oligotrophic nature of this strain it may simply not possess such a system. This suggests that a *Synechococcus* sp. WH8102 *zur* mutant might indeed show defects in metal uptake and hence metal content, as well as differences in zinc tolerance.

Construction of a *Synechococcus* sp. WH8102 *zur*-mutant will also help to identify components of the SynZur-regulon, since in the absence of Zur the SynZur-regulon would be de-repressed even at sufficient zinc levels. Such an approach has been used to identify members of the Zur-regulon of other bacteria, including cyanobacteria e.g. in *Anabaena* sp. strain PCC 7120⁵⁹ and *Synechococcus* sp. PCC 7002⁵⁴.

Interestingly, only four genes were predicted bioinformatically by the RegPrecise database to be regulated by SynZur in *Synechococcus* sp. WH8102 (three *znuABC* genes and *bmtA*)³⁸ compared to 23 Zur-regulated genes in *Anabaena* sp. strain PCC 7120⁵⁹. Thus, construction of the *zur*-mutant should help to identify new members of the Zur-regulon in *Synechococcus* sp. WH8102, if such are present.

Herein, I describe the construction and phenotypic characterisation of a *Synechococcus* sp. WH8102 *zur* mutant. Phenotypic characterisation compared to the WT included UV-vis absorption properties, SDS-PAGE analysis of the extracellular and intracellular proteome, determination of growth and carbon fixation rates at different zinc levels, metal accumulation by ICP-MS, RNA-seq analysis for the identification of differentially expressed genes and hence discovery of Zur-regulon components, and RT-qPCR to confirm RNA-seq results.

4.2 MATERIALS AND METHODS

4.2.1 Bacterial strains and growth conditions

4.2.1.1 *Escherichia coli* strains and growth conditions

E. coli cells were grown in liquid Luria-Bertani medium or streaked onto solid LB agar plates with either kanamycin (Km, 50 $\mu\text{g mL}^{-1}$), ampicillin (Amp, 100 $\mu\text{g mL}^{-1}$) and chloramphenicol (Cm, 30 $\mu\text{g mL}^{-1}$) added whenever appropriate as described in section 2.1. *E. coli* strains used for the *zur* knockout mutant generation were S17-1 λPir and DH5- α (see Table 2.1).

4.2.1.2 *Synechococcus* sp. WH8102 growth conditions

Synechococcus sp. WH8102 cells were cultured in 100 mL of artificial seawater medium (ASW)¹³⁵ in 250 mL glass conical flasks. ASW medium was used without addition of ZnSO_4 which was added separately for growth experiments. The components of ASW- Zn are shown in Table 4.1. Cultures were maintained at 23 °C with continuous illumination (10 $\mu\text{E m}^{-2}\text{sec}^{-1}$ white light) in a Multitron Standard INFORS HT incubator and sub-cultured once a month by 10 fold dilution into fresh ASW- Zn medium accompanied by checking for contamination onto contamination plates (see section 4.2.1.2.1) in order to ensure that cultures were axenic.

Table 4.1 Composition of ASW-zn medium *

| Macronutrient | Final concentration |
|--|--|
| NaCl | 428 mM (25 g L ⁻¹) |
| MgCl ₂ •6H ₂ O | 10 mM (2 g L ⁻¹) |
| KCl | 6.7 mM (0.5 g L ⁻¹) |
| NaNO ₃ | 8.8 mM (0.75 g L ⁻¹) |
| K ₂ HPO ₄ •3H ₂ O | 0.172 mM (3*10 ⁻² g L ⁻¹) |
| MgSO ₄ •7H ₂ O | 14 mM (3.5 g L ⁻¹) |
| CaCl ₂ •2H ₂ O | 3.4 mM (0.5 g L ⁻¹) |
| Tris HCl buffer | 9.08 mM (1.1 g L ⁻¹) |
| Trace metals | Final concentration |
| H ₃ BO ₃ | 46.1 μM (2.86*10 ⁻³ g L ⁻¹) |
| MnCl ₂ •4H ₂ O | 9.1 μM (1.81*10 ⁻³ g L ⁻¹) |
| Na ₂ MoO ₄ •2H ₂ O | 1.8 μM (4.0*10 ⁻⁴ g L ⁻¹) |
| CuSO ₄ •5H ₂ O | 0.032 μM (8.0*10 ⁻⁶ g L ⁻¹) |
| Co(NO ₃) ₂ •6H ₂ O | 0.17 μM (5.0*10 ⁻⁵ g L ⁻¹) |
| FeCl ₃ •7H ₂ O | 10.4 μM (3.0*10 ⁻³ g L ⁻¹) |
| EDTA (Na ₂ Mg) | 1.4 μM (5.0*10 ⁻⁴ g L ⁻¹) |

*The original concentration of ZnSO₄•7H₂O in ASW is 0.772 μM (2.2*10⁻⁴ g L⁻¹)¹³⁵.

The concentrated stock solutions of all the major nutrients except for NaCl were prepared separately in MilliQ water (18.2 MQ) and kept at room temperature. A MilliQ 185 Plus water system (Millipore) was used to generate ultrapure water for all experiments. The trace metal stock solution (1000 X) was prepared separately by dissolving all components in MilliQ water except for FeCl₃•6H₂O, which was dissolved in concentrated HCl and added last to 1 L of the trace metal solution. The trace metal stock solution was kept at 4 °C. To prepare 1 L of ASW medium, 25 g of NaCl was dissolved in 500 mL ultrapure water. Aliquots of macronutrient stock solutions were then added followed by pH adjustment to pH 8.1 with concentrated HCl. MilliQ water was then added up to 1 L followed by addition of 1 mL of trace metal stock which further changed the pH to 8.0. Subsequently, 100 mL ASW aliquots were transferred into 250 mL clean Erlenmeyer glass flasks, plugged with a cotton wool plug and covered with foil. Flasks were then autoclaved at 121 °C for 15 minutes. All chemicals used for preparation of ASW medium were of the highest available

grade and were ordered from either Sigma-Aldrich or Fisher Scientific. *Synechococcus* sp. WH8102 *zur* knockout mutant cells were grown in ASW_{-Zn} with kanamycin (50 µg mL⁻¹).

4.2.1.2.1 Contamination check

In order to check that *Synechococcus* sp. WH8102 WT and the *zur* knockout mutant remained axenic, following each sub-culturing cultures were streaked on ASW contamination plates, containing 8 g L⁻¹ yeast extract, and 10 g L⁻¹ Bacto Agar in ASW. The plates were incubated at 23 °C for 2 weeks. If no contamination was observed, it was assumed that cultures were axenic.

4.2.1.3 *Ruegeria pomeroyi* growth conditions

Sucrose intolerant *Ruegeria pomeroyi* DSS-3 (pBBR-MCSI Km^r pKNG101) was provided by Dr. J. Christie-Oleza, School of Life Sciences, University of Warwick. *R. pomeroyi* was grown in liquid Marine Broth medium (MB) with kanamycin added (25 µg mL⁻¹) or streaked onto solid MB agar plates with Km (50 µg mL⁻¹). Both liquid cultures and plates were incubated for 48 hours at 30 °C. Liquid medium was prepared by dissolving 37.7 g of MB (Difco) per 1 L of MilliQ water and autoclaved for 15 minutes at 121 °C. Plates were prepared by adding 10 g L⁻¹ of Bacto Agar to the appropriate amount of un-autoclaved MB followed by autoclaving. When the solution was cooled down to about 37 °C, Km stock was added. The mixture was then poured onto disposable plates. Agar plates were kept at 4 °C and warmed in the incubator for 1 hour prior to streaking. Liquid MB medium was kept at room temperature.

4.2.1.4 Monitoring bacterial growth and cell number calculation

Optical density (OD) at 750 nm was used for monitoring cyanobacterial growth. Growth of *E. coli* and *R. pomeroyi* cells was monitored by OD at 600 nm. Optical density was measured using a UV-vis spectrophotometer (Ultrospec 3000 pro, Biochrom Ltd).

Synechococcus sp. WH8102 cell numbers were calculated by a haemocytometer. Aliquots of the cell cultures diluted 100- and 1000-fold were placed into an ethanol washed and dried haemocytometer. The haemocytometer was then

covered by a cover slip so that the Newton's Rings appeared evidencing that the coverslip was installed properly. Cells were counted in the grids numbered 1-5 using a microscope.

The haemocytometer grid is shown in Figure 4.1. Each of the grids 1-5 has a volume of 4 nL. For cells present on lines separating grids, only the cells on the top and left lines were then counted. The mean number of cells for 5 grids was then calculated and divided by the 4 nL volume to calculate the cell number per 1 mL. 3 replicates were used for calculations.

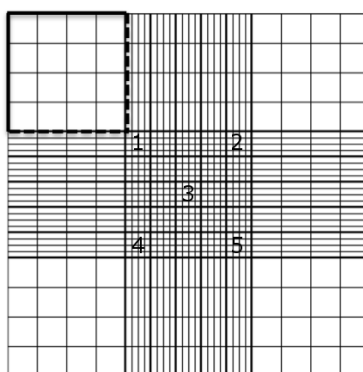


Figure 4.1 Haemocytometer grid.

E. coli cell numbers were calculated using an online *E. coli* concentration calculator¹³⁶.

4.2.1.5 Plasmids

The list of plasmids is shown in Table 4.2.

Table 4.2 List of plasmids used in this work.

| Plasmid | Description | Supplier |
|---------------|--|----------------------------------|
| pGP704CmKm | Ap ^r Km ^r Cm ^r ; pGP704-derivative containing the Km ^r and Cm ^r gene at the <i>EcoRI</i> site of its Multiple Cloning Sites (MCS) | Dr. Christie-Oleza ⁹¹ |
| pGP704CmKmZur | pGP704CmKm-derivative containing the <i>zur</i> (21-314) insert at the <i>XbaI</i> and <i>SalI</i> sites of its MCS | This study |

4.2.2 Mutant construction

4.2.2.1 DNA extraction from cyanobacteria

The protocol used to extract total genomic DNA from *Synechococcus* sp. WH8102 was a modified method from¹³⁷.

30 mL of late log phase *Synechococcus* culture was centrifuged at 6000 rpm for 10 minutes (Centrifuge 5810 R, Eppendorf). Following discarding the medium by pipetting and washing the pellet with 2 mL TE buffer (Tris 10 mM, EDTA 1 mM (pH 8), cells were centrifuged again at 6000 rpm for 10 minutes. The wash was then decanted and cells lysed by addition of a mixture of 567 μ L TE, 30 μ L 10% SDS, 3 μ L Proteinase K (20 mg mL⁻¹, Boehringer Mannheim) and 3 μ L RNase. The solution was then mixed thoroughly by re-suspending the cells without vortexing and incubated at 60 °C for 4 hours. Subsequently, 100 μ L 5 M NaCl and 80 μ L CTAB solution (cetyl trimethyl ammonium bromide, 10 % in 0.7 M NaCl) were added to the lysis mixture which was then incubated at 65 °C for another 10 minutes. The CTAB protein complexes were then extracted with 1 volume of a mixture of phenol/chloroform/isoamyl alcohol (~0.8 mL, 25:24:1, v/v/v) twice. Both times the mixture was centrifuged at 6000 rpm for 10 minutes. The aqueous phase containing DNA was transferred into a fresh tube. Following that step, remaining CTAB complexes were removed by extraction with 1 volume of chloroform/isoamyl alcohol (24:1, v/v), centrifuged as before and the supernatant transferred to a fresh tube. Nucleic acids were precipitated by the addition of 0.6 volumes of 100% isopropanol. The isopropanol mixture was left at -20 °C overnight and in the morning centrifuged at 6000 rpm for 45 minutes at 4 °C leaving a creamy coloured pellet. The pellet was then washed in progressive ethanol concentrations: 50%, 70%, 100% (v/v). On the removal of 100% (v/v) ethanol, the tubes were left to stand open to allow evaporation of the remaining ethanol. Finally, the pellet was re-suspended in 100 μ L nuclease-free water. The resulting DNA concentration in the solution as well as the presence of phenol were checked by NanoDrop as described in section 2.3.7.

4.2.2.2 PCR primers

Primers used for PCR were designed using Clone Manager Professional Suite 7.04 software. The list of primers used in this study is given in Table 4.3. Primers used

for cloning contained 18-20 nt sequences with the appropriate restriction sites added on 5' ends and 2-4 additional bases. All other primers were 18-20 nt in length. The template was *Synechococcus* sp. WH8102 genomic DNA in water (see section 4.2.2.1.), or plasmids or *Synechococcus* sp. WH8102 or *E. coli* cells boiled in water for 5 minutes.

Table 4.3 PCR primers used in this study. Restriction sites are highlighted.

| Name | Sequence | Restriction sites | Used for: |
|--------|----------------------------------|-------------------|---|
| Zur_F | ATGTCGACTTTGAACGCCCGTCAA CAGG | <i>SalI</i> | Amplifying <i>zur</i> ₂₁₋₃₁₅ insert |
| Zur_Re | ATTCTAGAGTCAGCAGGCACGTCG ATG | <i>XbaI</i> | |
| 3533_F | GTCTCATGAGCGGATAACATA | | Checking the plasmid pGP704CmKmZur |
| 73_Re | ACTTAACGGCTGACATGGCC | | |
| A_F | CTGGCCAATGTGATGTTC | | Checking <i>zur</i> mutant and WT after conjugation |
| B_Re | TTTCATCGCTCTGGAGTG | | |
| C_Re | GCAGGAGATCAAGACTTTCG | | |
| D_F | CGTCAACACGGGATAAATACC | | |

4.2.2.3 Plasmid construction and preparation of *E. coli* for conjugation

For generation of a single crossover *Synechococcus* sp. WH8102 *zur* interposon mutant plasmid pGP704CmKmZur was constructed by ligation of a 295 bp internal fragment of the *Synechococcus* sp. WH8102 *zur* gene using primers Zur_F and Zur_Re and WT genomic DNA as a template into plasmid pGP704CmKm. Plasmid pGP704CmKm was digested with *SalI* and *XbaI* (Promega) in 10x MultiCore Buffer (Promega) and de-phosphorylated using calf intestinal phosphatase in order to prevent self-ligation as described in section 2.3.2. The internal *zur* PCR product was also digested with the same restriction enzymes. The Promega Wizard® SV Gel and PCR Clean-Up System was used to purify the plasmid and the insert after digestion. PCR amplification, restriction enzyme digestion, dephosphorylation and ligation were performed as described in sections 2.2.1, 2.3.1, 2.3.2 and 2.3.3, respectively.

Subsequently, plasmid pGP704CmKmZur was transformed into *E. coli* S17-1 λ Pir as described in section 2.3.5. LB agar plates with kanamycin (50 μ g mL⁻¹) and chloramphenicol (30 μ g mL⁻¹) were used for selection of transformants.

4.2.2.4 *Synechococcus* sp. WH8102 conjugation

Conjugation was performed as described in Brahamsha (1996)¹³⁸ with minor modifications:

10 mL of an overnight culture of *E. coli* S17-1 λ Pir containing plasmid pGP704CmKmZur ($\sim 10^9$ cells mL⁻¹) was centrifuged at 3000 rpm for 10 minutes and washed twice with LB medium in order to remove antibiotics. Cells were then re-suspended in 500 μ L ASW containing 10% (w/v) LB. 50 mL of late log *Synechococcus* sp. WH8102 culture containing about 5×10^8 cells mL⁻¹ was then centrifuged at 3000 rpm for 15 minutes at room temperature and re-suspended in 150 μ L ASW. *Synechococcus* sp. WH8102 and *E. coli* cells were subsequently mixed at different ratios (50 μ L:50 μ L, 25 μ L:50 μ L and 10 μ L:50 μ L, respectively). Mixtures were spotted on at least a one week old 0.6% (w/v) agarose ASW plate. Following adsorption of the liquid from the spots, plates were incubated at low light (5 μ E m⁻² sec⁻¹) for 48 hours at 23 °C. Subsequently, spots were cut out from the plates with a sterile spatula and transferred into sterile 20 mL universal tubes. 2 mL ASW was added and cells re-suspended with gentle shaking.

10 mL of *R. pomeroyi* culture at the stationary stage of growth in MB was spun down and washed twice with 10% (w/v) LB in ASW, and re-suspended in 500 μ L ASW. 200 μ L *Synechococcus* + *E. coli* mixture was then mixed with 100 μ L *R. pomeroyi* (a helper strain that increases *Synechococcus* plating efficiency) and added into the empty sterile disposable petri dish. Warm (37 °C) 0.18% (w/v) agarose in ASW with Km (50 μ g mL⁻¹) was pour-plated on top of the petri dishes, shaken with gentle circle movements to uniformly spread the cells and left to solidify at low light overnight. Agarose plates were incubated at a light intensity of 5 μ E m⁻² sec⁻¹ at 23 °C until colonies appeared (\sim 4-5 weeks). Colonies that were large enough were picked by pipetting with sterile tips into 1 mL ASW containing 25 μ g mL⁻¹ Km. Following growth of these colonies PCR was performed using cells as a template followed by agarose gel electrophoresis (see sections 2.2.1 and 2.2.2) to check for successful plasmid integration into the *Synechococcus* sp. WH8102 genome. When DNA

fragments of the expected size were observed after PCR and agarose gel electrophoresis, putative *zur* knockout mutant *Synechococcus* sp. WH8102 cultures were checked for the presence of WT DNA and gradually transferred into larger volumes of ASW_{-Zn} with increasing Km concentrations. Concurrently, cultures were streaked onto contamination plates (see section 4.2.1.2.1.), LB plates and MB plates in order to check for the presence of *R. pomeroyi* or other heterotrophic bacteria contaminants.

4.2.2.5 Purification of the *Synechococcus* sp. WH8102 *zur*-mutant from *R. pomeroyi*

To 10 mL of the *Synechococcus* sp. WH8102 *zur* mutant culture, 10 mL of a sterile solution of ASW_{-Zn} containing kanamycin, ampicillin and sucrose was added giving final concentrations of 50 µg mL⁻¹ Km, 100 µg mL⁻¹ Amp and 10% (w/v) sucrose. The culture was incubated overnight and then pour-plated using a serial dilution series down to 10⁻⁷ 0.22 % (w/v) agarose ASW_{-Zn} medium containing Km (50 µg mL⁻¹). When colonies appeared (~ 3-4 weeks) they were picked with sterile pipette tips into 1 mL ASW_{-Zn} liquid medium containing 50 µg mL⁻¹ Km and incubated at 5 µE m⁻² sec⁻¹ at 23 °C. Cultures were gradually transferred into larger volumes of ASW_{-Zn} and checked for the presence of heterotrophs by streaking onto contamination plates, as well as LB and MB plates without antibiotics. If no contamination was observed within 2 weeks cultures were assumed to be axenic.

4.2.3 Phenotypic comparison of the *Synechococcus* sp. WH8102 *zur*-mutant with wild type

4.2.3.1 Growth rate comparison

4.2.3.1.1 General conditions

Cultures of the *Synechococcus* sp. *zur* knockout mutant and WT were grown in ASW_{-Zn} medium (see section 4.2.1.2). Cultures were grown until late-log phase (OD₇₅₀ >1) and experimental flasks were inoculated to a similar final OD₇₅₀. OD₇₅₀ measurements were taken every 48-72 hours until cultures reached stationary phase. Cultures were checked for contamination at each time point as well as at the end of each experiment.

4.2.3.1.2 Medium preparation

Before adding the trace metal stock, ASW was treated with Chelex 100 resin in order to remove any zinc impurities (see section 4.2.3.1.3). Trace metal stock without zinc was then added followed by ZnSO₄ to the desired final concentration. ASW medium was autoclaved at 121 °C for 20 minutes and when required Km added to a final concentration of 50 µg mL⁻¹. Growth experiments were performed in a volume of 250 mL with three replicates in acid-washed polycarbonate conical flasks. Acid-washing involved filling flasks with 6% (v/v) HCl overnight followed by rinsing flasks three times with RO grade water and once with Milli-Q water. Growth rates were compared at the following zinc concentrations: 0, 77.2 nM, 772 nM, 1.5 µM, 2.5 µM and 5 µM.

4.2.3.1.3 Chelex resin preparation

Chelex 100 (200–400 mesh: Bio-Rad Laboratories, Richmond, CA) was used and prepared as described in ¹³⁹. First, 5g of Chelex 100 resin was soaked in methanol for 6 hours at room temperature at a ratio of 1:5 (w/v) and rinsed with 750 mL Milli-Q water using a sintered glass filter funnel and a side-arm flask. It was then soaked in 1 M HCl overnight at room temperature and washed with 1 L of Milli-Q water. This was followed by leaving the resin for one week in 3 M NH₄OH and rinsing it with 1 L Milli-Q water. Following that, Chelex was soaked in 0.1 M HCl for 10 min, rinsed with 2 L Milli-Q water and rinsed with 200 mL medium (major ASW components without trace metals). The resin was then re-suspended in 200 mL medium solution titrating slowly to pH 8.1 with 1 M NaOH. Chelex was then rinsed with medium and transferred as a slurry to a chromatography column. The first 500 mL medium passed through the Chelex column were discarded. The pH of Chelexed medium was checked before use. The minimum amount of Chelex resin required was calculated as described in ¹⁴⁰, based on estimating the concentration of contaminating metals.

4.2.3.2 UV-vis absorption spectra comparison

To compare full range UV-vis spectra (300-750 nm), 1 mL of WT and *zur* knockout mutant cell cultures at the early stationary phase were scanned using a Varian Cary 50 Bio UV-vis spectrophotometer, Agilent Technologies.

4.2.3.3 Total cell lysate analysis by SDS-PAGE

To assess whether there were any visible differences in SDS-PAGE profiles of WT and *zur*-mutant at different zinc concentrations (0 and 772 nM) total protein cell lysates were prepared as described¹⁴¹ and run on a gel as described in section 2.4.1. 50 mL cell culture in mid-log phase (OD₇₅₀ 0.3-0.5) was centrifuged at 6000 rpm for 30 minutes at room temperature (Centrifuge 5810 R, Eppendorf). For every 0.1 g of cell pellet 10 mL 0.5 M Tris-HCl (pH 6.8) buffer was added and cells re-suspended. Following that, the suspension was mixed with 6x Laemmli buffer (see section 2.4.1) and samples boiled for 5 min and cooled down before further use. To assess whether there was a difference in extracellular proteomes, 1 mL of the same cultures was filtered through 0.22 µm pore size Minisart syringe filters (Sartorius) and supernatants mixed with 6x Laemmli buffer and prepared in the same way as described above.

4.2.3.4 Carbon fixation rate measurements

2 kBq of NaH¹⁴CO₃ was added to each mL of mid-log culture (Mutant and WT grown at 0 and 772 nM zinc). The total concentration of NaH¹⁴CO₃ was 0.866 µM. 100 µL culture was aliquoted into wells of black 96 well plates (Lumox®, Sarstedt Ltd). Each row was left at a different light intensity (1541, 719, 387, 196, 121, 58, 23, 0 µE m⁻²s⁻¹) for 1 h. Following 1 hour, 50 µL of the samples was transferred to clear 96-well microplates (SpectraPlate-96, PerkinElmer). 25 µL 6 M HCl was then added to remove ¹⁴CO₂ from NaH¹⁴CO₃ which was not consumed by the cells and left for 1 hour in a fume hood. No HCl was added to the total carbon control. Subsequently, 25 µL 6 M NaOH was added to neutralize HCl and 50 µL liquid Ultima Gold Scintillator, PerkinElmer added. Plates were then left overnight to equilibrate. The next day activity was measured using a 1450 MicroBeta TriLux Liquid Scintillation Counter (Wallac). All measurements were performed with three biological replicates and three technical replicates per sample. CO₂ fixation rate was normalised to chlorophyll *a* content as measured below.

4.2.3.4.1 Chlorophyll content measurement

1 mL of culture (2 replicates per sample) was centrifuged at 13000 rpm on a table top centrifuge (Heraeus Biofuge Pico, Kendro Laboratory Products). Supernatant

was removed by pipetting, and 1 mL methanol added and thoroughly mixed with the pellet. The samples were then left for 30 minutes and absorbance at 645 nm and 663 nm measured.

The content of chlorophyll *a* was calculated using the formula¹⁴²:

$$\text{Chlor}_a(\text{mg mL}^{-1}) = 12.7 \times A_{663} - 2.7 \times A_{645} \quad \text{Equation 4.1}$$

4.2.3.5 Inductively Coupled Plasma – Mass Spectrometry (ICP-MS)

Cells were grown in Chelexed ASW as described in section 4.2.3.1.2. For the *zur* knockout mutant only 0 and 772 nM zinc conditions were used and for the WT zinc concentrations were 0, 772 nM and 2.5 μM . Three biological replicates were used for each sample. At mid-log stage of growth ($\text{OD}_{750} = 0.4-0.5$) cells were harvested by centrifugation of 50 mL of cell culture at 4000 rpm for 30 minutes. Cells were then re-suspended in 10 mL ASW-Zn with 1 mM EDTA to chelate trace metal weakly bound to the cell surface, transferred into 15 mL Falcon tubes and centrifuged for 15 minutes at the same speed. The last step was repeated twice. Finally, the cell pellet was washed with 10 mL MilliQ water without re-suspension and centrifuged which was repeated twice. Cell pellets were then snap-frozen in liquid nitrogen until further use. Subsequently, cells were carefully transferred in 0.5 mL MilliQ water into weighed 1.5 mL Eppendorf tubes. Samples were then freeze-dried (lyophilised) overnight using VirTis BenchTop™ “K” Freezedryer at $-65\text{ }^\circ\text{C}$ until the weight was stable and taken as total dry weight. Dry pellets were digested in 300 μL 72% ultrapure HNO_3 overnight at $65\text{ }^\circ\text{C}$ and further dissolved in 5.7 mL MilliQ water. Those solutions were used as samples for ICP-MS measurements of Mn, Co, Ni, Cu, Zn, Cd. The samples were then diluted 10 times to measure P and Fe. Standards for ICP-MS were prepared from 1000 ppm standards (Fisher Scientific) by gravimetric dilution in 3.6% HNO_3 . Final standard concentrations were: 0, 0.1, 0.25, 0.5, 1, 2.5, 5, 10, 25, 50, 100, 250, 500 and 1000 ppb. The ICP-MS measurements were performed using an Agilent 7900 ICP-MS instrument in He gas mode (31-P, 55-Mn, 59-Co, 60-Ni, 63-Cu, 66-Zn, 111-Cd) and H_2 collision gas mode (56-Fe only) with typical 1.000 sec integration time. Data were acquired and processed using Mass Hunter 4.3 for Windows.

4.2.4 Transcriptomics

4.2.4.1 Sample preparation

ASW was prepared as described in section 4.2.3.1.2 and major ASW components were Chelexed to remove trace metal contaminants. Kanamycin ($50 \mu\text{g mL}^{-1}$) was added to the autoclaved $\text{ASW}_{-\text{Zn}}^{\text{Ch}}$ medium. The experimental design is shown in Figure 4.2.

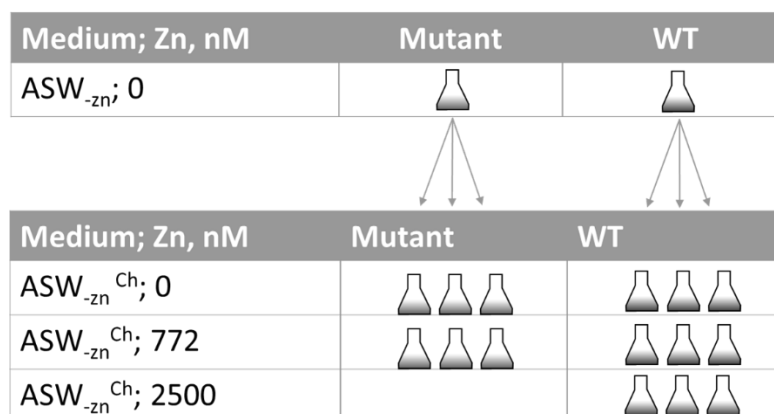


Figure 4.2 Experimental design of the RNA-seq experiment.

$\text{ASW}_{\text{Zn}}^{\text{Ch}}$: medium Chelexed prior to addition of trace metal stocks.

WT *Synechococcus* sp. WH8102 and *zur* knockout mutant were grown to the late-log phase of growth in non-Chelexed $\text{ASW}_{-\text{Zn}}$ and sub-cultured into flasks with different zinc content (0, 772 nM and $2.5 \mu\text{M}$). Cultures were then grown to mid-exponential phase ($\text{OD}_{750} \sim 0.3-0.4$). Each condition had three biological replicates. Since the *Synechococcus* sp. WH8102 *zur* mutant did not grow at $2.5 \mu\text{M}$ zinc this concentration was not used for the mutant (see section 4.3.2.1). When the desired OD_{750} was reached, 250 mL of each culture was separated into 5 sterile 50 mL Falcon tubes, centrifuged at 4000 rpm at 4°C , separated from the supernatant by decantation and snap-frozen in liquid nitrogen. One set of the samples was used for RNA-seq analysis (see section 4.2.4.2), one set for RT-qPCR (see section 4.2.4.4), another for ICP-MS (see section 4.2.3.6), and two sets for optimising RNA extraction (see below).

4.2.4.2 RNA isolation

Total RNA was extracted using a modified phenol-chloroform protocol¹⁴³. Cell pellets obtained in section 4.2.4.1 were thawed on ice in 1.5 mL of Z buffer (8 M

guanidinium hydrochloride; 50 mM β -mercaptoethanol; 20 mM EDTA) at room temperature for 30 min. Two volumes of acidified phenol (pH 4.5) were added and mixtures were heated to 65°C for 30 min, followed by the addition of a mixture of chloroform and isoamyl alcohol (24:1, v/v) for 15 min. The tubes were then centrifuged at 4000 rpm and the aqueous phase transferred to separate microcentrifuge tubes which were centrifuged again at 10,000 rpm and transferred into fresh tubes. RNA was precipitated by addition of 1 vol. of RNase-free isopropanol followed by incubation overnight at -20 °C. The tubes were then centrifuged at 13,000 rpm in a benchtop centrifuge (Biofuge Pico, Heraeus) at 4°C for 30 mins. The pellets were then washed with 70% (v/v) RNase-free ethanol and further centrifuged for 15 min. Following decanting of the supernatant, pellets were dried at room temperature for 15 min and then re-suspended in 50 μ L RNase-free water. DNA was removed using the TURBO DNA-free™ kit (Ambion®/Life Technologies). The following changes were made to the manufacturer’s instructions: double treatment for 30 minutes at 37 °C was applied with 2 units of DNase each time instead of one application with 1 unit. DNase was then inactivated according to the protocol and the samples were additionally purified using Zymo RNA Clean & Concentrator-5 using the manufacturer’s instructions. The concentration of RNA and the purity of the samples were assessed using NanoDrop (see section 2.3.7). The presence of DNA contamination was assessed by PCR with 16S rRNA gene primers (see Table 4.4). If no product was observed, the samples were diluted 1000 times and the concentration and RNA integrity of the samples assessed using an Agilent Bioanalyzer with an Agilent RNA 6000 Pico Kit following the manufacturer’s instructions.

Table 4.4 16S rRNA gene primers used for assessing DNA contamination.

| | |
|-----------|----------------------|
| 27F | AGAGTTTGATCMTGGCTCAG |
| 1492R (s) | ACCTTGTTACGACTT |

4.2.4.3 RNA-seq analysis

RNA samples were sent to the Centre for Genomic Research, Institute of Integrative Biology, at the University of Liverpool for library preparation and sequencing. Samples were re-purified using a Qiagen RNeasy Kit to remove 5S rRNA

(5S). Without this step library preparation failed. A RiboZero kit, Illumina was then used for rRNA depletion and a NEB Next Ultra Directional RNA library preparation kit used for preparation of dual-indexed, strand-specific RNA-seq libraries. When libraries were ready, they were sequenced using an Illumina HiSeq 4000 (paired-end, 2x150 bp). The raw data files were trimmed for the presence of Illumina adapter sequences using Cutadapt version 1.2.1¹⁴⁴. Option -O 3 was used, so the 3' end of any reads which matched the adapter sequence for 3 bp or more were trimmed. The reads were further trimmed using Sickle version 1.200 with a minimum window quality score of 20. Reads shorter than 20 bp after trimming were removed. If only one of a read pair passed this filter, it is included in the R0 file. The output files from Cutadapt and Sickle as well as resulting FASTQ files are available at http://cgr.liv.ac.uk/illum/LIMS16056_259a058713a41d9b/.

4.2.4.3.1 RNA-seq data analysis

An overview of the data analysis is shown in Figure 4.3.

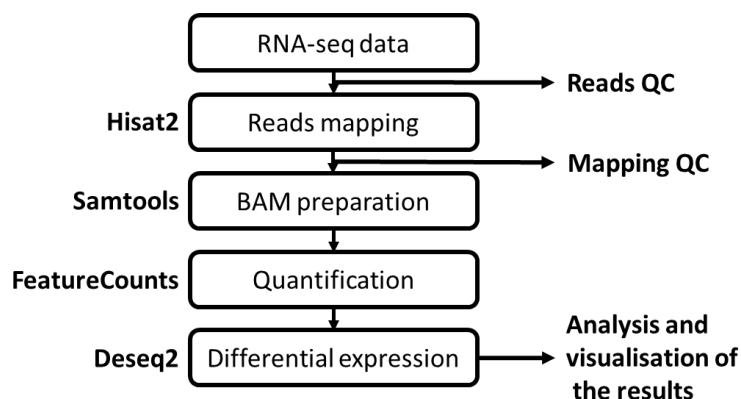


Figure 4.3 The overview of the RNA-seq data analysis.

Hisat2¹⁴⁵ software is used to map FASTQ reads onto the genome.

Index building

Index building based on the sequenced genome of *Synechococcus* sp. WH8102 was prepared with this code:

```
hisat2-build
/media/alevtina/home/khu/genome/Syn_WH8102.fa
/media/alevtina/home/khu/8102index/8102index
```

The complete genome of *Synechococcus* sp. WH 8102 (Syn_WH8102.fa) was downloaded at

<https://www.ebi.ac.uk/ena/data/view/BX548020>

Mapping

Mapping of the FASTQ files was performed with the default parameters using the following command:

```
hisat2 -p 8 --dta -x /mnt/safe/Alko/index/index -1  
/mnt/safe/Alko/Trimmed/Sample_16/16_CCGCGGTT-  
CTAGCGCT_L005_R1_001.fastq.gz -2  
/mnt/safe/Alko/Trimmed/Sample_16/16_CCGCGGTT-  
CTAGCGCT_L005_R2_001.fastq.gz -S  
/mnt/safe/Alko/sam/16.sam
```

Converting SAM to BAM

This is required to convert SAM (Sequence Alignment Map) into binary BAM (Binary Alignment Map) files so that the files are readable by the FeatureCounts software. This was done using Samtools¹⁴⁶.

```
samtools view -S -b /mnt/safe/Alko/sam/16.sam >  
/mnt/safe/Alko/sam/16.bam
```

Sorting BAM files

```
samtools sort -n /mnt/safe/Alko/sam/16.bam -o  
/mnt/safe/Alko/sam/16.sorted.bam
```

Quantification of the reads and gene assignment

FeatureCounts software was used to identify mapped reads as genes¹⁴⁷.

```
featureCounts -T 8 -t gene -g ID -a  
/mnt/safe/Alko/Syn_WH8102_2.gff -o  
/mnt/safe/Alko/sam/16.txt  
/mnt/safe/Alko/sam/16.sorted.bam
```

In order to optimise and speed up the data analysis, a bash-script combining mapping, converting SAM to BAM, sorting BAM and quantifying the reads was used for all the samples. The script is presented in Figure 8.8, Appendix.

Finally, DESeq2 as a R-package in R-studio software was used to normalize raw reads and calculate statistics¹⁴⁸. The example of DESeq2 script for a comparison of one condition containing three replicates with another condition containing three replicates is presented in Figure 8.9, Appendix.

4.2.4.4 Reverse transcription-quantitative polymerase chain reaction (RT-qPCR)

RNA samples were prepared as described in section 4.2.4.2. Reverse transcription was performed using the GoScript™ Reverse Transcription System from Promega according to the manufacturer's instructions. The resulting cDNA was used as a template without purification from GoScript reagents. The RT-qPCR mixtures were prepared in 96-well MicroAmp microplates (Applied Biosystems) and covered by MicroAmp adhesive film (Applied Biosystems). Each well contained 5 µL PowerUp™ SYBR Green Master Mix (Applied Biosystems), 1 µL template, a mixture of primers at the concentration needed, and water to a final volume of 10 µL. All the reactions had three technical replicates for each of three biological replicates. In addition, each plate had the same controls without reverse transcriptase added, three wells with negative controls (water as template) and three wells as positive control with DNA as template. RT-qPCR was run on a 7500 Fast Real-Time PCR System (Applied Biosystems). Reaction mixtures were heated to 50 °C for 2 minutes then to 95 °C for 10 min and underwent 40 cycles (95 °C for 15 sec and 60 °C for 1 min). Straight after qPCR cycles were finished, for melting curves the mixtures were heated to 95 °C for 15 sec, cooled to 60 °C for 1 min then slowly (1%) heated to 95 °C for 30 seconds when detection was performed and finally cooled down to 60 °C for 15 sec. Data were analysed using 7500 Software, v2.3 (Applied Biosystems) and Microsoft Excel. The amplification efficiency was calculated via Equation 4.2:

$$E = -1 + 10^{\frac{-1}{\text{slope}}} \quad \text{Equation 4.2}$$

which was derived from $N_t = N_0 \times (1 + E_N)^{Ct}$

Where: Ct is the cycle number at which the fluorescence reaches a fixed threshold; N_t is the threshold number of the amplified molecules; N_0 – initial number of the molecule; E_N – efficiency of amplification; The slope is a plot of $\log(N_0)$ versus Ct.

The $2^{-\Delta\Delta Ct}$ method was used to calculate differential expression of the genes¹⁴⁹:

$$\Delta\Delta Ct = Ct(\text{gene}) - Ct(\text{housekeeping}) - Ct(\text{normalised}) \quad \text{Equation 4.3}$$

And $-\Delta\Delta Ct$ is equal to log2fold change in case of 100% efficiency of amplification.

Primers for qPCR were designed using PrimerQuest Tool¹⁵⁰ from IDT with the parameters: qPCR, 2 Primers, Intercalating Dyes; so that the amplicon length was about 100 base pairs. The list of primers used is given in Table 4.5.

Table 4.5 qPCR primers used in this study.

| Gene | Primer sequences (F and Re) |
|----------------|---|
| ZnuA, SYNW0971 | AGCAAGGAGATCCAATCTTCAG GGCAATCCGGTTGACTTACT |
| BmtA, SYNW0359 | GCGCAATTGTTCTTCAAGGTAA TCCGCACTTACATCCACAAG |
| pepC, SYNW2047 | CGCAGTGATCTGGAAGTATGAT CGAAGATGTGGACCTGATTGA |

The housekeeping gene used for normalisation in marine cyanobacteria was phosphoenol pyruvate carboxylase¹⁵¹.

Calibration curves were built based on known concentrations of *Synechococcus* sp. WH8102 genomic DNA. The concentration of DNA was determined using NanoDrop (see section 2.3.7). The molecular weight of genomic DNA was calculated using DNA calculator¹⁵² from Molbiotools. DNA was diluted 1:1000 and then serially diluted 1:10 five times. Since each gene tested has only one copy per genome, the number of gene copies per 1 μ L template was calculated.

4.3 RESULTS AND DISCUSSION

4.3.1 Generation of a *Synechococcus* sp. WH8102 *zur* interposon mutant

In order to knock out the *zur* gene in *Synechococcus* sp. WH8102 a single crossover approach was chosen. The homologous recombination process for constructing a single crossover *zur* mutant is shown in Figure 4.4.

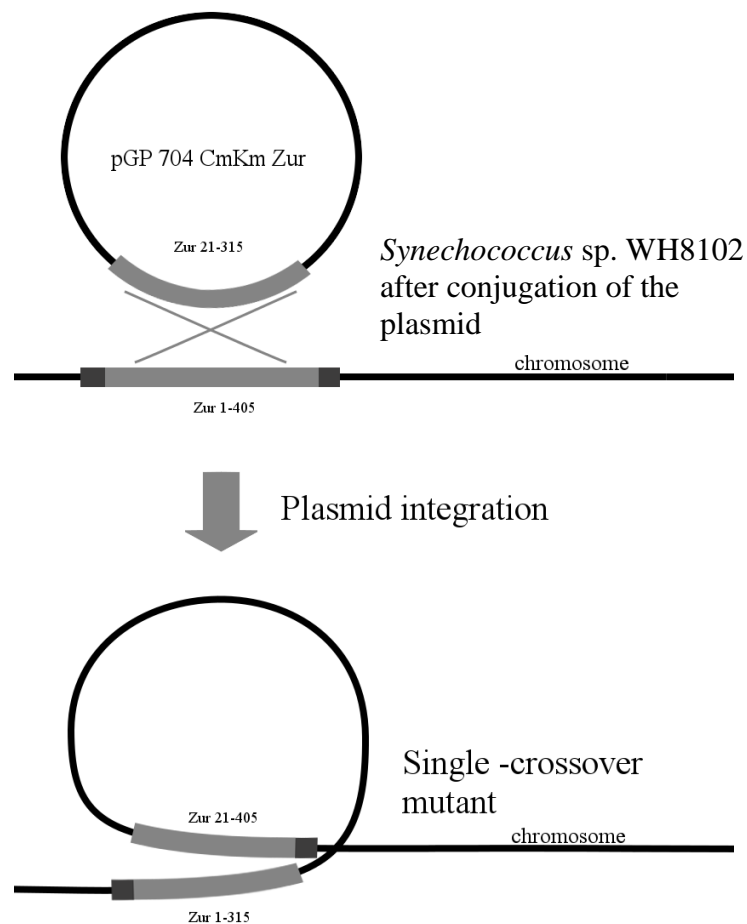


Figure 4.4 Schematic representation of a single crossover for generation of a *zur* knockout mutant of *Synechococcus* sp. WH8102 after conjugation.

The experimental design scheme of obtaining WH8102 *zur* knockouts by single crossover is presented in Figure 4.5. Steps are described below in sections 4.3.1.1, 4.3.1.2 and 4.3.1.3 in detail.

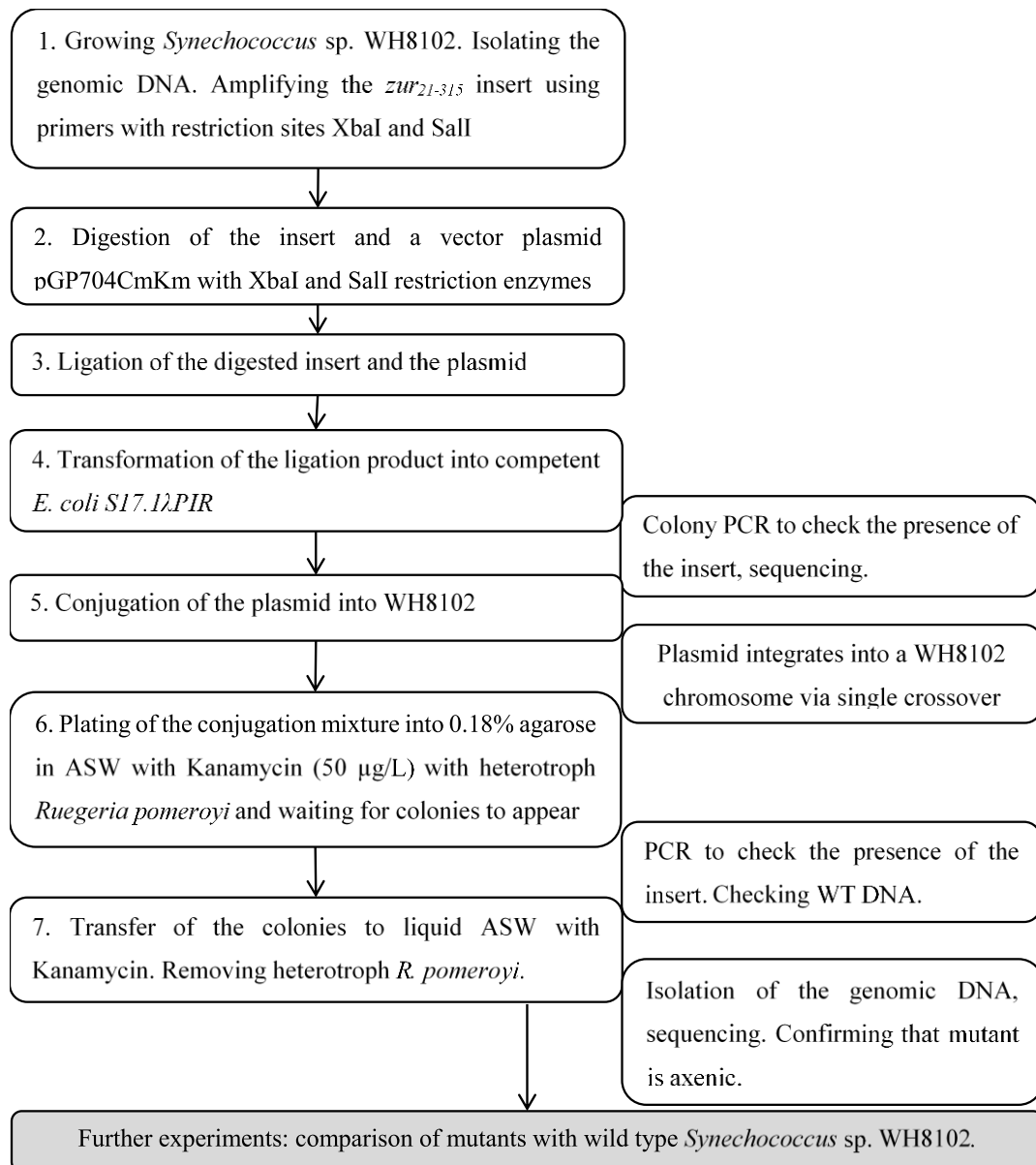


Figure 4.5 Steps needed for generation of a WH8102 *zur* knockout mutant by single crossover.

WT *Synechococcus* sp. WH8102 was grown to stationary phase (OD_{750} 1.385 = $\sim 5 \times 10^8$ cells mL⁻¹), genomic DNA was isolated (see section 4.2.2.1.) and used as template for PCR to amplify a 311 bp internal fragment of the *zur* gene (see Figure 4.6). This fragment was subsequently cloned into plasmid pGP704CmKm using *Xba*I and *Sa*II to yield plasmid pGP704CmKmZur (see Fig. 4.7).

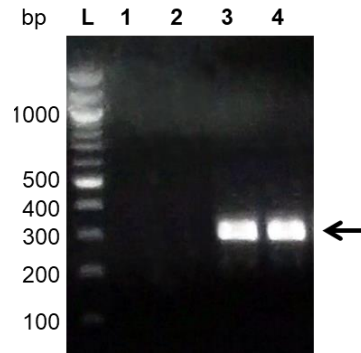


Figure 4.6 PCR amplification of the internal fragment of the *zur* gene from *Synechococcus* sp. WH8102.

L - 100 bp NEB DNA ladder. Lane 1 – DNA was used as a template without dilution, lane 2 – 10-fold dilution, lane 3 – 100-fold dilution and lane 4 – 1000-fold dilution. The 311 bp PCR product is marked with an arrow.

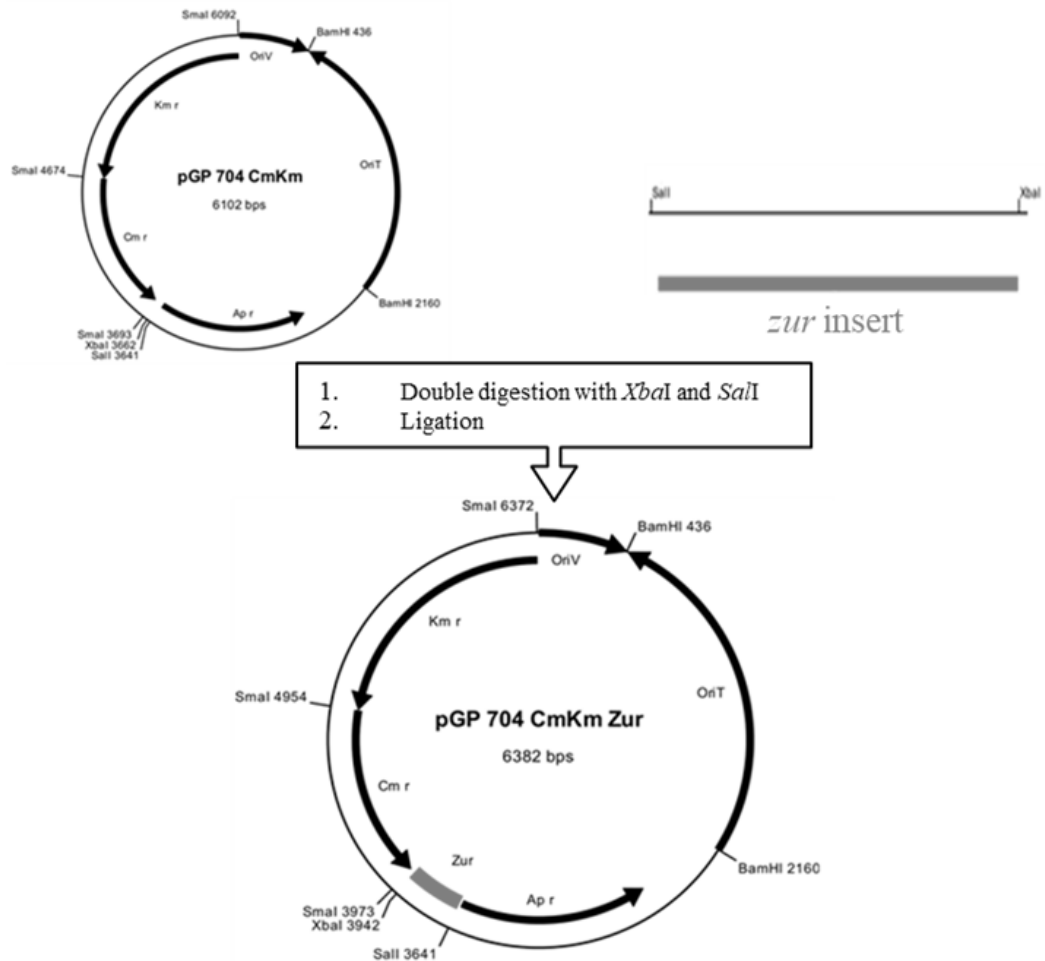


Figure 4.7 Construction of plasmid pGP704CmKmZur.

pGP704CmKm is a suicide vector with a R6K origin of replication, RP4 origin of transfer and three antibiotic resistance cassettes: ampicillin, kanamycin and chloramphenicol. Sequencing of pGP704CmKmZur confirmed the presence of the correct *zur* insert (data not shown). Conjugation into *Synechococcus* sp. WH8102 was carried out as described in section 4.2.2.4. Colonies appeared after 4-5 weeks and were checked for the presence of the *zur* mutation by PCR using primers A and B (see Figure 4.8 and Figure 4.9B for the explanation of the choice of primers).

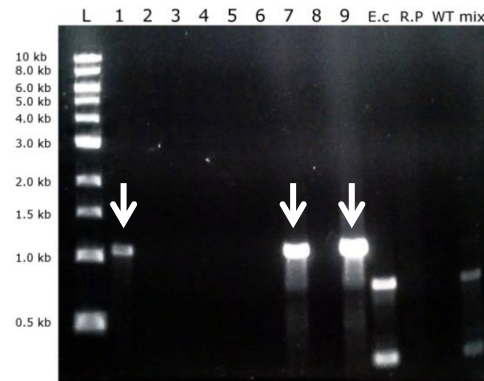


Figure 4.8 Analysis of plasmid integration into the genome of *Synechococcus* sp. WH8102 by PCR with A and B primers with cells as template.

L – 1kb ladder. **1-9** – different colonies of potential *Synechococcus* mutants after conjugation. Bands of the correct size (~1.1 kb) are marked with white arrows for colonies **1, 7 and 9** indicating potential successful plasmid integration (see Figure 4.9B for the primer disposition). **E.c** – *E. coli* S17-1 λ pir with plasmid pGP704CmKmZur giving non-specific products of the incorrect size. **R.P** - *Ruegeria pomeroyi*. **WT** – wild type *Synechococcus* sp. WH8102 (no product was observed as predicted). **Mix** – mixture of E.c., R.P. and WT.

Synechococcus colonies from lanes 1 and 7 were used for further experiments. Sucrose was used to remove the helper bacterium *Ruegeria pomeroyi* (see section 4.2.2.5), with new *Synechococcus* colonies appearing in 3-4 weeks and colonies transferred into 1 mL ASW_{-zn} containing 50 $\mu\text{g mL}^{-1}$ kanamycin. Once growth was observed, cultures were gradually transferred into larger volumes of ASW_{-zn} containing kanamycin until the volume reached 100 mL. Cultures were checked for contamination and shown to be axenic. DNA was isolated from one of these *zur* knockout mutants and the correct mutation assessed by PCR (see sections 2.2.1 and 4.2.2.2, Figure 4.9A) and sequencing of PCR products. The primers used are shown in Figure 4.9B. PCR and subsequent sequencing of PCR products confirmed that a fully segregated axenic *zur* knockout mutant of *Synechococcus* sp. WH8102 had been generated.

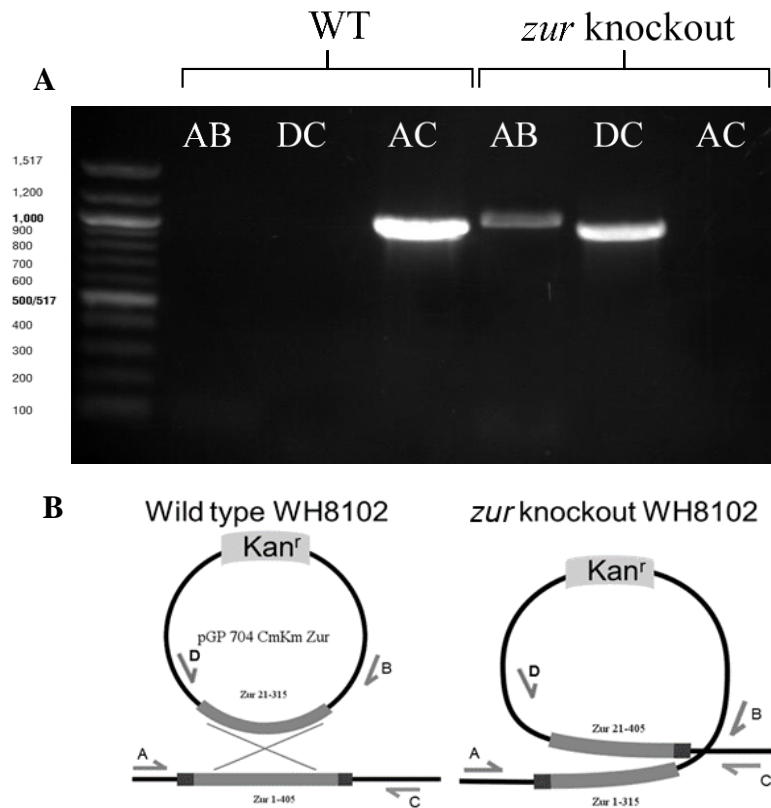


Figure 4.9 PCR confirmation of the *Synechococcus* sp. WH8102 mutant construct.

A. Gel after PCR with different primers. **B.** Explanation of the choice of primers for checking the successful *zur* single crossover mutant.

4.3.2 Phenotypic characterization of the *Synechococcus* sp. WH8102 *zur* mutant

4.3.2.1 Growth rate assessment

Cells were transferred from non-Chelexed ASW- Z_n into Chelexed medium with different zinc concentrations as described in section 4.2.3.1. Growth rate comparisons are shown in Figure 4.10.

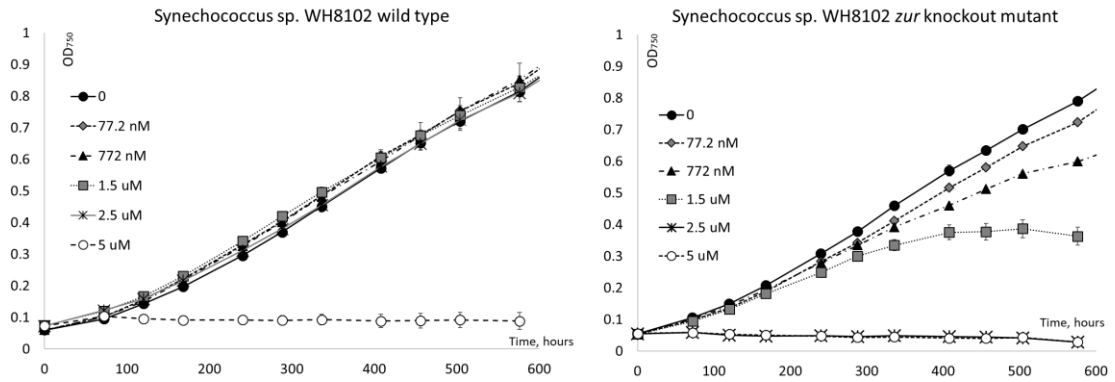


Figure 4.10 Growth curves of the WT and *Synechococcus* sp. WH8102 *zur* knockout mutant at different zinc concentrations.

Error bars indicate standard errors between three biological replicates.

Specific growth rates are shown in Figure 4.11.

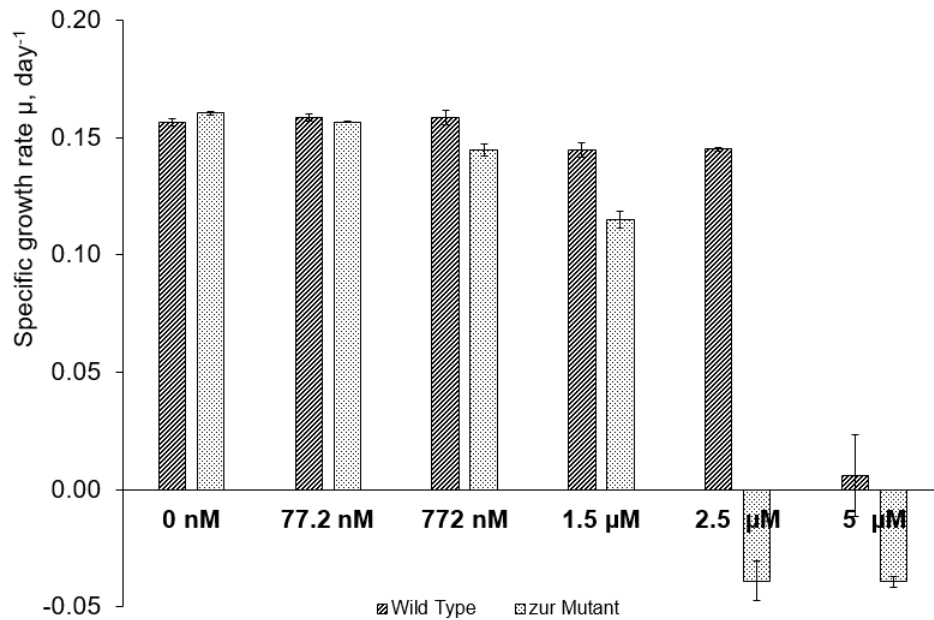


Figure 4.11 Specific growth rates for *Synechococcus* sp. WH8102 WT and *zur*-mutant at different zinc concentrations.

Error bars indicate errors between three biological replicates.

The absence of added zinc did not cause inhibition of cell growth, in agreement with previous reports^{32,88,153}. It is possible that open ocean *Synechococcus* are adapted to very low zinc concentrations and are extremely efficient at zinc uptake. However, it is also possible that despite using Chelex a small amount of zinc contamination occurred which was sufficient for cell growth at ‘zero’ zinc.

In contrast, whilst WT *Synechococcus* sp. WH8102 cultures did not grow at 5 $\mu\text{M Zn}^{2+}$, the *zur* mutant failed to grow even at 2.5 $\mu\text{M Zn}^{2+}$. Furthermore, *zur* mutant

growth gradually decreased with increasing zinc concentration whilst the wild type was only slightly affected even at 2.5 $\mu\text{M Zn}^{2+}$.

These observations are in agreement with the idea that SynZur in *Synechococcus* sp. WH8102 is a deficiency sensor which controls zinc uptake through *znuABC* as a zinc-dependent transcriptional repressor. Thus, in the *zur* knockout mutant, *znuABC* is probably always transcribed even at the highest zinc concentrations leading to accumulation of toxic zinc levels in cells which impairs growth (see section 4.3.3.2.1 for the RNA-seq data).

The toxicity data and specific growth rates (μ) for the WT *Synechococcus* sp. WH8102 are similar to those obtained previously¹⁵³. However μ was found to be in the range 0.115-0.125 day^{-1} unlike 0.145-0.159 day^{-1} in this study. This difference could be due to the different culture medium used: ASW (this study) versus Aquil¹⁵³ which contain different phosphate concentrations (0.172 mM and 0.0875 mM, respectively) which could explain the faster growth of *Synechococcus* sp. WH8102 in ASW in the current study. Noteworthy is that in Aquil medium the EDTA concentration is 0.113 mM, higher than the total trace metal concentration¹⁵³. Thus, Zn^{2+} is mostly chelated and its free concentration $[\text{Zn}^{2+}]$ is expected to be several orders of magnitude lower than the total zinc concentration. In ASW the concentration of EDTA is 1.4 μM ¹³⁵. In the absence of other complexing agents, Zn^{2+} is not strongly chelated and the free $[\text{Zn}^{2+}]$ concentration is very similar to the total Zn^{2+} concentration. Even so, in both studies, WT growth was similar across a range of Zn^{2+} concentrations. One possible explanation is that the *Synechococcus* sp. WH8102 zinc uptake system has a higher affinity for Zn^{2+} than EDTA. Therefore, it might be irrelevant for this cyanobacterium whether it deals with free or EDTA-chelated Zn^{2+} . In *Salmonella enterica*, colony growth on LB agar plates was not affected even at 2 mM EDTA, whereas the growth of the *znuA* deletion mutant was severely impaired¹⁵⁴. The addition of TPEN, another chelator which has higher affinity to Zn^{2+} than EDTA at physiological pH¹¹⁶, in turn impaired growth of the *S. enterica* WT.

Alternatively, *Synechococcus* sp. WH8102 cells could effectively capture zinc in both forms: as free zinc or as an EDTA-Zn complex. The ability to take up free EDTA or weaker EDTA-metal complexes and metabolise them using a specific monooxygenase/NADH2:FMN oxidoreductase enzyme system was observed for *Chelatobacter* and *Chelatococcus*¹⁵⁵. However, no such system has so far been

annotated for *Synechococcus* sp. WH8102. In addition, EDTA-Zn is rather a strong complex with a $K_{D \text{ Zn-EDTA}} = 2.3 \times 10^{-14}$ M at pH 7.4¹¹⁶. Which form of zinc is actually being acquired in marine *Synechococcus* will require further work.

4.3.2.2 Metal accumulation

The accumulation of trace metals by WT *Synechococcus* sp. WH8102 and the *zur*-mutant was compared following growth at different zinc concentrations using ICP-MS. Since the SynZur repressor is absent in mutant cells, zinc uptake components like *znuABC* are expected to be always up-regulated compared to WT. Thus, it may be expected that mutant cells would accumulate more zinc. In addition, comparison with accumulation patterns of other trace metal might show if the Zur-regulon is zinc specific or not. Trace metal and phosphorus content normalized to cell dry weight are shown in Table 4.6.

Table 4.6 ICP-MS (moles/kg) assessment of trace metal accumulation.

Data are normalised to the dry weight of cells. Errors refer to three biological replicates.

| | Mutant, 0 Zn | Mutant, 772 nM Zn | WT, 0 Zn | WT, 772 nM Zn | WT, 2.5 μ M Zn |
|-----------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|
| P | $(2.51 \pm 0.06) \times 10^{-1}$ | $(2.30 \pm 0.10) \times 10^{-1}$ | $(1.72 \pm 0.10) \times 10^{-1}$ | $(1.89 \pm 0.07) \times 10^{-1}$ | $(2.05 \pm 0.20) \times 10^{-1}$ |
| Mn | $(1.10 \pm 0.06) \times 10^{-3}$ | $(1.03 \pm 0.02) \times 10^{-3}$ | $(1.04 \pm 0.05) \times 10^{-3}$ | $(1.05 \pm 0.02) \times 10^{-3}$ | $(8.92 \pm 0.63) \times 10^{-4}$ |
| Fe | $(2.80 \pm 0.16) \times 10^{-2}$ | $(2.11 \pm 0.13) \times 10^{-2}$ | $(2.39 \pm 0.23) \times 10^{-2}$ | $(2.55 \pm 0.06) \times 10^{-2}$ | $(2.25 \pm 0.22) \times 10^{-3}$ |
| Co | $(1.55 \pm 0.04) \times 10^{-4}$ | $(1.01 \pm 0.07) \times 10^{-4}$ | $(1.63 \pm 0.02) \times 10^{-4}$ | $(1.41 \pm 0.07) \times 10^{-4}$ | $(1.54 \pm 0.64) \times 10^{-4}$ |
| Ni | $(4.90 \pm 0.95) \times 10^{-5}$ | $(3.41 \pm 0.49) \times 10^{-5}$ | $(2.45 \pm 0.16) \times 10^{-5}$ | $(3.32 \pm 0.55) \times 10^{-5}$ | $(2.59 \pm 0.18) \times 10^{-5}$ |
| Cu | $(2.49 \pm 0.49) \times 10^{-4}$ | $(1.97 \pm 0.05) \times 10^{-4}$ | $(2.28 \pm 0.09) \times 10^{-4}$ | $(2.51 \pm 0.05) \times 10^{-4}$ | $(2.08 \pm 0.04) \times 10^{-4}$ |
| Zn | $(2.53 \pm 0.23) \times 10^{-4}$ | $(1.51 \pm 0.08) \times 10^{-3}$ | $(4.90 \pm 2.39) \times 10^{-5}$ | $(2.28 \pm 0.16) \times 10^{-3}$ | $(7.24 \pm 0.73) \times 10^{-3}$ |
| Cd | $(2.97 \pm 1.07) \times 10^{-7}$ | $(2.00 \pm 1.04) \times 10^{-7}$ | $(2.25 \pm 1.02) \times 10^{-7}$ | $(4.43 \pm 4.64) \times 10^{-8}$ | $(1.22 \pm 1.44) \times 10^{-7}$ |

These values were then normalised to the phosphorus content (per replicate) and the WT and mutant were compared to assess metal accumulation patterns. The results are shown in Figure 4.12.

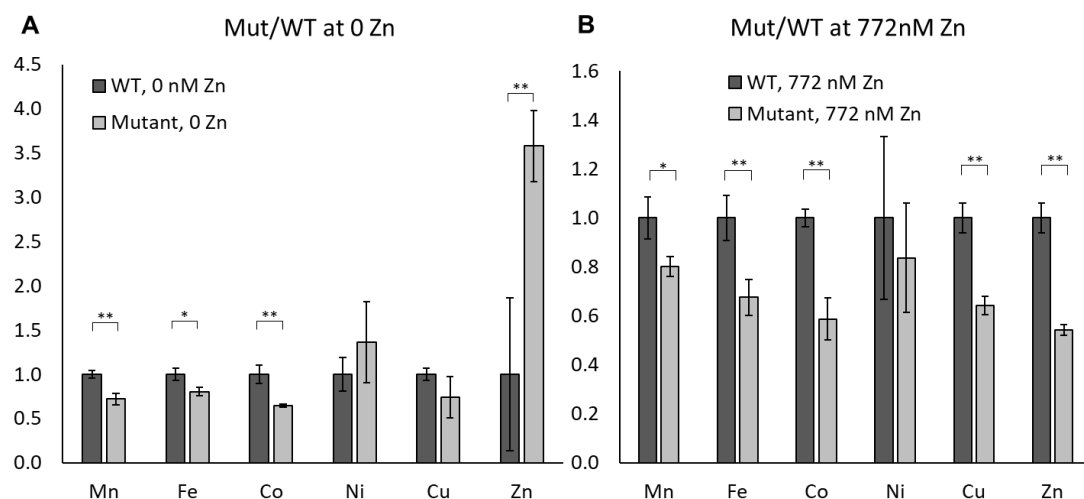


Figure 4.12 Normalised elemental content of the *zur* mutant compared to the *Synechococcus* sp. WH8102 wild type.

* indicates a p-value <0.05; ** indicates a p-value <0.01 (two-tailed t-test, two-sample equal variance). A. Comparison at 0 zinc added. B. Comparison at 772 nM zinc. Error bars indicate st. dev. values between three biological replicates.

Table 4.6 and Figure 4.12A show that even at the lowest possible zinc concentrations achieved following Chelex treatment, both the WT and the mutant acquired measurable quantities of zinc. The mutant accumulated 3.6 times more zinc than the WT, which is likely due to complete de-repression of zinc uptake components (*znuABC*). Lower zinc uptake by the WT suggests that at this “0 Zn” condition *znuABC* is not fully de-repressed in the WT. This was subsequently confirmed (see sections 4.3.3.2 and 4.3.3.3). Interestingly, at 0 added zinc the mutant accumulated less Mn^{2+} , Fe^{3+} and Co^{2+} . It is unclear what could be causing this. However, the opposite patterns of zinc accumulation by the mutant compared to other metals could mean that the de-repressed zinc uptake system is highly selective for zinc despite multi-fold excess of other metals in the medium.

Surprisingly, at 772 nM zinc the opposite situation occurred: the *zur* mutant contained almost half the amount of zinc compared to the WT. A similar picture was observed for other trace metals such as Mn, Fe, Co and Cu. However, the most significant difference was for zinc. This result was unexpected and it was initially unclear why zinc accumulation patterns differed at different zinc concentrations. However later RNA-seq data clarified this (see section 4.3.3.3).

The P-normalised zinc content obtained for WT *Synechococcus* sp. WH8102 and *zur* mutant is shown in Figure 4.13. These values differ to those previously

obtained for this organism following growth in Aquil medium where at 0 and 80 nM zinc these ratios were in the range of 4.5-5 mmol mol⁻¹ and there was no significant difference between cells grown at different zinc concentrations³². Here, I observed a significant difference in zinc content in the WT at different zinc concentrations (0.28 and 12.09 mmol mol⁻¹ for 0 and 772 nM zinc respectively, p-value = 1.18x10⁻⁵). Whilst the value obtained for cells grown in 80 nM zinc in Aquil by Barnett *et al.* (2014)³² (5 mmol mol⁻¹) is in good agreement with the current data, the value observed for 0 nM zinc in Aquil is significantly higher than for 0 nM zinc in ASW (4.5 and 0.28 mmol mol⁻¹ respectively), despite the higher EDTA concentration in Aquil. However, this may simply be due to the fact that Barnett *et al.* (2014) used cell lysates in their ICP-MS analysis compared to whole cell digests used here. Phosphorus content would be lower in the cell lysates due to the absence of membrane phospholipids. Therefore the ratio of Zn/P is expected to be higher for the lysates. In addition, chronic zinc limitation in the previous study could lead to cell adaptation and better efficiency for zinc accumulation. Also, *Synechococcus* cells would grow faster in ASW and the accumulated zinc could be distributed between a larger number of cells giving lower Zn/P ratios. Finally, Aquil-grown cells were harvested at the end of the growth cycle (early-mid stationary phase) while ASW-grown cells were collected at the mid-exponential phase potentially allowing for higher accumulation of zinc in the Aquil-grown cells.

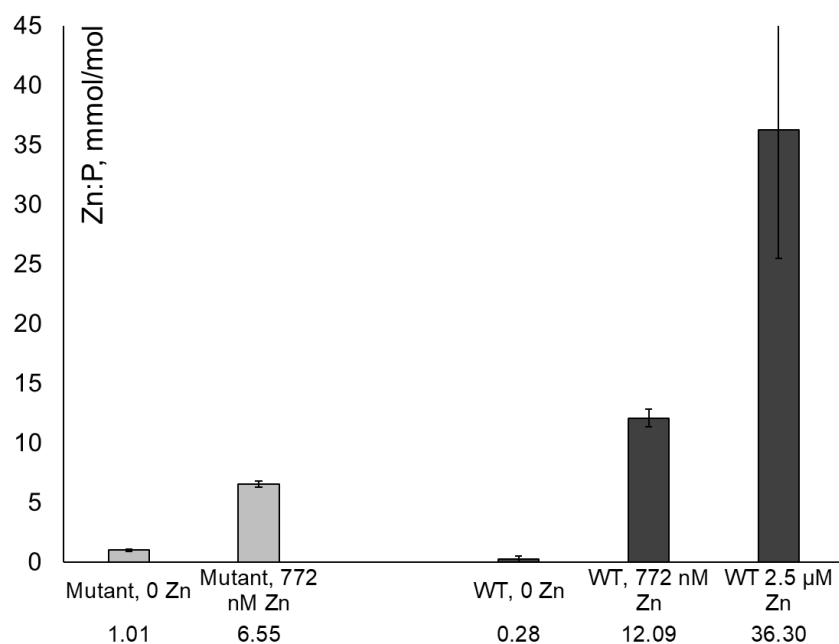


Figure 4.13 Comparison of P-normalised zinc content by ICP-MS.

Interestingly, it was noticed that P-normalised zinc content values for the WT grown at 2.5 mM and 772 μM zinc differed by ~3-fold (36.3 and 12.09 mmol mol⁻¹ P, respectively) similar to the 3.23 fold difference in zinc concentration in the medium. As a result the zinc concentrating factor for these conditions was calculated.

If the cell volume of *Synechococcus* sp. WH8102 cells is assumed to be similar to *Synechococcus* sp. WH7803 which has comparable linear size and shape, then the single cell size is 1.8x10⁻¹⁸ m³ or 1.8x10⁻¹⁵ L¹⁵⁶. Since samples of known OD₇₅₀ were used as well as actual cell counts calculated using a haemocytometer (see section 4.2.1.4), cellular zinc concentrations were then possible to be calculated. These values are given in Table 4.7 together with concentrating factors for the WT grown at 772 nM and 2.5 μM zinc.

Table 4.7 Cellular zinc content and calculation of zinc concentrating factors.

| Sample | Cellular zinc content, M | St. error | ASW zinc conc., M | Concentrating factor |
|--------------------|--------------------------|-----------------------|-----------------------|----------------------|
| Mut, 0 | 1.51*10 ⁻⁴ | 0.12*10 ⁻⁴ | ? | |
| Mut, 772 nM | 8.64*10 ⁻⁴ | 0.49*10 ⁻⁴ | 7.72*10 ⁻⁷ | 1.12*10 ³ |
| WT, 0 | 2.74*10 ⁻⁵ | 1.33*10 ⁻⁵ | ? | |
| WT, 772 nM | 1.27*10 ⁻³ | 0.08*10 ⁻³ | 7.72*10 ⁻⁷ | 1.64*10 ³ |
| WT, 2.5 μM | 4.12*10 ⁻³ | 0.47*10 ⁻³ | 2.50*10 ⁻⁶ | 1.65*10 ³ |

The concentrating factors were surprisingly similar at both concentrations of zinc. If this relationship held for the WT at “0 Zn”, the actual zinc concentration in “0 Zn” ASW could then be estimated as ~17 nM. However it is likely that at lower zinc levels cells are much more efficient in their zinc accumulation. Apparently, cells are able to bio-accumulate zinc from seawater by a factor of at least 10^5 ³².

If accumulation of other metals is compared at different zinc concentrations in the WT and the *zur* mutant, most elements change insignificantly (p-values > 0.05). However, Co uptake was slightly increased at 0 zinc compared to 772 nM zinc in both the WT (1.27 ± 0.09 times higher, p-value = 2.3×10^{-2}) and the mutant (1.40 ± 0.02 , p-value = 9.6×10^{-3}). In addition, iron content was also slightly higher in the mutant at 0 nM zinc compared to 772 nM zinc (1.21 ± 0.04 , p-value = 4.7×10^{-2}). It is possible that excess zinc causes lower uptake of these trace metals due to a selectivity issue and/or competitive uptake. For example, some sensors could respond to different metals as was observed in¹⁵⁷. These authors showed that regulation of *sitABCD* (Mn ABC-type transporter) by *Salmonella enterica* occurs via both MntR and Fur and that its expression is also Co-regulated since Co could mis-metallate Fur. Also, it was suggested that metal sensors are generally tuned to work across a narrow range of metal concentrations and cobalt or zinc shock could cause aberrant responses of the sensors¹⁵⁸. Finally, metal transporters themselves may not entirely be specific and could facilitate uptake of the ‘incorrect’ metal if that metal is present in excess.

4.3.2.3 Assessment of pigment composition via UV-vis absorption spectra

In addition to the different growth behaviour at elevated zinc concentrations shown in section 4.3.2.1, it was noticed that the WT and *zur*-mutant have slightly different pigmentation: at all zinc concentrations the *zur*-mutant appeared slightly more ‘reddish’ (see Figure 4.14A).

Cultures in the early stationary phase of growth were thus analysed by UV-vis spectrophotometry (300-750 nm). The OD₇₅₀ and OD₆₇₀ values were similar, and ultimately normalised, for both WT and mutant cultures (see Figure 4.14B).

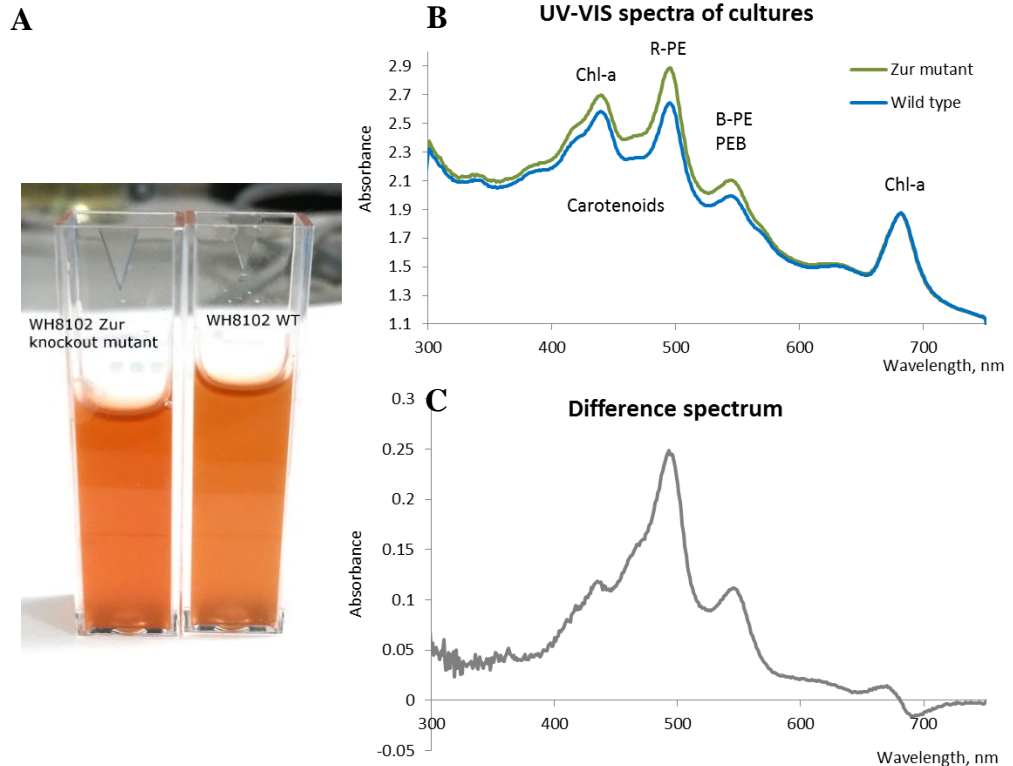


Figure 4.14 Comparison of UV-vis absorption spectra for WT *Synechococcus* sp. WH8102 and the *zur*-mutant.

A. Culture image. **B.** UV-vis spectra of the cultures normalised by OD_{750} . **C.** Difference spectrum of normalized WT and mutant spectra. Chl-*a*: chlorophyll-*a*; R-PE: R-phycoerythrin; B-PE: B-phycoerythrin; PEB: phycoerythrobilin.

Based on the difference spectrum (see Figure 4.14C) the *zur*-mutant contains either more R-phycoerythrin or more carotenoids. Interestingly, it was shown that the Δ_{zur} mutant of *Anabaena* sp. PCC 7120 contained less carotenoids than WT while other photosynthetic pigments measured (chlorophyll *a* and phycobiliproteins) were similar¹⁵⁹. However, it is unclear whether pigment content is regulated by zinc levels and a more detailed investigation of this phenomenon is required.

Given the observed difference in pigment content between mutant and wild type and that carbonic anhydrase, required for CO₂ fixation, is a potential zinc metalloenzyme¹⁶⁰ carbon fixation rates were thus assessed (see section 4.3.2.4.).

4.3.2.4 Comparison of CO₂ fixation rates in the *Synechococcus* sp. WH8102 WT and *zur* mutant

CO₂ fixation rates were measured by calculating the amount of ¹⁴C accumulated by WT and mutant cells at different zinc concentrations and light

intensities and normalised to chlorophyll *a* content. Results are presented in Figure 4.15.

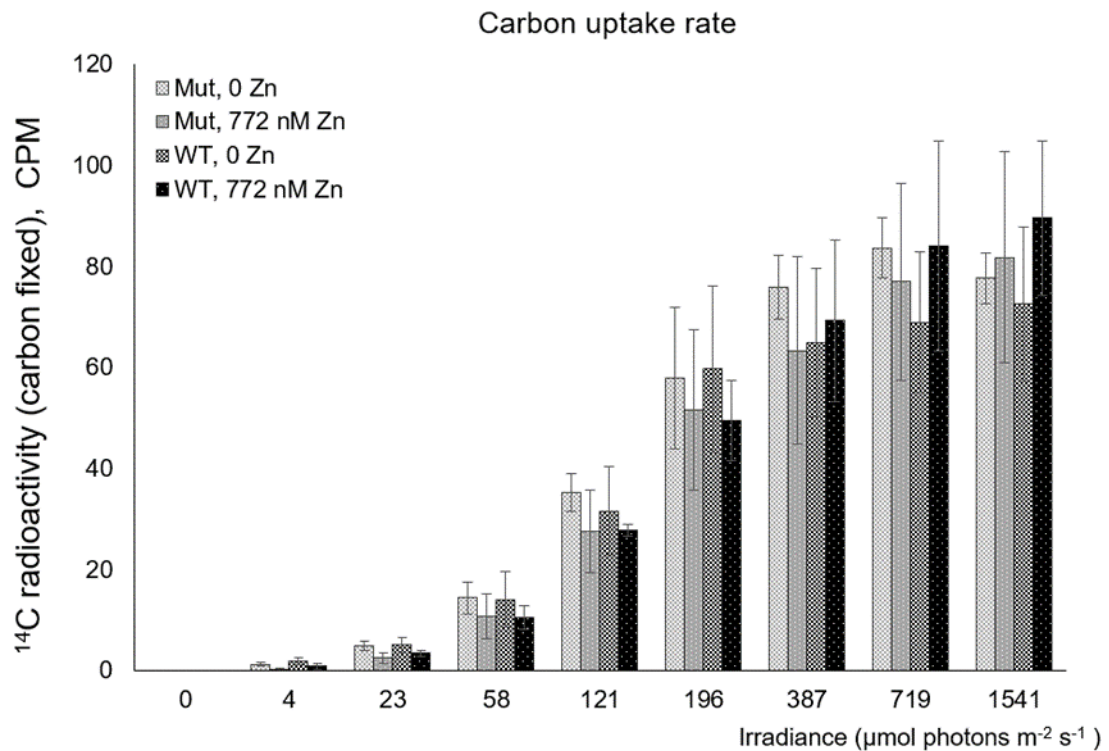


Figure 4.15 CO₂ fixation rate measurements at different light intensities.

Error bars indicate standard errors between three biological replicates.

No significant difference in CO₂ fixation rates between WT and mutant cultures was observed across all light intensities and zinc concentrations. Thus, under the conditions tested CO₂ fixation is independent of zinc or SynZur. However, it is possible that even at the lowest zinc concentration it was sufficient.

4.3.2.5 SDS-PAGE of total cell lysates

SDS-PAGE profiles of whole cell lysates of the WT and *zur* mutant were compared at different zinc concentrations (see Figure 4.16). Previously, it was shown that cell lysates of WT *Synechococcus* sp. WH8102 grown in Aquil medium at 0 and 80 nM zinc added, showed at least two differentially expressed proteins visible on SDS-PAGE gels: a putative outer membrane porin SYNW2224 and the large subunit of RuBisCO (SYNW1718) (see Figure 4.16B, boxes 1 and 2 respectively)³². SDS-PAGE profiles of whole cell lysates and supernatants (for extracellular proteins) obtained in the current study are shown in Figure 4.16A.

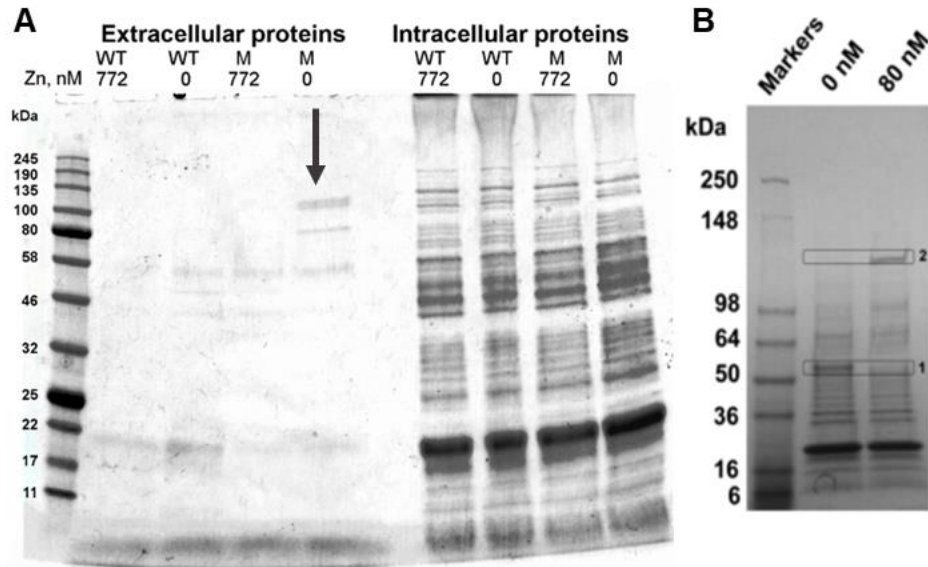


Figure 4.16 SDS-PAGE of total cell lysates and extracellular proteins.

A. Differentially expressed extracellular proteins are marked with a black arrow. **B.** Cell lysates of the WT grown in Aquil medium at different zinc concentrations (taken from Barnett *et al.*, 2014³²). Box 1 corresponds to a putative outer membrane protein SYNW2224, and box 2 the large subunit of RuBisCO.

Interestingly, no visible differences were observed for cell lysates of the WT at 0 and 772 nM zinc in contrast to what was previously observed despite the more significant change in zinc concentration for growth in the current study. It is possible that subtle differences in the experimental design and growth conditions used between the two studies are responsible:

- As mentioned previously Aquil¹⁶¹ contains significantly less phosphorus and nitrogen than ASW (10 μ M/172 μ M and 0.1 mM/8.8 mM respectively) and hence the previously observed zinc dependent differential expression of SYNW2224 and RbcL could also be due to the nutrient-poor conditions used. Indeed, both of these proteins were previously shown to be up-regulated under early P-stress (SYNW2224) and as one of the most strongly repressed after extended P stress (SYNW1718)¹⁶². Therefore, it is possible that their differential expression under different zinc levels is observed only under P-stress due to a possible requirement of zinc for phosphorus utilisation.
- Secondly, cells were grown at 0 and 80 nM Zn²⁺ in Aquil medium for 12 weeks and then sub-cultured into the same Zn²⁺ concentration for the final experiment. In the current study all cells were grown in Zn-depleted medium (non-Chelexed ASW with 0 zinc added) for at least 6 months and for the final

experiment they were transferred from the same flask into Chelexed 0 zinc or 772 nM zinc ASW. Thus, it is possible that protein expression differs under chronic zinc-deprivation and acute Zn-depletion shock. In addition, no distinctly visible difference was observed for either the mutant cell lysates at 0 and 772 nM zinc compared to each other or to the WT.

In the exoproteome, proteins of 80 and 110 kDa were induced at 0 Zn²⁺ in the mutant (see Figure 4.16A) but their identification requires further work.

4.3.3 Zur/Zn regulated transcriptomics of *Synechococcus* sp. WH8102

4.3.3.1 RNA-seq analysis

To identify Zur-regulated genes in *Synechococcus* sp. WH8102, transcriptomics analysis of RNA isolated from the WT and *zur* knockout mutant grown at different zinc concentrations was conducted by RNA-seq. Mutant and WT cells were transferred from non-Chelexed ASW into Chelexed medium with different zinc concentrations as described in section 4.2.4.1. Comparison of WT mRNA profiles at 0 zinc and 772 nM zinc would allow identification of Zn-responsive and potentially Zur-regulated genes. If SynZur is a true zinc sensor, comparison of the *zur*-mutant mRNA profile with the WT would show differential expression of Zur-regulated genes which should be similar to the comparison of the WT at 0 and 772 nM zinc. For example, if *znuABC* belongs to the Zur regulon, in the WT this gene cluster should be up-regulated under zinc-starvation and down-regulated under zinc excess. In the absence of Zur, *znuABC* should always be up-regulated compared to the WT even at sufficient zinc. Transcriptomics at 2.5 μM zinc compared to zinc sufficiency (772 nM) was used to identify genes involved in zinc detoxification as well as genes responding to zinc toxicity. In addition, it is possible that in the mutant differential expression analysis of genes at 0 zinc compared to 772 nM would reveal zinc regulated genes which do not belong to the Zur-regulon.

RNA was extracted from WT and the *zur* mutant cultures as described in section 4.2.4.2 and its integrity analysed using a Bioanalyzer (see Figure 4.17).

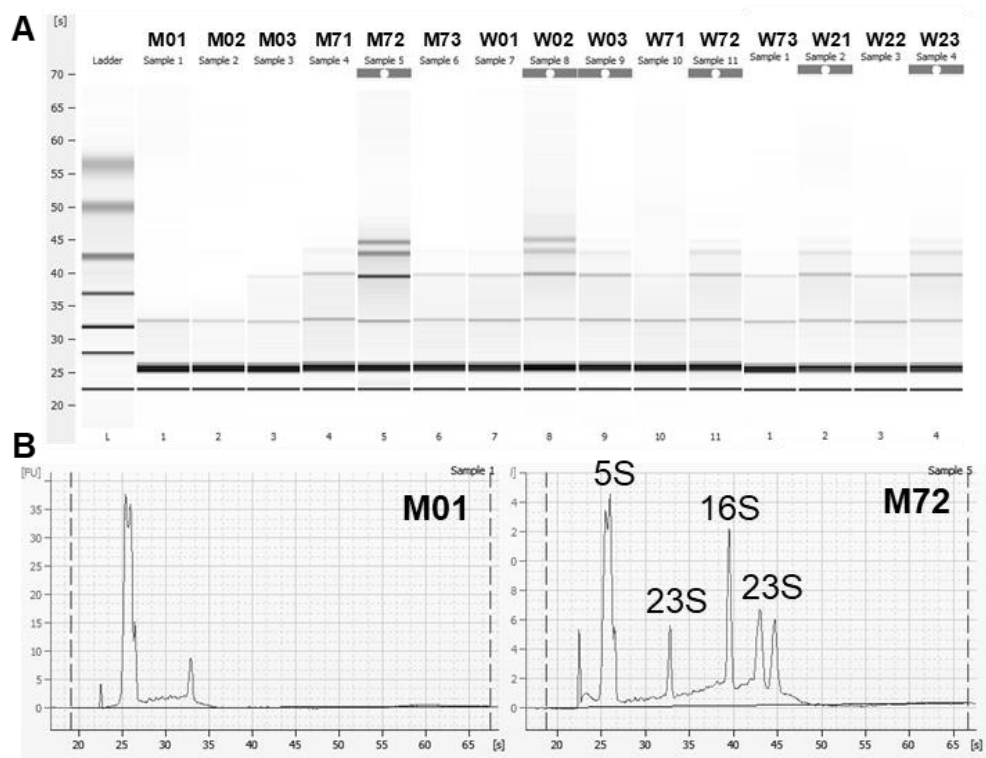


Figure 4.17 Visualisation of total RNA profiles on an Agilent 2100 Bioanalyzer using a RNA 6000 Pico LabChip.

A. Visualisation of all samples as a gel-like image. B. Visualisation of the fluorescence profiles (electrophoretograms) of samples M01 and M72.

Despite the low RNA integrity numbers (RIN) values obtained, 2.3-2.9 instead of typical values of 8 or above, given the general good quality of the mRNA profiles as evidenced from the Bioanalyzer traces and by NanoDrop (see below), samples were sent for sequencing at the Centre for Genomic Research (CGR), Liverpool. The concentration and purity of samples was assessed using NanoDrop and the results are shown in the Appendix (see Table 8.2).

4.3.3.1.1 Preparation of the libraries for RNA-sequencing

At the CGR Liverpool, RNA samples were further purified using a Qiagen RNeasy Kit column which aided library preparation. Agilent Bioanalyzer profiles of the samples following this step are shown in the Appendix (see Figure 8.10). Subsequently, RNA samples were treated with the RiboZero kit (Illumina) to deplete rRNA in the mRNA samples (see the Appendix Figure 8.11). The NEBNext Ultra Directional RNA library preparation kit was then used for preparation of dual-indexed, strand-specific RNA-seq libraries (see Figure 8.12, Appendix) and sequenced on an

Illumina HiSeq 4000 (Paired-end, 2x150 bp). The read data were post processed: low quality bases were removed as well as sequencing adapters. Resulting FASTQ files used for differential expression analysis were uploaded at http://cgr.liv.ac.uk/illum/LIMS16056_259a058713a41d9b/.

Post-processing statistics including total number of reads, amount of paired, singlet and discarded reads are shown in the Appendix (Figure 8.13).

4.3.3.1.2 Data processing and analysis

Original FASTQ files were checked with FastQC¹⁶³ software to confirm read quality. All the reads satisfied the quality score criteria (data not shown). The analysis was then performed as described in section 4.2.4.3.1. RNA-seq mapping and analysis statistics are shown in Table 8.3, Appendix.

4.3.3.2 Identifying the *Synechococcus* sp. WH8102 Zur-regulon

Under sufficient zinc it is expected that SynZur would be in its fully metallated form with high DNA-binding affinity repressing transcription. Under this condition Zur-regulated genes are therefore expected to be de-repressed in the *zur* knockout mutant since SynZur is absent and unable to affect gene transcription. The DESeq2 tool allowed identifying genes whose expression changed significantly (modulus of $\log_2\text{fold} > 2$, $p\text{-value} < 0.05$). The volcano plot of differentially expressed genes in the mutant compared to WT at sufficient zinc is presented in Figure 4.18. The list of genes with $|\log_2\text{Fold}| > 2$ with $p\text{-value} < 0.05$ is given in Table 4.8.

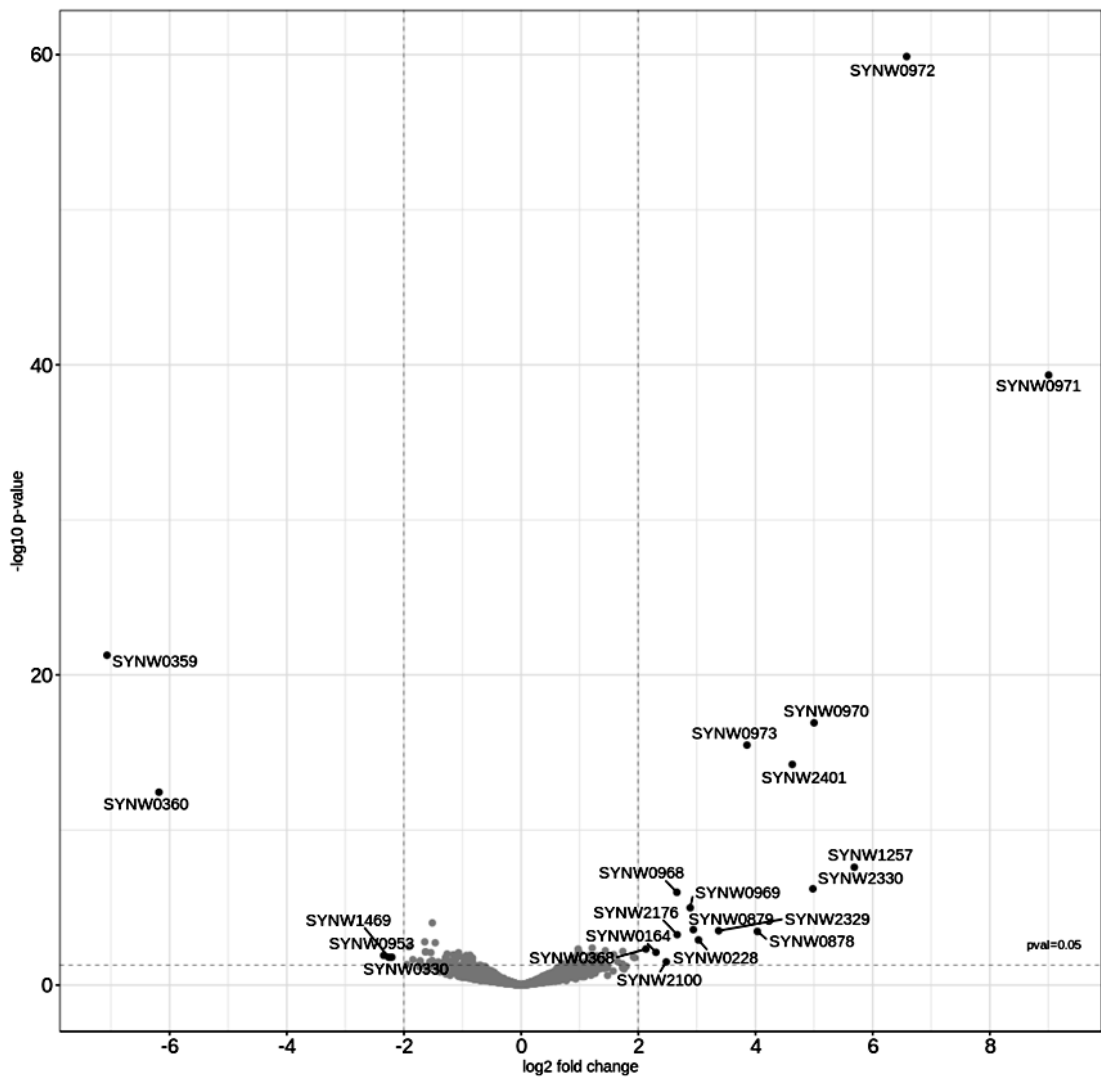


Figure 4.18 Genes differentially expressed in the *Synechococcus* sp. WH8102 *zur* mutant compared to wild type grown under 772 nM zinc.

Table 4.8 Genes differentially expressed in the *zur* knockout mutant of *Synechococcus* sp. WH8102 compared to the wild type under 772 nM zinc.

Only genes with $|\log_2\text{FoldChange}| > 2$ and $p\text{-value} < 0.05$ are given. Adjusted p -values calculated by DESeq2 software (padj) are also given. The colour scheme separates genes into different groups as explained in sections 4.3.3.2.1 and 4.3.3.2.2. If the gene has at least one predicted *zur*-box, its position and the number of mismatches with the predicted *zur*-box of *Synechococcus* sp. WH8102 are given (see below for the explanation).

| Gene | Log2 Fold Change | p-value | padj | gene name | Function | Zur-box position, mismatches |
|----------|------------------|----------|----------|-----------|--|---|
| SYNW0971 | 9.00 | 4.61E-40 | 5.68E-37 | ZnuA | ABC transporter, substrate binding protein, possibly Mn | 37, 0 |
| SYNW0972 | 6.58 | 1.32E-60 | 3.25E-57 | | Uncharacterized protein | 289, 2 |
| SYNW1257 | 5.69 | 2.52E-08 | 7.76E-06 | | Uncharacterized protein | - |
| SYNW0970 | 5.00 | 1.23E-17 | 7.55E-15 | ZnuC | ABC transporter, ATP binding domain, possibly Mn transport | -83, 0 |
| SYNW2330 | 4.98 | 6.22E-07 | 1.70E-04 | | Uncharacterized protein | - |
| SYNW2401 | 4.63 | 5.90E-15 | 2.42E-12 | Zur | Ferric uptake regulator family | - |
| SYNW0878 | 4.03 | 3.48E-04 | 5.71E-02 | | Uncharacterized protein | - |
| SYNW0973 | 3.86 | 3.34E-16 | 1.65E-13 | | Uncharacterized protein | - |
| SYNW2329 | 3.37 | 3.14E-04 | 5.52E-02 | pgk | Phosphoglycerate kinase (EC 2.7.2.3) | - |
| SYNW0228 | 3.03 | 1.24E-03 | 1.79E-01 | | Conserved hypothetical | - |
| SYNW0879 | 2.94 | 2.66E-04 | 5.04E-02 | | Uncharacterized protein | - |
| SYNW0969 | 2.89 | 1.04E-05 | 2.32E-03 | ZnuB | ABC transporter component, possibly Mn transport | Co-transcribed, SYNW0970 |
| SYNW2176 | 2.66 | 5.64E-04 | 8.68E-02 | | Possible serine protease (EC 3.4.21.-) | - |
| SYNW0968 | 2.66 | 1.04E-06 | 2.57E-04 | | Uncharacterized protein | - |
| SYNW2100 | 2.48 | 3.22E-02 | 9.99E-01 | | Uncharacterized protein | -663, 3 |
| SYNW0164 | 2.30 | 7.71E-03 | 6.70E-01 | | Uncharacterized protein | -124, 3 |
| SYNW0368 | 2.13 | 4.73E-03 | 5.06E-01 | | Uncharacterized protein | - |
| SYNW0330 | -2.21 | 1.60E-02 | 8.57E-01 | hli2 | Possible high light inducible protein | - |
| SYNW1469 | -2.26 | 1.60E-02 | 8.57E-01 | | Uncharacterized protein | - |
| SYNW0953 | -2.34 | 1.21E-02 | 7.91E-01 | swmB | SwmB-cell surface protein required for swimming motility | 14 boxes with 2 mismatches in the coding region |
| SYNW0360 | -6.18 | 3.67E-13 | 1.29E-10 | | Weak similarity to phage Integrase family | -545, 2 |
| SYNW0359 | -7.07 | 5.36E-22 | 4.40E-19 | bmtA | Bacterial metallothionein | -79, 0 |

Disruption of *zur* may lead to other consequences such as incorrect expression of other Zur-regulated transcription factors and/or disturbance of metal and redox homeostasis affecting expression of Zur-independent genes³⁵. Therefore, to identify members of the Zur-regulon it is also important to show binding of Zur to the promoter regions of the genes³⁵.

Here, binding of SynZur to promoter regions was only tested experimentally for *pznuABC* and *pbmtA* (see sections 3.3.2.7 and 5.3.1). However, bioinformatics was also used to scan the *Synechococcus* sp. WH8102 genome for other potential zur-boxes using PRODORIC¹⁶⁴ bacterial regulon analyser server. From this analysis a sequence, ANAATNANNATNATTNT, was derived based on three known boxes (see Figure 8.14B, Appendix). Positions of some of the newly identified *zur* boxes with a number of mismatches are given in Table 4.8 together with the differential expression analysis results.

However, the absence of a predicted zur-box in the promoter region of a gene does not necessarily mean that the gene does not belong to the Zur-regulon. Bioinformatics approaches often do not reveal all of the binding boxes. For example, in Gilston *et al.* (2014)⁴⁸ only analysis of the crystal structure of *pznuABC* - Zur allowed identification of a zur-box in the promoter region of the periplasmic lysozyme inhibitor *pliG* which previously was not identified bioinformatically. A gene could also be co-transcribed together with a gene which has a zur-box. Thus, SYNW0969 (*ZnuB*) does not have a predicted zur-box but it is likely to be co-transcribed together with SYNW0970 which possesses one³⁸. Finally, a transcriptional interference mechanism might be involved where the binding of RNA-polymerase to a target promoter is affected by an interfered promoter located elsewhere¹⁶⁵.

4.3.3.2.1 Genes up-regulated in the mutant compared to the WT

zur (SYNW2401) itself (see Table 4.8) is up-regulated in the mutant compared to WT, though reads map only to the first 315 bp of the gene, consistent with the single crossover nature of this mutant as checked by the Integrative Genomic Viewer¹⁶⁶ software (see Figure 4.19). The first copy of *zur* is incomplete, thus no complete mRNA/protein product is expected whilst the second copy lacks the first 21 bp of the gene and any promoter and Shine-Dalgarno sequences, so should not be expressed.

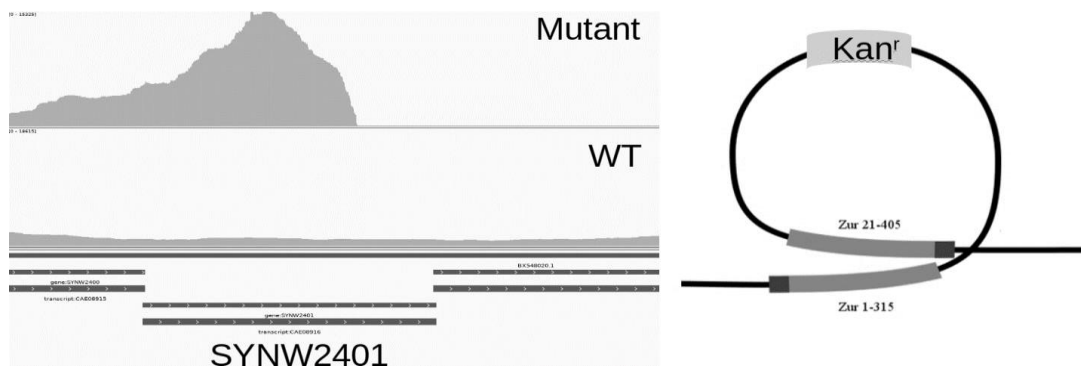


Figure 4.19 High expression of SYNW2401 in the *zur* knockout mutant compared to the wild type.

A. Profiles of mapped reads checked by the IGV software. B. Schematic of the single crossover *zur* disruption.

It is unclear why SYNW2401 is up-regulated in the mutant. One possibility is that it could regulate its own expression as reported for *Streptomyces coelicolor*¹¹⁴. However, no potential Zur binding boxes have been found in the operator-promoter region of SYNW2401 nor was any binding by EMSA observed⁸⁸. SynZur could also be a part of a larger regulatory network where transcription of *zur* is regulated by some other transcription factor(s). It should be noted however, that *zur* expression is not zinc regulated since SYNW2401 was not up-regulated when comparing WT growth at 0 and 772 nM zinc (see section 4.3.3.3).

A second group of genes (shown in red letters in Table 4.8) are three up-regulated genes **SYNW0969**, **SYNW0970**, **SYNW0971** encoding components of a *znuABC* high affinity zinc transporter. These genes were previously suggested to be Zur-regulated because they contain a cyanobacterial Zur-binding box in their operator-promoter region³⁸. Moreover, this region was confirmed experimentally (see section 3.3.2.7) to bind SynZur (see Ksibe (2016)⁸⁸). These genes are expected to be up-regulated in the mutant compared to the WT since under sufficient zinc levels cells do not need to produce the machinery for specific zinc uptake and SynZur with zinc bound represses their transcription. However, when SynZur is absent regulation is lacking and cells continue to produce *znuABC*.

Another group of genes up-regulated in the *zur* mutant (highlighted in orange in Table 4.8) were located in the same cluster as the *znuABC* system: **SYNW0968**, **SYNW0972** and **SYNW0973** (see Figure 4.20).

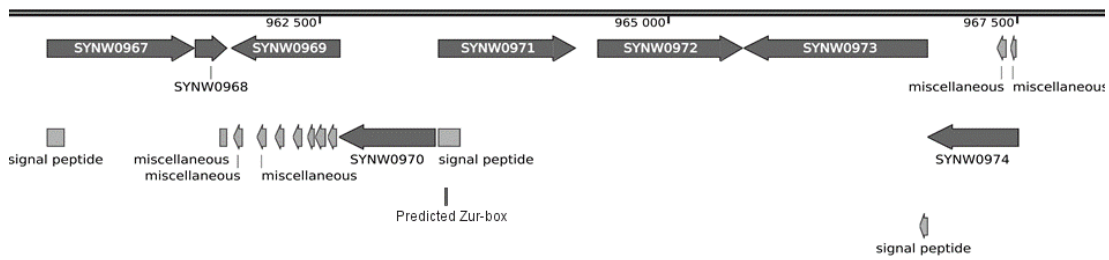


Figure 4.20 Location of SYNW0968, SYNW0972 and SYNW0973 on the genome.

The figure was generated using SnapGene4.3.3¹⁶⁷ software. Miscellaneous features shown with light grey arrows correspond to probable transmembrane helices predicted by TMHMM2.0.

All three genes are annotated as hypothetical proteins with no known function. However, further bioinformatics analysis gave potential clues to their function (see section 4.3.3.4). It is unclear how these genes are regulated since no zur-boxes are found in their operator-promoter regions. Several potential zur-boxes with a few mismatches were identified in the coding region of SYNW0972 (see Figure 8.14A, Appendix). However, it is unclear if they actually play a role in transcriptional repression or how that regulation occurs. It is possible that a transcriptional interference mechanism is involved¹⁶⁵. In addition, based on read coverage analysis using IGV software, SYNW0972 could be co-transcribed with SYNW0971 (*znuA*) since there is no gap in coverage downstream of SYNW0971 and upstream of SYNW0972 (see Figure 8.15, Appendix).

The last group of genes up-regulated ($\log_2\text{fold} > 2$, $p\text{-value} < 0.05$) in the mutant, were several genes located at random places along the genome: **SYNW1257**, **SYNW2330**, **SYNW0878**, **SYNW2329**, **SYNW0228**, **SYNW0879**, **SYNW2176**, **SYNW2100**, **SYNW0164**, **SYNW0368**. It should be noted that most of these genes had adjusted p -values > 0.05 : namely SYNW0878, SYNW2329, SYNW0228, SYNW0879, SYNW2176, SYNW2100, SYNW0164, SYNW0368 (see Table 4.8). The adjusted p -values are calculated by DESeq2 software automatically using the Benjamin-Hochberg method to cut off false positive results. Thus, only two genes from this group (SYNW1257, SYNW2330) are probably differentially expressed with high level of significance. However, these genes were retained to see if they were also found to be differentially expressed in comparisons of other growth conditions (see section 4.3.3.3).

The function of most of these genes is unknown except for SYNW2329 annotated as phosphoglycerate kinase and SYNW2176 annotated as a putative serine

protease. It is not clear why these might be regulated by *zur* if at all. One of the possibilities is that these enzymes are zinc-free alternatives which could be beneficial for cells under conditions of zinc deficiency³⁵.

4.3.3.2.2 *Genes down-regulated in the mutant compared to the WT*

Surprisingly, several genes were also found to be down-regulated in the *zur*-mutant at sufficient (772 nM) zinc. This means that in the WT under conditions of sufficient zinc metallated SynZur binds DNA and activates these genes (see section 5.3.1). In the absence of SynZur, such activation does not take place in the *zur* mutant causing down-regulation of these genes compared to the WT. Activation by Zur in cyanobacteria has not been shown before. However, previous reports show that Zur can be an activator in other phyla, including *Xanthomonas campestris*⁷² and *Streptomyces coelicolor*⁷³. In addition, González *et al.* (2016) showed that Fur can work as an activator in *Anabaena* sp. PCC 7120¹⁶⁸.

One of the most interesting down-regulated genes was **SYNW0359** encoding a predicted bacterial metallothionein (*bmtA*) (see Table 4.8). This is the first time that a metallothionein has been experimentally shown to be regulated, and specifically activated, by Zur. MTs were previously predicted to be Zur-regulated in this strain and for *Cyanothece* sp. PCC 8801³⁸, leading to speculation that bacterial metallothioneins might be produced in response to zinc starvation³³. Here, we confirm regulation by Zur and show (see sections 4.3.3.3 and 5.3.1) that metallated SynZur binds to *pbmtA* and activates expression of *bmtA* under conditions of zinc sufficiency and excess.

As indicated previously, in several other species cyanobacterial metallothioneins are regulated by the zinc excess sensor SmtB³⁸: *Cyanothece* sp. PCC 8801, *Synechococcus elongatus* PCC 7942, *Synechococcus* sp. PCC 7002 and *Thermosynechococcus elongatus* BP-1. No predicted SmtB homologues were identified in several marine *Synechococcus* species: WH8102, WH8109, WH7803, CC9311, CC9605³³. Based on my results in this section it is clear that SynZur can play a dual role: as a repressor of genes involved in zinc uptake (*znuABC*) and as an activator of genes dealing with zinc toxicity (*bmtA*).

Previously it was shown that *pbmtA* has two *zur*-boxes very similar to the single *pznuABC* box³⁸ (see Figure 4.21).

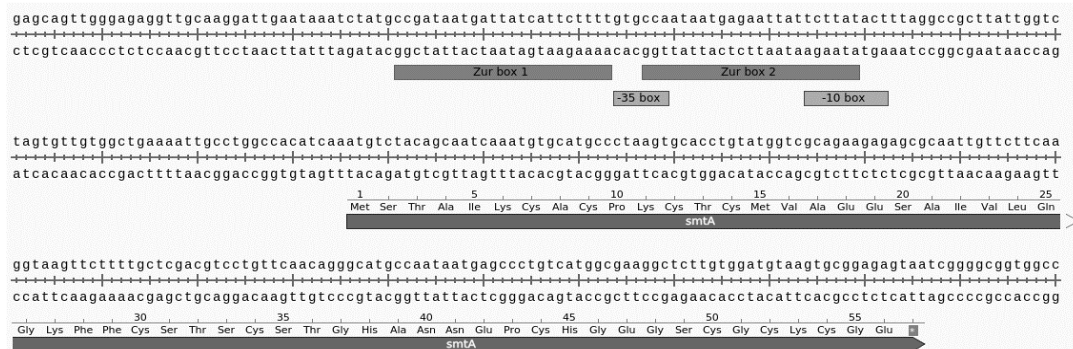


Figure 4.21 Location of zur-boxes in *pbmtA* and their position towards RNA-polymerase binding site (-10 and -35 elements).

It can be seen in Figure 4.21 that the zur-boxes cover the -10 and -35 elements of the promoter region and it was previously suggested that the metallated SynZur would rather block transcription of *bmtA* upon binding so that *bmtA* is de-repressed as a response to zinc limitation³³. However, the RNA-seq analysis revealed that *bmtA* is being repressed in the mutant and activated in the WT under high zinc conditions (see section 4.3.3.3).

The mechanism of how this activation may work is unclear. However, as suggested in section 3.2.4.1, possibly formation of a tetramer initiated by higher zinc concentrations could play a role here and the presence of two zur-boxes might be crucial for binding a tetramer and activation. A somewhat similar mechanism was observed in *S. coelicolor* when activation of a zinc efflux pump was observed upon binding of several Zur dimers in the promoter region of that gene⁷³. Interestingly, an alternative way for Zur activation takes place in *Xanthomonas campestris*: Zur binds to a 59-bp GC-rich sequence that is different to the canonical zur-box, and activates XC2976, a putative CDF-type cation efflux pump⁷². Moreover, a transcriptional interference mechanism could also be involved for activation¹⁶⁵.

It should be mentioned that the presence of two zur-boxes located next to each other is not a necessary attribute for activation. For example, the RegPrecise database shows that four *Streptomyces* species contain two zur-boxes in the upstream region of *znuABC* and *zur*³⁸. However, the location of the boxes is -17, -39 (22 nt difference) for *Streptomyces* while it is -79, -105 (26 nt) for *bmtA* in *Synechococcus* sp. WH8102. Several species from the *Bacillaceae* family (including *Bacillus amyloliquefaciens* and *Bacillus subtilis*) also have 2 zur-boxes in the promoter region of *adcA* (analogous

to the *znuA* component of a zinc uptake system) and *Paenibacillus* sp. JDR-2 even contains four *zur*-boxes³⁸. Even some cyanobacteria, e.g. *Cyanothece* sp. ATCC 51142 and *Anabaena* sp. PCC 7120, contain two *zur*-boxes upstream of *znuABC* (-46, -72 (26 nt) and -80, -110 (30 nt) respectively)³⁸. The latter has *zur*-boxes at positions that are very similar to *pbmtA* *zur*-boxes. Moreover, transcriptional repression of *znuA* in *Anabaena* sp. PCC 7120 by *Zur* was even confirmed experimentally: all0833 was de-repressed in the *zur*-mutant⁵⁹.

Interestingly, in *Anabaena* sp. PCC 7120 the position of one *zur*-box in *pznuA* in the region of the promoter elements (-10 and -35) is similar to the *Synechococcus* sp. WH8102 *pbmtA* box: it is located between two promoter elements (-10 and -35) (see Figure 8.16, Appendix). However, in *Anabaena* this *Zur* box completely covers the -10 element and does not intersect with the -35 box at all while in *Synechococcus* sp. WH8102 the *zur*-box is centred in between both promoter elements and slightly intersects both (see Figure 4.21). However, the most dissimilar feature is the location of a second *zur*-box which in *Anabaena* is located between the promoter elements and the start codon of *znuA*. Thus, it could probably block transcription. In *Synechococcus* sp. WH8102 the second *zur*-box is located even more upstream compared to the promoter elements (see Figure 4.21). Thus, it possibly serves some function other than blocking. However, more detailed analyses and experimental confirmation are required.

Another gene, which was strongly down-regulated in the mutant was **SYNW0360**. This gene is annotated as showing “Weak similarity to phage Integrase family” and located downstream of SYNW0359. The only *zur*-box which was identified in the operator-promoter region of SYNW0360 by the PRODORIC¹⁶⁴ web service was with two mismatches (CAATTGATGATGATTGT) and is located at -545 bp upstream of the gene (see Figure 8.17, Appendix). Two possible sets of promoter elements were identified by the BPRM Sofberry¹⁶⁹ web service at -30 and -346 and probably the -30 one is more likely to be the real promoter. However, further work is needed to clarify if this *zur*-box is relevant for SYNW0360 and how the regulation occurs.

Analysis of the coverage of SYNW0360 and SYNW0359 by IGV¹⁶⁶ (see Figure 8.18, in the Appendix) showed that SYNW0360 and SYNW0359 have similar expression profiles. However, the fact that these genes are convergently located with

independent promoter regions makes their co-expression unlikely. The possible function of SYNW0360 is discussed in section 4.3.3.4.

Finally, three other genes were down-regulated in the mutant: **SYNW0330** (a putative high light inducible protein), **SYNW1469** (an uncharacterized protein) and **SYNW0953** (encoding the SwmB-cell surface protein required for swimming motility). However, the adjusted p-values for all three genes are very high (0.79-0.86) indicating that these might be false-positive results. Secondly, log₂fold values of all three genes are only slightly lower than the threshold values of -2 indicating only moderate down-regulation. No zur-boxes were identified in the operator-promoter regions of these genes with three or less mismatches. In contrast, SYNW0953 contained 14 boxes with two mismatches in its coding region. However, this gene is very large (32376 nt) and hence these boxes could merely be a result of this large size. Nonetheless, potentially, if *swmB* was being activated by SynZur, it is possible that this is a protective chemotaxis mechanism, with cells producing a swimming motility protein to move away from higher zinc levels.

4.3.3.3 Validating the Zur-regulon

To further validate the Zur regulon, comparisons of gene expression between the mutant and WT grown at 0 zinc were also made. The Volcano plot showing differential expression of genes in the *zur* knockout mutant compared to the *Synechococcus* sp. WH8102 WT at 0 zinc is given in Figure 8.19 in the Appendix and the list of genes with |log₂fold| values >2 and p-values <0.05 is given in Table 4.9. Most genes identified above (see section 4.3.3.2) were also identified here with the *znuABC* cluster and associated genes (SYNW0968-SYNW0973) up-regulated, and SYNW0359/SYNW0360 down-regulated. It should be noted that in this comparison the absolute fold change values for these genes were lower than for the comparison at sufficient zinc. Interestingly, for SYNW2401 the fold change value was higher at 0 zinc than at 772 nM zinc.

Several new differentially expressed genes were identified in this comparison. However, it is unclear why they were not identified previously (see section 4.3.3.2) and indeed if they are regulated by SynZur at all. Possibly, the presence of kanamycin in the mutant culture could also affect their expression. No zur-boxes with three or less mismatches were identified in the promoter region of these genes except

SYNW1698 where one zur-box was found at position -302 upstream of the gene and one in the intergenic region. It is unclear whether this box is functional or not.

Table 4.9 Genes differentially expressed in the *zur* knockout mutant of *Synechococcus* sp. WH8102 compared to the wild type grown under 0 zinc.

If the gene has at least one predicted zur-box (RegPrecise³⁸ or PRODORIC¹⁶⁴), its position and the number of mismatches are given.

| Genes | Log2Fold Change | p-value | padj | Gene name | Gene function | Zur-box |
|----------|-----------------|-----------|-----------|-----------|--|--------------------------|
| SYNW0971 | 5.76 | 2.85E-101 | 3.57E-98 | ZnuA | ABC transporter, substrate binding protein, possibly Mn | 37, 0 |
| SYNW2401 | 5.38 | 1.05E-274 | 2.63E-271 | Zur | Ferric uptake regulator family | - |
| SYNW0972 | 4.84 | 9.34E-63 | 3.91E-60 | | Uncharacterized protein | 289, 2 |
| SYNW0970 | 4.00 | 1.64E-83 | 1.37E-80 | ZnuB | ABC transporter, ATP binding domain, possibly Mn transport | -83, 0 |
| SYNW0973 | 3.64 | 2.02E-52 | 7.24E-50 | | Uncharacterized protein | - |
| SYNW1559 | 2.49 | 1.14E-14 | 5.49E-13 | | Uncharacterized protein | 544, 2 |
| SYNW1560 | 2.48 | 2.22E-07 | 2.88E-06 | | Uncharacterized protein | -447, 2 |
| SYNW1137 | 2.42 | 1.76E-13 | 7.13E-12 | | Uncharacterized protein | - |
| SYNW1382 | 2.38 | 1.33E-39 | 3.34E-37 | | Conserved hypothetical | - |
| SYNW0969 | 2.38 | 1.73E-31 | 3.62E-29 | ZnuC | ABC transporter component, possibly Mn transport | Co-transcribed, SYNW0970 |
| SYNW0968 | 2.27 | 2.89E-28 | 4.84E-26 | | Uncharacterized protein | - |
| SYNW0483 | 2.09 | 2.60E-02 | 6.44E-02 | apcC | Phycobilisome 7.8 kDa linker polypeptide, allophycocyanin-associated, core | - |
| SYNW1698 | 2.09 | 6.92E-12 | 2.12E-10 | | Uncharacterized protein | -302, 2 |
| SYNW1859 | -2.02 | 2.87E-03 | 1.06E-02 | ctaB | Protoheme IX farnesyltransferase (EC 2.5.1.-) (Heme B farnesyltransferase) | - |
| SYNW1319 | -2.03 | 2.02E-31 | 3.89E-29 | | Uncharacterized protein | - |
| SYNW1453 | -2.08 | 8.84E-08 | 1.28E-06 | | Uncharacterized protein | - |
| SYNW0156 | -2.17 | 1.94E-06 | 1.92E-05 | | Alpha-1,4 glucan phosphorylase (EC 2.4.1.1) | - |
| SYNW1374 | -2.83 | 1.98E-64 | 9.92E-62 | | Uncharacterized protein | - |
| SYNW1373 | -2.88 | 1.52E-71 | 9.55E-69 | | Conserved hypothetical | - |
| SYNW0519 | -2.91 | 8.05E-05 | 4.99E-04 | gpml | 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (EC 5.4.2.12) | - |
| SYNW0359 | -3.70 | 4.96E-52 | 1.55E-49 | smtA | Bacterial metallothionein | -79, 0 |
| SYNW0360 | -3.89 | 6.24E-45 | 1.74E-42 | | Weak similarity to phage integrase family | -545, 2 |

Further confirmation of the Zur-regulon compared wild type *Synechococcus* sp. WH8102 grown at 0 and 772 nM zinc. The volcano plot showing the differential expression of genes in WT *Synechococcus* sp. WH8102 at 0 and 772 nM zinc is given

in Figure 8.20 in the Appendix, and the list of up/down-regulated genes is given in Table 4.10.

Table 4.10 Genes differentially expressed in WT *Synechococcus* sp. WH8102 at 0 added zinc compared to 772 nM zinc.

Only genes with $|\text{Log}_2\text{Fold}| > 2$ and p-values < 0.05 are given. If the gene has at least one predicted zur-box (RegPrecise³⁸ or PRODORIC¹⁶⁴), its position and the number of mismatches are given.

| Gene | Log2Fold Change | p-value | padj | Gene name | Gene function | Zur-box |
|----------|-----------------|----------|----------|-----------|--|---------|
| SYNW0971 | 3.99 | 7.40E-13 | 9.15E-10 | ZnuA | ABC transporter, substrate binding protein, possibly Mn | 37, 0 |
| SYNW0972 | 3.45 | 7.73E-19 | 1.91E-15 | | Uncharacterized protein | 289, 2 |
| SYNW0203 | 2.06 | 1.51E-03 | 1.43E-01 | psbF | Cytochrome b559 subunit beta | - |
| SYNW0205 | 2.05 | 1.32E-02 | 5.04E-01 | | Ycf48-like protein | - |
| SYNW1424 | 2.04 | 7.75E-06 | 6.39E-03 | | Uncharacterized protein | - |
| SYNW0121 | 2.03 | 5.63E-03 | 3.66E-01 | ssb | Single-stranded DNA-binding protein | - |
| SYNW0798 | -2.02 | 2.04E-02 | 5.15E-01 | | Putative transcriptional regulator, ArsR family | - |
| SYNW1898 | -2.06 | 3.25E-02 | 5.31E-01 | | CysQ protein homolog | - |
| SYNW0330 | -2.08 | 2.35E-02 | 5.15E-01 | hli2 | Possible high light inducible protein | - |
| SYNW2414 | -2.10 | 1.81E-02 | 5.15E-01 | | Uncharacterized protein | - |
| SYNW1469 | -2.20 | 1.61E-02 | 5.05E-01 | | Uncharacterized protein | - |
| SYNW2406 | -2.20 | 2.22E-02 | 5.15E-01 | | Uncharacterized protein | - |
| SYNW1407 | -2.24 | 2.98E-02 | 5.31E-01 | | Uncharacterized protein | - |
| SYNW0216 | -2.44 | 1.40E-02 | 5.05E-01 | folK | Possible 2-amino-4-hydroxy-6-hydroxymethylidihydropteridine pyrophosphokinase (EC 2.7.6.3) | - |
| SYNW2028 | -2.50 | 1.07E-03 | 1.21E-01 | rlpA | Probable endolytic peptidoglycan transglycosylase RlpA (EC 4.2.2.-) | - |
| SYNW0721 | -2.83 | 1.61E-02 | 5.05E-01 | petG | Cytochrome b6-f complex subunit 5 | - |
| SYNW0359 | -2.88 | 8.55E-05 | 3.25E-02 | smtA | Bacterial metallothionein | -79, 0 |
| SYNW0360 | -2.92 | 1.73E-05 | 1.07E-02 | | Weak similarity to phage integrase family | -545, 2 |
| SYNW0916 | -3.18 | 1.14E-02 | 4.93E-01 | | Uncharacterized protein | - |

In this comparison, only two members of the previously found cluster were identified as up-regulated at 0 zinc (SYNW0971 and SYNW0972). Both previously found down-regulated genes, SYNW0359 and SYNW0360, are also in this list. For all of these, absolute values of log2fold change were lower compared to the data which compared the *zur*-mutant with the WT, as predicted. These three comparisons Mut772/WT772, Mut0/WT0 and WT0/WT772 together prove that these genes are regulated by SynZur in a zinc-dependent manner.

Interestingly, two genes, SYNW0330 and SYNW1469, were found to be down-regulated when comparing growth of the WT at 0 and 772 nM zinc in the same way as comparing the mutant and WT at 772 nM. However, adjusted p-values are >0.05 for these.

Venn diagram plots for the three comparisons above are shown in Figure 4.22. The modulus of fold change for differentially expressed genes goes down in the order Mut/WT at 772 nM zinc > Mut/WT at 0 nM zinc > WT at 0/WT at 772 nM zinc. For example, for *znuA* (SYNW0971) it is 9.00, 5.76 and 3.99 respectively. Thus, it is possible that some genes could have the modulus of log₂fold change lower than two and be missing from the previous comparisons. To avoid this, it was decided to use all the genes which were up-regulated or down-regulated with p-values <0.05.

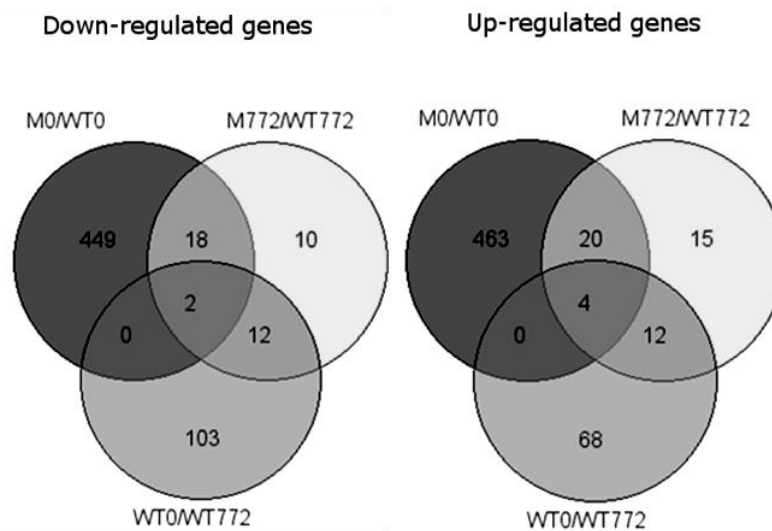


Figure 4.22 Venn diagrams of three different comparisons: Mutant/Wild Type at 0 zinc, Mutant/Wild Type at 772 nM zinc and Wild Type at 0/Wild Type at 772 nM.

The list of genes common to all three comparisons is shown in the Appendix (see Table 8.4). Noteworthy is that for the M0/WT0 comparison, the number of differentially expressed genes with p-values <0.05 was significantly higher than for all the other comparisons. This means that M0 replicates and WT0 replicates had more similarity within their groups compared to all the other groups. It is unclear why this is the case. Interestingly, despite a large number of differentially expressed genes in the M0/WT0 group, its intersection with the WT0/WT772 group gave only a few genes which were common to the M772/WT772 group. No new genes with lower fold

change were consistently differentially expressed in all three groups. Thus, the Zur-regulon of *Synechococcus* sp. WH8102 likely comprises two activated genes: SYNW0359 and SYNW0360 and a cluster of repressed genes: SYNW0968, SYNW0969, SYNW0970, SYNW0971, SYNW0972 and SYNW0973. Previously, only SYNW0359, SYNW0969, SYNW0970 and SYNW0971 were identified by bioinformatics as Zur-regulated and all were assumed to be repressed³⁸.

4.3.3.3.1 Confirmation of RNA-seq data by RT-qPCR

Given that the RNA-seq data revealed potential components of the *Synechococcus* sp. WH8102 zur regulon, the expression of a subset of these genes was verified independently using RT-qPCR. The genes used for comparison were *znuA* (SYNW0971) which was up-regulated in the mutant, and *bmtA* (SYNW0359) down-regulated in the mutant. The housekeeping gene phosphoenol pyruvate carboxylase (SYNW2047) was used for normalisation as described in section 4.2.4.4.

A concentration of primers equal to 400 nM was found to give similar amplification efficiencies for each gene (94.6, 91.6 and 92.5% for *znuA*, *bmtA* and SYNW2047, respectively (see Figure 8.21 in the Appendix). In order to use the $2^{-\Delta\Delta C_t}$ method, ΔC_t (the C_t difference between the target and reference gene) values are plotted against the logarithm of template copy number¹⁴⁹. The modulus of the slope of the graph should not exceed 0.1. This condition was satisfied (see Figure 8.21C, in the Appendix). Since SYBR green was used for detection of the PCR products formed, it was important to assess whether only a single product was formed. This was assessed by analysing the dissociation of double-stranded products (melting curves). Only single peaks were observed for *znuA*, *bmtA* and SYNW2047 (see Figure 8.21D in the Appendix) evidencing a single product.

Following the testing of the standards, RNA samples were used for cDNA synthesis and subsequently as templates for RT-qPCR as described in section 4.2.4.4. The resulting $-\Delta\Delta C_t$ values were calculated separately for each replicate and average values are shown in Figure 4.23. $-\Delta\Delta C_t$ values are equal to the log₂fold change of genes given 100% amplification efficiency.

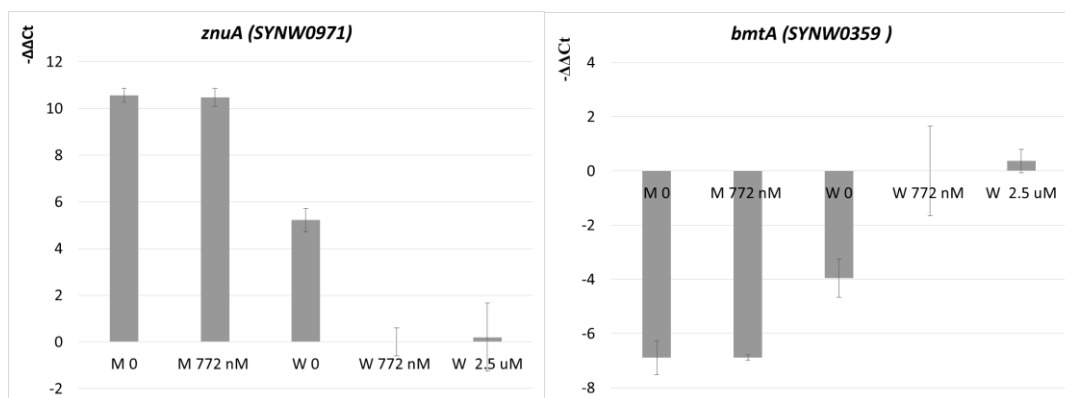


Figure 4.23 RT-qPCR analysis of *znuA* and *bmtA* gene expression.

-ΔΔCt values are normalised to WT 772 nM. Error bars represent errors between three biological replicates.

RT-qPCR results are in very good agreement with the RNA-seq data for these genes and confirm that *znuA* is up-regulated in the mutant and in the WT at low zinc concentration, and *bmtA* is down-regulated in the mutant and in the WT at low zinc concentration.

4.3.3.4 Functions and properties of newly identified SynZur-regulated genes

Based on section 4.3.3.3, it was suggested that the following genes are regulated by SynZur: SYNW0359 and SYNW0360 are activated, and a cluster of genes, SYNW0968, SYNW0969, SYNW0970, SYNW0971, SYNW0972, SYNW0973, is repressed.

4.3.3.4.1 Zur-regulon: Activation

SYNW0359 is a bacterial metallothionein which apparently binds excess cellular Zn^{2+} in order to detoxify/sequester it under high zinc and/or to store excess zinc for future deficiency scenarios. A related hypothesis is that SynBmtA could be a zinc chaperone supplementing zinc for some essential enzymes such as alkaline phosphatase or carbonic anhydrase. For example, transfer of zinc to carbonic anhydrase by a mouse metallothionein was observed¹⁷⁰. However, it is unclear how activation of SYNW0359 occurs.

Previously, Cox and Saito (2013) showed using proteomics that BmtA was up-regulated under high zinc conditions in *Synechococcus* sp. WH8102¹⁷¹. This is likely a consequence of transcriptional activation of *bmtA* by SynZur, which is first reported

here. If SynZur is an activator of *bmtA*, this could explain the result of the ICP-MS analysis (see section 4.3.2.2). Whilst the *zur*-mutant accumulated more zinc under 0 zinc conditions compared to the WT, consistent with up-regulation of *znuABC*, the opposite pattern at 772 nM zinc had been unexpected. The lower zinc content in the mutant may be due to *bmtA* not being activated and hence SynBmtA not being available for storing intracellular zinc. The presence of sufficient amounts of SynBmtA could help deal with zinc excess, safely sequestering more zinc which otherwise would be toxic.

However, in natural environments occupied by strains like *Synechococcus* sp. WH8102, zinc concentrations are very low. In the Central North Pacific Ocean total zinc is estimated to be in the low nanomolar range and free Zn^{2+} in the pico- to femtomolar range⁵⁶. Thus, rather than SynBmtA being used to deal with zinc toxicity it is perhaps more likely required to accumulate zinc on the rare occasions sufficient zinc concentrations are reached in order to store it safely within cells for future use.

SYNW0360 is annotated as showing “Weak similarity to phage integrase family” proteins. Protein NCBI BLAST¹³¹ searches showed results for either “weak similarity to phage integrase family”, “site-specific integrase” or “hypothetical protein” in marine *Synechococcus* and *Prochlorococcus*. Potentially, the presence of a phage integrase could be evidence of the horizontal gene transfer of genes located next to it i.e. the SynZur-activated metallothionein gene (SYNW0359). This genomic region is part of a large genomic island (ISL01, 140420 bp) in *Synechococcus* sp. WH8102. Interestingly, a tmRNA (transfer-messenger RNA, *ssrA*) located upstream of the ISL01, which prevents stalled ribosomes from becoming dysfunctional, served as the integration site for this genomic island^{3,172}. Conserved sequences at the 3' end of the tmRNA gene are recognised by phage integrases and hence these regions can play a role as the point of integration for phage-derived genes during which the the 3' end of tmRNA duplicates. In *Synechococcus* sp. WH8102 three 3' ends of the tmRNA are present potentially indicating multiple insertions³. One of them is located directly upstream of SYNW0360 (see Figure 8.22, Appendix).

Why **SYNW0360** is being up-regulated in response to high zinc is unclear. Phage integrase could be up-regulated as a response to high zinc in order to increase the probability of insertion events which bring new genes to deal with zinc stress. Since this strain inhabits an oligotrophic environment, the most beneficial system

apparently included integration of not a zinc efflux system in case of rare events of zinc excess but a metallothionein which would help keeping acquired zinc for times of deficiency.

4.3.3.4.2 *Zur-regulon: Repression*

As mentioned above, genes encoding members of the high affinity zinc uptake system (ZnuABC) are repressed: **SYNW0969**, **SYNW0970** and **SYNW0971** encoded by *znuB*, *znuC* and *znuA*, respectively. Their repression by SynZur under higher zinc conditions is logical: when there is enough zinc, cells do not need to produce high-affinity machinery for zinc uptake. Zur regulation of *znuABC* systems has been reported for several different bacterial phyla. For example, all three genes were de-repressed in an *Anabaena* sp. strain PCC 7120 *zur*-mutant⁵⁹.

The three other genes repressed by SynZur in *Synechococcus* sp. WH8102: **SYNW0968**, **SYNW0972** and **SYNW0973** are all uncharacterized proteins. NCBI Protein BLAST¹³¹ searches for SYNW0968 and SYNW0973 using the non-redundant protein sequence database only revealed *Synechococcus* hypothetical proteins with E-values <1. SYNW0972 also showed only hits to hypothetical proteins or conserved hypothetical proteins except for one hit (Total score 387, Query cover 99%, E value 2^{-130} and 59.27 percent identity) for a protein from *Synechococcus* sp. WH8101 (SynWH8101_1256) which is annotated as a “Zinc ABC transporter substrate-binding protein ZnuA”. However, Blast searches using SynWH8101_1256 as the query retrieved only hypothetical proteins. An EMBOSS Needle¹⁷³ pairwise sequence alignment also did not reveal any similarity of SYNW0972 to *Synechococcus* sp. WH8102 ZnuA (SYNW0971, data not shown), so this seems a case of mis-annotation.

All three proteins were then analysed for potential zinc binding residues (see Figure 4.24).

>SYNW0968

MTEIELNRLRAKLRRLAFTTEEAIDRTRLRRRFQRDLETFASASRIIEGLEARNNEEQQVLQALACGLIAAV
LGGWFATVY

>SYNW0972

MAHGSHGGGDEELEAGEFDFTPLITIEGHGGFETNLDNPKHYAIDGLFGGVFQWGLNGGSLTVEAA
VGPSVWVGEAEHFYGVVHVDHGHEDDHHHADAHAGHEDDHHHDEHAEDDHHDEEHADHDDH
DHHDDHDEHAEDDHHEHAEHAHDEHGHGDPPELKRITVVRGFFSVRYEPNDRLSLTVDWMPYYVTRDQGD
DIQGLKNELGAEVWVALGDDVDVDFALGDGLENILDGVFLSVMHRQGWESDGTWWMGNYTDPRLGVGFNI
DQINVTIDAGPRFYTPGSYSGLSQRTDFAGEVEVSIPVGDVAVLFAHWKPTYSPDDAPGWGEGWQHHLG
TGVTFSS

>SYNW0973

MTRRLTLLLLLAAATVVVLQQQLLRRRAPRLLGVEPQPLQSGRAAVDLSFSRSMHRNNLAAESSLDPAL
PHRWLGQGNRLRLILEGMTPIDAPIALTLAGRDQRMQPMTPQKRWWDPWPWLLVTRQVEEGEQQLQD
RQQQWHPVPLSPVWTSLSQSLVPLGNGSGVAMVSSNGKGEIWIWRKRLTPRNALSQQQLGVPVQGVVEPL
SKGDLVLFHISSNINGDLLVQTGGLKPGSESLLELLANGEQRMLKLPSSGPMQLLPAGGGLVVPGYNG
ISIQPLKDNKGPPQVLPGSRLELGAFCATSGRAVLIRHWPDYRRSIELVIPGLAPRQLHLGEGQAVLAVS
CSGSGDRIWAVLVGIWQGRRSQHELVQFNSEGTVLQRRNLAPWSLSPGTGMEHNPVGNLLMMLTLPDL
KGGRAALIEADTLRLRKVMDESIKEARWLPAG

Figure 4.24 Sequences of SYNW0968, SYNW0972 and SYNW0973.

His, Asp, Glu and Cys residues are highlighted with the corresponding colours.

Only one cysteine residue is present in SYNW0968 and no histidine residues. However, this protein contains eight glutamate and two aspartate residues, which is rather a lot for a short protein (77 aa). SYNW0972 contains a very long stretch of 82 amino acid residues containing almost only histidine, aspartate and glutamate residues with only a few occasional alanine and glycine residues. Histidine-rich loops are often involved in metal binding in membrane metal ion transporters, especially in zinc transporters¹⁷⁴. It is possible that this protein is involved in zinc acquisition/scavenging and has high affinity to Zn²⁺. Blast analyses using only this region revealed SYNW0972 itself and a homologue from the closely related strain *Synechococcus* sp. WH8103. SYNW0973 contains several histidine residues, two cysteine and multiple aspartate and glutamate residues which could in principle be involved in metal/zinc binding.

Analyses searching for the presence of a signal peptide using the SignalP server revealed none for SYNW0968 and SYNW0972 as annotated. However, NCBI BLAST¹³¹ searches showed two different Genbank entries for SYNW0972: one is 347 aa (CAE07487.1) and another is 375 aa in length (WP_042503314.1) containing an additional 28 residues: MFVFSKAPRLRLLLLLALLLPCTAARPV. Both protein sequences do not have experimental confirmation. If the WP_042503314.1 sequence is checked by the SignalP server, these additional 28 aa are identified as a signal peptide sequence (Sec/SPI, see Figure 8.23A in the Appendix). The presence of a Sec

signal peptide indicates that SYNW0972 could be an inner membrane, periplasmic, outer membrane or extracellular protein¹⁷⁵. SYNW0973 was also predicted to have a Sec signal peptide but with low probability (see Figure 8.23B in the Appendix).

Phyre2¹⁷⁶ analysis of the secondary structure of SYNW0968 revealed similarity with proteins of different function, however with low confidence. The predicted secondary structure based on superposition of Phyre2 hits is shown in Figure 4.25A. It is possible that SYNW0968 contains a four- α -helix bundle.

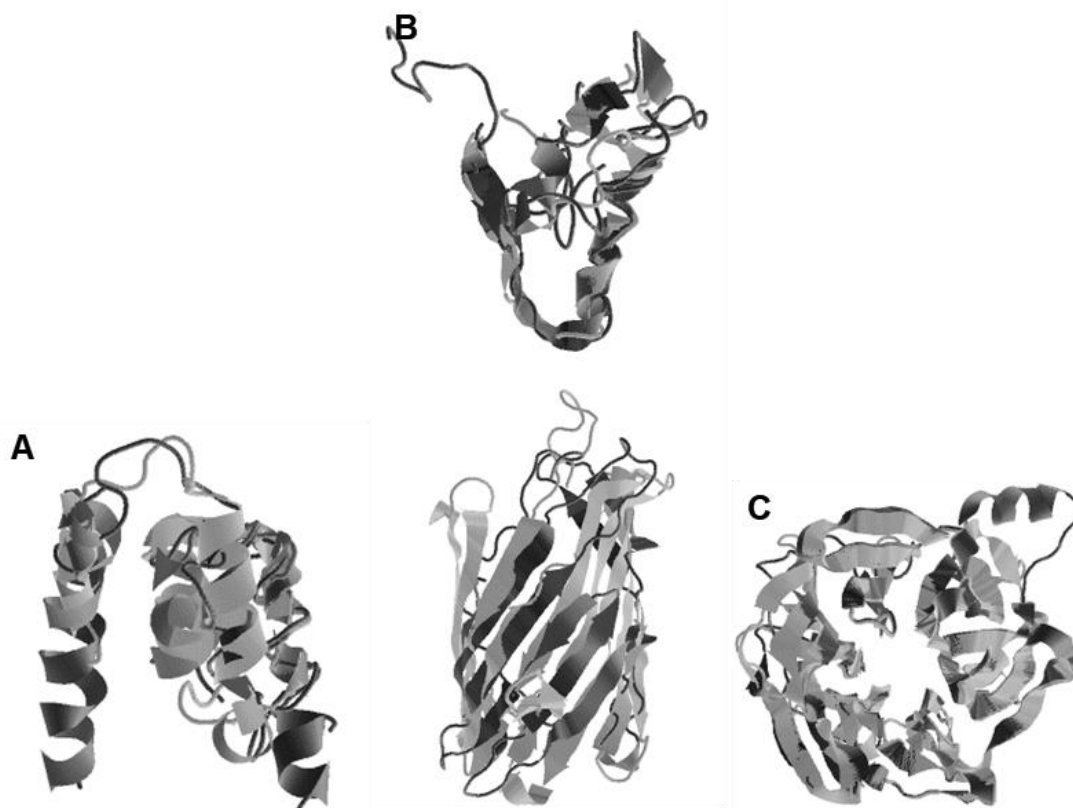


Figure 4.25 Predicted secondary structure of SYNW0968 (A), SYNW0972 (B) and SYNW0973 (C) proteins based on superposition of a selection of hits using Phyre2¹⁷⁶ analysis.

The best Phyre2 hit for WP_042503314.1 (SYNW0972) revealed similarity of 11% to the sequence at the N-terminus of a probable translation initiation factor IF-2 with 96.7% confidence. However, no similarity with this translation initiation factor was observed for SynWH8101_1256 (the homologous protein from *Synechococcus* sp. WH8101 mentioned above). Importantly, for both proteins there were two more hits for 45% of the sequence at the C-terminus which showed similarity with 78.9 and 72.2% confidence to β -barrels of membrane proteins such as *E. coli* outer membrane

porin OmpG and *Pseudomonas* outer membrane amyloid transporter FapF, respectively. Histidine-rich loops are usually intrinsically disordered¹⁷⁷ and no structured proteins with similarities to this region were found. Therefore the suggested structure is incomplete: only 11% N-terminus hits and 45% C-terminus hits corresponding to a β -barrel are shown (see Figure 4.25B).

Finally, Phyre2¹⁷⁶ analysis of SYNW0973 showed similarity of 74% to a sequence with 99.4% confidence to a thiocyanate dehydrogenase (oxidoreductase) from *Thioalkalivibrio paradoxus* (an alkaliphilic autotrophic sulfur-oxidizing bacterium). Interestingly, there were two more hits with just 0.1% lower confidence but with higher percentages of identity (both were 15% *versus* 10% for thiocyanate dehydrogenase) for 62% of the sequence as *E. coli* β -propeller TolB. The Tol system is a five-protein assembly which helps to stabilize the Gram-negative outer membrane¹⁷⁸. The predicted structure of SYNW0973 is shown in Figure 4.25C.

Thus, intriguingly both SYNW0972 and SYNW0973 could be outer membrane components involved in zinc metabolism! It has remained unknown how metals are transported across the outer membrane in marine cyanobacteria. However, clearly further experimental evidence is required to confirm these findings and clarify the function of these newly identified components.

To assess whether these genes are present in other marine cyanobacteria, the Cyanorak¹⁷⁹ database was interrogated, which is a database of marine picocyanobacterial genomes. Analysis of SYNW0968 in Cyanorak¹⁷⁹ revealed that this gene is present only in *Synechococcus* sp. sub-cluster 5.1 clades II, III, IV, VIII, IX and WPC1 (all are marine clades known to occupy a wide variety of habitats from oligotrophic ocean to hypersaline and coastal environments¹³). SYNW0972 is present in all marine *Synechococcus* sp. except clade 5.2B while all the *Synechococcus* species in the Cyanorak¹⁷⁹ database have a conserved homologue of SYNW0973. Both SYNW0972 and SYNW0973 are also present in LL *Prochlorococcus* (LLIV).

Interestingly, genomic context analysis of SYNW0972 in Cyanorak¹⁷⁹ showed that its homologues are very often located downstream of ZnuA and that the whole cluster of genes (<ZnuC<ZnuB ZnuA>SYNW0972><SYNW0973<SYNW0974 <SYNW0975) is very well conserved in marine cyanobacteria (see Figure 4.26). More intriguingly, it was found that in several *Prochlorococcus* strains, this arrangement is

located directly upstream of *zur* which increases the probability that these genes are involved in zinc metabolism.

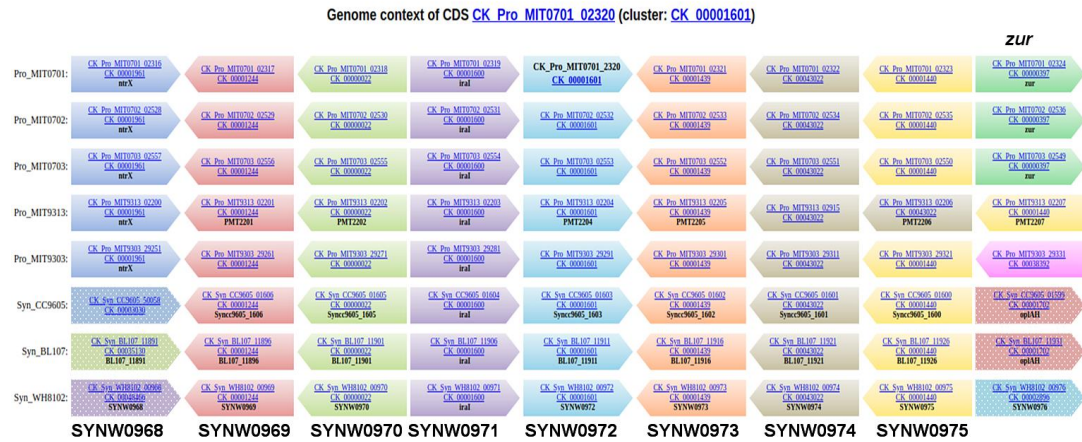


Figure 4.26 Cyanorak¹⁷⁹ genome context analysis of SYNW0972.

Only a selection of cyanobacterial species is shown. Gene homologues in different species are shown in the same colour.

In summary, whilst it was not possible to predict a function for SYNW0968, this protein probably consists of several α -helices, one of which at the C-terminus might be located in the inner membrane since it is predicted to be transmembrane by the Protter server (data not shown) while no signal peptide was revealed. The rest of the protein apparently is located inside the cytosol and contains a significant amount of E and D residues which could be metal binding. However, only one Cys residue was found and cysteines are often hallmarks of strong zinc binding sites.

SYNW0972 and SYNW0973 are widely distributed genes in marine *Synechococcus* and some *Prochlorococcus* strains where this cluster is sometimes located next to the *zur* gene. For both proteins Phyre2¹⁷⁶ analyses suggested they may be outer-membrane proteins. In addition, SYNW0972 has a very long region of metal binding residues which are mostly histidines, strongly reminiscent of other His-rich loops in zinc and nickel binding proteins from a variety of phyla^{174,177,180}. It is unclear if this region is located in the periplasm or outside the cell. Considering both of these genes were up-regulated in the *zur*-mutant together with *znuABC*, it is hypothesised here that these might be components of a novel outer membrane zinc transport mechanism.

4.3.3.5 SynZur-independent genes potentially regulated by zinc

To find which genes could be Zur-independent but zinc-responsive, differentially expressed genes in the mutant at 0 zinc compared to 772 nM zinc were analysed. The volcano plot is given in Figure 8.24 in the Appendix and the list of genes with absolute values of log₂fold change >2 and p-values <0.05 shown in Table 4.11.

Table 4.11 Genes differentially expressed in the *Synechococcus* sp. WH8102 *zur*-mutant at 0 added zinc compared to 772 nM zinc.

Only genes with $|\text{Log2Fold}| > 2$ and p-values < 0.05 are given.

| Gene | Log2Fold Change | p-value | padj | Gene name | Function |
|----------|-----------------|----------|----------|--------------|--|
| SYNW0483 | 2.86 | 1.39E-04 | 9.91E-03 | apcC | Phycobilisome 7.8 kDa linker polypeptide, allophycocyanin-associated, core |
| SYNW1560 | 2.31 | 1.33E-04 | 9.77E-03 | | Uncharacterized protein |
| SYNW1303 | 2.12 | 3.06E-04 | 1.73E-02 | petN | Cytochrome b6-f complex subunit 8 |
| SYNW0495 | 2.00 | 1.49E-04 | 1.02E-02 | atpG atpC | ATP synthase gamma chain (ATP synthase F1 sector gamma subunit) |
| SYNW1665 | -2.03 | 2.48E-03 | 6.06E-02 | | Uncharacterized protein |
| SYNW1128 | -2.05 | 1.04E-02 | 1.37E-01 | | Uncharacterized protein |
| SYNW1133 | -2.12 | 3.04E-03 | 6.88E-02 | | Uncharacterized protein |
| SYNW1594 | -2.14 | 6.81E-04 | 2.65E-02 | | Uncharacterized protein |
| SYNW2114 | -2.15 | 3.58E-03 | 7.59E-02 | | Uncharacterized protein |
| SYNW0357 | -2.19 | 2.57E-02 | 2.09E-01 | | Uncharacterized protein |
| SYNW0164 | -2.24 | 5.62E-03 | 9.47E-02 | | Uncharacterized protein |
| SYNW0808 | -2.42 | 6.11E-05 | 6.09E-03 | | Two-component response regulator |
| SYNW2100 | -2.88 | 1.34E-02 | 1.52E-01 | | Uncharacterized protein |
| SYNW0809 | -2.89 | 1.01E-04 | 8.42E-03 | | Uncharacterized protein |
| SYNW2329 | -3.07 | 2.09E-04 | 1.34E-02 | pgk | Phosphoglycerate kinase (EC 2.7.2.3) |
| SYNW0228 | -3.16 | 5.40E-05 | 5.61E-03 | | Conserved hypothetical |
| SYNW0878 | -4.05 | 3.42E-05 | 4.26E-03 | | Uncharacterized protein |
| SYNW2330 | -4.21 | 6.77E-06 | 1.55E-03 | | Uncharacterized protein |
| SYNW1257 | -4.27 | 8.46E-06 | 1.62E-03 | | Uncharacterized protein |

Venn analysis performed as described in section 4.3.3.3 but for comparison of Mut0/Mut772 with WT0/WT772 is shown in Figure 4.27 and in Table 4.12.

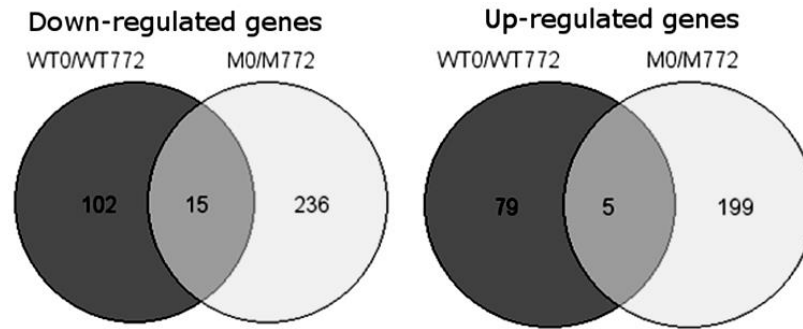


Figure 4.27 Venn diagrams for comparison of M0/M772 with WT0/WT772.

Table 4.12 Differentially expressed genes in common for comparisons M0/M772 and W0/W772.

| Common down-regulated genes | Common up-regulated genes |
|-----------------------------|---------------------------|
| SYNW0071 | SYNW0203 |
| SYNW0106 | SYNW0972 |
| SYNW0216 | SYNW0973 |
| SYNW0403 | SYNW1018 |
| SYNW0466 | SYNW2340 |
| SYNW0526 | |
| SYNW0633 | |
| SYNW1042 | |
| SYNW1066 | |
| SYNW1180 | |
| SYNW1248 | |
| SYNW1407 | |
| SYNW1665 | |
| SYNW2182 | |
| SYNW2235 | |

Genes in common were only mildly differentially expressed in the mutant at 0 zinc compared to the mutant at 772 nM zinc ($|\log_2\text{fold change}| < 2$, $p\text{-value} < 0.05$). Interestingly, both genes SYNW0972 and SYNW0973 previously found to be up-regulated by SynZur were also found here. It is possible that either that is a coincidence since $\log_2\text{fold change}$ is below two for both SYNW0972 and SYNW0973 for M0/M772. Alternatively, they could be co-regulated with some other so far unknown zinc-dependent transcription factor.

4.3.3.6 SynZur-independent genes responding to zinc toxicity.

Finally, to understand which genes are differentially expressed in the WT in response to the highest zinc concentration (2.5 μ M) mRNA profiles were compared to the WT grown at 772 nM zinc. The volcano plot of the analysis is given in Figure 8.25 in the Appendix and the list of genes with absolute values of $|\log_2\text{fold change}| > 2$ and p-values < 0.05 is given in Table 4.13.

Table 4.13 Genes differentially expressed in the wild type *Synechococcus* sp. WH8102 at 772 nM zinc compared to 2.5 μ M.

Only genes with $|\text{Log}_2\text{Fold}| > 2$ and p-values < 0.05 are given.

| Gene | Log2Fold Change | p-value | padj | Gene name | Function |
|----------|-----------------|----------|----------|-----------|--|
| SYNW1397 | -2.67 | 2.16E-03 | 1.61E-01 | | Conserved hypothetical |
| SYNW1424 | -2.24 | 1.00E-04 | 2.71E-02 | | Uncharacterized protein |
| SYNW1415 | -2.09 | 4.12E-03 | 1.96E-01 | | ABC transporter, nitrate-like substrate binding protein |
| SYNW1946 | -2.06 | 2.26E-02 | 3.23E-01 | | PhoH family protein |
| SYNW1425 | -2.03 | 1.48E-03 | 1.37E-01 | | Probable nitrilase (EC 3.5.5.7) |
| SYNW1423 | -2.00 | 2.81E-04 | 5.74E-02 | | Uncharacterized protein |
| SYNW1511 | 2.00 | 3.53E-05 | 1.44E-02 | | Conserved hypothetical |
| SYNW1070 | 2.04 | 1.59E-03 | 1.42E-01 | | Uncharacterized protein |
| SYNW2028 | 2.35 | 5.33E-04 | 8.38E-02 | rlpA | Probable endolytic peptidoglycan transglycosylase RlpA (EC 4.2.2.-) |
| SYNW0216 | 2.45 | 5.35E-03 | 2.11E-01 | folK | Possible 2-amino-4-hydroxy-6-hydroxymethylidihydropteridine pyrophosphokinase (EC 2.7.6.3) |
| SYNW0721 | 2.52 | 8.35E-02 | 4.54E-01 | petG | Cytochrome b6-f complex subunit 5 |
| SYNW2288 | 2.69 | 5.85E-06 | 3.99E-03 | | Probable pseudouridine synthase (EC 4.2.1.70) |

Comparison of these genes (see Table 4.13) with those differentially expressed genes in the mutant 0/772 nM should allow identification of genes which are differentially expressed in response to zinc toxicity whilst not being Zur-regulated. This is because in the mutant 772 nM zinc is probably at toxic levels for cells due to insufficient amount of SynBmtA sequestering Zn^{2+} . Venn diagrams for those two comparisons are shown in Figure 4.28. Common differentially expressed genes for comparisons M0/M772 and W772/W2500 are given in Table 4.14.

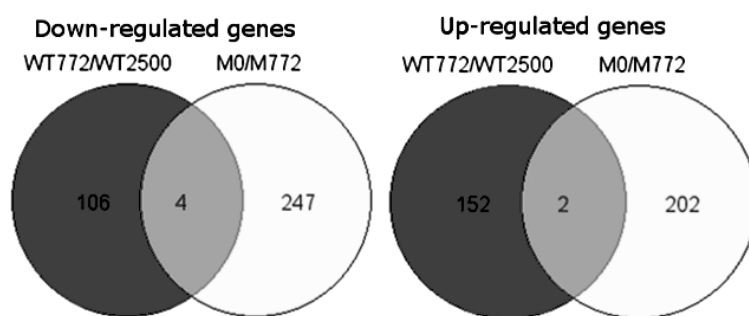


Figure 4.28 Venn diagrams for comparison of WT772/WT2500 nM zinc with Mut0/Mut772.

Table 4.14 Differentially expressed genes in common for comparisons M0/M772 and W772/W2500 which are probably involved in zinc toxicity.

| Down-regulated | Up-regulated |
|--|----------------------|
| SYNW1025 SYNW1257 SYNW2176 SYNW2330 | SYNW0735 SYNW0994 |

Interestingly, SYNW2176 is a Deg-protease involved in envelope stress responses and in organisation of outer membrane proteins¹⁸¹. SYNW2176 was up-regulated as a response to Cu stress and it is possible that it would be Zn-responsive in the presence of high zinc as well. SYNW2330 was also amongst the genes, up-regulated during Cu shock¹⁸², whilst SYNW1025 (annotated as *trpD/trpG*, a putative anthranilate synthase component II) and SYNW1257 (conserved hypothetical protein) were down-regulated under Ni-deprived growth using urea and ammonia respectively, as a nitrogen source¹⁸³. All four down-regulated genes could be used to deal with zinc excess, possibly non-specifically if they are also utilised by cells for dealing with Cu and Ni.

SYNW0735 and SYNW0994 are both conserved hypothetical proteins down-regulated under the early/late stages of P-stress and under late P-stress, respectively¹⁶². Thus, it is unclear whether these are zinc-responsive. However, zinc and phosphorus metabolism were reported to be linked in *Synechococcus* sp. WH8102^{171,184} and these genes could be zinc-responsive members of phosphorus metabolism.

4.4 CONCLUSION AND FUTURE WORK.

In this chapter a *Synechococcus* sp. WH8102 *zur* knockout mutant was constructed and phenotypically characterised. Growth rate comparison showed that the mutant has a lower tolerance to zinc compared to WT, whilst ICP-MS analysis revealed different accumulation behaviour for trace metals and most significantly for zinc.

RNA-seq analysis of the WT and *zur* mutant facilitated experimental identification of the *Synechococcus* sp. WH8102 Zur regulon: SYNW0359, SYNW0360, SYNW0968, SYNW0969, SYNW0970, SYNW0971, SYNW0972 and SYNW0973. Genes SYNW0359 (metallothionein) and SYNW0969, SYNW0970, SYNW0971 (components of a zinc uptake system *znuABC*) were previously predicted to be regulated by SynZur. However, all were assumed to be repressed by zinc-bound SynZur since they have similar Zur-binding boxes in their operator-promoter regions. Here, I show for the first time that SYNW0359 is transcriptionally activated by SynZur while SYNW0969, SYNW0970 and SYNW0971 are repressed as predicted. Indeed, this is the first experimental evidence that a bacterial metallothionein can be regulated by Zur.

Genes SYNW0968, SYNW0972 and SYNW0973 are annotated as hypothetical proteins. The entire cluster is well-conserved in marine cyanobacteria, and from structural predictions using Phyre2, I suggest that SYNW0972 and SYNW0973 are components of a novel outer membrane zinc uptake mechanism. The role of SYNW0968 remains unknown.

If normalized gene expression for *znuA* and *bmtA* from WT *Synechococcus* sp. WH8102 determined by RT-qPCR is plotted against cellular zinc content as found by ICP-MS, a clear pattern of Zn-dependent repression/activation is observed (see Figure 4.29): *znuA* is activated (de-repressed) under low zinc and *bmtA* is activated under high zinc.

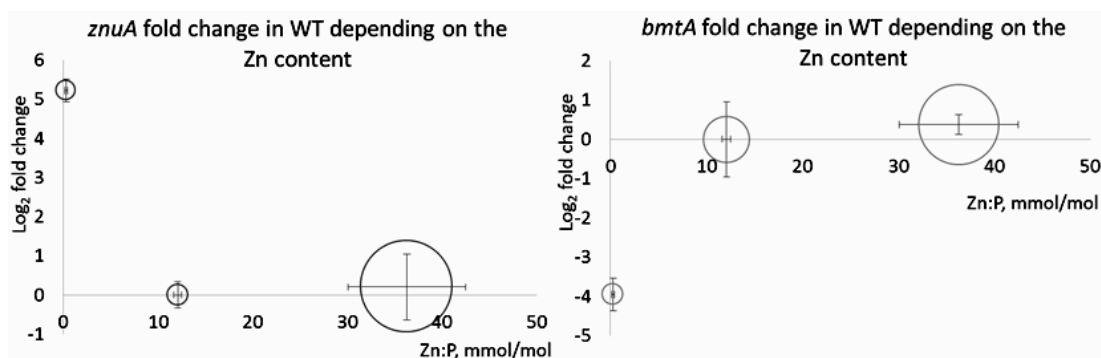


Figure 4.29 A plot of normalised gene expression for *znuA* and *bmtA* measured by RT-qPCR against normalised zinc content.

Values for gene abundance in three WT samples were normalised to gene expression at 772 nM zinc, and zinc content was normalised to P content. Error bars represent errors between three biological replicates.

Most importantly, the transcriptomic profile of WT as a response to changing zinc concentrations matches the transcriptomic profile of the *zur*-mutant (see sections 4.3.3.2 and 4.3.3.3), clearly indicating that SynZur is a zinc sensor and a true zinc uptake regulator, which confirms conclusions based on the biophysical data obtained in Chapter 3.

Future work

Firstly, the mechanism of activation of *bmtA* is unclear. In addition to the proposed experiments modifying the SynZur protein mentioned in section 3.3, investigation of the *bmtA* operator-promoter region could clarify this. For example, transcription assays of a mutant promoter with one of the two *zur*-boxes removed or mutant promoters with modified locations of the *zur* boxes. Structural characterisation of the $(Zn_4Zur_2)_2$ -*pbmtA* complex could also be attempted.

Additionally, to finalise proof of SynZur as a zinc specific sensor, it would be interesting to compare growth rates of the *zur*-mutant and the WT at different concentrations of other trace metals (Co, Cd, Cu, Fe, Mn and Ni). If SynZur is solely a Zn-specific sensor, the mutant should not show increased sensitivity towards other trace metals compared to the WT unless deleting SynZur indirectly affects its whole trace metal metabolism. For example, up-regulation of outer membrane proteins in the mutant could potentially change OM structure which would then have an effect on uptake/efflux of other metals/nutrients. Comparison of transcriptomic profiles of the WT as a response to different trace metals could also be performed.

Understanding/proving the role of members of the Zur-regulon in zinc metabolism is also very important. Thus, creating *bmtA* or *znuA* mutants and exploring their zinc sensitivity would also prove their involvement in zinc metabolism: a *bmtA*-mutant should be more sensitive to higher zinc concentrations similar to the *zur*-mutant but without influence of de-repressed components of zinc uptake. In contrast, a *znuA*-mutant should be more sensitive to lower zinc concentrations due to insufficient uptake.

Probably the most intriguing line of investigation would be to understand what role the conserved uncharacterised Zur-regulated genes play in zinc metabolism of *Synechococcus* sp. WH8102, especially SYNW0972 and SYNW0973. Their over-expression in *E. coli* and construction of deletion/knockout mutants in *Synechococcus* sp. WH8102 may clarify whether they are true outer membrane components of zinc uptake. For most of these proteins their mode of Zur-regulation is also unclear since they do not contain predicted zur-boxes in their promoter regions. However, EMSA experiments or transcription assays could shed light on that.

Chapter 5.
Properties of a Zur-activated metallothionein
from *Synechococcus* sp. WH8102

5.1 INTRODUCTION

SYNW0359 is a predicted bacterial metallothionein (BmtA) in *Synechococcus* sp. WH8102 (hereafter named SynBmtA). Metallothioneins (MTs) are low-molecular weight proteins with high cysteine content and significant metal ion binding capacity (usually for Zn^{2+} , Cu^+ and Cd^{2+})⁷⁸. Bacterial MTs, including some cyanobacterial MTs, are usually expressed under zinc and cadmium excess conditions, likely in order to sequester and detoxify these metals inside cells. Transcription of MTs in cyanobacteria is usually repressed in zinc-depleted conditions by zinc excess sensors like SmtB^{37,39}.

In the absence of an SmtB homologue in *Synechococcus* sp. WH8102, it has remained unclear what the function of SynBmtA is and under which circumstances it is expressed. Based on the presence of two Zur boxes in the promoter region of *synbmtA*, it was predicted to be repressed by SynZur at high zinc concentrations³³. However, in Chapter 4 I showed that SynZur activates transcription of *synbmtA* under elevated zinc concentrations. Thus, SynBmtA apparently acts as a classical MT serving to sequester excess zinc in order to detoxify it or to act as a storage protein for future zinc use^{5,79}. This latter function would be very beneficial for *Synechococcus* sp. WH8102 that occupies an extremely oligotrophic environment. Another proposed role for SynBmtA is as a metallochaperone: potentially transferring zinc to other proteins or subcellular locations such as carboxysomes.

In this chapter I show that SynZur binds to the operator-promoter region of SynBmtA (*pbmtA*). I also determined the molecular weight of the SynZur-*pbmtA* complex by Ferguson plot analysis, and calculated the DNA binding affinity of SynZur for *pbmtA*.

Furthermore, I over-expressed SynBmtA in *E. coli* and following purification identified SynBmtA by ESI-MS and characterised its metal content by ICP-OES. Subsequently, the MT was pH-titrated in order to find the pH at which half the zinc is displaced and this value was compared with those for other MTs. Finally, demetallation of SynBmtA was investigated.

To better understand what particular role SynBmtA plays, its zinc affinity was determined by titration of apo-SynBmtA with zinc in the presence of the fluorescent metallochromic dye Quin-2, which allowed a direct comparison with the zinc affinity

determined for SynZur (Chapter 3). How zinc transfer between SynZurZn₂ and apo-SynBmtA may occur is also shown.

Finally, SynBmtA was analysed by 3D (¹H, ¹⁵N, ¹H) Nuclear Magnetic Resonance (NMR) spectroscopy including 3D heteronuclear NMR (¹⁵N TOCSY-HSQC [Total correlation spectroscopy - Heteronuclear single quantum coherence] and ¹⁵N NOESY-HSQC [Nuclear Overhauser effect spectroscopy - Heteronuclear single quantum coherence]) in an attempt to pave the way for future structural characterisation.

5.2 MATERIALS AND METHODS

5.2.1 Zur-*pbmtA* EMSA

EMSA experiments for SynZur binding to *pbmtA* including the gel running conditions and visualisation were done in the same way as described in section 3.2.2.1. The promoter region was amplified by PCR using the primers shown in Table 5.1.

Table 5.1 Primers used for amplification of a 143 bp region of *pbmtA*.

The overall length of the amplified region was 163 bp. The underlined letters represent the 5' and 3' end of the amplified *pbmtA*.

| Name | Sequence, 5'-3' |
|----------|------------------------------|
| pbmtA_F | AGAGGAGCTCGTGTCTAGAGCAGTTGGG |
| pbmtA_Re | AGAGTCTAGACCAGGCAATTTTCAGCC |

Determination of *pbmtA*-Zur stoichiometry (Ferguson plots) was performed as described in section 3.2.2.1. To confirm the difference in size between *pbmtA*-Zur *pznuA*-Zur complexes, both were run together on the same gel at 150 nM Zur. To equalise the DNA length, 250 bp *pbmtA* and 256 bp *pznuA* sequences were used as designed previously⁸⁸.

5.2.2 SynBmtA over-expression and purification

E. coli strains used in this study for construction of a SynBmtA over-expressing vector and for its over-expression are shown in Table 2.1. Rosetta2 (DE3) pLysS and DH5- α strains were used. Conditions for SynBmtA over-expression are given in section 2.1. Cells were grown on LB plates with antibiotics where appropriate (35 $\mu\text{g mL}^{-1}$ chloramphenicol for Rosetta2 (DE3) pLysS) and made competent as described in section 2.3.4.

5.2.2.1 Plasmid construction for SynBmtA over-expression

The plasmid for SynBmtA over-expression was pET24a(+) (Novagen), a T7 promoter-driven high-efficiency protein expression and sequencing vector containing a kanamycin resistance cassette and multiple cloning sites. A map of pET24 is shown in Figure 8.26 in the Appendix.

pET-24a(+) vector was cut with restriction enzymes *NdeI* and *SacI* in MultiCore Buffer (Promega) and de-phosphorylated as described in sections 2.3.1 and 2.3.2. The insert (*bmtA*) was amplified by PCR as described in section 2.2.1 using primers given in Table 5.2 and *Synechococcus* sp. WH8102 genomic DNA as a template. Genomic DNA was isolated as described in section 4.2.2.1.

Table 5.2 Primers used for amplification of *synbmtA* for subsequent cloning into pET24a(+).

Restriction sites are underlined and parts of the coding region of the gene are shown with bold letters.

| Name | Sequence | Restriction site |
|---------|--|------------------|
| BmtA_F | GCGCC <u>CATATG</u> TCTACAGCAATCAAATG | <i>NdeI</i> |
| BmtA_Re | CGAC <u>GAGCTCTT</u> ACTCTCCGCACTTAC | <i>SacI</i> |

The insert was cut with restriction enzymes *NdeI* and *SacI* in MultiCore Buffer (Promega) but without de-phosphorylation, and ligated into the *NdeI* and *SacI* cut vector pET24a(+) as described in sections 2.3.1 and 2.3.3. The ligation mixture was transformed into competent *E. coli* DH5- α as described in section 2.3.5. Cells were then streaked onto LB plates with kanamycin (50 $\mu\text{g mL}^{-1}$) and left overnight at 30 °C. Next day, colonies were checked by PCR using cells as template with T7 primers for the presence of the correct size product (see section 2.2.1). T7 primers are shown in Table 5.3.

Table 5.3 T7 primers used for checking the successful ligation of the *synbmtA* insert.

| Name | Sequence 5'-3' |
|-------|----------------------|
| T7_F | TAATACGACTCACTATAGGG |
| T7_Re | GCTAGTTATTGCTCAGCGGT |

For colonies where a correct size product was observed (382 bp), plasmids were isolated using a QIAprep® Spin Miniprep Kit as described in section 2.3.6 and the correct insert confirmed by Sanger sequencing using T7 primers (forward and reverse separately).

5.2.2.2 BmtA over-expression

The protocol was adapted from previous works on bacterial MTs^{82,153}. For over-expression of SynBmtA only freshly transformed cells were used. Transformation of pET-24a(+) BmtA into *E. coli* Rosetta2 (DE3) pLysS was performed as described in section 2.3.5. After transformation, several single colonies were picked simultaneously from the plates and transferred into 100 mL LB media with kanamycin (50 µg mL⁻¹) and chloramphenicol (35 µg mL⁻¹) and incubated overnight at 30 °C.

The next day, overnight *E. coli* cultures were transferred at a ratio 1:100 into 1 L of fresh LB medium containing the same antibiotics and left shaking at 180 rpm at 30 °C until the OD₆₀₀ was in the range 0.5-0.8. Then, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the culture to induce protein over-expression (0.5 mM final concentration). In addition, ZnSO₄ was also added to a final concentration of 0.5 mM. To the control culture no IPTG or ZnSO₄ was added. Cultures were left overnight at 30 °C shaking at 180 rpm.

Subsequently, cells were harvested by centrifugation (5000 rpm for 15 minutes at 4 °C) and the cell pellet was transferred into 50 mL Falcon tubes and frozen at -20 °C. The presence of SynBmtA was monitored by SDS-PAGE as described in section 2.4.1, with protein visualisation achieved by silver staining (see section 5.2.2.3.3).

5.2.2.2.1 Over-expression of ¹⁵N labelled SynBmtA

SynBmtA over-expression in minimal medium (MM, see Table 5.4.) was carried out in a similar manner to that described above: overnight cultures in LB medium were used to inoculate MM at a 1:100 ratio. Then cultures were incubated at 37 °C until the OD₆₀₀ reached ~1. After that IPTG and ZnSO₄ were added to a final concentration of 0.5 mM and ~0.102 mM respectively. Induction was carried out at 23 °C for 24 h. Then cultures were centrifuged at 5000 rpm for 15 minutes and the resulting cell pellets were transferred into 50 mL Falcon tubes and frozen at -20 °C. The presence of SynBmtA in the cell pellet was monitored by SDS-PAGE (see section 2.4.1) visualised by silver-staining (see section 5.2.2.3.3).

Table 5.4 M9 Minimal Medium recipe.Adapted from Sambrook and Russell, (2001)¹⁸⁵.

| Main components for autoclaving | Concentration |
|---|---------------------------|
| Na ₂ HPO ₄ | 12.8 g L ⁻¹ |
| KH ₂ PO ₄ | 3 g L ⁻¹ |
| NaCl | 0.5 g L ⁻¹ |
| MgSO ₄ •7H ₂ O | 0.2465 g L ⁻¹ |
| FeCl ₃ | 7.5 mg L ⁻¹ |
| Yeast extract | 0.05 g L ⁻¹ |
| ¹⁵ NH ₄ Cl | 1 g L ⁻¹ |
| CaCl ₂ •2H ₂ O | 14.7 mg L ⁻¹ |
| ZnSO ₄ •7H ₂ O | 0.575 mg L ⁻¹ |
| MnSO ₄ •5H ₂ O | 0.0845 mg L ⁻¹ |
| H ₃ BO ₃ | 0.145 mg L ⁻¹ |
| CuSO ₄ •5H ₂ O | 0.0875 mg L ⁻¹ |
| Components added after autoclaving | Concentration |
| Glucose | 5 g L ⁻¹ |
| Biotin | 10 mg L ⁻¹ |
| Thiamine | 0.5 mg L ⁻¹ |
| Kanamycin | 50 mg L ⁻¹ |
| Chloramphenicol | 34 mg L ⁻¹ |

The main components were mixed from stock solutions and autoclaved. When the mixture was cooled to room temperature after autoclaving, filter-sterilised stocks of the heat-intolerant components were subsequently added.

5.2.2.3 Purification of SynBmtA

5.2.2.3.1 Chemical precipitation

Cell pellets obtained as described in section 5.2.2.2. were re-suspended in lysis buffer (50 mM Tris-HCl (pH 8.5), 0.1 M KCl, 1 mM ZnSO₄, 3 mM dithiothreitol (DTT), 0.1 mM phenylmethylsulfonyl fluoride (PMSF)) at a ratio of 5 mL buffer per 1 g wet cell pellet.

The mixture was sonicated at 4 °C for 5 minutes with amplitude 60 and pulse 4 (Ultrasonic Processor, SONICS). Then it was centrifuged at 18,000×g for 10 minutes at 4 °C and the supernatant retained.

For chemical precipitation ice-cold ethanol/chloroform (100:8, v/v) mixture was very slowly added to the lysate while stirring on ice. After adding each volume of the precipitant, the suspension was centrifuged at 8,000×g for 10 minutes at 4 °C and the pellet separated from the supernatant. This step was repeated after every

centrifugation and pellet separation until 8 volumes of the precipitant had been added. Each fraction was dissolved in 5 mL 20 mM NH_4HCO_3 and analysed by SDS-PAGE followed by silver staining (see section 5.2.2.3.3).

5.2.2.3.2 Size Exclusion Chromatography of SynBmtA

Fractions from chemical precipitation containing SynBmtA were pooled and size exclusion chromatography (SEC) was used to further purify SynBmtA. The same FPLC instrumentation was used for SEC as for IMAC (see section 3.2.1.2) except for the column which was a HiLoad 16/60 Superdex 75 column (120 mL, Amersham Pharmacia). The system was equilibrated with elution buffer (20 mM NH_4HCO_3 , pH 7.9) prior to sample injection.

2 mL of the SynBmtA containing fraction was filtered through a 0.22 μm pore size filter and loaded onto the loop. Proteins were then eluted by size at a rate of 1 mL min^{-1} during 2 column volumes. Absorbance was monitored at 215 nm and 280 nm.

5.2.2.3.3 Silver staining

After SDS-PAGE (see section 2.4.1) gels were rinsed with Milli-Q water and incubated for 15 minutes in fixing solution (50% (v/v) acetone, 1.25% (v/v) trichloroacetic acid and 0.015% (v/v) formaldehyde) on a rocker at room temperature. After that, gels were rinsed with water thrice followed by 5 minutes washing with water and triple rinsing again. Then, incubation in 50% (v/v) acetone for 5 minutes was followed by washing the gel in 0.017% (w/v) sodium thiosulfate during 1 minute and triple rinsing with water. Then gels were stained for 8 minutes in a solution of 0.27% (w/v) silver nitrate, 0.37% (w/v) formaldehyde with further quick washings with water. After that, gels were developed for 10-20 seconds using 2% (w/v) sodium carbonate, 0.015 % (v/v) formaldehyde and 0.004% (w/v) sodium thiosulfate. Bands appeared very fast and it was important to be ready to quickly remove and wash the gels with stopping solution (1% (v/v) acetic acid) as soon as bands became visible. After that gels were rinsed with water and were ready to be scanned. All the solutions were freshly prepared in advance on the same day when the silver staining protocol was used.

5.2.3 Characterisation of SynBmtA

5.2.3.1 ESI-MS and ICP-OES

ESI-MS and ICP-OES were used as described in sections 2.4.3 and 2.4.4 to assess molecular mass of the protein, its metal stoichiometry and to calculate its molar absorption coefficient (ϵ_{230}).

5.2.3.2 pH titration SynBmtA

Samples of native SynBmtA (20 μ M) were adjusted to a range of pH values between 8.1 and 1.9 with 0.01, 0.1 and 1 M NaOH and HCl. UV-vis spectra in the range of 200-800 nm were recorded after each pH adjustment using a Varian Cary 50 Bio UV-vis spectrophotometer, Agilent Technologies. Absorption was normalised to the total volume.

5.2.3.3 Investigation of SynBmtA de-metallation

To investigate de-metallation of native SynBmtA samples containing 10 μ M BmtA in 20 mM NH_4HCO_3 were incubated with (i) 1 mM EDTA; (ii) 1 mM TPEN and 0.5 mM DTT; (iii) 6 M urea, 1 mM EDTA and 0.5 mM DTT final concentration.

Samples were left overnight and separated from inorganic and low molecular mass components using PD-10 columns (see section 2.4.2) and checked by ESI-MS for successful de-metallation.

The final experiment on BmtA de-metallation was performed as follows. All the ratios were measured in advance so that pH measurements were not needed at each step.

1. 2.5 mL 30 μ M SynBmtA in 10 mM NH_4HCO_3 was acidified to pH 3.3 with 4 μ L concentrated formic acid.
2. The PD-10 column was equilibrated with 25 mL 10 mM NH_4HCO_3 buffer, pH 3.3 (120 μ L concentrated formic acid was added per 100 mL buffer).
3. 2.5 mL acidified SynBmtA was loaded onto the PD-10 column and eluted with 3.5 mL 10 mM NH_4HCO_3 buffer, pH 3.3. Only the first 2.5 mL were collected.

4. The PD-10 column was then eluted and re-equilibrated with 25 mL 10 mM NH_4HCO_3 buffer (pH 7.9).
5. 2.5 mL of the acidified Zn-free BmtA sample was loaded onto the column, 3.5 mL 10 mM NH_4HCO_3 buffer (pH 7.9) added and 2.5 mL collected.

All the above steps were performed inside a glove bag filled with N_2 (see section 3.2.2.5). All the buffers were purged beforehand with either nitrogen or argon and all empty tubes/bottles were filled with N_2 prior to loading inside the glove bag. De-metallation was monitored by ESI-MS as described in section 2.4.4.

5.2.3.4 Titration of apo-BmtA with zinc in the presence of Quin-2

Upon successful de-metallation, $\sim 15 \mu\text{M}$ SynBmtA was titrated with ZnSO_4 in the presence of $30 \mu\text{M}$ Quin-2 as described in section 3.2.2.5.

5.2.3.5 SynBmtA/SynZur competition for Zn

SynZur was mixed with apo-BmtA at a ratio of 4:1 ($20 \mu\text{M}$ and $5 \mu\text{M}$, respectively). The mixture was set up inside a glove bag filled with nitrogen and left for 30 minutes to equilibrate. Thereafter, the reaction mixture was filtered with $0.22 \mu\text{m}$ pore size Minisart® syringe filters (Sartorius) and analysed by ESI-MS (see section 2.4.4).

5.2.3.6 Nuclear Magnetic Resonance Spectroscopy

5.2.3.6.1 1D ^1H -NMR Spectroscopy

A ^{15}N labelled protein sample of BmtA over-expressed from 1 L of MM as described in section 5.2.2.2.1 and purified as described in section 5.2.2.3 was prepared in 20 mM NH_4HCO_3 with 10% D_2O . A Bruker Avance 700 Ultrashield spectrometer fitted with a cryoprobe was used for measurements. The operating frequency for 1D ^1H NMR spectrum was 700.24 MHz. Excitation sculpting with gradients was used for water suppression¹⁸⁶. Spectra were recorded at 25 °C with 64 k data points, a spectral width of 20.5972 ppm and with 64 scans (program zgsgp). TopSpin 3.2 software was used to process the NMR data. The water resonance peak was used for chemical shift referencing¹⁸⁷.

5.2.3.6.2 2D Heteronuclear NMR Spectroscopy (HSQC)

A 2D (^1H , ^{15}N) heteronuclear single quantum coherence NMR spectrum (HSQC) was recorded with 128 data points in F1 (^{15}N) and with 2048 data points in F2 (^1H). Spectra were recorded at operating frequencies of 700.24 MHz for ^1H and 70.95 MHz for ^{15}N . The spectral width values were 40 ppm in F1 and 18 ppm in F2 dimensions, and 8 scans were used. The data were then Fourier transformed into 2048 \times 512 data points in the F2 \times F1 dimensions. The spectrum was recorded using the `hsqcfpf3gpplwg` program with 1 s recycle delay time and a $^1\text{J}_{\text{NH}}$ coupling constant of 90 Hz.

5.2.3.6.3 3D Heteronuclear NMR Spectroscopy

3D (^1H , ^{15}N , ^1H) heteronuclear NMR spectra (TOCSY-HSQC and NOESY-HSQC) were recorded with 2048 in F3 (^1H), 200 data points in F1 (^1H), and 34 in F2 (^{15}N). Spectra were recorded at the same operating frequencies of 700.24 MHz for ^1H and 70.95 MHz for ^{15}N . The spectral width values were 17 ppm in F1 and F3 and 34 ppm in F2 dimensions, and 4 scans were used. The data were Fourier transformed into 2048 \times 64 \times 512 data points in the F3 \times F2 \times F1 dimensions. The TOCSY-HSQC spectrum was recorded using the `dipsihsqcf3gpsi3d` program (60 ms mixing time) and the program for NOESY-HSQC was `noesyhsqcfpf3gpsi3d` (100 ms mixing time).

Bruker Topspin v. 3.2 software was used to process 3D NMR spectra. After phase and baseline corrections, spectra were analysed in Sparky v3.114¹⁸⁸.

5.3 RESULTS AND DISCUSSION

5.3.1 Binding stoichiometry and affinity of *pbmtA* to SynZur

To assess the affinity of SynZur to *pbmtA* without EDTA and zinc added, EMSA experiments were conducted as described for *pznuA* in section 3.2.2.1. The *pbmtA* insert was 163 bp and contained 143 bp of the operator-promoter region of *synbmtA*. It contains two predicted zur-boxes (see Figure 8.27, Appendix). The resulting EMSA gel is shown in Figure 5.1.

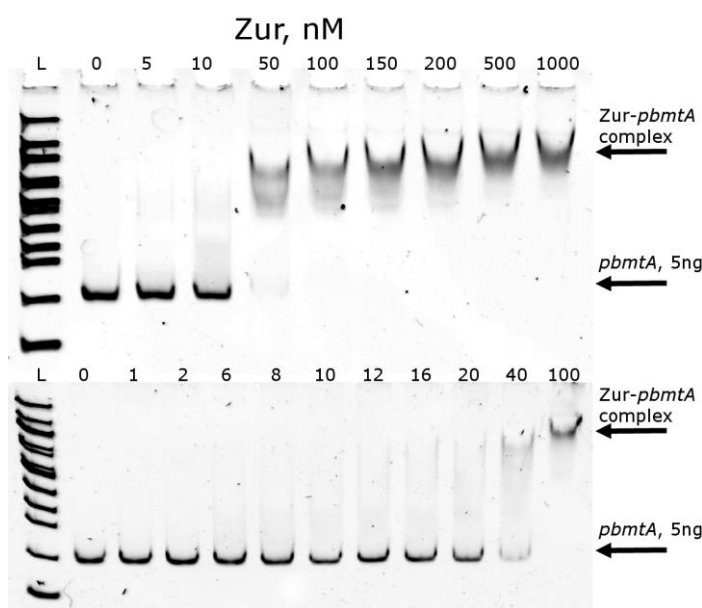
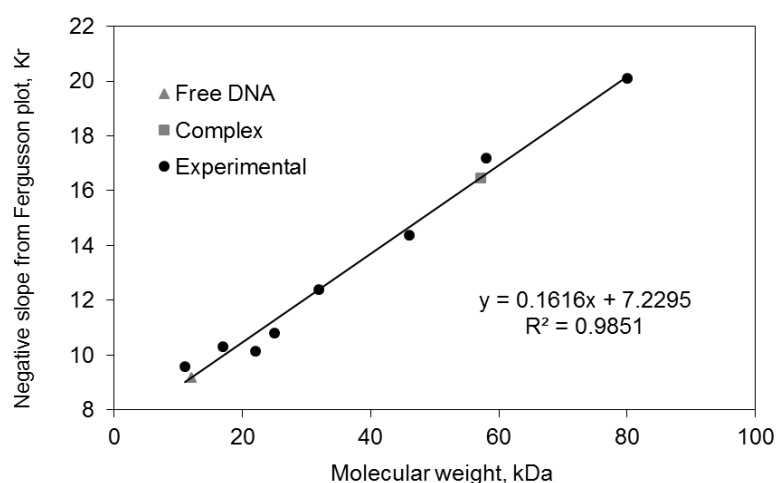


Figure 5.1 EMSA of *pbmtA* with SynZur.

Free DNA and the shifted SynZur-DNA complex are indicated with black arrows. Tris-glycine running buffer was used for the experiment. The specificity of binding of SynZur was tested previously⁸⁸. EMSA conditions are described in sections 5.2.1 and 3.2.2.1. Total DNA concentration is 4.96 nM. Zur concentration is given for the monomer.

If the gel figure is compared to the figure obtained previously for *pbmtA*⁸⁸, the shift takes place at significantly lower concentrations: here almost no free DNA is observed at 50 nM SynZur, while previously no free DNA was visible only at 200 nM SynZur. Possibly, the presence of EDTA in the running buffer affected the apparent affinity of SynZur for *pbmtA* in a similar way as for *pznuA* (see section 3.3.2.7). However, it has also to be taken into consideration that in the present study the *pbmtA* insert was significantly shorter (163 bp vs 250 bp in the previous study). Therefore, it is also possible that the shorter sequence could bind SynZur with slightly higher affinity due to for example lower steric or electrostatic restrictions.

The molecular weight of the Zur-DNA complex was calculated as for *pznuABC* (see section 3.3.2.7.1), and also as described⁴⁸. Briefly, the Rf of the ladder proteins, free DNA and the complex were calculated for 8% and 12% polyacrylamide gels based on the travel distances in gels after PAGE. Then, Ferguson plots were built for Rf ($100 \times \log(\text{Rf} \times 100)$) vs gel percentage. Finally, negative values of the slopes of Ferguson plots were plotted against molecular weight of the ladder proteins and the found “molecular weight” of the free DNA subtracted from that of the complex. The results are given in Figure 5.2.



| | "MW" | "MW" _{-DNA} |
|---------|-------|----------------------|
| Complex | 57.18 | 45.10 |
| DNA | 12.08 | |

Figure 5.2 Determination of the molecular weight of the SynZur-*pbmtA* complex.

The values found are given in the Table below the plot. Experimental values marked with black circles correspond to different ladder bands.

Interestingly, the observed value of SynZur bound to *pbmtA* at 150 nM SynZur was ~45 kDa, which roughly corresponds to a trimer. However, this value also matches values found by ND SDS-PAGE and analytical SEC for native SynZur which then crystallized as a tetramer (see sections 3.3.2.3, 3.3.2.4 and 3.3.3). Binding a tetramer or two dimers is in agreement with the fact that *pbmtA* contains two predicted zur-boxes (see Figure 4.21 and Figure 8.27, Appendix). Therefore, the predicted complex most likely corresponds to *pbmtA*-Zur₄. Interestingly, for the *pznuABC*-Zur₂ complex the apparent molecular weight matched the predicted MW. It is unclear why in the case of *pbmtA* such a significant difference was observed.

To confirm that the complex for *pbmtA*-Zur₄ runs more slowly than for *pznuA*-Zur₂, both were run together on the same gel at 150 nM Zur (Figure 5.3) and the result

broadly confirms the conclusion from Ferguson plot analysis. To equalise the DNA length, 250 bp *pbmtA* and 256 bp *pznuA* sequences were used as designed previously⁸⁸.

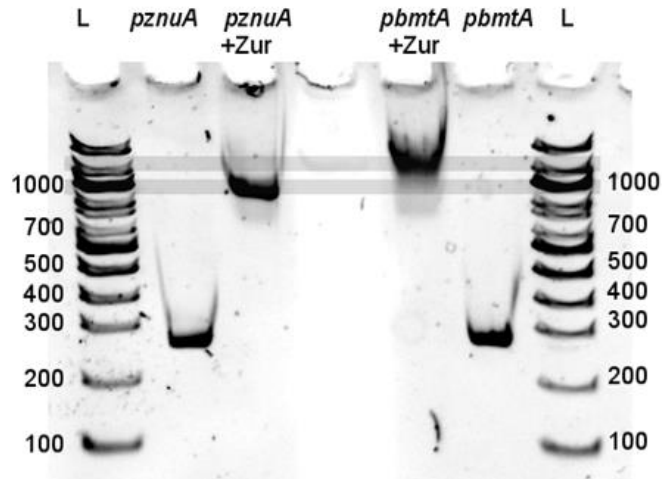
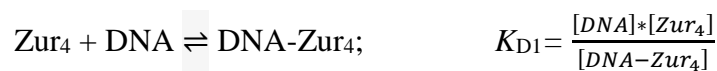


Figure 5.3 Comparison of *pznuA*-Zur and *pbmtA*-Zur complexes by EMSA.

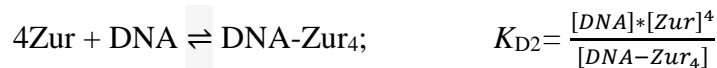
Semi-transparent lines connect matching ladder bands to indicate different levels. 250 bp *pbmtA* and 256 bp *pznuA* sequences were used as designed previously⁸⁸. EMSA conditions are described in sections 5.2.1 and 3.2.2.1. Total DNA concentration is 3.16 and 3.23 nM for *pznuA* and *pbmtA* respectively. Zur concentration is 150 nM.

Finally, to calculate the K_D for the suggested stoichiometry of *pbmtA*-Zur₄, three models were explored:

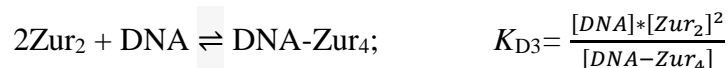
Model 1: A simple 1:1 interaction of a SynZur tetramer with DNA



Model 2: Interaction of four SynZur monomers with DNA



Model 3: Interaction of two SynZur dimers with DNA



Bands from two EMSA experiments were quantified by ImageJ. Models were fitted using DynaFit¹⁰⁰ software based on the amount of free DNA. The scripts for all three models are given in Table 8.5 in the Appendix. The resulting plot is shown in Figure 5.4.

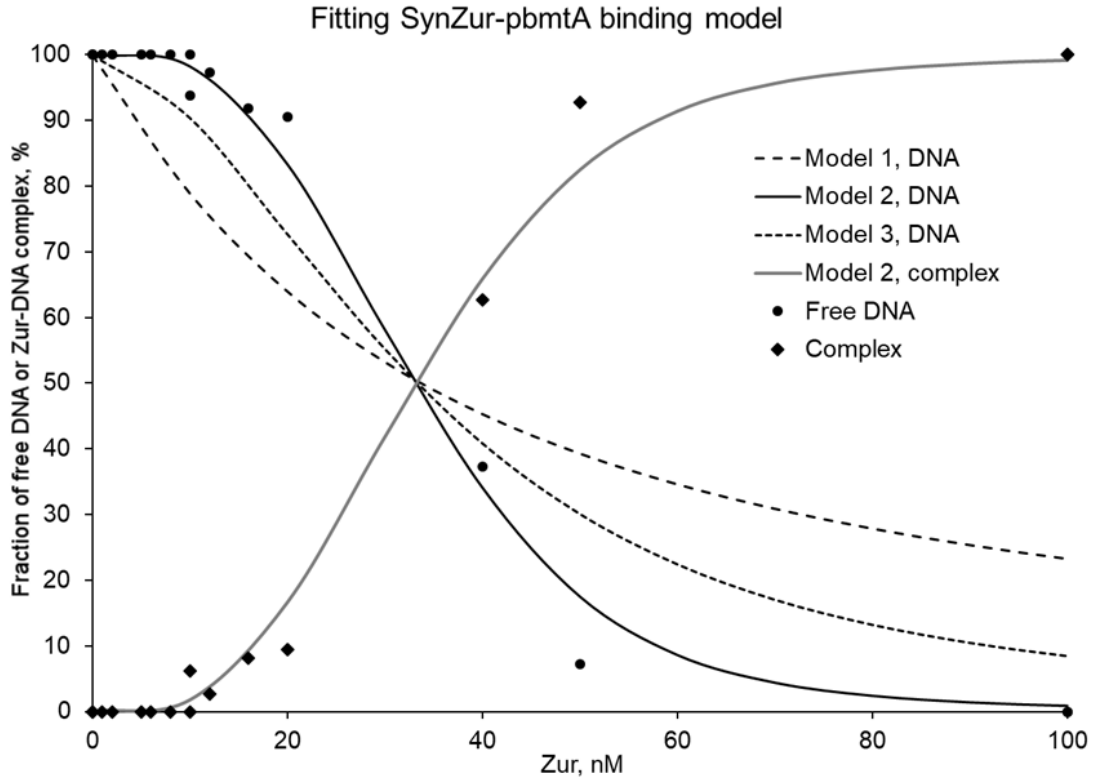


Figure 5.4 Visualisation of the EMSA data and fitted models.

Fitting was performed to the values of free DNA at each SynZur concentration. Points represent data from two replicates. For the dimer and the tetramer, double and quadruple concentrations are plotted for the X-axis to represent the concentration of the monomer. $K_{D1} = 6.8 \pm 2.0$ nM; $K_{D2} = 5.07 \pm 0.68 \times 10^6$ nM⁴ or corresponding to 26.7 nM per monomer; $K_{D3} = 181 \pm 41$ nM² or 13.5 nM per dimer.

As can be seen from Figure 5.4, the best fitting model is model 2. However as was concluded in section 3.3.2.7.2, there could be an issue with using either too low SynZur concentrations or not knowing the dimerization/tetramerisation K_{DS} .

Unfortunately, only a few points were present at the intermediate fraction values for the DNA-protein complex. Therefore, no conclusion could be drawn about cooperativity of binding using Hill equations^{189,190}. Due to time limitations these experiments were not repeated in the required range.

5.3.2 SynBmtA over-expression and purification

5.3.2.1 Plasmid construction for SynBmtA purification

Synechococcus sp. WH8102 *bmtA* (SYNW0359) was amplified by PCR with primers BmtA_F and BmtA_Re as described in section 5.2.2.1. The product size was

188 bp as predicted (see Figure 5.5). The sequence of the *bmtA* PCR product is shown in the Appendix, Figure 8.28.

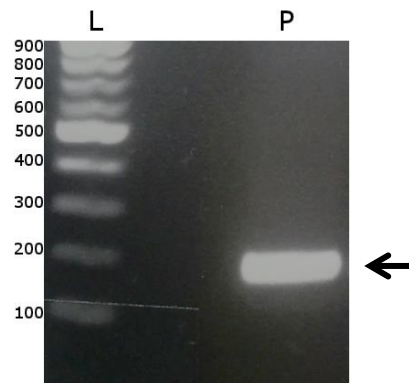


Figure 5.5 Agarose gel electrophoresis of PCR products obtained with primers *BmtA_F* and *BmtA_Re* using *Synechococcus* sp. WH8102 genomic DNA as template.

L – 100 bp NEB ladder, P – PCR product. The product matches the theoretical size of 188 bp.

The plasmid construction work flow (section 5.2.2.1) was similar as for the construction of the pGP704CmKmZur plasmid (see section 4.3.1.1.). PCR products using *T7_F* and *T7_Re* primers were confirmed by agarose gel electrophoresis (see Figure 5.6; the theoretical size of the product was predicted to be ~380 bp) and subsequent Sanger sequencing (not shown).

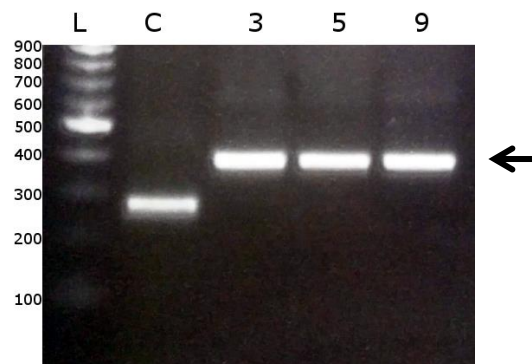


Figure 5.6 Agarose gel electrophoresis of pET24a(+)BmtA plasmids after PCR with *T7* primers.

The theoretical product size is 380 bp. L – 100 bp ladder, NEB, C – control PCR with *T7_F* and *T7_Re* primers using pET-24+(a) vector without the insert as a template (The theoretical product size is 259 bp). 3, 5 and 9 are colony-coding numbers.

The plasmid for SynBmtA over-expression was named pET24a(+)BmtA. This construct was used for protein over-expression as described below.

5.3.2.2 SynBmtA over-expression

SynBmtA over-expression was performed using similar conditions as for SynZur with the exception of adding ZnSO₄ at the point of induction, as described in section 5.2.2.2. Visualisation of the gel by both silver staining (see section 5.2.2.3.3) and SimplyBlue SafeStain after SDS-PAGE (see section 2.4.1) is shown in Figure 5.7.

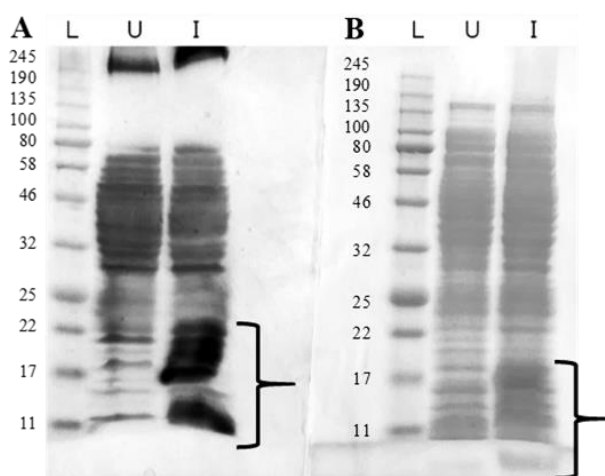


Figure 5.7 SDS-PAGE of BmtA over-expression.

A. Visualization by silver staining. B. Visualization by SimplyBlue SafeStain. L-protein ladder, NEB, U – un-induced culture and I – culture induced with 0.5 mM IPTG, 0.5 mM ZnSO₄. Areas where the SynBmtA protein is likely observed are highlighted with black brackets (0-22 kDa).

The predicted size of the protein (~6 kDa) differs from what can be seen on the gel. However, this phenomenon often occurs with metallothioneins and was previously observed in our group^{153,191}. The reason might be a very high affinity of metallothioneins to zinc resulting in a very stable fold and incomplete denaturation. Zinc is probably not removed by the standard denaturation protocol with SDS. Therefore, SynBmtA has a higher mass/charge ratio than the corresponding ladder proteins and travels more slowly on the gel compared to them. For example, native SmtAZn₄ was shown to behave similarly appearing at a significantly higher molecular weight (>70 kDa); however, apo-SmtA showed up at its expected molecular weight¹⁹².

5.3.2.2.1 Chemical precipitation

Chemical precipitation of SynBmtA was performed as described in section 5.2.2.3.1. SDS-PAGE of the fractions obtained is shown in Figure 5.8.

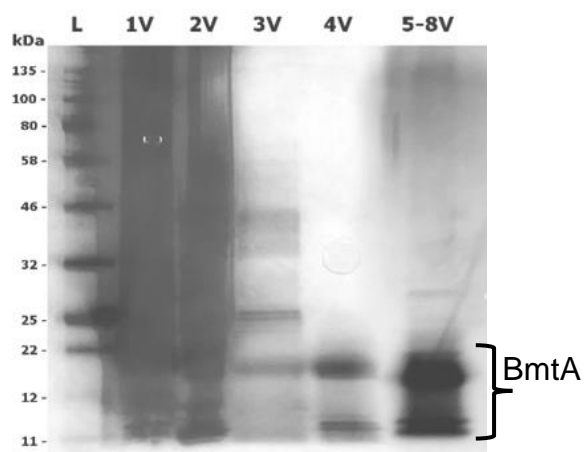


Figure 5.8 SDS-PAGE of BmtA fractions following chemical precipitation visualized by silver-staining.

L – ladder (#P7712S, New England Biolabs). Fractions representing relative amounts of the precipitant (ethanol/chloroform (100:8 v/v)) added are shown on the x-axis as volumes (V).

As can be seen in Fig. 5.8 most proteins precipitated after the addition of two volumes of precipitant. Potential SynBmtA precipitates obtained during the addition of 4 and more volumes of precipitant look quite pure. Precipitates from fractions 4 to 8 volumes were dissolved in 20 mM NH_4HCO_3 , combined and filtered through a 0.22 μm pore size syringe filter before Size Exclusion Chromatography (SEC).

5.3.2.2.2 Size Exclusion Chromatography

SEC was performed as described in section 5.2.2.3.2. The chromatogram is shown in Figure 5.9.

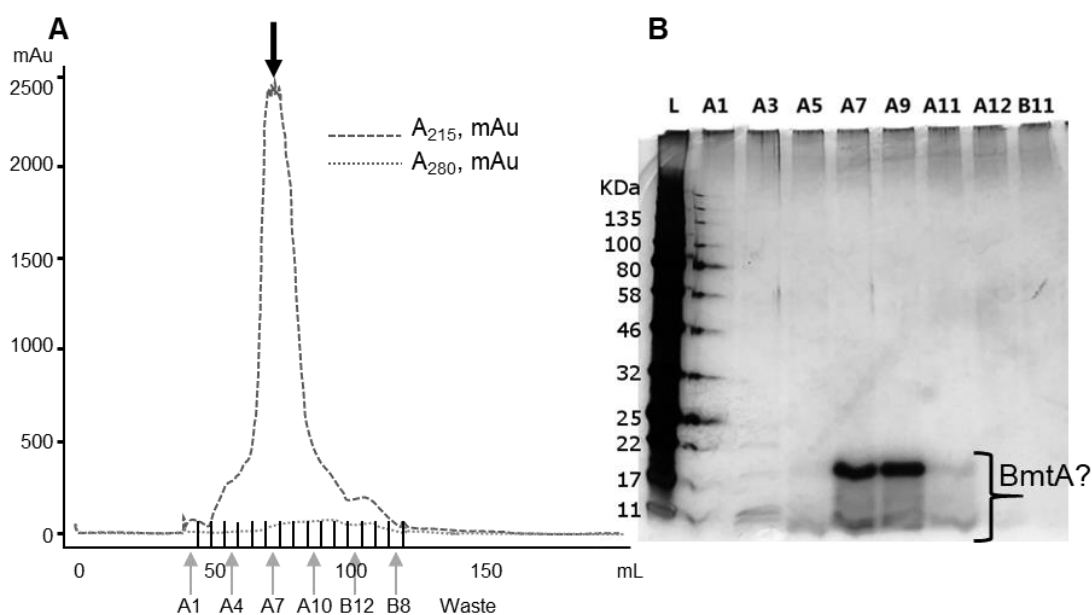


Figure 5.9 Size exclusion chromatography of SynBmtA.

A. SEC visualisation of SynBmtA obtained after the chemical precipitation step. **B.** SDS-PAGE of the fractions obtained after SEC, visualised with silver-staining. L – ladder (#P7712S, New England Biolabs) and fractions are shown on the x-axis. Potential SynBmtA is marked with a bracket.

The chromatogram mostly matched the theoretical predictions based on previous work with cyanobacterial MTs (*sync_1081* from *Synechococcus* sp. CC9311): one large peak is observed which is consistent with the relative purity expected based on SDS-PAGE¹⁵³. As shown in Figure 5.9, the eluted protein has a high absorbance at 215 nm and almost no absorbance at 280 nm. This is typical for MTs since they usually lack amino acid residues contributing to A₂₈₀ (Trp, Tyr or Cys in disulfide bonds)¹⁰⁸.

As can be seen in Figure 5.9, fractions A7-A9 contained SynBmtA. Fractions A6 and A10 were checked for the presence of the protein as well (data not shown), with fraction A6 also containing a sufficient amount of SynBmtA, whereas in A10 the protein concentration was low. Based on the SDS-PAGE and chromatogram data, the protein was considered pure enough for further use. Fractions A6-A9 were collected and concentrated using an Amicon filter (see section 2.4.2).

It is interesting that two bands were observed in SDS-PAGE (Figure 5.9). It is possible that the lower band corresponds to unfolded, possibly partially denatured SynBmtA while the upper band represents the folded protein showing up at higher apparent molecular weight as discussed in section 5.3.2.2. Alternatively, it could be an unknown impurity.

ESI-MS of SynBmtA at pH 2 showed that the molecular weight of the protein was in agreement with the one theoretically predicted (see section 5.3.3.1 for details). This confirmed that SynBmtA had been over-expressed and purified.

5.3.3 Characterisation of SynBmtA

5.3.3.1 ESI-MS

ESI-MS was used to identify SynBmtA by its molecular weight. At pH 2 many zinc metalloproteins appear in their apo-form. Thus, ESI-MS measurement at this pH was performed. A full ESI-MS spectrum of SynBmtA at pH 2 together with the deconvoluted spectrum are shown in Figure 5.10.

The theoretical molecular mass of SynBmtA calculated with ExPASy¹⁰⁷ is 5751.61 Da without any post-translational modifications or 5620.42 Da with the removal of the first methionine. Comparison of calculated and found molecular mass values of apo-SynBmtA without the first methionine is shown in Table 5.5. Based on the found mass, it was assumed that SynBmtA undergoes removal of the amino-terminal methionine by methionine aminopeptidase¹⁹³.

Table 5.5 Values of theoretical and found molecular masses for apo-SynBmtA without the initiator methionine.

Average neutral mass was calculated using the ExPASy Compute pI/Mw tool¹⁰⁷.

| | Theoretical | Found | Difference |
|-----------------|-------------|---------|------------|
| Apo-BmtA (pH 2) | 5620.42 | 5619.32 | 1.10 |

A single charge state (+5) of SynBmtA is shown in Figure 5.10B. The main peak, corresponding to a deconvoluted neutral mass of 5619.70 Da, agrees with the mass for apo-SynBmtA. Also, several minor peaks corresponding to 5641.75, 5663.5, 5684.25 and 5706.35 Da are present. The differences in mass with apo-SynBmtA are 22.05, 43.8, 64.55 and 86.65 respectively, and likely correspond to sodium adducts, BmtANa_n, where n = 1-4. Several more peaks with the same mass difference corresponding to Na adducts are also seen in the spectrum. It is unclear why the sample was contaminated with sodium. Possibly, either the buffer was contaminated during pH adjustment or the ESI-MS instrument was the source of sodium.

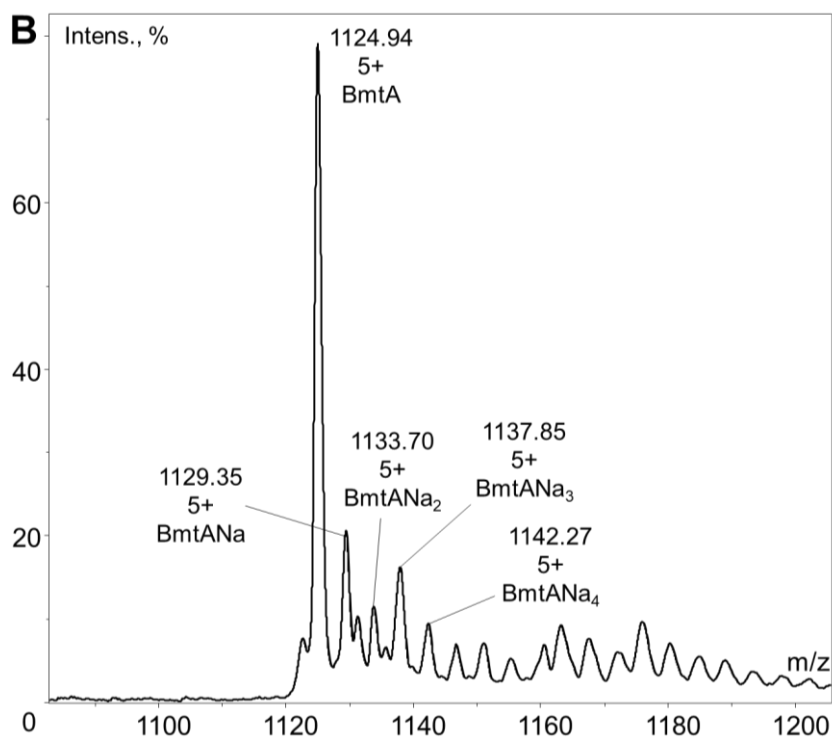
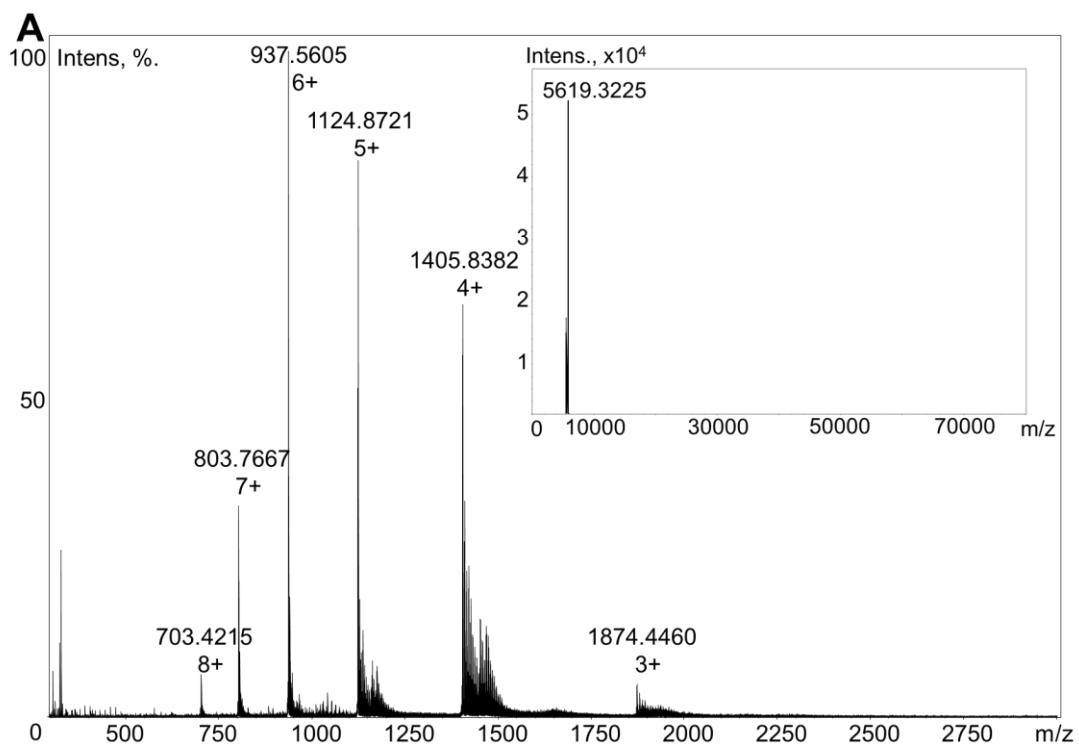


Figure 5.10 ESI-MS spectrum of SynBmtA at pH 2.

The spectrum was recorded on a Bruker MaXis Plus, with 20 μM SynBmtA in 20 mM NH_4HCO_3 , 2% (v/v) HCOOH . A. Full range spectrum. The deconvoluted spectrum is shown in the top right corner. B. Zoomed in and smoothed +5 charge state. Smoothing resulted in a slight change of the peak values.

At neutral pH proteins may be observed in their native (metallated) form. A full ESI-MS spectrum of BmtA at pH 7.9 together with the deconvoluted spectrum is shown in Figure 5.11A.

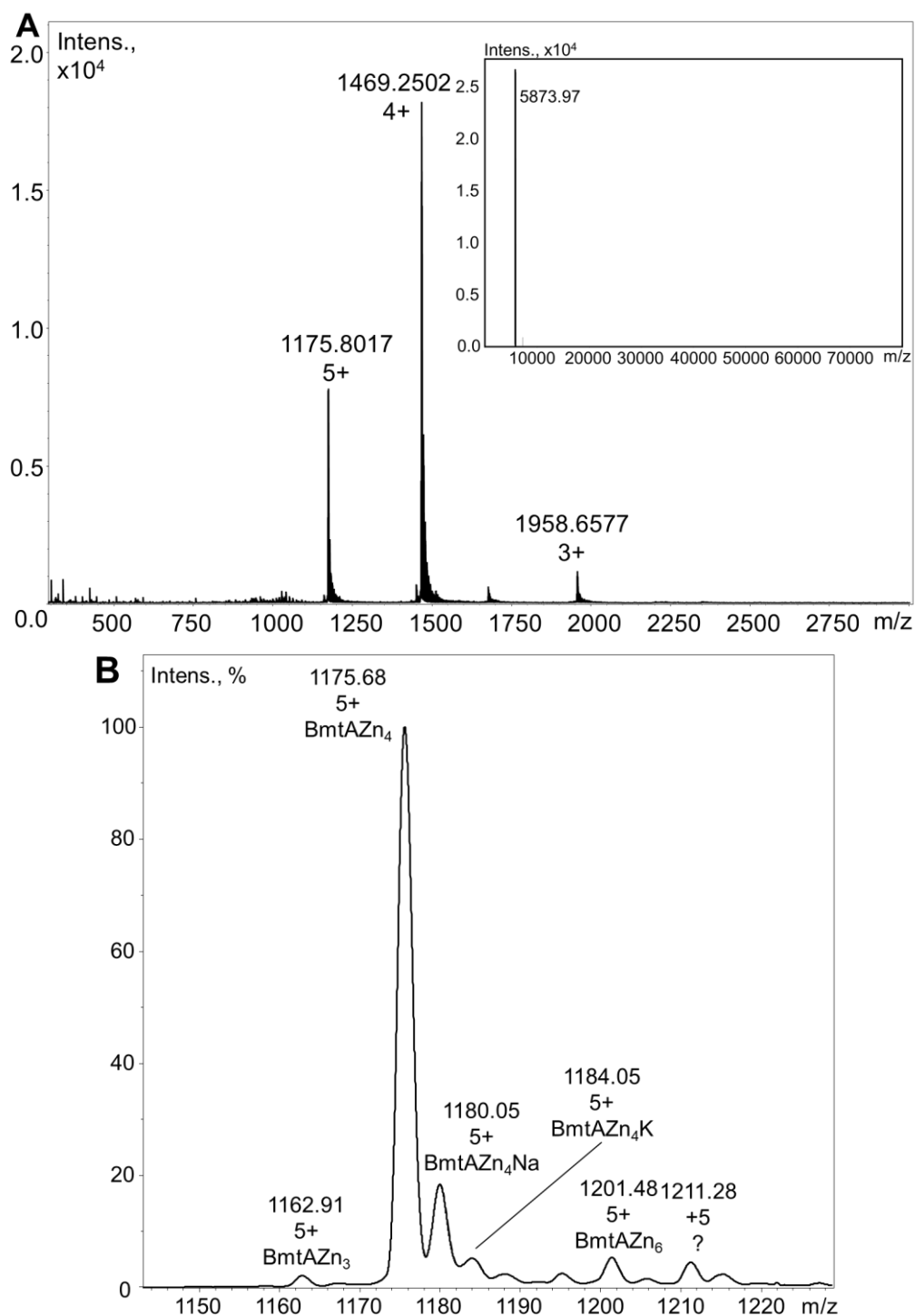


Figure 5.11 ESI-MS spectrum of native SynBmtA at pH 7.9.

The spectrum was recorded on a Bruker MaXis Plus, with 20 μ M SynBmtA in 20 mM NH_4HCO_3 . **A.** Full range spectrum. The deconvoluted spectrum is shown in the top right corner. **B.** +5 charge state in the smoothed spectrum. Smoothing resulted in a slight change of the peak values.

The main peak of native SynBmtA corresponds to a species containing 4 Zn²⁺ ions. The molecular weight of SynBmtAZn₄ was calculated and compared with the results of ESI-MS. Comparison of the found molecular mass with the theoretical mass is shown in Table 5.6. The found value is in excellent agreement with the calculated one.

Table 5.6 Found and calculated molecular mass values for SynBmtA at pH 7.9.

Average neutral mass was calculated using the ExPASy Compute pI/Mw tool¹⁰⁷.

| | Theoretical | Found | Difference |
|----------------------------|-------------|---------|------------|
| BmtAZn ₄ (pH 8) | 5873.97 | 5873.94 | 0.03 |

A single charge state (+5) of the native SynBmtA is shown in Figure 5.11B. The main peak with mass 5873.40 Da corresponds to the species BmtAZn₄. Several minor peaks corresponding to 5809.55, 5895.25, 5915.25, 6002.40 and 6051.40 Da are also present. The differences in mass with BmtAZn₄ are -63.85, 21.85, 41.85, 129.00 and 178.00, respectively and likely correspond to BmtZn₃, BmtAZn₄Na, BmtAZn₄K and/or BmtAZn₄Na₂, BmtAZn₆ and possibly BmtAZn₆Na₂, respectively. However, the mass difference for the last peak differs from the predicted mass by 5.0 Da. Alternatively, it could be an oxidized BmtAZn₇ species or a copper containing species. BmtAZn₆ could also be a BmtAZn₄ species where the initiator methionine was not cleaved by methionine aminopeptidase¹⁹³. This was previously observed for a recombinantly expressed MT encoded by *sync_2426* in *Synechococcus* sp. CC9311¹⁵³.

5.3.3.2 ICP-OES

The next step was to find a way to quickly assess the SynBmtA concentration. Since it has negligible A₂₈₀, the A₂₃₀ was measured prior to ICP-OES analysis. To calculate the concentration of SynBmtA, first its sulfur content had to be established (see Figure 5.12).

STAIK**C**AC**P**K**C**T**C**MVAEESAIVLQGKFF**C**STS**C**STGHANNE**P**CHGEGS
CG**C**K**C**GE

Figure 5.12 Sulfur containing residues of SynBmtA without the first methionine residue.

Thus, each SynBmtA molecule contains 10 cysteines and 1 methionine since it was shown in section 5.3.3.1 that SynBmtA has lost the initiator methionine. Based on sulfur, protein concentration in the samples was calculated, and ϵ_{230} was calculated to be $37374 \pm 2178 \text{ M}^{-1}\text{cm}^{-1}$ in the same way as was described for SynZur (see section 3.3.2.1). The calculation is shown in Table 5.7. Knowing ϵ_{230} allowed calculation of the yield of the protein which was 4-5 mg per litre of *E. coli* culture.

Table 5.7 Calculation of the ϵ_{230} for SynBmtA.

| # | S, ppm | S, M | Protein conc., M | A ₂₃₀ | ϵ_{230} |
|--|--------|----------|------------------|------------------|------------------|
| 1 | 9.149 | 2.85E-04 | 2.59E-05 | 1.063 | 40981 |
| 2 | 2.734 | 8.53E-05 | 7.75E-06 | 0.264 | 34059 |
| 3 | 1.127 | 3.51E-05 | 3.20E-06 | 0.106 | 33175 |
| 4 | 1.6319 | 5.09E-05 | 4.63E-06 | 0.191 | 41282 |
| Average ϵ_{230} | | | | | 37374 |
| St. error | | | | | 2178 |

The metal content of SynBmtA analysed by ICP-OES is given in Table 5.8: zinc content was found to be 4.6 ± 0.13 per molecule. This is in agreement with the previously found ratio for the homologous BmtA encoded by *sync_1081* from *Synechococcus* sp. CC9311 which also had 4 zinc ions per molecule¹⁵³ and SmtA from *Synechococcus* sp. PCC7942⁸¹.

Table 5.8 ICP-OES analysis of SynBmtA:

The content of trace metals in the analysed SynBmtA samples. Four replicates were used to calculate errors.

| Element | Ratio to BmtA | St. error |
|---------|---------------|-----------|
| Zn | 4.60 | 0.13 |
| Ni | 0.104 | 0.043 |
| Cu | 0.168 | 0.061 |
| Cd | 0.057 | 0.024 |
| Mn | 0.127 | 0.053 |

It is likely that extra zinc in the protein came from the lysing solution which contained 1 mM ZnSO₄ (see section 5.2.2.3.1.). A slight excess of zinc found by ICP-OES was confirmed by ESI-MS previously: a minor peak corresponding to BmtAZn₆ was observed. ICP-OES showed that SynBmtAZn₄ did not contain significant impurities of other metals.

5.3.3.3 pH titration of SynBmtA

pH-dependent displacement of zinc in native MTs is a traditionally used indicator of the metal-binding affinity of metallothioneins. A lower pH value of half displacement ($\text{pH}_{1/2}$) indicates higher metal binding affinity. For example, by this method *Synechococcus* sp. PCC7942 SmtA showed higher zinc affinity in comparison to equine MT¹⁹⁴ (Table 5.9).

The value of $\text{pH}_{1/2}$ is monitored spectrophotometrically: ligand to metal charge transfer for Zn-thiolate bonds in metallothioneins is observed at ~230 nm¹⁹⁵. Upon lowering the pH and protonating, cysteine thiols release zinc and the spectrum would show a decrease at this wavelength. The titration curve of SynBmtA is shown in Figure 5.13.

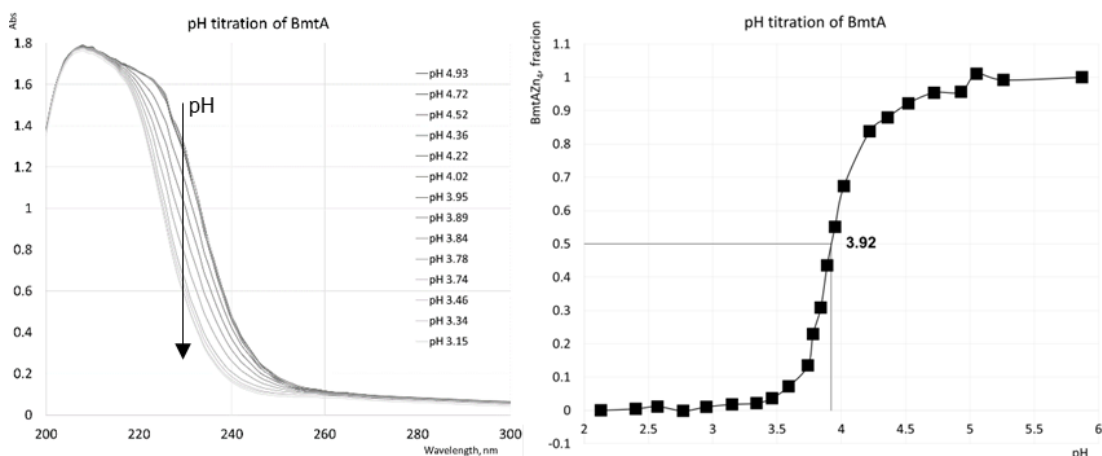


Figure 5.13 pH titration of SynBmtA.

A. Titration curves in the range of 200-300 nm. B. Normalised A₂₃₀ change.

Based on the titration curve, pK_{1/2} was found to be 3.92 for SynBmtAZn₄. This value is in overall agreement with other values for MT from cyanobacteria (see Table 5.9). The data in Table 5.9 also suggest that MTs from other organisms usually have higher values of pK_{1/2} in comparison to BmtAs, which indicates lower zinc affinity. This fact might be related to dealing with lower intracellular zinc levels for bacteria in comparison to plants or mammals.

Table 5.9 pH values of half displacement for metallothioneins from different organisms.

| Organism | pH _(1/2) | Source |
|----------------------------------|---------------------|-----------|
| Pea | 5.25 | 196 |
| Wheat | 4.7 | 196 |
| Human | 4.6-4.9 | 196 |
| Equine renal MT | 4.50 | 196 |
| <i>Synechococcus</i> sp. PCC7942 | 4.10 ~4 | 194 82 |
| <i>Anabaena</i> PCC 7120 | >4.1 | 82 |
| <i>Pseudomonas aeruginosa</i> | <4.1 | 82 |

5.3.3.4 Investigation of SynBmtA de-metallation

Section 3.3.2.5 showed that SynZur was de-metallated at the sensory site by reaction with a significant excess of EDTA overnight. This approach was used to investigate the removal of zinc from native SmtA¹⁹⁷. Thus, SynBmtA de-metallation was attempted using 100-fold excess of EDTA. The spectrum of SynBmtA treated with EDTA (SynBmtA_E) overnight after the buffer exchange step with a PD-10 column is shown in Figure 5.14 (third spectrum from the top).

Although a significant proportion of apo-SynBmtA was obtained, in principle demonstrating that SynBmtA-bound zinc ions are accessible to EDTA, it was found that de-metallation was not complete even with 100-fold excess of EDTA: significant amounts of BmtAZn₄ were still present. A very similar picture was shown previously for the homologous *sync_1081*: only the same two species were observed and de-metallation was incomplete after 24 h¹⁵³. It was suggested that upon zinc removal from BmtAZn₄, as soon as a single ion is removed, dissociation of the remaining zinc ions occurs more easily.

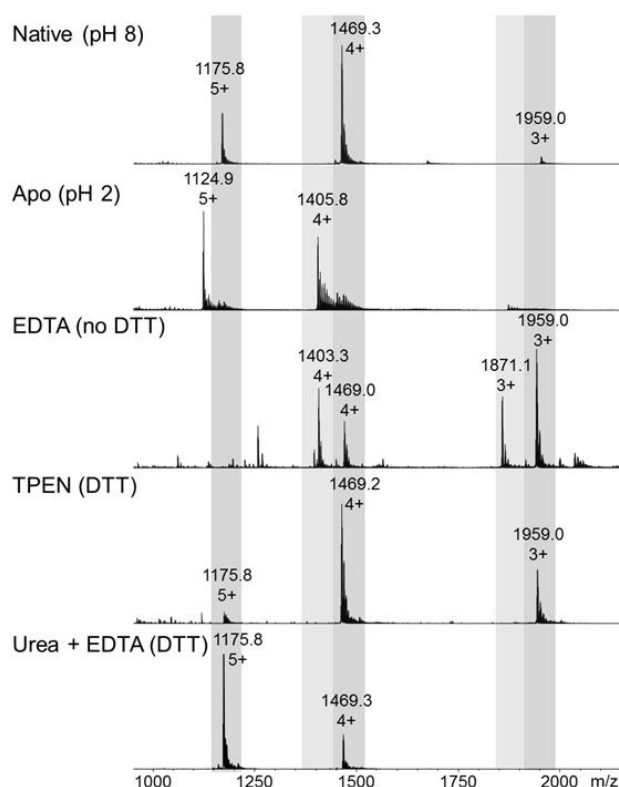


Figure 5.14 Investigation of SynBmtA de-metallation.

The first two spectra are given for comparison. Conditions were 10 μ M SynBmtA in 20 mM NH₄HCO₃, pH 7.9, except for the pH 2 condition, where 2% (v/v) formic acid was added. Other conditions were 1 mM EDTA; 1 mM TPEN, 0.5 mM DTT; and 6 M urea, 1 mM EDTA, 0.5 mM DTT, respectively.

A comparison of SynBmtA_E and SynBmtA at pH 2 showed that the molecular weight of SynBmtA_E was only 5609.75 compared to 5619.32 Da for SynBmtA. The difference of 9.57 Da likely corresponds to 10 oxidized cysteine residues forming disulfide bonds since no reducing agent was used, and the reaction had not been conducted under N₂. Under the conditions tested this oxidation is probably irreversible. Thus, oxidized apo-SynBmtA is slowly being accumulated since the mixture is not thermodynamically equilibrated.

To increase the yield of apo-SynBmtA and to get it in reduced form, it was decided to use a stronger chelator (TPEN) together with a reducing agent (DTT). The conditional $K_{D(Zn)}$ of TPEN is 6.4×10^{-16} M vs $K_{D(Zn)}$ of EDTA 2.3×10^{-14} M at pH 7.4¹¹⁶.

Surprisingly, in the presence of DTT after overnight treatment of SynBmtA with 100-fold excess of TPEN no apo-SynBmtA was observed (see Figure 5.14, TPEN (DTT)). Possibly, in the presence of a reducing agent, oxidation of the MT does not happen and de-metallation stops being irreversible and oxidized apo-protein does not accumulate. This could be due to either a significantly higher affinity of SynBmtA for zinc, which is unlikely considering very high affinity of TPEN to zinc and its 100 fold excess. Alternatively, this could be due to the very slow kinetics of zinc release from SynBmtA or due to structural/electrostatic hindrances, where zinc might be inaccessible to chelators.

If Zn-binding sites of SynBmtA are inaccessible in the folded protein, unfolding SynBmtA using urea in the presence of EDTA and DTT should be a viable approach to produce apo-SynBmtA. Urea is known to be a fast and strong denaturing agent for proteins¹⁹⁸. Surprisingly, after the overnight treatment and buffer exchange using a PD-10 column, the ESI-MS spectrum again showed unchanged BmtAZn₄ species (see Figure 5.14. Urea + EDTA (DTT)). It appears that the protein fold (along with the Zn-binding sites) of SynBmtA is extremely stable. Indeed, SynBmtA cannot be unfolded by urea or SDS leading to its slower traveling speed in SDS-PAGE (see section 5.3.2.2).

Interestingly, unsuccessful de-metallation of SynBmtA by EDTA is reminiscent of previous observations for SynZur. Its treatment by EDTA did not cause zinc release from the Cys₄ structural site (see section 3.3.2.5). The main zinc binding residues of SynBmtA are cysteines as well (see section 5.3.3.7).

However, as was shown in sections 5.3.3.1 and 5.3.3.3 and in Figure 5.14 (pH 2), it is possible to make apo-SynBmtA by acidification. Therefore, I tried this approach rather than using chelating/denaturing agents. Since at acidic pH Zn^{2+} is released, it was decided to use a PD-10 column to separate the MT from metal. However, if an eluting buffer at neutral pH is used, SynBmtA might immediately bind zinc as soon as the pH is increased. This led to the idea of using an eluting buffer at acidic pH. A value of pH 3.3 was chosen based on the pH titration curve where most of SynBmtA was de-metallated (see Figure 5.13). A lower pH should be avoided since the working range of a PD-10 column is 2-13¹⁹⁹. A similar approach of generating apo-proteins by acidification and using gel-filtration chromatography was developed previously²⁰⁰.

After this step, ESI-MS showed the formation of apo-SynBmtA (data not shown). Upon pH adjustment of the sample to pH 8 with NH_4OH and checking by ESI-MS, the sample contained significant amounts of potassium adducts probably from the pH-electrode (not shown). An alternative method was thus required. For pH adjustment of the acidified Zn-free apo-SynBmtA I thus decided to add another buffer exchange step with a PD-10 column into a neutral eluting buffer as described in section 5.2.3.3. The resulting pH of apo-SynBmtA sample was 7.8 and the ESI-MS spectrum is shown in Figure 5.15.

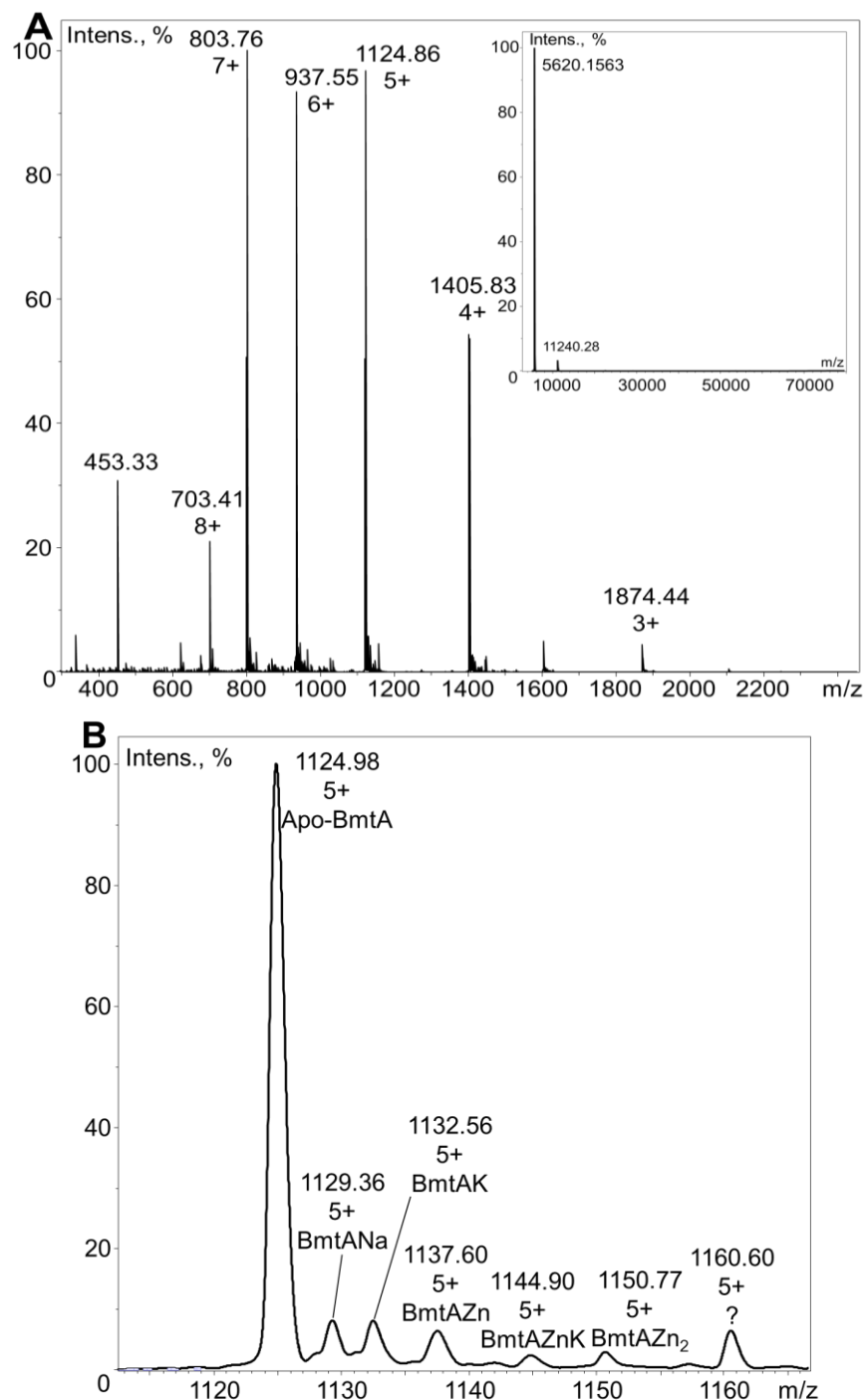


Figure 5.15 ESI-MS spectrum of apo-SynBmtA at pH 7.8.

The spectrum was recorded on a Bruker Compact with ~10 μ M apo-SynBmtA in 10 mM NH_4HCO_3 . **A.** Full range spectrum. The deconvoluted spectrum is shown in the top right corner. **B.** A single charge state (+5) for de-metallated apo-SynBmtA. Smoothing resulted in a slight change of the peak values.

The single charge state spectrum shows that the main peak corresponding to a molecular mass of 5619.90 Da is reduced apo-SynBmtA. Additional minor peaks with MW 5641.80, 5657.80, 5683.00, 5719.50, 5748.85 and 5798.00 giving a mass difference with the main peak of 21.90, 37.90, 63.10, 99.60, 128.95 and 178.10

correspond to BmtANa, BmtAK, BmtAZn, BmtAZnK, BmtAZn₂ and a similar difference for the last peak that was observed in section 5.3.3.1 (see Figure 5.11). This is probably a partially oxidized BmtAZn₃ species or an analogous copper containing one.

5.3.3.5 Zinc affinity of SynBmtA

Before titration of apo-SynBmtA with Zn²⁺, it was verified whether the protein returns to its original state upon addition of Zn²⁺. A four-fold excess of Zn²⁺ was added to apo-SynBmtA and the mixture checked by ESI-MS. The spectrum is shown in Figure 5.16.

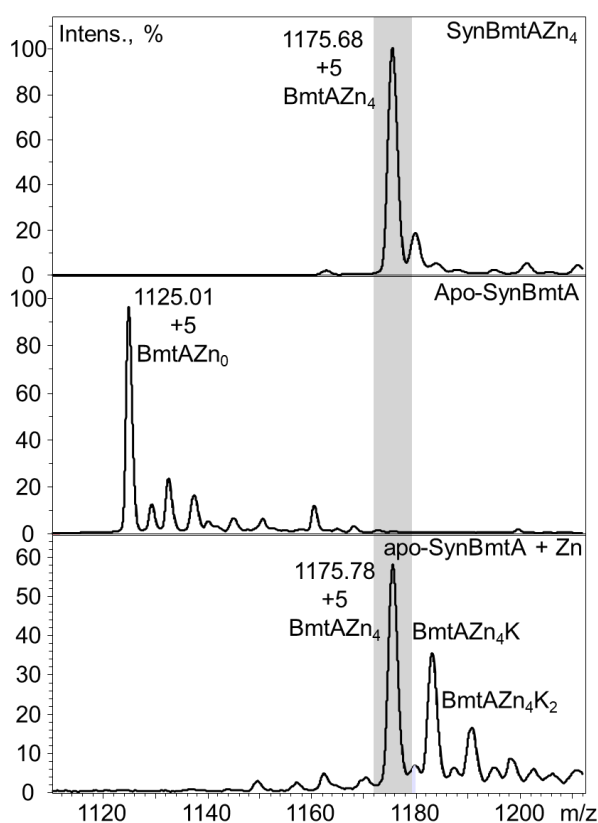
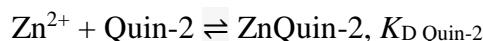


Figure 5.16 Reconstitution of apo-SynBmtA with zinc, pH 7.8.

The spectrum was recorded on a Bruker Compact with ~10 μM SynBmtA samples in 10 mM NH₄HCO₃. The +5 charge state is shown. The original native SynBmtAZn₄ is shown for comparison at the top of the image. Apo-SynBmtA spectrum is shown in the middle. Upon addition of Zn²⁺ solution to apo-SynBmtA, SynBmtAZn₄ is formed (the spectrum shown at the bottom). The protein was apparently contaminated with potassium during zinc addition which shows up as two adducts of BmtAZn₄K (5911.50 Da) and BmtAZn₄K₂ (5949.95 Da).

Thus, based on the ESI-MS data, apo-SynBmtA can be re-constituted with zinc and could be titrated in the presence of Quin-2 for calculation of a dissociation constant. Apo-SynBmtA was titrated under similar conditions to SynZurZn₁ (see

section 3.3.2.6). The script for calculation of a dissociation constant used in DynaFit¹⁰⁰ is shown in Figure 8.29 in the Appendix. A representative Quin-2 competition titration is shown in Figure 5.17, with parameter values obtained from a simple 1:1 binding model:



where BmtA' is the concentration of zinc binding sites in the SynBmtA protein and not the protein concentration itself. A molar extinction coefficient for apo-SynBmtA was not experimentally determined and total apo-SynBmtA concentration was not calculated independently, but derived from fitting as explained below.

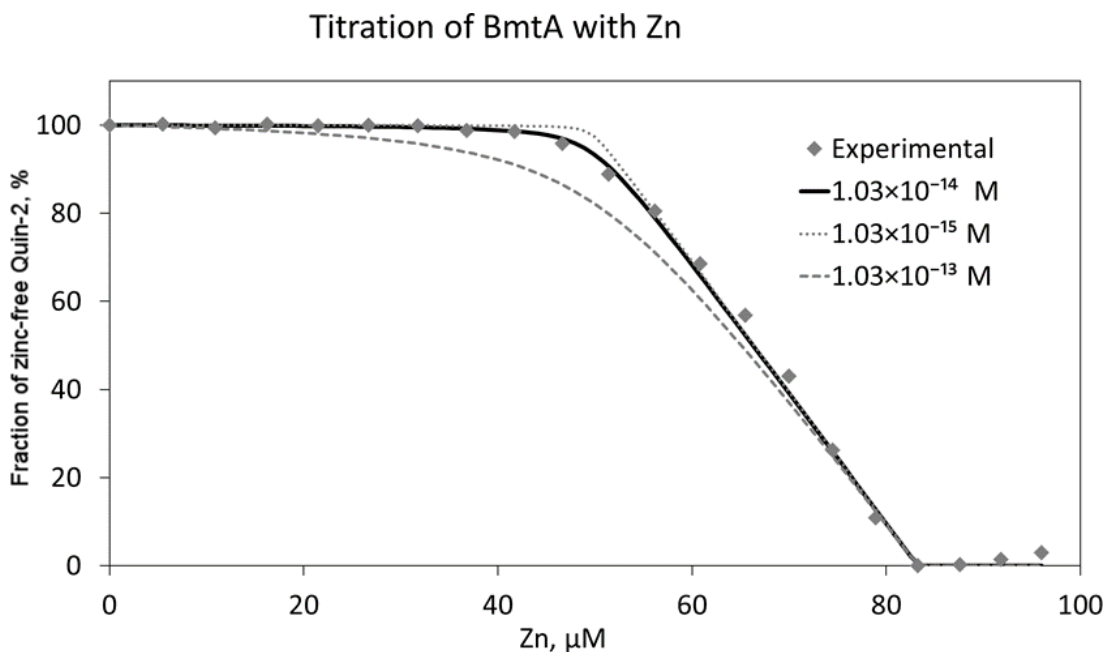


Figure 5.17 A representative titration curve of SynBmtA with Zn^{2+} in the presence of Quin-2.

Experimental values of normalised $A_{261 \text{ nm}}$ are shown with grey diamonds. Zn^{2+} was titrated into 33.67 μM Quin-2 and 49.59 μM apo-SynBmtA equivalents of zinc binding sites. The solid line represents the curve of best fit using DynaFit¹⁰⁰ software. Dashed lines are simulated curves describing the K_{D} of one order higher and lower values. The K_{D} was $1.13 \pm 0.07 \times 10^{-14} \text{ M}$ based on three replicates.

In order to assess the zinc affinity of SynBmtA, the concentration of zinc binding sites was calculated by subtracting the concentration of Quin-2 found by

spectrophotometry from the total zinc concentration added. As the Quin-2 concentration can be determined with high accuracy from absorbance, and because the resulting zinc protein binding site concentration refers to those actually available, the resulting K_D is meaningful. For a simple 1:1 interaction, the apparent SynBmtA-Zn K_D was $1.13 \pm 0.07 \times 10^{-14}$ M based on three replicates.

This value is only ~ half that of the K_D of EDTA (2.3×10^{-14} M) and ~18 times higher than the K_D of TPEN (6.4×10^{-16} M). Thus, it still does not explain why a 100-fold excess of EDTA or TPEN did not fully de-metallate SynBmtA.

However, as seen from the titration curve (Figure 5.17) the K_D for zinc of SynBmtA is significantly lower than that of Quin-2 (3.7×10^{-12} M). Given the almost double number of binding sites for SynBmtA, in addition to the more than two orders of magnitude difference in K_D s, it is possible that the K_D of SynBmtA cannot be reliably assessed using titration with Quin-2. Under the current conditions binding appears not competitive since Quin-2 fluorescence did not change until about half of the zinc was added (~40 μ M). In addition to the very high relative standard deviation values of residuals during the curve fitting (>40%), this leads to the conclusion that the observed K_D represents rather an upper limit, and the actual affinity of SynBmtA for zinc could be higher still. Perhaps titrating a more diluted apo-SynBmtA sample under a strict anaerobic atmosphere with Quin-2 or another metallo-chromic dye with higher zinc affinity would help to determine an even lower K_D value.

Moreover, titration of SynBmtA was different to SynZur. For SynZur, after addition of each zinc aliquot, equilibrium was reached within 10 minutes when the A_{261} was constant. However, for SynBmtA this equilibrium was sometimes only achieved after about 1 hour (see Figure 5.18). This was especially the case for the points next to the inflection on the titration curve which defines the K_D (see Figure 5.17, 40-60 μ M range).

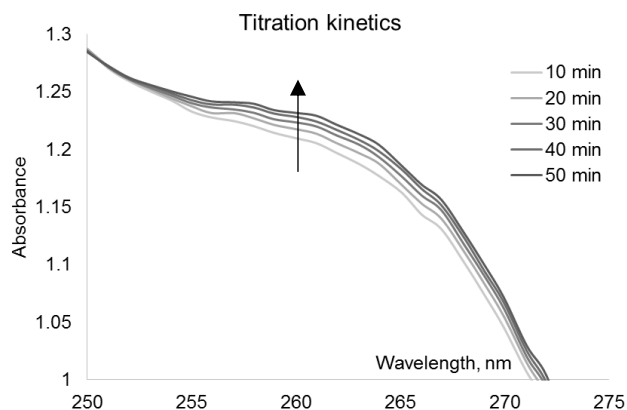


Figure 5.18 Absorbance equilibration of SynBmtA upon titration with zinc in the presence of Quin-2.

Conditions are the same as for Figure 5.17. The example shows absorbance equilibration after addition of $47.52 \mu\text{M Zn}_{\text{tot}}$.

The probable explanation is that zinc binding kinetics for SynBmtA are slower than for Quin-2. Thus, upon addition of Zn^{2+} , at first Quin-2 quickly binds Zn^{2+} , giving a lower A_{261} value. Subsequently, due to the higher affinity of SynBmtA for zinc but slower binding kinetics, Zn^{2+} is then slowly re-distributed to SynBmtA from Quin-2. This causes the value of A_{261} , representing the concentration of free Quin-2, to increase again.

Alternatively, increasing values at A_{261} could indicate oxidation of the sample since disulfide bonds absorb weakly at 260 nm^{201} . However, the reducing agent TCEP was present in the mixture and for addition of zinc aliquots the cuvette was opened only under nitrogen flow to reduce oxidation of free thiols. Also, the effect was most significant at the inflection of the titration curve and negligible on the flat part ($0\text{-}40 \mu\text{M Zn}_{\text{tot}}$). Thus, probably the formation of disulfide bonds can be neglected and the increase in A_{261} value could be explained by slow zinc binding kinetics of SynBmtA.

While all efforts were made to reach equilibrium, especially at the inflection of the titration curve, it was physically impossible to wait for more than 1 hour after addition of each aliquot. Thus, it is possible that some points on the curve show up below the actual values, which would overall lead to a less steep bend on the curve (see Figure 5.17). This in turn probably further underestimates the Zn-binding affinity of SynBmtA, and the value determined should be taken as an upper limit: $K_{\text{D}(\text{Zn})}$ of SynBmtA $\leq 1.13 \pm 0.07 \times 10^{-14} \text{ M}$.

This very low K_D value is also in agreement with the very low value of $pH_{1/2}$ equal to 3.92, which is one of the lowest observed for metallothioneins even from other cyanobacteria (see section 5.3.3.3). This extremely high affinity of SynBmtA for zinc might be a consequence of this strain living in extremely oligotrophic open ocean environments with very low zinc levels.

5.3.3.6 Competition for zinc between SynZur and SynBmtA

The calculated $K_{D(Zn)}$ for SynZur was $8.27 \pm 2.50 \times 10^{-13}$ M (section 3.3.2.6), which is higher than the K_D of SynBmtA ($\leq 1.13 \pm 0.07 \times 10^{-14}$ M) by almost two orders of magnitude. Given that I previously showed (section 4.3.3.5.3) that SynZur activates transcription of *synbmtA* under elevated zinc levels (therefore with zinc bound) producing apo-SynBmtA protein, this begs the question whether apo-SynBmtA is able to remove zinc from metallated SynZur as might be predicted based on their K_D values.

This was checked by ESI-MS: protein samples were run separately and as a mixture of 4:1 for SynZurZn₂ to apo-SynBmtA, respectively (see Figure 5.19).

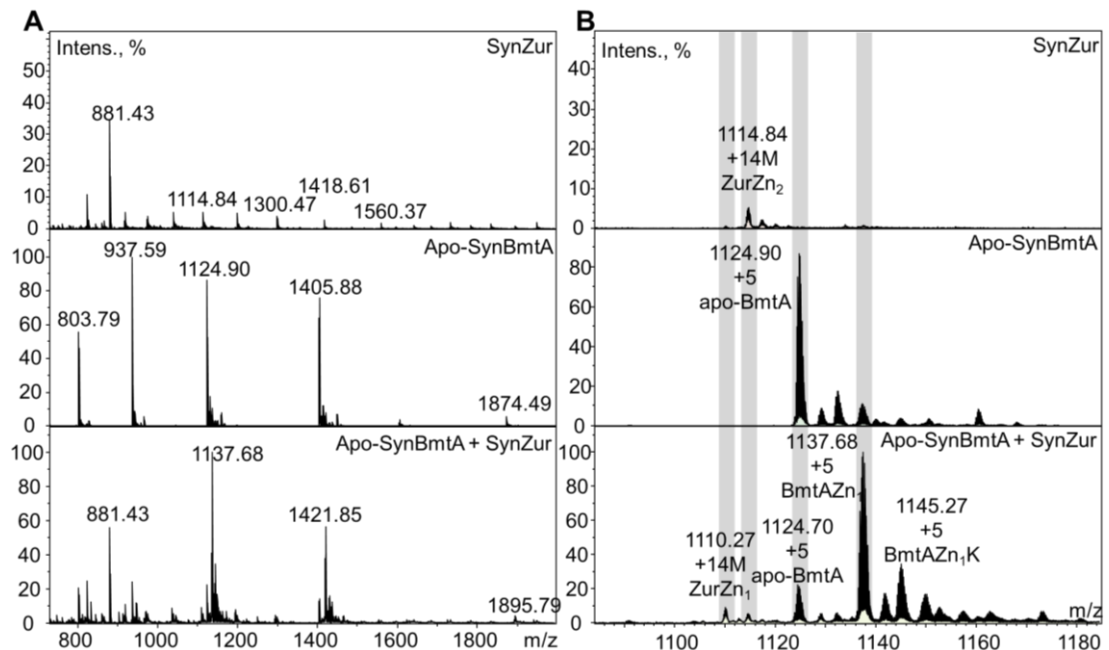


Figure 5.19 ESI-MS spectra of native SynZur, apo-SynBmtA and their mixture.

The spectrum was recorded on a Bruker Compact with 20 μ M SynZur and \sim 5 μ M apo-SynBmtA in 10 mM NH_4HCO_3 . A. full range spectra. The peak with $m/z = 881.43$ represents a background impurity. B. +5 charge state of SynBmtA and +14 for SynZur (monomer). The absolute intensity of SynBmtA was significantly higher than that of SynZur, likely due to different ionisation efficiencies of the proteins and/or due to the distribution of SynZur charge-state envelopes between higher numbers of peaks. A minor peak of SynZur₂ is still present after the reaction as well as peaks for SynBmtAZn₂ and SynBmtAZn₃.

As was expected (Figure 5.19), the transfer of zinc occurs from SynZurZn₂ to apo-SynBmtA forming a metallated species of SynBmtAZn₁ and de-metallated SynZurZn₁. It is unclear whether the SynBmtAZn₁ species was formed due to a potentially incorrect ratio of the proteins or due to K_D for the second zinc binding site in SynBmtA being lower than for SynZurZn₂. There might be kinetic issues as well. However, after leaving the mixture for 20 minutes and checking the spectrum again, the distribution of the peaks did not change. Moreover, leaving the mixtures longer without a reducing agent could lead to the oxidation of apo-proteins since they both have unprotected thiol groups.

The detected transfer of Zn²⁺ from metallated SynZur to apo-SynBmtA might represent an intracellular process of zinc regulation: in the presence of excess zinc SynZurZn₂ activates transcription of *synbmtA* which is then translated to apo-SynBmtA. Apo-SynBmtA binds excess zinc from the cytosol and removes zinc from SynZurZn₂. De-metallated SynZurZn₁ then loses its ability to activate transcription of *synbmtA*.

5.3.3.7 Towards a three-dimensional structure of SynBmtA

In order to structurally characterise SynBmtA, native MT with zinc bound was analysed by NMR spectroscopy. Upon successful obtaining of 1D spectra (not shown), ¹⁵N-labelled protein was analysed by 3D (¹H, ¹⁵N, ¹H) NMR spectroscopy as described in section 5.2.3.6. The assigned ¹H-¹⁵N heteronuclear single quantum coherence spectrum (¹H-¹⁵N HSQC) is shown in Figure 5.20.

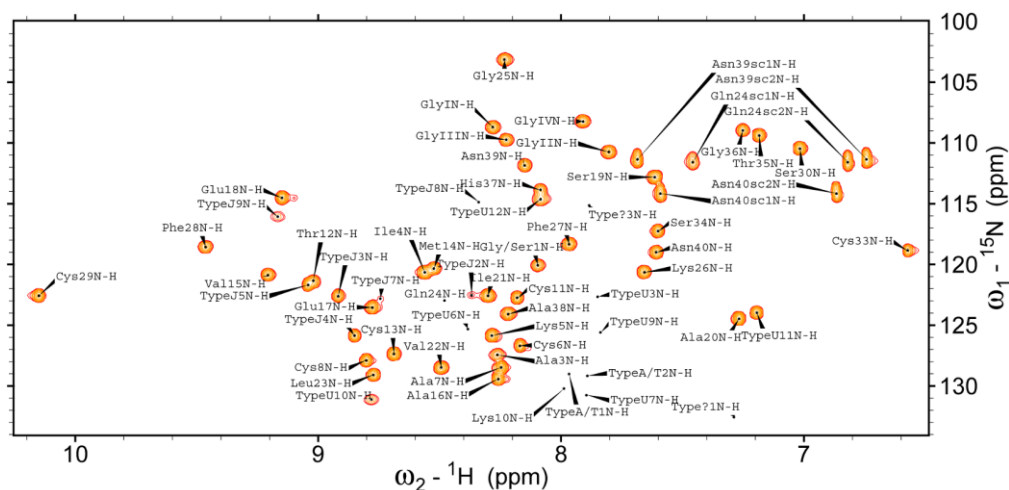


Figure 5.20 ^1H - ^{15}N HSQC spectrum of SynBmtAZn₄.

Peaks with weaker intensities giving diagonal signals on NOESY and TOCSY spectra are also labelled.

Each cross-peak on the ^1H - ^{15}N HSQC spectrum represents a separate backbone N-H group from each amino-acid residue, plus Gln and Asn side chain NH_2 groups. Peaks that are well spread on the ^1H axis and well-resolved indicate that the protein is well-folded. This is typical for zinc-loaded bacterial metallothioneins^{82,153}.

The characteristic low-field chemical shift for the H_N proton of the Cys29 residue previously observed by Blindauer *et al.* (2002) for bacterial MTs was used as a starting point for the assignment⁸². To identify this residue, SynBmtA was sequentially aligned with the resolved structure of SmtA (a MT from *Synechococcus elongatus* strain PCC 7942)⁸¹. The alignment is shown in Figure 5.21.

```
SynBmtA --STAIKCACPKCTCMVAEESAIVLQGGKFFCSTSCSTGHANNEPC-H-GEGSCGCKCGE 55
SmtA     MTSTTLVKCACEPCLCNVDPSKAIDRNGLYYCSEACADGHTGGSGKCGHT---GCNCHG 56
```

Figure 5.21 Sequential alignment of SynBmtA with SmtA from *Synechococcus elongatus* strain PCC 7942.

Proteins were manually aligned by Cys (yellow) and His (cyan) residues. Characteristic Cys residues shifted low-field on NMR spectra are highlighted in red: Cys29 in SynBmtA (10.16 ppm), Cys32 in SmtA (10.09 ppm). Other key residues (Ser34 in SynBmtA and Ala37 in SmtA) are highlighted in brown (see explanation below).

This extreme shift of Cys32 occurs due to the formation of a hydrogen bond between its H_N proton and the sulfur of Cys9 in SmtA⁸². The strips of the corresponding residues from SynBmtA based on the sequential alignment, namely Cys29 and Cys6, are shown in Figure 5.22A. The presence of inter-residue NOESY

cross-peaks between the backbone H_N and the CH_β protons of the side chains indicates the proximity of Cys29 and Cys6 also in SynBmtA.

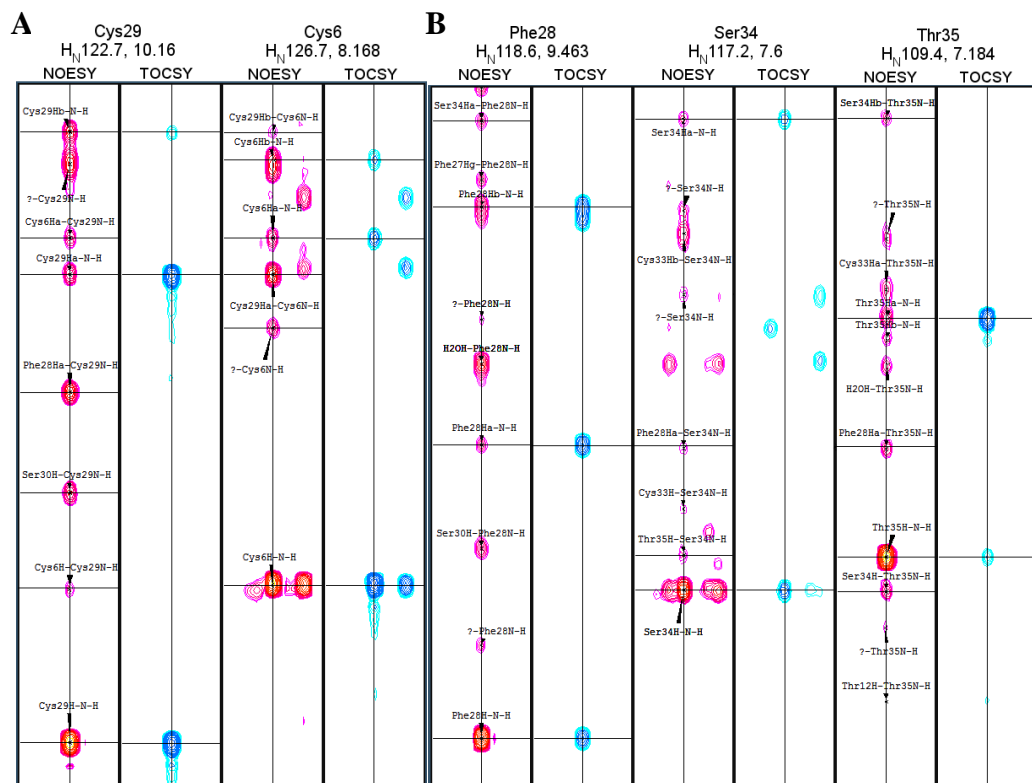


Figure 5.22 Strips of 3D NMR spectra corresponding to selected characteristic residues from SynBmtA.

A. Cys29 and Cys6; B. Phe28, Ser34 and Thr35. NOESY peaks are shown with red-yellow colours and TOCSY peaks are blue-cyan.

Another key feature of the SmtA-like MTs is the interaction of the aromatic ring of Tyr31 with the CH_α of Ala37⁸². This results in a very high-field chemical shift of CH_α of Ala37 (1.73 ppm). Interestingly, based on the sequence alignment (see Figure 5.21), the corresponding residues in SynBmtA are aromatic Phe28 instead of the Tyr31 and Ser34 instead of Ala37. It is unclear whether this “exchange” of a hydroxo-group between the residues plays any role. However, previously it was shown that a similar interaction is observed for the MTs from *Synechococcus* sp. CC9311 encoded by *sync_2426* and *sync_0853*¹⁵³. Both contain Phe residues instead of Tyr31, but Ala37 is conserved in both. Thus, no hydroxo-group is involved in the formation of this interaction in these BmtAs. The strips of Phe28 and Ser34 from SynBmtA are shown in Figure 5.22B.

An example of sequential assignment of two residues is also shown in Figure 5.22B. The NOESY strip of Ser34 contains a signal from H_N of Thr35 apart from intra-residual cross-peaks to its own hydrogen atoms. In turn, the NOESY strip of Thr35 contains signals from H_N and H_α of Ser34. Chemical shift values for the residues of preliminary partially assigned SynBmtA are given in Table 5.10.

Table 5.10 Preliminary 3D [¹H, ¹⁵N, ¹H] TOCSY-HSQC and NOESY-HSQC resonance assignments of SynBmtAZn₄ at 298 K and pH 7.9 recorded at 700 MHz.

| Residue | N | H _N | H _α | H _β | H _γ , H _δ , H _ε |
|---------|-------|----------------|----------------|----------------|--|
| Ala3 | 127.5 | 8.264 | 4.408 | 1.204 | |
| Ile4 | 120.7 | 8.562 | 4.253 | 1.645 | 0.887 |
| Lys5 | 125.9 | 8.284 | 4.35 | 1.581 | |
| Cys6 | 126.6 | 8.168 | 3.854 | 2.878, 2.538 | |
| Ala7 | 128.5 | 8.244 | 4.034 | 0.792 | |
| Cys8 | 128.0 | 8.803 | 4.882 | 3.997 | |
| Lys10 | 130.3 | 7.986 | 3.995 | 2.229 | 1.805 |
| Cys11 | 122.8 | 8.18 | 4.231 | 3.373 | |
| Thr12 | 121.3 | 9.022 | 4.48 | | -0.4996 |
| Cys13 | 127.4 | 8.691 | 4.19 | 3.592, 2.984 | |
| Met14 | 120.2 | 8.527 | 4.973 | 1.83 | 2.512 |
| Val15 | 120.8 | 9.207 | 4.252 | 1.792 | 0.7353 |
| Ala16 | 129.4 | 8.257 | 4.531 | 1.37 | |
| Glu17 | 123.5 | 8.787 | 3.44 | 1.864 | |
| Glu18 | 114.5 | 9.148 | 4.164 | 2.005 | 2.268 |
| Ser19 | 112.8 | 7.61 | 4.726 | 3.907 | |
| Ala20 | 124.5 | 7.269 | 4.184 | 1.278 | |
| Ile21 | 122.6 | 8.296 | 4.185 | 1.7 | 1.285, 0.726 |
| Val22 | 128.5 | 8.493 | 4.952 | 1.755 | 0.4871 |
| Leu23 | 129.1 | 8.775 | 4.655 | 1.743 | 0.796, 1.386 |
| Gln24 | 122.9 | 8.481 | 4.236 | 1.911 | 2.227 |
| Gly25 | 103.2 | 8.241 | 3.935, 3.349 | | |
| Lys26 | 120.7 | 7.653 | 4.256 | 1.797 | 0.975, 1.345, 2.997 |
| Phe27 | 118.4 | 7.965 | 4.781 | 2.886, 2.442 | |
| Phe28 | 118.6 | 9.455 | 5.783 | 2.779 | |
| Cys29 | 122.6 | 10.16 | 4.297 | 2.52 | |
| Ser30 | 110.4 | 7.019 | 4.573 | 3.647 | |
| Cys33 | 118.7 | 6.575 | 3.817 | 3.239, 2.844 | |
| Ser34 | 117.3 | 7.602 | 1.682 | | |
| Thr35 | 109.3 | 7.18 | 4.188 | 4.456 | -0.594 |
| Gly36 | 109.0 | 7.252 | 3.956 | | |
| His37 | 113.9 | 8.086 | 3.835 | 3.422, 2.988 | |
| Ala38 | 124.0 | 8.219 | 4.034 | 1.355 | |
| Asn39 | 111.8 | 8.151 | 4.346 | 3.007, 2.734 | |
| Asn40 | 119.1 | 7.606 | 4.304 | 2.924, 2.492 | |

Additional side chain signals of the Gln and Asn residues as well as data for non-assigned residues are given in Table 8.6 in the Appendix. Further work is required to fully assign SynBmtA.

Confirmation of some of the assigned peaks could be made based on the existing structure of SmtA. Hydrogen bonds responsible for the tertiary structure of the protein are formed between the residues located close in space but distant from each other in the sequence. Sometimes those proximities could be detected by NOESY if the H_N of one of the residues is involved in the formation of a hydrogen bond. Its orientation towards the other residue allows this interaction to be detectable by NOESY NMR. Analysis of the H-bonds in the SmtA structure with Swiss PDB viewer revealed several pairs of such distant residues. Some of those pairs from SmtA are shown in Table 5.11 together with the NMR strips of the corresponding residues from SynBmtA based on sequential alignment (see Figure 5.21).

Table 5.11 Correlational peaks from SynBmtA confirming formation of hydrogen bonds based on the SmtA structure.

NOESY peaks are shown with red-yellow colours and TOCSY peaks are blue-cyan. Structural visualisation of SmtA (1jjd) was performed using the Swiss PDB viewer.

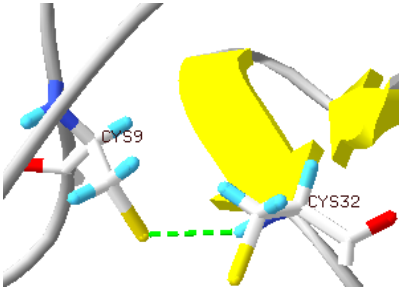
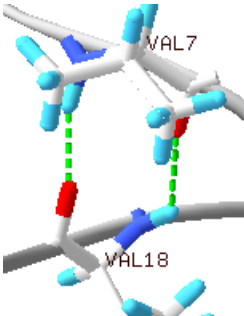
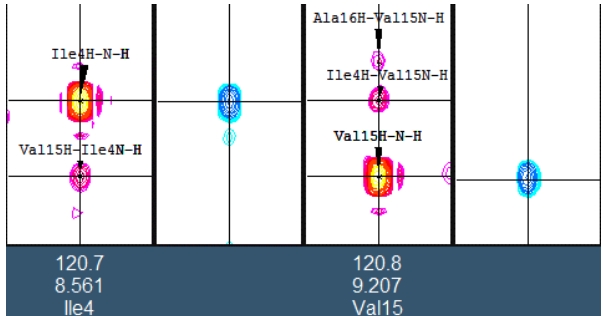
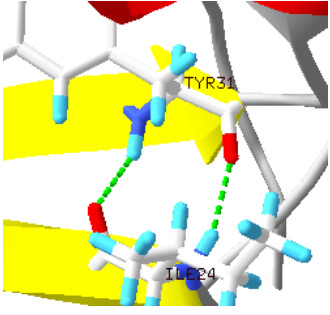
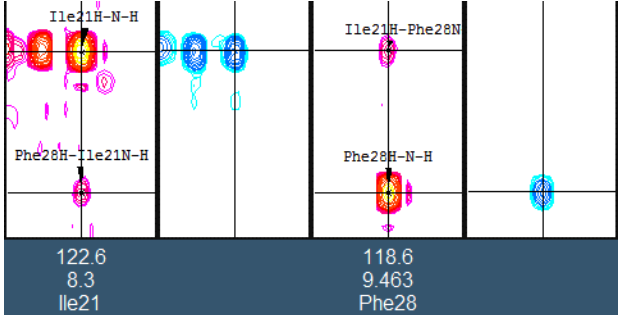
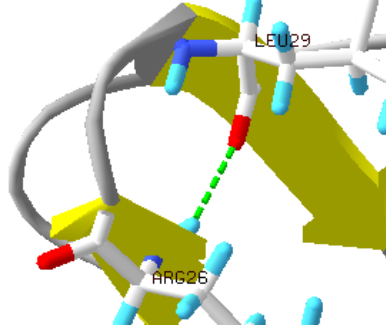
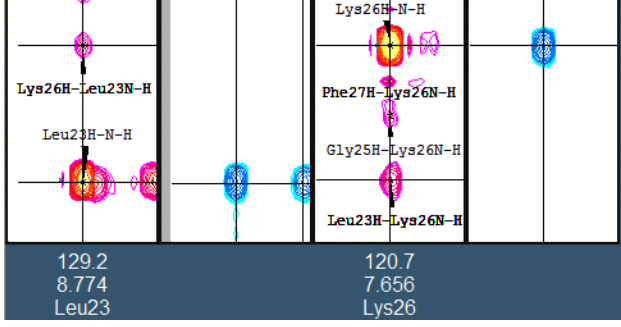
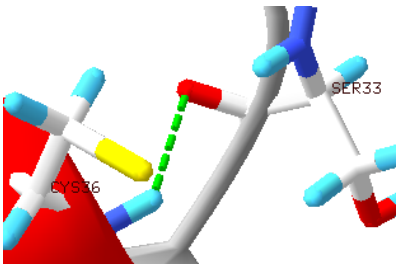
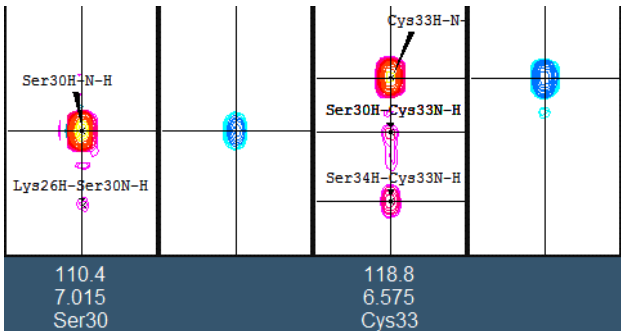
| H-bonds in SmtA | Paired NMR strips for the corresponding residues from SynBmtA |
|--|---|
| <p>Cys9-Cys32</p>  | <p>See Figure 5.21, A for Cys6 - Cys29 interaction.</p> |
| <p>Val7-Val18</p>  | <p>Ile4-Val15</p>  |

Table 5.11. Continuation.

| H-bonds in SmtA | Paired NMR strips for the corresponding residues from SynBmtA |
|---|--|
| <p>Ile24- Tyr31</p>  | <p>Ile21-Phe28</p>  <p>122.6 8.3 Ile21</p> <p>118.6 9.463 Phe28</p> |
| <p>Arg26-Leu29</p>  | <p>Leu23-Lys26</p>  <p>129.2 8.774 Leu23</p> <p>120.7 7.656 Lys26</p> |
| <p>Cys33-Cys36</p>  | <p>Ser30-Cys33</p>  <p>110.4 7.015 Ser30</p> <p>118.8 6.575 Cys33</p> |

Since the spectrum was not fully assigned and no additional experiments were performed due to time limitation, it was impossible to fully resolve the structure of SynBmtA. However, based on the values of the chemical shifts of H α protons, it is possible to predict the secondary structure of the protein. Those values are expected to be shifted up-field in α -helices and downfield in β -sheets compared to the values for random coils²⁰². Values for random coil chemical shifts were taken from Poulsen IDP/IUP random coil chemical shifts webpage²⁰³.

Positive values of the H_{α} proton chemical shift differences with random coil shifts (Chemical Shift Indices, CSI or $\Delta\delta$) were set to +1 if >0.1 ppm, and negative values were set to -1 in $\Delta\delta < -0.1$ ppm. The rest were set to 0. Stretches of three or more positive values of CSI in a row were assigned as β -strands. Stretches of four or more negative values of CSI were assigned as α -helices, respectively²⁰². The resulting plot for predicted secondary structure is shown in Figure 5.23.

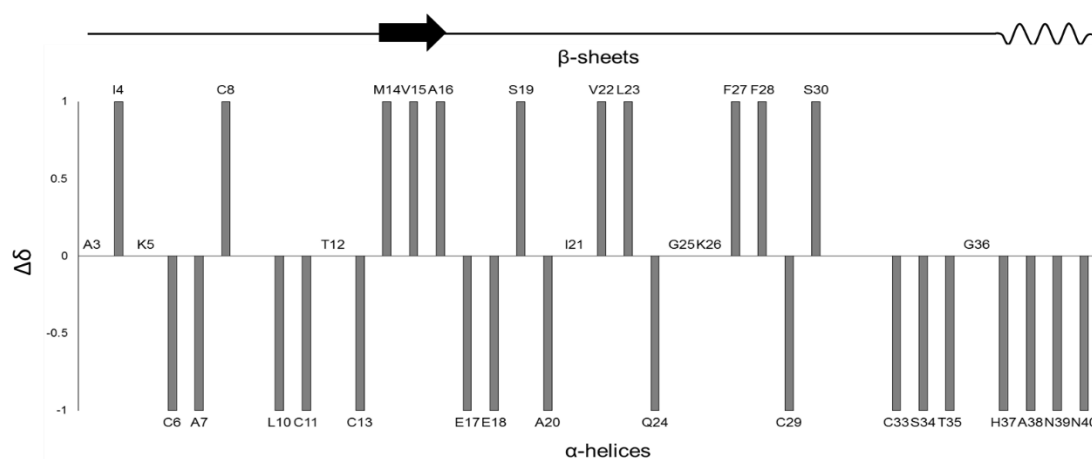


Figure 5.23 Secondary structure prediction for SynBmtA based on the CSIs of H_{α} .

Random coils are shown with a straight line above the CSI plot, α -helices are shown with a wavy line and β -strands are shown with a black arrow.

Comparison of the CSI based structure with JPred4¹³⁰ and Phyre2¹⁷⁶ predicted secondary structures and the SmtA structure revealed dissimilarities except for one β -strand in the region of M14-A16 (see Figure 5.24).

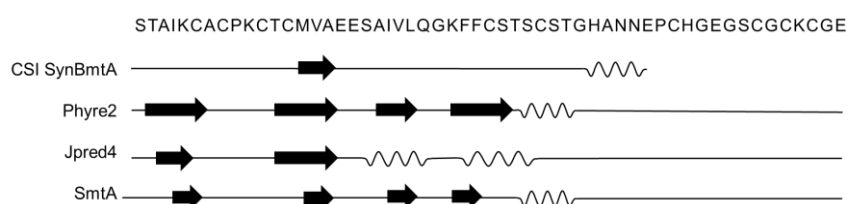


Figure 5.24 Prediction of the secondary structure based on CSI and its comparison with Phyre2¹⁷⁶ and Jpred4¹³⁰ predictions and with the known SmtA structure (1jjd).

Random coils are shown with a straight line, α -helices are shown with wavy lines and β -strands are shown with black arrows. The secondary structure of SmtA was visualised based on the automatic assignment provided by the Swiss PDB viewer software.

It is known however that the CSI prediction method is not completely accurate²⁰². Therefore, CSI (H_{α}) values were also directly compared with those of SmtA. This comparison is shown in Figure 5.25.

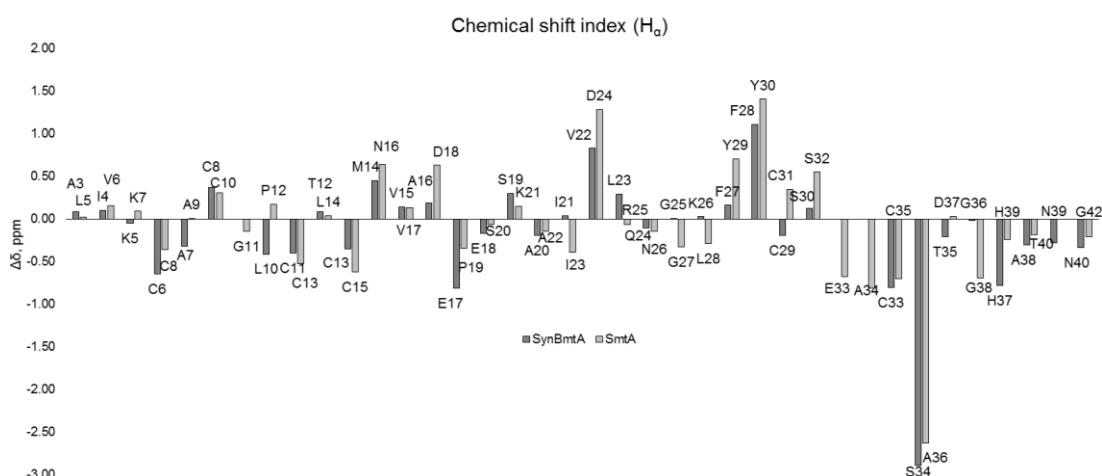


Figure 5.25 Comparison of CSI values for H_{α} of SynBmtA and SmtA.

A somewhat similar approach was used previously^{82,153}. The values of H_N and H_{α} chemical shifts were similar for the majority of the sequentially corresponding residues for several SmtA-like metallothioneins.

This comparison (see Figure 5.25) revealed that for most of the assigned residues, the trends in CSI values are similar to those of the corresponding residues of SmtA. Therefore, the structures of both MTs are expected to be similar for the first 40 residues (zinc-finger fold) as well which is also confirmed by the interactions shown in Table 5.11. The model of SynBmtA as predicted with the Phyre2¹⁷⁶ server based on the structures of other SmtA-like MTs is shown in Figure 5.26.

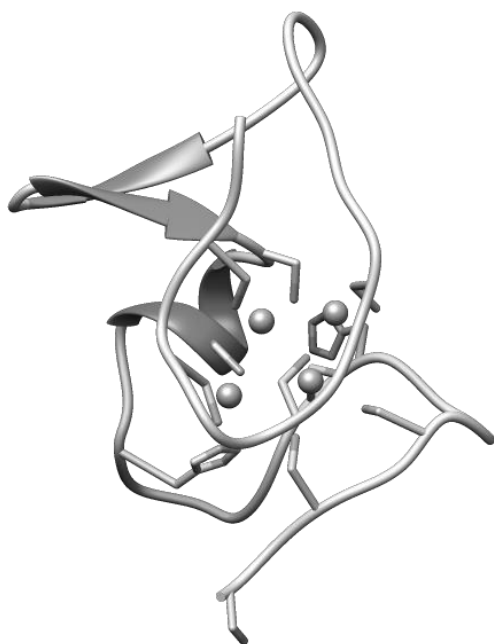


Figure 5.26 Structure of SynBmtA predicted by the Phyre2¹⁷⁶ server.

The location of zinc ions is taken from SmtA after manual structural alignment with SynBmtA.

At this moment, the NMR spectrum has not been assigned fully and the whole C-terminus (Glu41-Glu55) has not been resolved. It is likely that after full assignment the C-terminus of SynBmtA would be distinct from that of SmtA. Based on the sequence alignment (see Figure 5.21), SynBmtA contains one additional Cys residue in the C-terminus compared to SmtA. Therefore, it is unclear whether their zinc-binding residues are equivalent, especially in that region. Also, SmtA contains an extra (non-coordinating) His-residue at the C-terminus. Thus, an additional ¹¹¹Cd NMR analysis is required to assess if zinc-binding sites in SynBmtA are conserved compared to SmtA and whether the suggested alignment is correct.

If the Cys53 residue is not involved in zinc binding, the Zn₄ cluster of SynBmtA would be predicted to have the configuration given in Figure 5.27.

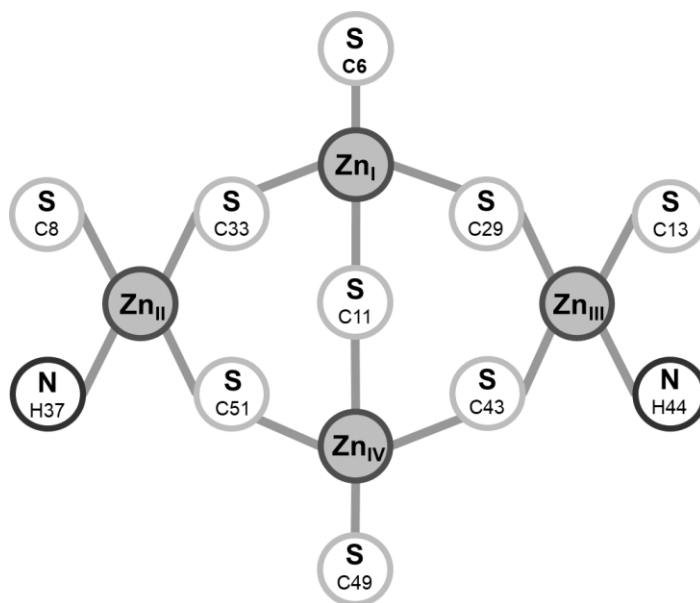


Figure 5.27 Predicted Zn₄ cluster in SynBmtA based on the model of SmtA and sequential alignment shown in Figure 5.21.

5.4 CONCLUSIONS AND FUTURE WORK

In this chapter I described how a Zur-activated metallothionein from *Synechococcus* sp. WH8102 was recombinantly over-expressed in *E. coli*, purified and characterised. In addition, SynZur-binding properties of *pbmtA* were assessed. Based on the results of this chapter, further conclusions can be drawn about SynBmtA:

1. SynZur binds to *pbmtA* in a 4:1 stoichiometry, with an affinity of 26.7 nM per monomer.
2. SynBmtA was over-expressed and purified with 4 zinc ions bound which was confirmed by ESI-MS and ICP-OES.
3. A method to generate apo-SynBmtA was developed based on acidification of the sample rather than using chelating or denaturing agents such as EDTA, TPEN or urea.
4. The SynBmtA zinc-binding affinity is extremely high with a $K_D \leq 1.13 \pm 0.07 \times 10^{-14}$ M. This is in agreement with the pH value of half displacement of zinc.
5. Apo-SynBmtA is able to remove Zn^{2+} from metallated SynZurZn₂. Under the conditions tested, SynZurZn₁ and SynBmtAZn₁ were formed. This transfer is in agreement with the calculated K_D values of SynZur and SynBmtA for Zn^{2+} . It is as yet unknown whether this transfer involves direct protein-protein interactions or requires dissociation of Zn^{2+} from SynZurZn₂.
6. NMR confirmed a well-defined fold for zinc-loaded SynBmtA. ¹⁵N-NOESY-HSQC and ¹⁵N-TOCSY-HSQC spectra were partially assigned. The current assignment is in agreement with the structure resembling closely that of SmtA from *Synechococcus elongatus* strain PCC 7942 except for the C-terminus of SynBmtA which was un-assigned and differs sequentially.

Future work:

1. Calculation of an appropriate molar absorption coefficient for apo-SynBmtA using ICP-OES.
2. Clarification of the species formed during titration of apo-SynBmtA with Zn^{2+} in the presence of Quin-2. This would allow determination of the stoichiometry of the complexes formed at various stages of the titration.

3. Since SynBmtAZn₁ was formed after the reaction of apo-SynBmtA with SynZurZn₂, it is possible that the remaining zinc ions bind SynBmtA with weaker affinity than the first one. Thus, it would be interesting to clarify if this occurred due to the low ratio of SynZurZn₂ to apo-SynBmtA or was a consequence of different affinity values of zinc-binding sites in apo-SynBmtA. Alternatively, it could be a consequence of slow binding kinetics.
4. Finishing the NMR assignment.
5. ¹¹¹Cd NMR spectroscopy including ¹H, ¹¹¹Cd correlation spectroscopy of ¹¹¹Cd-substituted SynBmtA to determine connectivities between metal ions and metal-binding cysteine and histidine residues. In addition, site directed mutagenesis experiments could also be utilised to clarify the participation of individual residues.
6. 2D (¹H, ¹H) full range homonuclear NOESY and TOCSY NMR spectroscopy to assign Pro residues which are invisible on ¹⁵N NMR. In addition, this would show aliphatic regions of the SynBmtA residues, could help with the assignment of unassigned residues and with assignment confirmation of the assigned residues and could provide important inter-residue distance restraints for structure determination.

Chapter 6.
General conclusions

6.1 INTRODUCTION

Zinc is typically the second most abundant trace metal after iron within cells and is predicted to operate as a co-factor for between 5-9% of proteomes, depending on kingdom²⁶. Zinc may be required for CO₂ fixation by cyanobacteria through zinc-requiring carbonic anhydrases²⁷. A previously formulated “zinc hypothesis” postulated that global atmospheric CO₂ levels are inversely correlated to oceanic zinc levels so that low available zinc in the ocean might limit the global CO₂ fixation rate²⁹. However, while a negative effect of zinc limitation on cell growth has been observed for some oceanic microalgae (coccolithophores²⁰⁴ and diatoms²⁰⁵), an absolute requirement for zinc is not known for the marine photosynthetic prokaryotes *Prochlorococcus* and *Synechococcus*⁴. Moreover, of the oceanic locations assessed, none have identified zinc as a limiting nutrient for marine phytoplankton growth²⁰⁶.

Despite these findings, marine cyanobacteria have developed a remarkable ability to accumulate zinc³². Surprisingly, cellular zinc quotas for some epipelagic prokaryotic picophytoplankton including *Synechococcus* sp. were shown to be at a similar level to iron, or sometimes higher, depending on the sample location and nutrient concentration^{18,207,208}. In addition, genes putatively encoding components of high affinity zinc uptake systems were identified in most marine cyanobacterial genomes³⁷ suggesting a necessity for zinc. Thus, it is possible that previous *in situ* studies⁷ did not cause sufficient zinc limitation to affect cyanobacterial growth. Indeed, zinc-limiting conditions are extremely hard to achieve even for model bacteria like *E. coli*. Graham *et al.*, (2009) showed only slight growth limitation for *E. coli* MG1655 under “severe zinc depletion”²⁰⁹ despite Chelex treatment of the growth medium and careful precautions to prevent zinc contamination (achieving a final zinc concentration in the medium of “only” ~60 nM). In this thesis, possibly due to the different nature of the seawater medium used, zinc-depleted medium was estimated to contain only ~17 nM zinc (see section 4.3.2.2). However, this value could be still almost two orders of magnitude higher than total zinc levels in oceanic waters e.g. in the Central North Pacific where total zinc concentration is ~ 0.5 nM in the euphotic zone⁵⁶. Perhaps the high macronutrient (nitrogen and phosphorus) concentrations in ASW medium, which support significant accumulation of the cellular biomass compared to natural seawater, are sufficient to cause a transcriptional response to zinc scarcity in culture. Alternatively, the use of chelators such as EDTA that decrease

“available” free Zn^{2+} can be used, though these chelators may not be applicable to some species due to their potential ability to remove zinc from such EDTA complexes¹⁵⁴ (see also section 4.3.2.1). Thus, it is possible that the precise conditions for achieving zinc limitation have not yet been generated, leading to erroneous conclusions like the fact that marine picocyanobacteria have no absolute requirement for zinc. Constructing a mutant with a disrupted zinc uptake machinery could potentially clarify this question. Indeed, disrupting *znuA* in *E. coli* almost completely inhibited growth under zinc-depleted conditions²⁰⁹.

In this thesis I sought to shed light on zinc regulation and uptake in the model open ocean cyanobacterium *Synechococcus* sp. WH8102. This strain belongs to clade III from sub-cluster 5.1, a clade largely restricted to oligotrophic environments particularly low in phosphorus. Bioinformatics predicted various zinc regulation and uptake components such as the transcriptional repressor Zur (zinc uptake regulator), at least one high affinity inner membrane zinc uptake transporter (ZnuABC), and a zinc metallothionein BmtA³⁸. Other potential components of the zinc homeostasis machinery include a zinc chaperone (a COG0523 family member)³³, an outer membrane porin and a second ZnuABC system³².

EMSA analysis previously showed that the *Synechococcus* sp. WH8102 Zur protein binds to zur-boxes predicted by the RegPrecise database in the operator promoter regions of *znuABC* (SYNW0969-SYNW0971) and *bmtA*⁸⁸. Ksibe (2016) also suggested that the mode of action of SynZur would be different from Zur sensors from other bacterial phyla: removal of zinc from the sensory site facilitated formation of monomers from a Zur dimer instead of (or in addition to) the formation of an open non-DNA binding configuration.

The current study continued this work on zinc homeostasis in *Synechococcus* sp. WH8102 since several questions had remained un-answered:

- Is the predicted SynZur a true zinc uptake regulator? What is the affinity of SynZur for zinc? Indeed, no experimentally determined structure of any cyanobacterial Zur protein was known while cyanobacterial Zurs are expected to be significantly different to published structures from other phyla³³.
- What genes comprise the Zur regulon in *Synechococcus* sp. WH8102? How does zinc cross the cyanobacterial outer membrane?

- What is the function of the metallothionein SynBmtA? Is SynBmtA repressed by the “deficiency sensor” SynZur? Alternatively, it was hypothesised that SynZur might work as a transcriptional activator of *synbmtA*. However, no experimental confirmation of this idea was obtained⁸⁸.

6.2 KEY FINDINGS

6.2.1 Zinc Uptake Regulator from *Synechococcus* sp. WH8102

SynZur is a representative of the Fur-family of sensor proteins. In Chapter 3, I showed that SynZur containing zinc in the sensory site binds to *pznuABC* selectively despite an ability to bind other divalent metals. Chapter 4 revealed gene expression profiles of the *Synechococcus* sp. WH8102 wild type and *zur* knockout mutant across a range of zinc concentrations, with growth of the mutant being more sensitive to higher zinc concentrations compared to WT, and cellular zinc quotas observed to be most pronouncedly different in mutant and WT. Together these data proved SynZur as a true zinc uptake regulator.

The structure of SynZur was solved by X-ray crystallography (Chapter 3), providing the first representative of the cyanobacterial Fur-family of proteins to be structurally characterised. Moreover, it was shown that SynZur represents a novel class of bacterial Zur proteins with a unique sensory site and, likely, mode of action. The mode of action of other Zur proteins is based on changing the mutual orientation of the dimerization and DNA-binding domains upon binding zinc, forming a “closed” conformation of the dimer, which is able to bind to DNA. This conformational change takes place probably due to binding zinc by a sensory site located in the inter-domain region. Although different, both *E. coli* Zur and *S. coelicolor* Zur have such an inter-domain sensory site. However, in the case of SynZur and, by inference, homologous cyanobacterial Zur proteins, the sensory site is located in the dimerization domain. Together with the ESI-MS experiments conducted previously and in the current work, this is consistent with suggestions that upon removal of zinc from the sensory site the SynZur dimer is more likely to dissociate into monomers which are unable to bind to zur-boxes with relevant affinity⁸⁸. However, additional experiments are required to provide conclusive evidence for this mechanism (see section 3.4).

EMSA and Ferguson plot analyses revealed that SynZur binds with a 2:1 ratio to *pznuABC* (see Chapter 3) and 4:1 to *pbmtA* (see Chapter 5) with affinity values K_D *pznuABC* ~10 nM and K_D *pbmtA* ~27 nM per SynZur monomer. This is in agreement with the number of identified zur-boxes: one Zur dimer binds one zur-box. This is distinct from *E. coli* Zur where two dimers were bound per box⁴⁸. It is as yet unclear whether the presence of two zur-boxes instead of one in *pbmtA* is relevant for its activation. Location of the second zur-box upstream of the promoter element possibly serves

some function other than blocking and could be related to the activation mechanism (see Chapter 4). However, more detailed analyses and experimental confirmation are required.

Previously, it was suggested that the SynBmtA gene (SYNW0359) might be expressed by cells in response to zinc limitation³². However, in Chapter 4 I showed for the first time, by RNA-seq and RT-qPCR, that expression of *synbmtA* was activated in the presence of excess zinc in WT *Synechococcus* sp. WH8102, and repressed in the *zur*-mutant. I also showed that metallated SynZur binds to both *pznuABC* and *pbmtA*. These data show that SynZur, with zinc bound in the sensory site, apparently plays a dual role i) working as a transcriptional repressor of *znuABC*, and ii) as a transcriptional activator of *bmtA*. This is the first time that a cyanobacterial Zur has been shown to be a transcriptional activator, a feature that could be an adaptation mechanism to oligotrophic environments.

Given the observed SynZur-DNA binding affinities to both the *pznuABC* and *pbmtA* promoters and the proposed dual mode of action of SynZur allows to propose three different scenarios for zinc regulation in *Synechococcus* sp. WH8102 depending on the zinc concentration:

- A. At high zinc levels:** zinc might permeate into cells passively through non-specific cation uptake channels or perhaps through the constitutively expressed “second” ZnuABC system⁸⁸. Zinc concentrations would be high enough to produce a fully metallated SynZur which would assemble as a dimer $(\text{SynZurZn}_2)_2$ and bind to both promoters activating transcription of *bmtA* and repressing transcription of *znuABC* (see Figure 6.1, high zinc).
- B. When zinc concentrations are intermediate (sufficient/replete):** only a fraction of SynZur molecules will be metallated. Since the affinity of SynZur for *pznuABC* is higher, the dimer formed would preferentially bind to *pznuABC*, leaving *pbmtA* un-occupied. Transcription of *znuABC* would be repressed whilst *bmtA* would not be activated leading to the absence/low levels of both ZnuABC and BmtA since no uptake or sequestration is required (see Figure 6.1, intermediate zinc). Presumably, the amount of zinc coming in through the “second” ZnuABC system/passively is enough.
- C. Under low zinc levels:** SynZur is de-metallated at the sensory site, no dimers are formed and both operator-promoter regions are un-occupied. This leads to de-

repression of *znuABC* transcription while transcription of *bmtA* is not activated. Thus, ZnuABC is translated facilitating zinc uptake at these low zinc concentrations, while BmtA is absent or at a very low level (see Figure 6.1, low zinc).

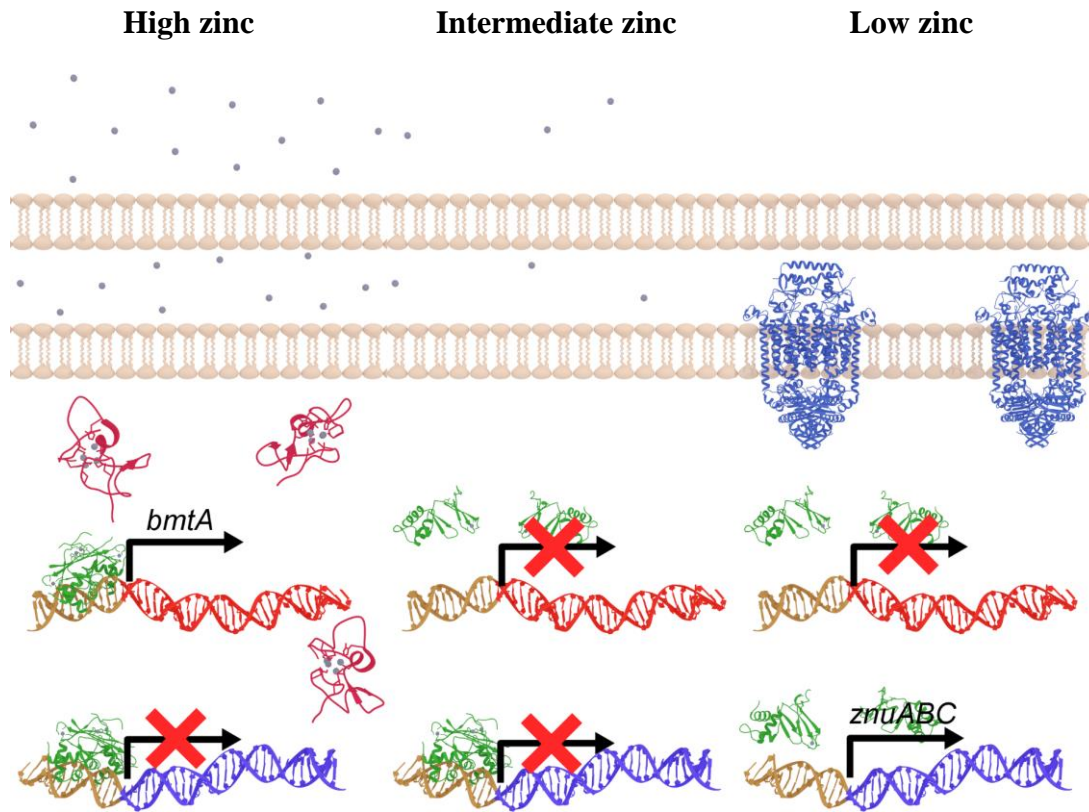


Figure 6.1 Proposed mode of zinc regulation in *Synechococcus* sp. WH8102.

6.2.2 *Synechococcus* sp. WH8102 metallothionein

WT *Synechococcus* sp. WH8102 expresses *synbmtA* as a response to zinc excess/toxicity and its expression is activated by metallated SynZur (see chapters 4 and 5). This suggests that SynBmtA is produced by cells to either protect them from zinc excess or as a zinc storage molecule, both being in agreement with roles of classical MTs.

It is somewhat puzzling why an open-ocean cyanobacterium occupying an extremely oligotrophic environment would possess a mechanism for zinc detoxification. Interestingly, *Synechococcus* sp. WH8102 also shows some tolerance to copper which was at an intermediate level between another oceanic *Synechococcus* (strain CC9605) and a coastal strain (*Synechococcus* sp. CC9311)¹⁸². *Synechococcus* sp. WH8102 also exhibits increased metal quotas for both iron and zinc compared to

other strains of *Synechococcus* cultured in the same medium²⁰⁷. This is also in agreement with the significantly higher zinc quotas observed in *Synechococcus* cells from anticyclonic eddies²⁰⁸. Given that *Synechococcus* sp. WH8102 possesses a unique mechanism for swimming motility⁸⁵, this might allow this organism to move towards higher concentrations of nutrients. In so doing, it is possible that this strain developed machinery to accumulate and store nutrients such as zinc as an adaptation to the overall nutrient poor environment it occupies. However, this has yet to be confirmed.

Initial NMR structural studies confirmed a well-defined fold for SynBmtAZn₄ conserved in other BmtAs e.g. SmtA from *Synechococcus elongatus* PCC 7942 (see chapter 5). Similarities in structure could be an additional confirmation of similarity in function such as zinc sequestering and/or storage. However, additional experiments are required to finalise sequential assignment and work towards a 3D structure for SynBmtA - particularly to understand its remarkable highly stable fold and high zinc affinity.

SynBmtA shows an extremely high affinity for zinc: $K_D \leq 1.13 \pm 0.07 \times 10^{-14}$ M (see chapter 5). It was impossible to de-metallate SynBmtA even using a variety of strong chelating or denaturing agents such as EDTA, TPEN or urea, and their combinations. Similar behaviour was observed for the structural site of SynZur which did not react with EDTA (see Chapter 3) and for the same site in *E. coli* Zur¹¹⁵. Interestingly, the structural Cys₄ site of Zur proteins is similar to the SynBmtA binding sites composed of Cys residues. Previously, it was suggested that the structural sites of Zur proteins were kinetically inert, even though they were not deeply buried³⁵. ESI-MS showed that Cd²⁺ could quickly replace Zn²⁺ in the structural site, confirming accessibility to cations (see Chapter 3); however, electrostatic or steric hindrances towards reactions with much bulkier negatively charged chelators are still conceivable.

Given the presence of specific zinc uptake machinery in this species (ZnuABC), which suggests a necessity for this metal, the very high zinc affinity of MT could be an indication of SynBmtA serving a specific purpose i.e. keeping the cytosolic free [Zn²⁺] very low and/or supplying zinc to only a selection of enzymes with an even higher affinity or to enzymes possessing a mechanism for zinc transfer from SynBmtA to them. SynBmtA potentially possesses the highest affinity for zinc

or the lowest observed $K_{D(\text{Zn})}$ value amongst other MTs: its $\text{pH}_{1/2}$ is the lowest (see chapter 5) and its K_D at pH 7.9 in 20 mM NH_4HCO_3 is lower than that of SmtA at pH 8.1, ionic strength = 4 mM ($\text{p}K_{\text{Dapp}} = 12\text{-}13$)²¹⁰. This could be a direct adaptation of this species to the oligotrophic environment it occupies in order to keep the cytosolic free $[\text{Zn}^{2+}]$ very low compared to other species.

Finally, I propose a model of zinc transfer from SynZur to SynBmtA, which stops *synbmtA* activation. Under conditions of zinc excess *synbmtA* is transcribed in response to high zinc (see Chapter 4), and once all available zinc is bound by SynBmtA, any remaining apo-SynBmtA could remove zinc from metallated SynZurZn₂ forming SynZurZn₁ and metallated SynBmtA based on comparisons of their K_{DS} ($K_{\text{D}(\text{ZurZn}_1)} = 8.27 \times 10^{-13}$ M, see Chapters 3 and 5). This process is shown in Figure 6.2. De-metallated SynZurZn₁ does not have the ability to activate transcription of *synbmtA* and its expression will stop. The possibility of zinc transfer from Zur to BmtA was indeed demonstrated by ESI-MS in Chapter 5, although the mechanism of transfer awaits elucidation.



Figure 6.2 Zinc transfer between SynZur and BmtA.

6.2.3 Discovery of potential new players in zinc homeostasis

Differential gene expression analysis in Chapter 4 revealed several members of the *Synechococcus* sp. WH8102 Zur-regulon both activated (SYNW0359 [*bmtA*], SYNW0360 [phage integrase]) and repressed (SYNW0968 [uncharacterised], SYNW0969 [*znuB*], SYNW0970 [*znuC*], SYNW0971 [*znuA*], SYNW0972 [uncharacterised] and SYNW0973 [uncharacterised]) (see Figure 6.3). Only genes SYNW0359 and SYNW0969, SYNW0970, SYNW0971 were previously predicted to be regulated by SynZur whilst SYNW0360, SYNW0968, SYNW0972 SYNW0973 were newly discovered.

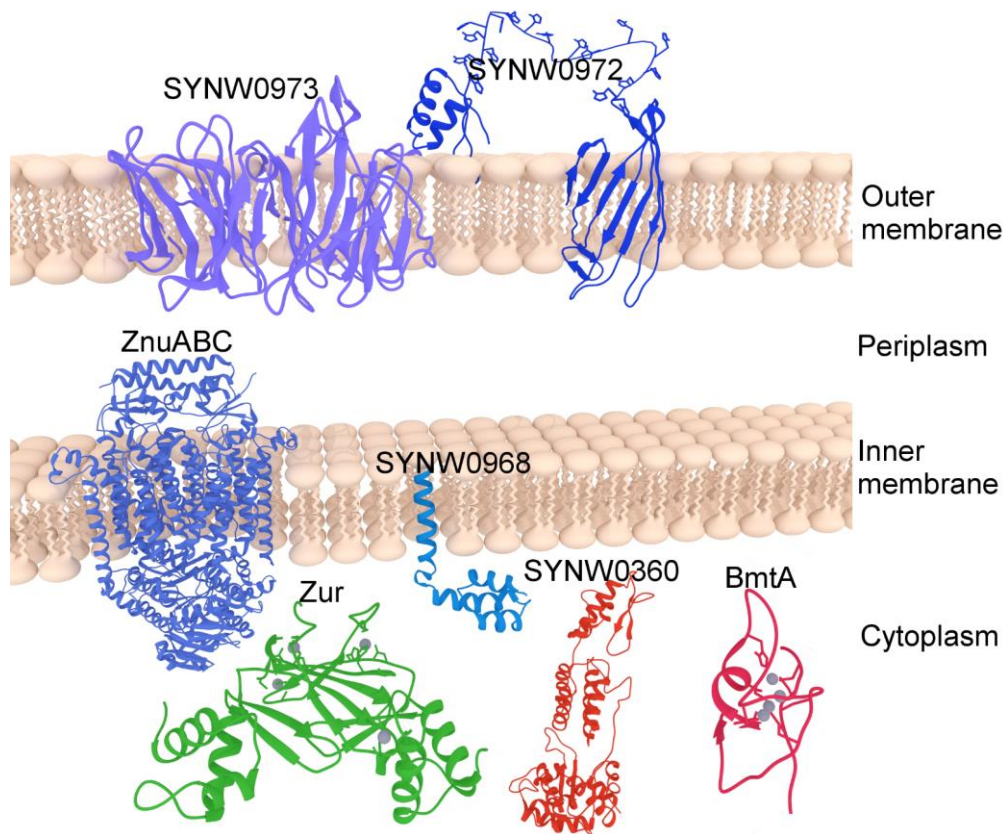


Figure 6.3 Components of the Zur (green) regulon in *Synechococcus* sp. WH8102 identified in Chapter 4 with their suggested locations.

Components repressed by SynZur are shown in shades of blue; components activated by SynZur are shown in shades of red. Zinc is shown with grey balls.

6.2.3.1 New putative outer membrane zinc uptake transporters

Interestingly, two newly identified Zur-regulated genes (see Chapter 4), SYNW0972 and SYNW0973, are well conserved in marine cyanobacteria, and based on structural prediction by the Phyre2 server likely to represent novel members of the outer membrane zinc uptake machinery in marine picocyanobacteria. This discovery calls for new studies exploring the role of these proteins in outer membrane zinc (and perhaps other metals) transport.

6.3 FUTURE WORK

Several questions directly related to the current study have arisen and are discussed in detail in the conclusion sections of each results and discussion chapter. Finalising the work on SynZur includes testing whether it is a zinc specific sensor, clarification of the novel suggested sensing mechanism and understanding SynZur-DNA interactions. The mechanism of activation of *synbmtA* is unclear; an intriguing newly identified His₄ site presumably involved in the formation of the SynZur tetramer could be responsible for transcriptional activation of *synbmtA*. The SynBmtA protein work also has to be completed including calculation of the ϵ_{230} of apo-SynBmtA, determining the stoichiometry of zinc binding during titration of SynBmtA with zinc in the presence of Quin-2, and finishing the NMR assignment and identification of zinc binding residues and their connectivities to metal ions for structural calculations. Finally, understanding/proving the role of members of the Zur regulon in zinc metabolism is also very important: construction of deletion/knockout mutants of *znuABC*, *bmtA* and newly identified members of the Zur regulon, especially those potentially involved in outer membrane zinc uptake (SYNW0972 and SYNW0973) could clarify that.

The apparent absence of a requirement for zinc in marine cyanobacteria needs to be re-considered. As was discussed in section 6.1, it is impossible to create actual zinc-free conditions for cyanobacteria. Potentially using the strongest zinc chelator TPEN might give better results rather than EDTA. However, this would need testing. In addition, using such chelators might cause limitation for other trace metals and a mis-leading cellular response. Possibly, construction of a deletion mutant of *Synechococcus* sp. WH8102 for the whole cluster of genes potentially involved in zinc uptake (SYNW0968-SYNW0973) could clarify zinc requirements in this species. However, the presence of a putative “second” ZnuABC system also has to be considered.

No enzymes were identified in the Zur regulon of *Synechococcus* sp. WH8102. Therefore, zinc subcellular location is still unclear in this species. Furthermore, the previously suggested link between phosphorus^{171,184}/carbon²⁹ metabolism and zinc was also not clarified. Considering the enormous contribution of cyanobacteria to global CO₂ fixation, investigation of whether zinc plays a role in this process is essential given that atmospheric CO₂ concentrations are increasing²¹¹. In order to

address this, CO₂ fixation rates at different zinc levels should be compared with conditions mimicking natural seawater including low concentrations of inorganic phosphate but sufficient concentrations of organic phosphates. Thus, much still remains to be done regarding the zinc biology of these globally important microbes.

7. References

- (1) Sinha, R. P.; Häder, D.-P. Photobiology and ecophysiology of rice field cyanobacteria. *Photochem. Photobiol.* **1996**, *64* (6), 887–896.
- (2) Six, C.; Thomas, J. C.; Garczarek, L.; Ostrowski, M.; Dufresne, A.; Blot, N.; Scanlan, D. J.; Partensky, F. Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. *Genome Biol.* **2007**, *8* (12), 1–22.
- (3) Scanlan, D. J.; Ostrowski, M.; Mazard, S.; Dufresne, A.; Garczarek, L.; Hess, W. R.; Post, A. F.; Hagemann, M.; Paulsen, I.; Partensky, F. Ecological genomics of marine picocyanobacteria. *Microbiol. Mol. Biol. Rev.* **2009**, *73* (2), 249–299.
- (4) Saito, M. A.; Sigman, D. M.; Morel, F. M. M. The bioinorganic chemistry of the ancient ocean: the co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean-Proterozoic boundary? *Inorg. Chim. Acta* **2003**, *356*, 308–318.
- (5) Cavet, J. S.; Borrelly, G. P. M.; Robinson, N. J. Zn, Cu and Co in cyanobacteria: selective control of metal availability. *FEMS Microbiol. Rev.* **2003**, *27* (2–3), 165–181.
- (6) Falcón, L. I.; Magallón, S.; Castillo, A. Dating the cyanobacterial ancestor of the chloroplast. *ISME J.* **2010**, *4* (6), 777–783.
- (7) Kutschera, U.; Niklas, K. J. Endosymbiosis, Cell evolution, and speciation. *Theory Biosci.* **2005**, *124* (1), 1–24.
- (8) Partensky, F.; Hess, W. R.; Vaultot, D. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol. Mol. Biol. Rev.* **1999**, *63* (1), 106–127.
- (9) Flombaum, P.; Gallegos, J. L.; Gordillo, R. a; Rincón, J.; Zabala, L. L.; Jiao, N.; Karl, D.; Li, W.; Lomas, M.; Veneziano, D.; *et al.* Present and future global distributions of the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc. Natl. Acad. Sci. U S A* **2013**, *110* (24), 9824–9829.
- (10) Karl, D.; Michaels, A.; Bergman, B.; Capone, D.; Carpenter, E.; Letelier, R.; Lipschultz, F.; Paerl, H.; Sigman, D.; Stal, L. Dinitrogen fixation in the world's oceans. *Biogeochemistry* **2002**, *57/58*, 47–98.
- (11) Jardillier, L.; Zubkov, M. V.; Pearman, J.; Scanlan, D. J. Significant CO₂ fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. *ISME J.* **2010**, *4* (9), 1180–1192.

- (12) Chisholm, S. W.; Falkowski, P. G.; Cullen, J. J. Dis-crediting ocean fertilization. *Science* **2001**, *294* (June), 309–310.
- (13) Ahlgren, N. A.; Rocap, G. Diversity and distribution of marine *Synechococcus*: multiple gene phylogenies for consensus classification and development of qPCR assays for sensitive measurement of clades in the ocean. *Front. Microbiol.* **2012**, *3* (JUN), 1–24.
- (14) Farrant, G. K.; Doré, H.; Cornejo-Castillo, F. M.; Partensky, F.; Ratin, M.; Ostrowski, M.; Pitt, F. D.; Wincker, P.; Scanlan, D. J.; Iudicone, D.; *et al.* Delineating ecologically significant taxonomic units from global patterns of marine picocyanobacteria. *Proc. Natl. Acad. Sci.* **2016**, *113* (24), E3365–E3374.
- (15) Fuller, N. J.; Marie, D.; Partensky, F.; Vaultot, D.; Post, A. F.; Scanlan, D. J. Clade-specific 16S ribosomal DNA oligonucleotides reveal the predominance of a single marine *Synechococcus* clade throughout a stratified water column in the Red Sea. *Appl. Environ. Microbiol.* **2003**, *69* (5), 2430–2443.
- (16) Xia, X.; Cheung, S.; Endo, H.; Suzuki, K.; Liu, H. Latitudinal and vertical variation of *Synechococcus* assemblage composition along 170° W transect from the south Pacific to the Arctic Ocean. *Microb. Ecol.* **2019**, *77* (2), 333–342.
- (17) Kranzler, C.; Rudolf, M.; Keren, N.; Schleiff, E. Iron in cyanobacteria. In *Advances in Botanical Research*; Chauvat, Franck; Cassier-Chauvat, C., Ed.; Elsevier Ltd., **2013**; *65*, 57–105.
- (18) Twining, B. S.; Baines, S. B. The trace metal composition of marine phytoplankton. *Ann. Rev. Mar. Sci.* **2013**, *5* (1), 191–215.
- (19) Shcolnick, S.; Keren, N. Metal homeostasis in cyanobacteria and chloroplasts. balancing benefits and risks to the photosynthetic apparatus. *Plant Physiol.* **2006**, *141* (3), 805–810.
- (20) Raven, J. A. Predictions of Mn and Fe use efficiencies of phototrophic growth as a function of light availability for growth and of C assimilation pathway. *BY. New Phytol.* **1990**, *116*, 1–18.
- (21) Morrissey, J.; Bowler, C. Iron utilization in marine cyanobacteria and eukaryotic algae. *Front. Microbiol.* **2012**, *3*, 43.
- (22) Nelson, N.; Junge, W. Structure and energy transfer in photosystems of oxygenic photosynthesis. *Annu. Rev. Biochem.* **2015**, *84* (1), 659–683.
- (23) López-Maury, L.; Giner-Lamia, J.; Florencio, F. J. Redox control of copper homeostasis in cyanobacteria. *Plant Signal. Behav.* **2012**, *7* (12), 1712–1714.

- (24) Bruland, K. W.; Middag, R.; Lohan, M. C. Controls of trace metals in seawater. In *Treatise on Geochemistry: Second Edition*; Elderfield, H., Holland, H. D., Turekian, K. K., Eds.; Elsevier Ltd., **2013**; 8, 19–51.
- (25) Menzel, D. W.; Ryther, J. H. Nutrients limiting the production of phytoplankton in the Sargasso Sea, with special reference to iron. *Deep Sea Res.* **1961**, 7 (4), 276–281.
- (26) Andreini, C.; Bertini, I.; Rosato, A. Metalloproteomes: a bioinformatic approach. *Acc. Chem. Res.* **2009**, 42 (10), 1471–1479.
- (27) Cannon, G. C.; Heinhorst, S.; Kerfeld, C. A. Carboxysomal carbonic anhydrases: structure and role in microbial CO₂ fixation. *Biochim. Biophys. Acta - Proteins Proteomics* **2010**, 1804 (2), 382–392.
- (28) Silverman, D. N.; Lindskog, S. The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water. *Acc. Chem. Res.* **1988**, 21, 30–36.
- (29) Morel, F. M. M.; Reinfelder, J. R.; Roberts, S. B.; Chamberlain, C. P.; Lee, J. G.; Yee, D. Zinc and carbon co-limitation of marine phytoplankton. *Nature* **1994**, 369 (6483), 740–742.
- (30) Coleman, J. E. Zinc proteins: enzymes, storage proteins, transcription factors, and replication proteins. *Annu. Rev. Biochem.* **1992**, 61, 897–946.
- (31) Mahaffey, C.; Reynolds, S.; Davis, C. E.; Lohan, M. C. Alkaline phosphatase activity in the subtropical ocean: insights from nutrient, dust and trace metal addition experiments. *Front. Mar. Sci.* **2014**, 1 (December), 1–13.
- (32) Barnett, J. P.; Scanlan, D. J.; Blindauer, C. A. Identification of major zinc-binding proteins from a marine cyanobacterium: insight into metal uptake in oligotrophic environments. *Metallomics* **2014**, 6 (7), 1254–1268.
- (33) Barnett, J. P.; Millard, A.; Ksibe, A. Z.; Scanlan, D. J.; Schmid, R.; Blindauer, C. A. Mining genomes of marine cyanobacteria for elements of zinc homeostasis. *Front. Microbiol.* **2012**, 3, 1–21.
- (34) Blindauer, C. A. Advances in the molecular understanding of biological zinc transport. *Chem. Commun.* **2015**, 51 (22), 4544–4563.
- (35) Mikhaylina, A.; Ksibe, A. Z.; Scanlan, D. J.; Blindauer, C. A. Bacterial zinc uptake regulator proteins and their regulons. *Biochem. Soc. Trans.* **2018**, 46 (4), 983–1001.
- (36) Morby, A. P.; Turner, J. S.; Huckle, J. W.; Robinson, N. J. SmtB is a metal-dependent repressor of the cyanobacterial metallothionein gene SmtA: identification of a Zn inhibited DNA-protein complex. *Nucleic Acids Res.* **1993**, 21 (4), 921–925.

- (37) Blindauer, C. A. Zinc-handling in cyanobacteria: an update. *Chem. Biodivers.* **2008**, 5 (10), 1990–2013.
- (38) Collection of manually curated inferences of regulons in prokaryotic genomes. <http://regprecise.lbl.gov/RegPrecise/> (accessed Jul 18, 2019).
- (39) Huckle, J. W.; Morby, A. P.; Turner, J. S.; Robinson, N. J. Isolation of a prokaryotic metallothionein locus and analysis of transcriptional control by trace metal ions. *Mol. Microbiol.* **1993**, 7 (2), 177–187.
- (40) Thelwell, C.; Robinson, N. J.; Turner-Cavet, J. S. An SmtB-like repressor from *Synechocystis* PCC 6803 regulates a zinc exporter. *Proc. Natl. Acad. Sci.* **1998**, 95 (18), 10728–10733.
- (41) Liu, T.; Nakashima, S.; Hirose, K.; Shibasaka, M.; Katsuhara, M.; Ezaki, B.; Giedroc, D. P.; Kasamo, K. A Novel cyanobacterial SmtB/ArsR family repressor regulates the expression of a CPx-ATPase and a metallothionein in response to Both Cu(I)/Ag(I) and Zn(II)/Cd(II). *J. Biol. Chem.* **2004**, 279 (17), 17810–17818.
- (42) Liu, T.; Golden, J. W.; Giedroc, D. P. A Zinc(II)/lead(II)/cadmium(II)-inducible operon from the cyanobacterium *Anabaena* is regulated by AztR, an R3N ArsR/SmtB metalloregulator. *Biochemistry* **2005**, 44, 8673–8683.
- (43) Pruteanu, M.; Neher, S. B.; Baker, T. A. Ligand-controlled proteolysis of the *Escherichia Coli* transcriptional regulator ZntR. *J. Bacteriol.* **2007**, 189 (8), 3017–3025.
- (44) Brocklehurst, K. R.; Hobman, J. L.; Lawley, B.; Blank, L.; Marshall, S. J.; Brown, N. L.; Morby, A. P. ZntR is a Zn(II)-responsive MerR-like transcriptional regulator of ZntA in *Escherichia coli*. *Mol. Microbiol.* **1999**, 31 (3), 893–902.
- (45) Reyes-Caballero, H.; Guerra, A. J.; Jacobsen, F. E.; Kazmierczak, K. M.; Cowart, D.; Koppolu, U. M. K.; Scott, R. A.; Winkler, M. E.; Giedroc, D. P. The metalloregulatory zinc site in *Streptococcus pneumoniae* AdcR, a zinc-activated MarR family repressor. *J. Mol. Biol.* **2010**, 403 (2), 197–216.
- (46) Feng, Y.; Li, M.; Zhang, H.; Zheng, B.; Han, H.; Wang, C.; Yan, J.; Tang, J.; Gao, G. F. Functional definition and global regulation of Zur, a zinc uptake regulator in a *Streptococcus suis* serotype 2 strain causing streptococcal toxic shock syndrome. *J. Bacteriol.* **2008**, 190 (22), 7567–7578.
- (47) Shin, J. H.; Jung, H. J.; An, Y. J.; Cho, Y.-B.; Cha, S.-S.; Roe, J.-H. Graded expression of zinc-responsive genes through two regulatory zinc-binding sites in Zur. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, 108 (12), 5045–5050.

- (48) Gilston, B. A.; Wang, S.; Marcus, M. D.; Canalizo-Hernandez, M. A.; Swindell, E. P.; Xue, Y.; Mondragon, A.; O'Halloran, T. V. Structural and mechanistic basis of zinc regulation across the *E. coli* Zur regulon. *PLoS Biol.* **2014**, *12* (11), 1–16.
- (49) Lucarelli, D.; Russo, S.; Garman, E.; Milano, A.; Meyer-klaucke, W.; Pohl, E. Crystal structure and function of the zinc uptake regulator FurB from *Mycobacterium tuberculosis*. *J. Biol. Chem.* **2007**, *282* (13), 9914–9922.
- (50) Hudek, L.; Pearson, L. A.; Michalczyk, A. A.; Neilan, B. A.; Ackland, L. M. Molecular and cellular characterisation of the Zinc uptake (Znu) system of *Nostoc punctiforme*. *FEMS Microbiol. Ecol.* **2013**, *86* (2), 149–171.
- (51) Patzer, S. I.; Hantke, K. The ZnuABC high-affinity zinc uptake system and its regulator Zur in *Escherichia coli*. *Mol. Microbiol.* **1998**, *28* (6), 1199–1210.
- (52) Sein-Echaluce, V. C.; González, A.; Napolitano, M.; Luque, I.; Barja, F.; Peleato, M. L.; Fillat, M. F. Zur (FurB) is a key factor in the control of the oxidative stress response in *Anabaena* sp. PCC 7120. *Environ. Microbiol.* **2015**, *17* (6), 2006–2017.
- (53) Spohn, M.; Wohlleben, W.; Stegmann, E. Elucidation of the zinc dependent regulation in *Amycolatopsis japonicum* enabled the identification of the ethylenediamine-disuccinate ([S,S]-EDDS) genes. *Environ. Microbiol.* **2015**, *18*, 1249–1263.
- (54) Ludwig, M.; Chua, T. T.; Chew, C. Y.; Bryant, D. A.; Lawrence, C. M.; Hanson, T. E.; Bryant, D. A. Fur-type transcriptional repressors and metal homeostasis in the cyanobacterium *Synechococcus* sp. PCC 7002. *Front. Microbiol.* **2015**, *6*, 1217.
- (55) Neff, J. M. Zinc in the ocean. In *Bioaccumulation in Marine Organisms*; Neff, J. M., Ed.; Elsevier, **2002**; 175–189.
- (56) Bruland, K. W. Complexation of zinc by natural organic ligands in the central north Pacific. *Limnol. Oceanogr.* **1989**, 269–285.
- (57) Jakuba, R. W.; Saito, M. A.; Moffett, J. W.; Xu, Y. Dissolved zinc in the subarctic north Pacific and Bering Sea: its distribution, speciation, and importance to primary producers. *Global Biogeochem. Cycles* **2012**, *26* (2), 1–15.
- (58) Noinaj, N.; Guillier, M.; Barnard, T. J.; Buchanam, S. K. TonB-dependent transporters: regulation, structure, and function. *Annu. Rev. Microbiol.* **2010**, No. 64, 43–60.
- (59) Napolitano, M.; Rubio, M. Á.; Santamaría-Gómez, J.; Olmedo-Verd, E.; Robinson, N. J.; Luque, I. Characterization of the response to zinc deficiency in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J. Bacteriol.* **2012**, *194* (10), 2426–2436.

- (60) Stork, M.; Bos, M. P.; Jongerius, I.; de Kok, N.; Schilders, I.; Weynants, V. E.; Poolman, J. T.; Tommassen, J. An outer membrane receptor of *Neisseria meningitidis* involved in zinc acquisition with vaccine potential. *PLoS Pathog.* **2010**, *6* (7), 1–10.
- (61) Hood, M. I.; Mortensen, B. L.; Moore, J. L.; Zhang, Y.; Kehl-Fie, T. E.; Sugitani, N.; Chazin, W. J.; Caprioli, R. M.; Skaar, E. P. Identification of an *Acinetobacter baumannii* zinc acquisition system that facilitates resistance to calprotectin-mediated zinc sequestration. *PLoS Pathog.* **2012**, *8* (12), 20–24.
- (62) Lim, C. K.; Hassan, K. A.; Penesyan, A.; Loper, J. E.; Paulsen, I. T. The effect of zinc limitation on the transcriptome of *Pseudomonas protegens* Pf-5. *Environ. Microbiol.* **2013**, *15* (3), 702–715.
- (63) Weston, B. F.; Brenot, A.; Caparon, M. G. The metal homeostasis protein, Lsp, of *Streptococcus pyogenes* is necessary for acquisition of zinc and virulence. *Infect. Immun.* **2009**, *77* (7), 2840–2848.
- (64) Grass, G.; Wong, M. D.; Rosen, B. P.; Smith, R. L.; Rensing, C. ZupT is a Zn(II) uptake system in *Escherichia coli*. *J. Bacteriol.* **2002**, *184* (3), 864–866.
- (65) Eide, D. J. Zinc transporters and the cellular trafficking of zinc. *Biochim. Biophys. Acta - Mol. Cell Res.* **2006**, *1763* (7), 711–722.
- (66) Zhang, T.; Liu, J.; Fellner, M.; Zhang, C.; Sui, D.; Hu, J. Crystal structures of a ZIP zinc transporter reveal a binuclear metal center in the transport pathway. *Sci. Adv.* **2017**, *3*, 1–9.
- (67) Hudek, L.; Pearson, L. A.; Michalczyk, A.; Neilan, B. A.; Ackland, M. L. Functional characterization of the twin ZIP/SLC39 metal transporters, NpunF3111 and NpunF2202 in *Nostoc punctiforme*. *Appl. Microbiol. Biotechnol.* **2013**, *97* (19), 8649–8662.
- (68) Haas, C. E.; Rodionov, D. A.; Kropat, J.; Malasarn, D.; Merchant, S. S.; de Crécy-Lagard, V. A Subset of the diverse COG0523 family of putative metal chaperones is linked to zinc homeostasis in all kingdoms of life. *BMC Genomics* **2009**, *10*, 470.
- (69) Gabriel, S. E.; Miyagi, F.; Gaballa, A.; Helmann, J. D. Regulation of the *Bacillus subtilis* YciC gene and insights into the DNA-binding specificity of the zinc-sensing metalloregulator Zur. *J. Bacteriol.* **2008**, *190* (10), 3482–3488.
- (70) Blaby-Haas, C. E.; Flood, J. A.; Crécy-Lagard, V. De; Zamble, D. B. YeiR: a metal-binding GTPase from *Escherichia coli* involved in metal homeostasis. *Metallomics* **2012**, *4* (5), 488–497.
- (71) Nies, D. H.; Silver, S. Ion efflux systems involved in bacterial metal resistances. *J. Ind. Microbiol.* **1995**, *14* (2), 186–199.

- (72) Huang, D.; Tang, D.; Liao, Q.; Li, H.; Chen, Q.; He, Y.; Feng, J.; Jiang, B.; Lu, G.; Chen, B. The Zur of *Xanthomonas campestris* functions as a repressor and an activator of putative zinc homeostasis genes via recognizing two distinct sequences within its target promoters. *Nucleic Acids Res.* **2008**, *36* (13), 4295–4309.
- (73) Choi, S.-H.; Lee, K.; Shin, J.-H.; Cho, Y.-B.; Cha, S.-S.; Roe, J.-H. Zinc-dependent regulation of zinc import and export genes by Zur. *Nat. Commun.* **2017**, *8*, 1–11.
- (74) Hudek, L.; Pearson, L.; Michalczyk, A. A.; Bräu, L.; Neilan, B. A.; Ackland, M. L. Characterization of two cation diffusion facilitators NpunF0707 and NpunF1794 in *Nostoc punctiforme*. *J. Appl. Microbiol.* **2015**, *119* (5), 1357–1370.
- (75) Teramoto, H.; Inui, M.; Yukawa, H. *Corynebacterium Glutamicum* Zur acts as a zinc-sensing transcriptional repressor of both zinc-inducible and zinc-repressible genes involved in zinc homeostasis. *FEBS J.* **2012**, *279* (23), 4385–4397.
- (76) Mazzon, R. R.; Braz, V. S.; da Silva Neto, J. F.; do Valle Marques, M. Analysis of the *Caulobacter crescentus* Zur regulon reveals novel insights in zinc acquisition by TonB-dependent outer membrane proteins. *BMC Genomics* **2014**, *15*, 734.
- (77) Nies, D. H. Zinc starvation response in a cyanobacterium revealed. *J. Bacteriol.* **2012**, *194* (10), 2407–2412.
- (78) Blindauer, C. A.; Leszczyszyn, O. I. Metallothioneins: unparalleled diversity in structures and functions for metal ion homeostasis and more. *Nat. Prod. Rep.* **2010**, *27* (5), 720–741.
- (79) Blindauer, C. A. Bacterial metallothioneins: past, present, and questions for the future. *J. Biol. Inorg. Chem.* **2011**, *16* (7), 1011–1024.
- (80) Turner, J. S.; Morby, A. P.; Whitton, B. A.; Gupta, A.; Robinson, N. J. Construction of Zn²⁺/Cd²⁺ hypersensitive cyanobacterial mutants lacking a functional metallothionein locus. *J. Biol. Chem.* **1993**, *268* (6), 4494–4498.
- (81) Blindauer, C. A.; Harrison, M. D.; Parkinson, J. A.; Robinson, A. K.; Cavet, J. S.; Robinson, N. J.; Sadler, P. J. A metallothionein containing a zinc finger within a four-metal cluster protects a bacterium from zinc toxicity. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *58* (17), 9593–9598.
- (82) Blindauer, C. A.; Harrison, M. D.; Robinson, A. K.; Parkinson, J. A.; Bowness, P. W.; Sadler, P. J.; Robinson, N. J. Multiple bacteria encode metallothioneins and SmtA-like zinc fingers. *Mol. Microbiol.* **2002**, *45* (5), 1421–1432.

- (83) Waterbury, J. B.; Watson, S. W.; Valois, F. W.; Franks, D. G. Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*. In *Photosynthetic Picoplankton. Can. Bull. Fish. Aquat. Sci.* 214; Platt, T. and Li, W. K. W., Ed.; Ottawa, **1986**; 71–120.
- (84) Palenik, B.; Brahamsha, B.; Larimer, F. W.; Land, M.; Hauser, L.; Chain, P.; Lamerdin, J.; Waterbury, J. The genome of a motile marine *Synechococcus*. *Nature* **2003**, 424 (6952), 1035–1037.
- (85) Brahamsha, B. An abundant cell-surface polypeptide is required for swimming by the nonflagellated marine cyanobacterium *Synechococcus*. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, 93 (13), 6504–6509.
- (86) McCarren, J.; Brahamsha, B. Swimming motility mutants of marine *Synechococcus* affected in production and localization of the S-layer protein SwmA. *J. Bacteriol.* **2009**, 191 (3), 1111–1114.
- (87) McCarren, J.; Brahamsha, B. SwmB, a 1.12-megadalton protein that is required for nonflagellar swimming motility in *Synechococcus*. *J. Bacteriol.* **2007**, 189 (3), 1158–1162.
- (88) Ksibe, A. Z. Zinc on the move: insights towards understanding zinc homeostasis in the open ocean cyanobacterium *Synechococcus* sp. WH8102, University of Warwick, Coventry, **2016**.
- (89) Conejo, M. C.; García, I.; Martínez-Martínez, L.; Picabea, L.; Pascual, A. Zinc eluted from siliconized latex urinary catheters decreases OprD expression, causing carbapenem resistance in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **2003**, 47 (7), 2313–2315.
- (90) Foster, A. W.; Osman, D.; Robinson, N. J. Metal preferences and metallation. *J. Biol. Chem.* **2014**, 289 (41), 28095–28103.
- (91) Christie-Oleza, J. A.; Brunet-Galmés, I.; Lalucat, J.; Nogales, B.; Bosch, R. MiniUIB, a novel minitransposon-based system for stable insertion of foreign DNA into the genomes of gram-negative and gram-positive bacteria. *Appl. Environ. Microbiol.* **2013**, 79 (5), 1629–1638.
- (92) Doyle, S. A. High throughput protein expression and purification. *Methods Mol. Biol.* **2009**, 498 (1), 297–307.
- (93) Pohl, E.; Haller, J. C.; Mijovilovich, A.; Meyer-Klaucke, W.; Garman, E.; Vasil, M. L. Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator. *Mol. Microbiol.* **2003**, 47 (4), 903–915.
- (94) Butcher, J.; Sarvan, S.; Brunzelle, J. S.; Couture, J.-F.; Stintzi, A. Structure and regulon of *Campylobacter jejuni* ferric uptake regulator Fur define apo-Fur regulation. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, 109 (25), 10047–10052.

- (95) Jacquamet, L.; Traoré, D. A. K.; Ferrer, J. L.; Proux, O.; Testemale, D.; Hazemann, J. L.; Nazarenko, E.; El Ghazouani, A.; Caux-Thang, C.; Duarte, V.; *et al.* Structural characterization of the active form of PerR: insights into the metal-induced activation of PerR and Fur proteins for DNA binding. *Mol. Microbiol.* **2009**, *73* (1), 20–31.
- (96) An, Y. J.; Ahn, B. E.; Han, A. R.; Kim, H. M.; Chung, K. M.; Shin, J. H.; Cho, Y. B.; Roe, J. H.; Cha, S. S. Structural basis for the specialization of Nur, a nickel-specific Fur homolog, in metal sensing and DNA recognition. *Nucleic Acids Res.* **2009**, *37* (10), 3442–3451.
- (97) Bellini, D.; Lebedev, A.; Keegan, R.; Todd, J.; Hemmings, A. M.; Johnston, A. W.; Walsh, M. A. Structure of a zinc-bound manganese uptake regulator, Mur. *To be Publ.*
- (98) Nowakowski, A. B.; Wobig, W. J.; Petering, D. H. Native SDS-PAGE: high resolution electrophoretic separation of proteins with retention of native properties including bound metal ions. *Metallomics* **2014**, *6* (5), 1068–1078.
- (99) Jefferson, J. R.; Hunt, J. B.; Ginsburg, A. Characterization of Indo-1 and Quin-2 as spectroscopic probes for Zn²⁺- protein interactions. *Anal. Biochem.* **1990**, *187* (2), 328–336.
- (100) Kuzmič, P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.* **1996**, *237*, 260–273.
- (101) Kabsch, W. XDS. *Acta Crystallogr. Sect. D* **2010**, *66* (2), 125–132.
- (102) Dodson, E. J.; Winn, M.; Ralph, A. Collaborative computational project, number 4: providing programs for protein crystallography. *Methods Enzymol.* **1997**, *277*, 620–633.
- (103) Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. Sect. A* **2008**, *64* (1), 112–122.
- (104) Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D* **2004**, *60* (12, 1), 2126–2132.
- (105) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53* (3), 240–255.
- (106) Langer, G.; Cohen, S. X.; Lamzin, V. S.; Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/WARP version 7. *Nat. Protoc.* **2008**, *3* (7), 1171–1179.
- (107) ExPASy - Compute pI/Mw tool. https://web.expasy.org/compute_pi/ (accessed Jul 18, 2019).

- (108) Pace, N. C.; Vajdos, F.; Fee, L.; Gerald, G.; Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **1995**, *4*, 2411–2423.
- (109) Ho, C. S.; Lam, C. W. K.; Chan, M. H. M.; Cheung, R. C. K.; Law, L. K.; Lit, L. C. W.; Ng, K. F.; Suen, M. W. M.; Tai, H. L. Electrospray ionisation mass spectrometry: principles and clinical applications. *Clin Biochem Rev* **2003**, *24* (February), 3–12.
- (110) Van Den Heuvel, R. H. H.; Heck, A. J. R. Native protein mass spectrometry: from intact oligomers to functional machineries. *Curr. Opin. Chem. Biol.* **2004**, *8* (5), 519–526.
- (111) Hernandez, J. A.; Bes, M. T.; Fillat, M. F.; Neira, J. L.; Peleato, M. L. Biochemical analysis of the recombinant Fur (ferric uptake regulator) protein from *Anabaena* PCC 7119: factors affecting its oligomerization state. *Biochem. J.* **2002**, *366* (1), 315–322.
- (112) Neupane, D. P.; Jacquez, B.; Sundararajan, A.; Ramaraj, T.; Bird, A. Zinc-dependent transcriptional regulation in *Paracoccus denitrificans*. *Front. Microbiol.* **2017**, *8* (April), 1–15.
- (113) Ma, Z.; Gabriel, S. E.; Helmann, J. D. Sequential binding and sensing of Zn(II) by *Bacillus subtilis* Zur. *PLoS Biol.* **2011**, *39* (21), 9130–9138.
- (114) Shin, J. H.; Oh, S. Y.; Kim, S. J.; Roe, J. H. The zinc-responsive regulator Zur controls a zinc uptake system and some ribosomal proteins in *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **2007**, *189* (11), 4070–4077.
- (115) Outten, C. E.; Tobin, D. A.; Penner-hahn, J. E.; Halloran, T. V. O. Characterization of the metal receptor sites in *Escherichia coli* Zur, an ultrasensitive zinc(II) metalloregulatory protein. *Biochemistry* **2001**, *40*, 10417–10423.
- (116) Krężel, A.; Maret, W. The biological inorganic chemistry of zinc ions. *Arch. Biochem. Biophys.* **2016**, *611* (May), 3–19.
- (117) Fillat, M. F. The FUR (ferric uptake regulator) superfamily: diversity and versatility of key transcriptional regulators. *Arch. Biochem. Biophys.* **2014**, *546*, 41–52.
- (118) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L. N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **1989**, *179* (1), 131–137.
- (119) Osman, D.; Piergentili, C.; Chen, J.; Chakrabarti, B.; Foster, A. W.; Lurie-Luke, E.; Huggins, T. G.; Robinson, N. J. Generating a metal-responsive transcriptional regulator to test what confers metal sensing in cells. *J. Biol. Chem.* **2015**, *290* (32), 19806–19822.

- (120) Patterson, C. J. *In vitro* properties and *in vivo* responses of CoaR, ZiaR & Zur (trans-family metal-sensing), Newcastle University, Newcastle, **2010**.
- (121) Ferguson, K. A. Starch-gel electrophoresis - application to the classification of pituitary proteins and polypeptides. *Metabolism* **1964**, *13* (10), 985–1002.
- (122) Malgieri, G.; Palmieri, M.; Esposito, S.; Maione, V.; Russo, L.; Baglivo, I.; de Paola, I.; Milardi, D.; Diana, D.; Zaccaro, L.; *et al.* Zinc to cadmium replacement in the prokaryotic zinc-finger domain. *Metallomics* **2014**, *6* (1), 96–104.
- (123) Blindauer, C. A.; Harrison, M. D.; Parkinson, J. A.; Robinson, N. J.; Sadler, P. J. Isostructural replacement of zinc by cadmium in bacterial metallothionein. In *Metal Ions in Biology and Medicine*; Collery, P., Maynard, I., Theophanides, T., Khassanova, L., Collery, T., Eds.; John Libbey Eurotext, **2008**; *10*, 167–173.
- (124) Kopera, E.; Schwerdtle, T.; Hartwig, A.; Bal, W. Co(II) and Cd(II) substitute for Zn(II) in the zinc finger derived from the DNA repair protein XPA, demonstrating a variety of potential mechanisms of toxicity. *Chem. Res. Toxicol.* **2004**, *17* (11), 1452–1458.
- (125) Shumilina, E.; Dobrovolska, O.; Del Conte, R.; Holen, H. W.; Dikiy, A. Competitive cobalt for zinc substitution in mammalian methionine sulfoxide reductase B1 overexpressed in *E. coli*: structural and functional insight. *J. Biol. Inorg. Chem.* **2014**, *19* (1), 85–95.
- (126) Irving, B. H.; Williams, R. J. P. The stability of transition-metal complexes. *J. Chem. Soc.* **1953**, 3192–3210.
- (127) Rubino, J. T.; Franz, K. J. Coordination chemistry of copper proteins: how nature handles a toxic cargo for essential function. *J. Inorg. Biochem.* **2012**, *107* (1), 129–143.
- (128) Prudent, M.; Girault, H. H. The role of copper in cysteine oxidation: study of intra- and inter-molecular reactions in mass spectrometry. *Metallomics* **2009**, *1* (2), 157–165.
- (129) Harding, M. M. Small revisions to predicted distances around metal sites in proteins. *Acta Crystallogr. Sect. D* **2006**, *62* (6), 678–682.
- (130) Drozdetskiy, A.; Cole, C.; Procter, J.; Barton, G. J. JPred4: A protein secondary structure prediction server. *Nucleic Acids Res.* **2015**, *43* (W1), W389–W394.
- (131) NCBI protein BLAST.
https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome (accessed Jul 18, 2019).

- (132) Newby, R.; Lee, L. H.; Perez, J. L.; Tao, X.; Chu, T. Characterization of zinc stress response in cyanobacterium *Synechococcus* sp. IU 625. *Aquat. Toxicol.* **2017**, *186*, 159–170.
- (133) Lindsay, J. A.; Foster, S. J. Zur: A Zn²⁺-responsive regulatory element of *Staphylococcus aureus*. *Microbiology* **2001**, *147* (5), 1259–1266.
- (134) Smith, K. F.; Bibb, L. A.; Schmitt, M. P.; Oram, D. M. Regulation and activity of a zinc uptake regulator, Zur, in *Corynebacterium diphtheriae*. *J. Bacteriol.* **2009**, *191* (5), 1595–1603.
- (135) Wilson, W. H.; Carr, N. G.; Mann, N. H. The effect of phosphate status on the kinetics of cyanophage infection in the oceanic cyanobacterium *Synechococcus* sp. WH7803. *J. Phycol.* **1996**, *32*, 506–516.
- (136) *E. coli* cell culture concentration calculator.
<https://www.chem.agilent.com/store/biocalculators/calcODBacterial.jsp>
(accessed Jul 18, 2019).
- (137) Murray, M. G.; Thompson, W. F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **1980**, *8* (19), 4321–4325.
- (138) Brahamsha, B. A genetic manipulation system for oceanic cyanobacteria of the genus *Synechococcus*. *Appl. Environ. Microbiol.* **1996**, *62* (5), 1747–1751.
- (139) Sunda, W. G.; Price, N. M.; Morel, F. M. M. Trace metal ion buffers and their use in culture studies. In *Algal Culturing Techniques*; Andersen, R. A., Ed.; Elsevier Academic Press, **2005**; 52–55.
- (140) Chelex® 100 and Chelex 20 chelating ion exchange resin instruction manual. www.bio-rad.com/webroot/web/pdf/lsr/literature/LIT200.pdf (accessed Jul 18, 2019).
- (141) Scanlan, D. J.; Silman, N. J.; Donald, K. M.; Wilson, W. H.; Carr, N. G.; Joint, I.; Mann, N. H. An immunological approach to detect phosphate stress in populations and single cells of photosynthetic picoplankton. *Appl. Environ. Microbiol.* **1997**, *63* (6), 2411–2420.
- (142) Bruinsma, J. The quantitative analysis of chlorophylls a and b in plant extracts. *Photochem. Photobiol.* **1963**, *2* (2), 241–249.
- (143) Logemann, J.; Schell, J.; Willmitzer, L. Improved method for the isolation of RNA from plant tissues. *Anal. Biochem.* **1987**, *163* (1), 16–20.
- (144) Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **2015**, *17* (1), 10–12.
- (145) Kim, D.; Langmead, B.; Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **2016**, *12* (4), 357–360.

- (146) Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, *25* (16), 2078–2079.
- (147) Liao, Y.; Smyth, G. K.; Shi, W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**, *30* (7), 923–930.
- (148) Love, M. I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15* (12), 1–21.
- (149) Livak, K. J.; Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **2001**, *25* (4), 402–408.
- (150) IDT PrimerQuest tool. <https://eu.idtdna.com/PrimerQuest/Home/Results> (accessed Jul 18, 2019).
- (151) Eriksson, V. I. The response of *Synechococcus* sp. CC9311 to iron stress, The University of Warwick, Coventry, **2013**.
- (152) Molbioltools DNA calculator. <http://www.molbioltools.com/dnacalculator.html> (accessed Jul 18, 2019).
- (153) Chu, J. Towards an understanding of multiple paralogues for metal-handling genes in a coastal cyanobacterium, University of Warwick, Coventry, **2012**.
- (154) Ammendola, S.; Pasquali, P.; Pistoia, C.; Petrucci, P.; Petrarca, P.; Rotilio, G.; Battistoni, A. High-affinity Zn^{2+} uptake system ZnuABC is required for bacterial zinc homeostasis in intracellular environments and contributes to the virulence of *Salmonella enterica*. *Infect. Immun.* **2007**, *75* (12), 5867–5876.
- (155) Egli, T. Biodegradation of metal-complexing aminopolycarboxylic acids. *J. Biosci. Bioeng.* **2001**, *92* (2), 89–97.
- (156) Brown, C. M.; Lawrence, J. E.; Campbell, D. A. Are phytoplankton population density maxima predictable through analysis of host and viral genomic DNA content? *J. Mar. Biol. Assoc. United Kingdom* **2006**, *86*, 491–498.
- (157) Ikeda, J. S.; Janakiraman, A.; Kehres, D. G.; Maguire, M. E.; Slauch, J. M. Transcriptional regulation of SitABCD of *Salmonella enterica* serovar typhimurium by MntR and Fur. *J. Bacteriol.* **2005**, *187* (3), 912–922.
- (158) Osman, D.; Foster, A. W.; Chen, J.; Svedaite, K.; Steed, J. W.; Lurie-Luke, E.; Huggins, T. G.; Robinson, N. J. Fine control of metal concentrations is necessary for cells to discern zinc from cobalt. *Nat. Commun.* **2017**, *8* (1), 1–12.

- (159) Sein-Echaluce, V. C.; González, A.; Napolitano, M.; Luque, I.; Barja, F.; Peleato, M. L.; Fillat, M. F. Zur (FurB) is a key factor in the control of the oxidative stress response in *Anabaena* sp. PCC 7120. *Environ. Microbiol.* **2015**, *17* (6), 2006–2017.
- (160) So, A. K.-C.; Espie, G. S. Cyanobacterial carbonic anhydrases. *Can. J. Bot.* **2005**, *83* (7), 721–734.
- (161) Morel, F. M. M.; Rueter, J. G.; Anderson, D. M.; Guillard, R. R. L. Aquil: a chemically defined phytoplankton culture medium for trace metal studies. *J. Phycol.* **1979**, *15*, 135–141.
- (162) Tetu, S. G.; Brahamsha, B.; Johnson, D. A.; Tai, V.; Phillippy, K.; Palenik, B.; Paulsen, I. T. Microarray analysis of phosphate regulation in the marine cyanobacterium *Synechococcus* sp. WH8102. *ISME J.* **2009**, *3* (7), 835–849.
- (163) Andrews, S. Babraham bioinformatics - FastQC a quality control tool for high throughput sequence data.
<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed Jul 18, 2019).
- (164) PRODORIC bacterial regulon analyzer: regulon prediction.
http://www.prodoric.de/vfp/vfp_regulon.php (accessed Jul 18, 2019).
- (165) Hao, N.; Palmer, A. C.; Dodd, I. B.; Shearwin, K. E. Directing traffic on DNA - how transcription factors relieve or induce transcriptional interference. *Transcription* **2017**, *8* (2), 120–125.
- (166) Robinson, J. T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E. S.; Getz, G.; Mesirov, J. P. Integrative Genomics Viewer. *Nat. Biotechnol.* **2011**, *29* (1), 24–26.
- (167) SnapGene software from GSL Biotech. <https://www.snapgene.com/> (accessed Jul 18, 2019).
- (168) González, A.; Bes, M. T.; Luisa Peleato, M.; Fillat, M. F. Expanding the role of FurA as essential global regulator in cyanobacteria. *PLoS One* **2016**, *11* (3), 1–22.
- (169) Solovyev, V.; Salamov, A. Automatic annotation of microbial genomes and metagenomic sequences. In *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies*; Li, R. W., Ed.; Nova Science Publishers, **2010**; 61–78.
- (170) Kraker, A. J.; Krakower, G.; Shaw, C. F.; Petering, D. H.; Garvey, J. S. Zinc metabolism in ehrlich cells: properties of a metallothionein-like zinc-binding protein. *Cancer Res.* **1988**, *48* (12), 3381–3388.

- (171) Cox, A. D.; Saito, M. A. Proteomic responses of oceanic *Synechococcus* WH8102 to phosphate and zinc scarcity and cadmium additions. *Front. Microbiol.* **2013**, *4*, 1–17.
- (172) Dufresne, A.; Ostrowski, M.; Scanlan, D. J.; Garczarek, L.; Mazard, S.; Palenik, B. P.; Paulsen, I. T.; de Marsac, N. T.; Wincker, P.; Dossat, C.; *et al.* Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol.* **2008**, *9* (5), R90.
- (173) EMBOSS Needle pairwise sequence alignment.
https://www.ebi.ac.uk/Tools/psa/emboss_needle/ (accessed Jul 18, 2019).
- (174) Blindauer, C. A.; Schmid, R. Cytosolic metal handling in plants: determinants for zinc specificity in metal transporters and metallothioneins. *Metallomics* **2010**, *2* (8), 510–529.
- (175) Pugsley, A. P. The complete general secretory pathway in gram-negative bacteria. *Microbiol. Rev.* **1993**, *57* (1), 50–108.
- (176) Kelley, L. A.; Mezulis, S.; Yates, C. M.; Wass, M. N.; Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **2015**, *10* (6), 845–858.
- (177) Martin, E. M.; Kondrat, F. D. L.; Stewart, A. J.; Scrivens, J. H.; Sadler, P. J.; Blindauer, C. A. Native electrospray mass spectrometry approaches to probe the interaction between zinc and an anti-angiogenic peptide from histidine-rich glycoprotein. *Sci. Rep.* **2018**, *8* (1), 1–13.
- (178) Bonsor, D. A.; Hecht, O.; Vankemmelbeke, M.; Sharma, A.; Krachler, A. M.; Housden, N. G.; Lilly, K. J.; James, R.; Moore, G. R.; Kleanthous, C. Allosteric β -propeller signalling in TolB and its manipulation by translocating colicins. *EMBO J.* **2009**, *28* (18), 2846–2857.
- (179) Cyanorak, a database of marine picocyanobacteria genomes.
<http://application.sb-roscoff.fr/cyanorak/cyanorak.html?jsessionid=6E056C9058DAC6013869C1DFD0447303?execution=e1s1> (accessed Jul 18, 2019).
- (180) Fu, C.; Olson, J. W.; Maier, R. J. HypB protein of *Bradyrhizobium Japonicum* is a metal-binding GTPase capable of binding 18 divalent nickel ions per dimer. *Proc Natl Acad Sci U. S. A.* **1995**, *92* (March), 2333–2337.
- (181) Stuart, R. K.; Brahamsha, B.; Palenik, B. A Deg-protease family protein in marine *Synechococcus* is involved in outer membrane protein organization. *Front. Mar. Sci.* **2014**, *1* (June), 1–12.
- (182) Stuart, R. K.; Dupont, C. L.; Johnson, D. A.; Paulsen, I. T.; Palenik, B. Coastal strains of marine *Synechococcus* species exhibit increased tolerance to copper shock and a distinctive transcriptional response relative to those of open-ocean strains. *Appl. Environ. Microbiol.* **2009**, *75* (15), 5047–5057.

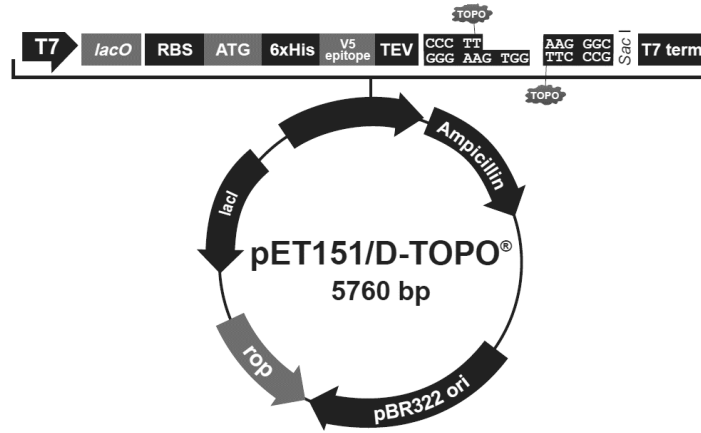
- (183) Dupont, C. L. Physiological, ecological and evolutionary studies of trace metal homeostasis in marine microbes, University of California, San Diego, **2008**.
- (184) Ostrowski, M.; Mazard, S.; Tetu, S. G.; Phillippy, K.; Johnson, A.; Palenik, B.; Paulsen, I. T.; Scanlan, D. J. PtrA is required for coordinate regulation of gene expression during phosphate stress in a marine *Synechococcus*. *ISME J.* **2010**, *4* (7), 908–921.
- (185) Sambrook, J.; Russell, D. W. *Molecular cloning: a laboratory manual*, 3rd ed.; Cold Spring Harbor Laboratory Press: New York, **2001**.
- (186) Hwang, T. L.; Shaka, A. J. Water suppression that works. Excitation sculpting using arbitrary wave-forms and pulsed-field gradients. *J. Magn. Reson. Ser. A* **1995**, *112* (2), 275–279.
- (187) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **1995**, *6* (2), 135–140.
- (188) Goddard, T. D.; Kneller, D. G. Sparky 3 - NMR assignment program. University of California: San Francisco, **2007**.
- (189) Hill, A. V. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *J. Physiol.* **1910**, 4–6.
- (190) Stefan, M. I.; Noverre, N. Le. Cooperative binding. *PLOS Comput. Biol.* **2013**, *9* (6), 2–7.
- (191) Imam, H. T. Understanding cluster topology and metal-cluster dynamics of two zinc binding plant metallothionein isoforms, University of Warwick, Coventry, **2015**.
- (192) Acharya, C.; Blindauer, C. A. Unexpected interactions of the cyanobacterial metallothionein SmtA with uranium. *Inorg. Chem.* **2016**, *55* (4), 1505–1515.
- (193) Flinta, C.; Persson, B.; Journvall, H.; Heijne, G. Sequence determinants of cytosolic N-terminal protein processing. *Eur. J. Biochem.* **1986**, *154*, 193–196.
- (194) Shi, J.; Lindsay, W. P.; Huckle, J. W.; Morby, A. P.; Robinson, N. J. Cyanobacterial metallothionein gene expressed in *Escherichia Coli*. Metal-binding properties of the expressed protein. *FEBS Lett.* **1992**, *303* (2), 159–163.
- (195) Vasák, M.; Kāgi, J. H. R.; Hill, H. A. O. Zinc(II), cadmium(II), and mercury(II) thiolate transitions in metallothionein. *Biochemistry* **1981**, *20* (10), 2852–2856.

- (196) Freisinger, E. Plant MTs—long neglected members of the metallothionein superfamily. *Dalt. Trans.* **2008**, 9226 (47), 6663–6675.
- (197) Leszczyszyn, O. I.; Evans, C. D.; Keiper, S. E.; Warren, G. Z. L.; Blindauer, C. A. Differential reactivity of individual zinc ions in clusters from bacterial metallothioneins. *Inorg. Chim. Acta* **2007**, 360 (1), 3–13.
- (198) Bennion, B. J.; Daggett, V. The molecular basis for the chemical denaturation of proteins by urea. *Proc Natl Acad Sci U. S. A.* **2003**, 100 (9), 5142–5147.
- (199) PD-10 desalting column manual.
http://wolfson.huji.ac.il/purification/PDF/dialysis/AMERSHAM_PD10Desalting.pdf (accessed Jul 18, 2019).
- (200) Vasak, M. Metal removal and substitution in vertebrate and invertebrate metallothioneins. *Methods Enzymol.* **1991**, 205, 452–458.
- (201) Antosiewicz, J. M.; Shugar, D. UV–Vis spectroscopy of tyrosine side-groups in studies of protein structure. Part 2: selected applications. *Biophys. Rev.* **2016**, 8 (2), 163–177.
- (202) Wishart, D. S.; Sykes, B. D.; Richards, F. M. The chemical shift index : a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **1992**, 31 (6), 1647–1651.
- (203) Poulsen IDP/IUP random coil chemical shifts.
https://spin.niddk.nih.gov/bax/nmrserver/Poulsen_rc_CS/ (accessed Aug 5, 2019).
- (204) Schulz, K. G.; Zondervan, I.; Gerringa, L. J. A.; Timmermans, K. R.; Veldhuis, M. J. W.; Riebesell, U. Effect of trace metal availability on coccolithophorid calcification. *Nature* **2004**, 430 (7000), 673–676.
- (205) Varela, D. E.; Willers, V.; Crawford, D. W. Effect of zinc availability on growth, morphology, and nutrient incorporation in a coastal and an oceanic diatom. *J. Phycol.* **2011**, 47 (2), 302–312.
- (206) Moore, C. M.; Mills, M. M.; Arrigo, K. R.; Berman-Frank, I.; Bopp, L.; Boyd, P. W.; Galbraith, E. D.; Geider, R. J.; Guieu, C.; Jaccard, S. L.; *et al.* Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* **2013**, 6 (9), 701–710.
- (207) Köbberich, M.; Vance, D. Zn isotope fractionation during uptake into marine phytoplankton: implications for oceanic zinc isotopes. *Chem. Geol.* **2019**, 523, 154–161.
- (208) Twining, B. B. S.; Nunez-Milland, D.; Nuñez-Milland, D.; Vogt, S.; Johnson, R. S.; Sedwick, P. N. Variations in *Synechococcus* cell quotas of phosphorus, sulfur, manganese, iron, nickel, and zinc within mesoscale eddies in the Sargasso Sea. *Limnol. Oceanogr.* **2010**, 55 (2), 492–506.

- (209) Graham, A. I.; Hunt, S.; Stokes, S. L.; Bramall, N.; Bunch, J.; Cox, A. G.; McLeod, C. W.; Poole, R. K. Severe zinc depletion of *Escherichia coli*. *J. Biol. Chem.* **2009**, *284* (27), 18377–18389.
- (210) Blindauer, C. A.; Razi, M. T.; Campopiano, D. J.; Sadler, P. J. Histidine ligands in bacterial metallothionein enhance cluster stability. *J. Biol. Inorg. Chem.* **2007**, *12* (3), 393–405.
- (211) Monastersky, R. Global carbon dioxide levels near worrisome milestone. *Nature* **2013**, *497* (7447), 13–14.

8. Appendix

A



B

```

ATGACCCGGTAGCAGTCCGGCACTGAATGCACGTCAGCAAGCACTGCTGACAGCCCTGAAT      60
GCATGTGGTGATGAAATGAGCGGTGACGAGCTGCATCGTAGCCCTGGATGATGAAGCAAGC      120
ATGGGTCTGGCAACCGTTTATCGTAATCTGCGTCAGCTGCAGCAGCGTGGTCTGGTTCGT      180
TGTCGTCATCTGCCGACCGGTGAAGCACTGTATGCACCGTTGATCGTGATCGTCATCAT      240
CTGACCTGTGTGATTTGTGGCACCACCAGGTTCTGGATCATTGTCCGATTTCATGGTATT      300
GATTGTTCCGGCAGATAGCCGTGGTGATTTTTGAACTGCTGTTTTCATACCTGGAATTTTTT      360
GGTTTTTTGTAGCAGCTGTCGTCCGCAGCGTAGCAGCAAACCGTAA                        405
  
```

Figure 8.1 Plasmid used for Zur over-expression and a codon-optimized synthetic SynZur sequence.

A: pET151/D-TOPO vector map. B: Sequence of the codon-optimized synthetic SynZur gene. Modified nucleotides are shown with bold underlined letters.

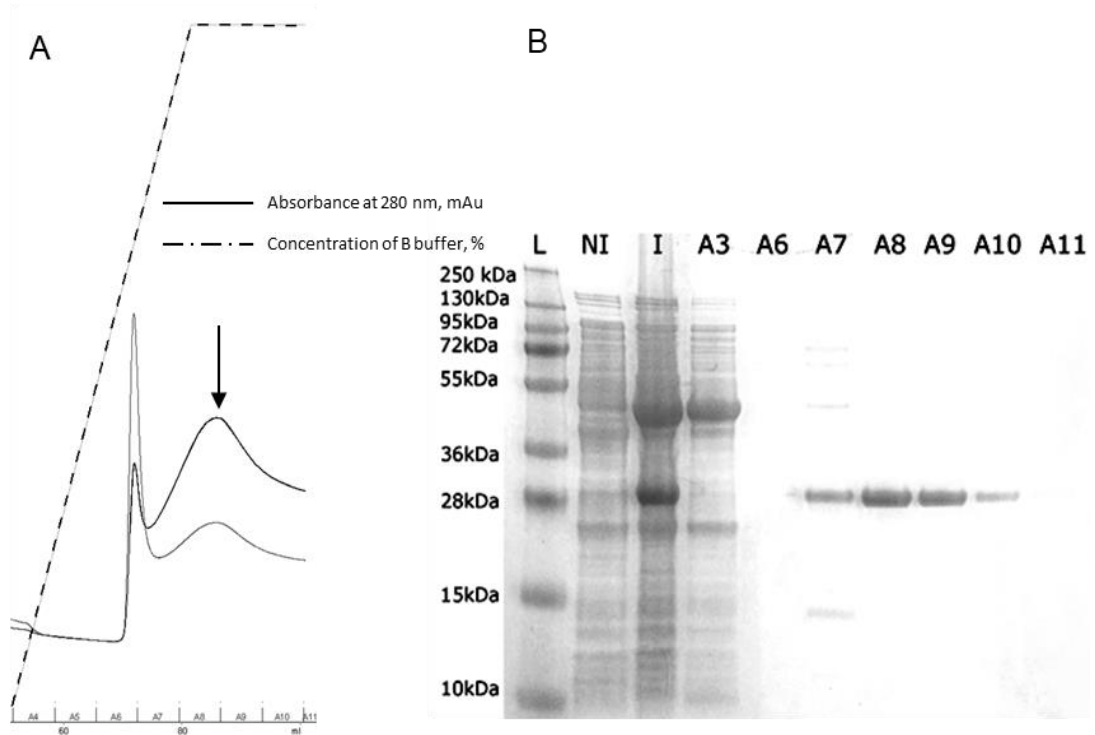


Figure 8.2 Over-expression and purification of TEV protease by FPLC.

A: Elution profile of TEV protease during FPLC. The absorbance at 280 nm is shown in black and the concentration of elution buffer is shown with a dashed line (100% at the top). Fractions collected are shown at the bottom. **B:** SDS-PAGE analysis of the fractions obtained during FPLC. NI: Culture without induction; I: Induced cell culture; A3-A11: Fractions collected. The TEV protease, with a theoretical size of 30 kDa, is marked with a black arrow.

```

[task]
    task = fit
    data = equilibria
[mechanism]
    Zn + Quin <==> ZnQuin : Kd1 dissoc
    Zn + Zur <==> ZnZur : Kd2 dissoc
[constants]      ; units: uM, %
    Kd1 = .0000037
    Kd2 = .00000083?
[concentrations] ; units: uM
    Quin = Q
    Zur = Z
[responses]      ; units: % Quin/uM Quin
    Quin = 5.72
[data]
    variable Zn
    set titration
[set:titration]
    Zn, uM          Quin, %
    X1             Y1
    X2...          Y2...
    Xn             Yn
[output]
    directory ./Zur/output/
[end]

```

Figure 8.3 Dynafit script used for calculation of $\text{SynZurZn}_1 K_D$.

Kd1: The Quin-2-Zn dissociation constant, μM . **Kd2:** The SynZurZn_1 dissociation constant, μM . This value is fitted by the software to match the experimental curve (parameters marked by a question mark are being fitted). **Q:** The total Quin-2 concentration, μM ; **Z:** The total Zur concentration, μM ; **R:** The Quin-2 response coefficient = $100\%/\text{Quin-2}$, μM ; **Xi:** The total zinc concentration after addition of each ZnSO_4 aliquot, μM , experimental value; **Yi:** The Free Quin-2 fraction calculated at 261 nm spectrophotometrically, experimental value, %.

```

A
[task]
    task = fit
    data = equilibria
[mechanism]
    ZurZur + DNA <==> ZurZurDNA
: Kd dissoc
[constants] ; units: 10-9 nM
    Kd = 4.78?
[concentrations] ; units: nM
    DNA = 3.16
[responses] ; units: DNA %/ DNA
nM
    DNA = 0.3165
[data]
    variable ZurZur
    set titration
[set:titration]
ZurZur, nM    DNA,%
    0         1
    0.5       0.959580754
    1         0.952724864
    3         0.786174137
    4         0.706384412
    5         0.663063606
    6         0.563620257
    8         0.371672265
    10        0.329551119
    20        0.019828884
    50        0
[output]
    directory ./Zur/output/
[end]

B
[task]
    task = fit
    data = equilibria
[mechanism]
    Zur + Zur + DNA <==>
ZurZurDNA : Kd dissoc
[constants] ; units: 10-18 nM2
    Kd = 99.9?
[concentrations] ; units: nM2
    DNA = 3.16
[responses] ; units: DNA %/ DNA
nM
    DNA = 0.3165
[data]
    variable Zur
    set titration
[set:titration]
    Zur total, nM;    DNA,%
    0         1
    1         0.959580754
    2         0.952724864
    6         0.786174137
    8         0.706384412
    10        0.663063606
    12        0.563620257
    16        0.371672265
    20        0.329551119
    40        0.019828884
    100       0
[output]
    directory ./Zur/output/
[end]

```

Figure 8.4 Scripts used in DynaFit software for fitting the K_D to the experimental EMSA data.

A: Model 1. B: Model 2. K_d : The zinc dissociation constant, 10^{-9} (nM) or 10^{-18} (nM²); parameters marked by a question mark are being fitted. DNA: % is calculated by dividing the intensity of the DNA band by the sum of intensities of the complex and DNA bands for each particular lane.

```

[task]
    task = fit
    data = equilibria

[mechanism]
    Zur + Zur <==> ZurZur : Kd dissociation
    ZurZur + DNA <==> ZurZurDNA : Kd1 dissociation
[constants] ; units: 10-9 nM
    Kd = 50
    Kd1 = 1.13?
[concentrations] ; units: nM
    DNA = 3.16
[responses] ; units: DNA % / DNA nM
    DNA = 0.3165
[data]
    variable Zur
    set titration
[set:titration]
Zur total, nM      DNA,%
    0      1
    1      0.959580754
    2      0.952724864
    6      0.786174137
    8      0.706384412
    10     0.663063606
    12     0.563620257
    16     0.371672265
    20     0.329551119
    40     0.019828884
    100    0
[output]
    directory ./Zur/output/
[end]

```

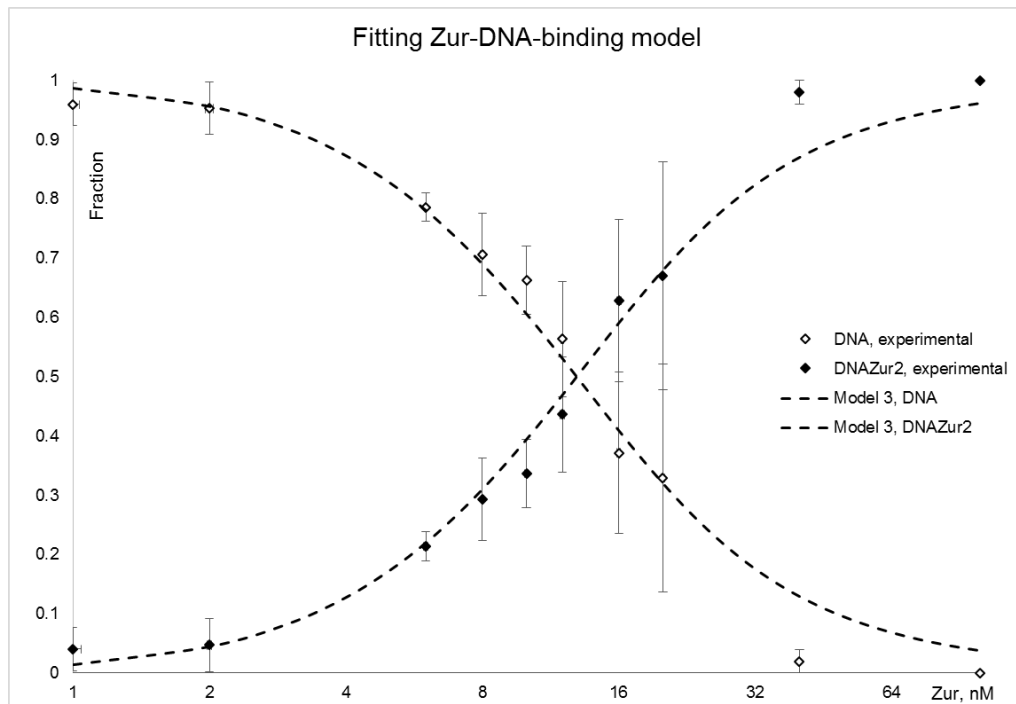


Figure 8.5 Dynafit script for Model 3 and the fitted curve.

Table 8.1 SynZur structural information**SynZur structural information:**

Resolution – 2.24 Å

Space group – P6₅

Rfactor – 0.20

Rfree – 0.25

RMSD bonds – 0.0145

Number of chains in asymmetric unit – 4 (A, B, C, D)

Number of domains in each monomer – 2 (domain 1 = residues 13 – 66, domain 2 = residues 67-132).

Model was built for residues 13-107 and 116-132. Residues 1-12, 108-115, and 133 to the end of the sequence were not possible to model due to poor electron density in these sections.

Number of co-ordinated zinc ions per monomer – 2

Symmetry related metal co-ordination of zinc ion between two histidine residues in chain C and two histidine residues in chain D.

PISA Analysis

Interfaces can be observed between chains A and C, and also between B and D.

Interfaces contain several hydrogen bonds (see Tables below) but no salt bridges, covalent bonds or disulfide bonds.

| Chain B residues | Distance/ Å | Chain D residues |
|-------------------------|--------------------|-------------------------|
| Pro71 (O) | 3.31 | Gln60 (NE2) |
| Asp91 (O) | 3.24 | Gln61 (NE2) |
| Phe128 (O) | 2.88 | Glu117 (N) |
| Phe126 (O) | 2.83 | Leu119 (N) |
| Phe126 (O) | 3.22 | Phe120 (N) |
| Glu124 (O) | 2.82 | Thr122 (N) |
| Thr122 (O) | 2.90 | Glu124 (N) |
| Phe120 (O) | 2.69 | Phe126 (N) |
| Glu117 (O) | 2.69 | Phe128 (N) |
| Asp115 (O) | 3.29 | Ser130 (N) |
| Phe128 (N) | 2.92 | Glu117 (O) |
| Phe126 (N) | 2.90 | Phe120 (O) |
| Glu124 (N) | 2.78 | Thr122 (O) |
| Thr122 (N) | 2.90 | Glu124 (O) |
| Phe120 (N) | 3.06 | Phe126 (O) |
| Leu119 (N) | 2.74 | Phe126 (O) |
| Glu117 (N) | 2.70 | Phe128 (O) |
| Chain A residues | Distance/ Å | Chain C residues |
| Pro71 (O) | 3.66 | His69 (NE2) |
| Asp91 (O) | 3.20 | Gln61 (NE2) |
| Leu70 (O) | 3.37 | His69 (NE2) |
| Phe128 (O) | 2.73 | Glu117 (N) |
| Phe126 (O) | 2.96 | Leu119 (N) |
| Phe126 (O) | 3.17 | Phe120 (N) |
| Glu124 (O) | 2.81 | Thr122 (N) |
| Thr122 (O) | 2.88 | Glu124 (N) |
| Phe120 (O) | 2.74 | Phe126 (N) |

| | | |
|-------------|------|-------------|
| Glu117 (O) | 3.03 | Phe128 (N) |
| His69 (NE2) | 3.37 | Pro71 (O) |
| Gln61 (NE2) | 3.12 | Asp91 (OD1) |
| Phe128 (N) | 2.90 | Glu117 (O) |
| Phe126 (N) | 2.91 | Phe120 (O) |
| Glu124 (N) | 2.91 | Thr122 (O) |
| Thr122 (N) | 2.79 | Glu124 (O) |
| Leu119 (N) | 2.76 | Phe126 (O) |
| Phe120 (N) | 3.16 | Phe126 (O) |
| Glu117 (N) | 3.12 | Phe128 (O) |

Metal geometry was checked using CheckMyMetal (CMM) server.

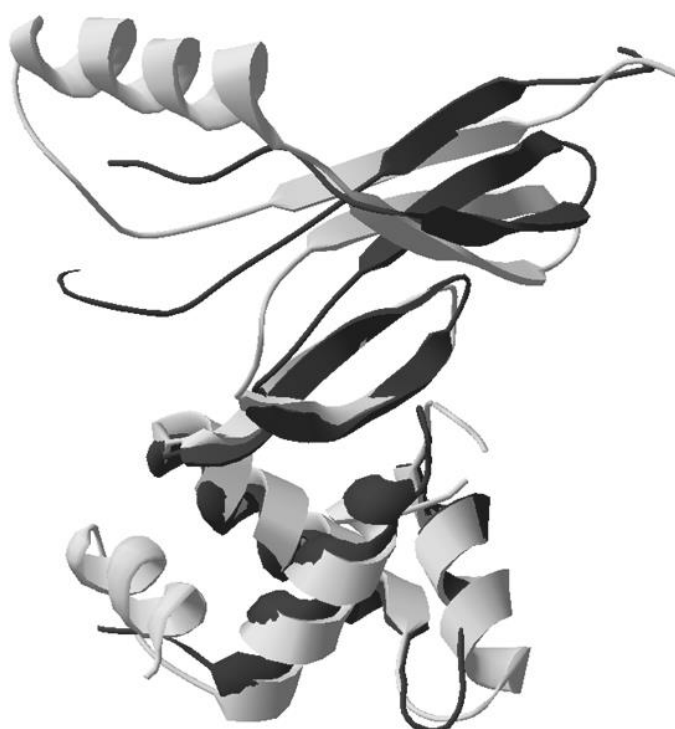


Figure 8.6 Superposition of SynZur (dark grey) and PaFur (light grey) monomers performed by the Swiss PDB viewer.

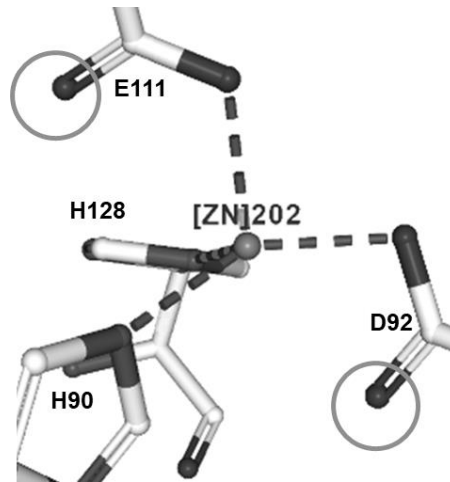


Figure 8.7 Site 3 in *Rhizobium leguminosarum* Mur.

This protein possesses a tetrahedrally coordinated zinc. Bonds are shown with dash lines. However, this could be a site with coordination number 6. Hence, two additional potential binding sites are highlighted with circles.

```
index=/media/alevtina/home/khu/8102index/8102index

for x in 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30;
do
    echo "Working on ${x}"
    cd /mnt/safe/Alko/Trimmed/Sample_${x}
    echo "Running Hisat2 ${x}"
    hisat2 -p 8 --dta -x "${index}" -1
    ${x}*R1_001.fastq.gz -2 ${x}*R2_001.fastq.gz -S
    "/mnt/safe/Alko/sam/${x}.sam" >
    "/mnt/safe/Alko/sam/${x}_log.txt" 2>&1
    echo "Running sam to bam ${x}"
    samtools view -S -b /mnt/safe/Alko/sam/${x}.sam >
    /mnt/safe/Alko/sam/${x}.bam
    echo "Sorting bam ${x}"
    samtools sort /mnt/safe/Alko/sam/${x}.bam -o
    /mnt/safe/Alko/sam/${x}.sorted.bam
    echo "Performing featureCounts ${x}"
    featureCounts -T 8 -t gene -g ID -a
    /mnt/safe/Alko/Syn_WH8102_2.gff -o
    /mnt/safe/Alko/sam/${x}.txt
    /mnt/safe/Alko/sam/${x}.sorted.bam
    echo "The work on ${x} is finished"
done
```

Figure 8.8 Bash-script used for RNA-seq data analysis.


```

library(DESeq2)
library(RColorBrewer)
library(ggplot2)
library(ggrepel)
library(dplyr)

sampleNames <- c('M01', 'M02', 'M03', 'W01', 'W02', 'W03')
filePath = '/media/alevtina/home/khu/for R/M0 over WT0.csv'
countData = read.table(file = filePath, header = TRUE, row.names = 1,
sep = ',')
condition <- c('mutant', 'mutant', 'mutant', 'wild type', 'wild type',
'wild type') #vector of column names for the data frame
colData <- data.frame(row.names=colnames(countData),
condition=factor(condition, levels=c('mutant','wild type')))
dataset <- DESeqDataSetFromMatrix(countData = countData,
colData = colData,
design = ~condition)

dds <- DESeq(dataset)
result <- results(dds, contrast=c('condition','mutant','wild type'))
result <- result[complete.cases(result),] #remove any rows with NA
head(result)

n = 50
resOrdered <- result[order(result$padj),]
topResults <- rbind( resOrdered[ resOrdered[,'log2FoldChange'] >
0, ][1:n,],
resOrdered[ resOrdered[,'log2FoldChange'] <
0, ][n:1,] )
topResults[c(1:10, (2*n-4):(2*n)),
c('baseMean','log2FoldChange','padj')]

plotMA(result, main='DESeq2: mutant vs. wild type', ylim=c(-2,2))
rld <- rlogTransformation(dds, blind=TRUE)
plotPCA(rld)
plotCounts(dds, gene=which.min(result$padj), intgroup='condition',
pch = 19)
hmcol <- brewer.pal(11,'RdBu')
nCounts <- counts(dds, normalized=TRUE)
heatmap(as.matrix(nCounts[ row.names(topResults), ]), Rowv = NA, col
= hmcol, mar = c(8,2))

par(mar=c(5,5,5,5), cex=1.0, cex.main=1.4, cex.axis=1.4, cex.lab=1.4)

#volcano plot
results=as.data.frame(resOrdered) # convert matrix in data frame
results$gene=row.names(resOrdered) #create a gene column
results
##Highlight genes that have an absolute fold change > 2 and a p-value
< Bonferroni cut-off
results$threshold = as.factor(abs(results$log2FoldChange ) > 2 &
results$pvalue < 0.05)

##Construct the plot object
ggplot(results, aes(x=log2FoldChange, y=-log10(pvalue),
colour=results$threshold)) +
theme_bw() +
ggtitle(label="Volcano plot of Deseq2 analysis for WT 772 nM Zn over
WT 2.5 uM Zn ") +
theme(plot.title = element_text(hjust = 0.5),
legend.position="none")+

```

```

#scale_x_continuous(breaks=seq(-4,6,2))+
geom_point(size=1.75) +
geom_hline(yintercept = -log(0.05, 10),color = "red",
linetype="dashed", size=0.2) +
geom_vline(xintercept = -2, color = "red", linetype="dashed",
size=0.2) +
geom_vline(xintercept = 2, color = "red", linetype="dashed",
size=0.2) +
xlab("log2 fold change") + ylab("-log10 p-value") +
geom_text_repel(data=filter(results, results$threshold==TRUE),
aes(label=gene), col="black")+
geom_text(data=data.frame( x = 2.5, y= 1.5), map=aes(x=x, y=y),
label="pval=0.05", col="black", size=3)
write.table(results, file = "/media/alevtina/home/khu/for
R/temporary.csv")

```

Figure 8.9 Example of the DESeq2 script in R software used for comparison of one condition containing three replicates with another condition containing three replicates (mutant at 0 zinc with WT at 0 Zn).

Table 8.2 RNA sample concentrations and purity assessment by NanoDrop.

Values above two for 260/280 nm and 260/230 nm indicates high purity of the samples.

| Sample ID | Nucleic Acid Conc. | Unit | 260/280 | 260/230 |
|-----------|--------------------|-----------------------|---------|---------|
| M01 (16) | 212.6 | ng μL^{-1} | 2.08 | 2.27 |
| M02 (17) | 87.6 | ng μL^{-1} | 2.09 | 2.29 |
| M03 (18) | 357.4 | ng μL^{-1} | 2.03 | 2.2 |
| M71 (19) | 238.6 | ng μL^{-1} | 2.08 | 2.15 |
| M72 (20) | 180.6 | ng μL^{-1} | 2.09 | 2.04 |
| M73 (21) | 322.9 | ng μL^{-1} | 2.06 | 2.2 |
| W01 (22) | 269.8 | ng μL^{-1} | 2.05 | 2.3 |
| W02 (23) | 61.1 | ng μL^{-1} | 2.08 | 2.23 |
| W03 (24) | 283.8 | ng μL^{-1} | 2.06 | 2.16 |
| W71 (25) | 54.8 | ng μL^{-1} | 1.95 | 1.98 |
| W72 (26) | 383 | ng μL^{-1} | 2.07 | 2.24 |
| W73 (27) | 60.7 | ng μL^{-1} | 1.98 | 2.03 |
| W21 (28) | 317.2 | ng μL^{-1} | 2.08 | 2.18 |
| W22 (29) | 251.6 | ng μL^{-1} | 2.05 | 2.02 |
| W23 (30) | 169 | ng μL^{-1} | 2.04 | 2.12 |

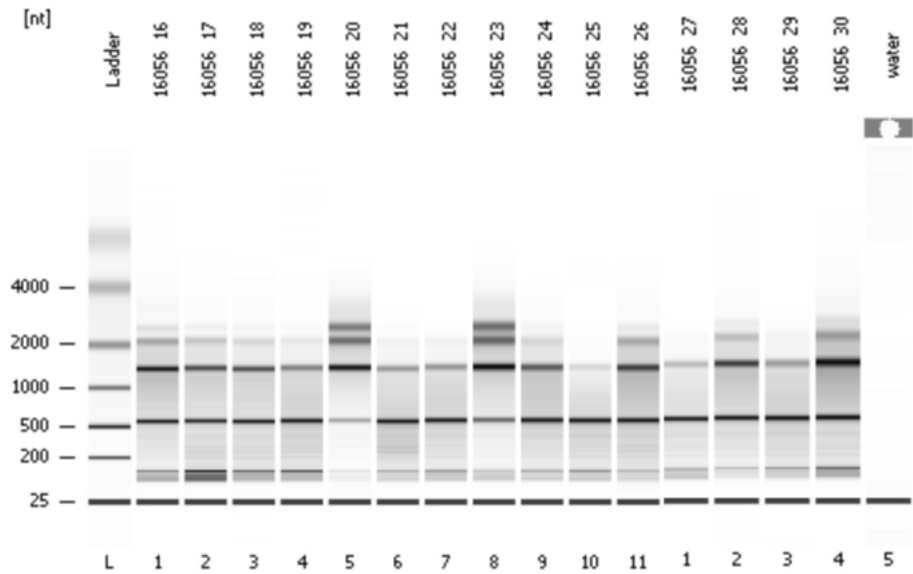


Figure 8.10 Agilent Bioanalyzer profiles of the RNA samples after Qiagen RNeasy columns.

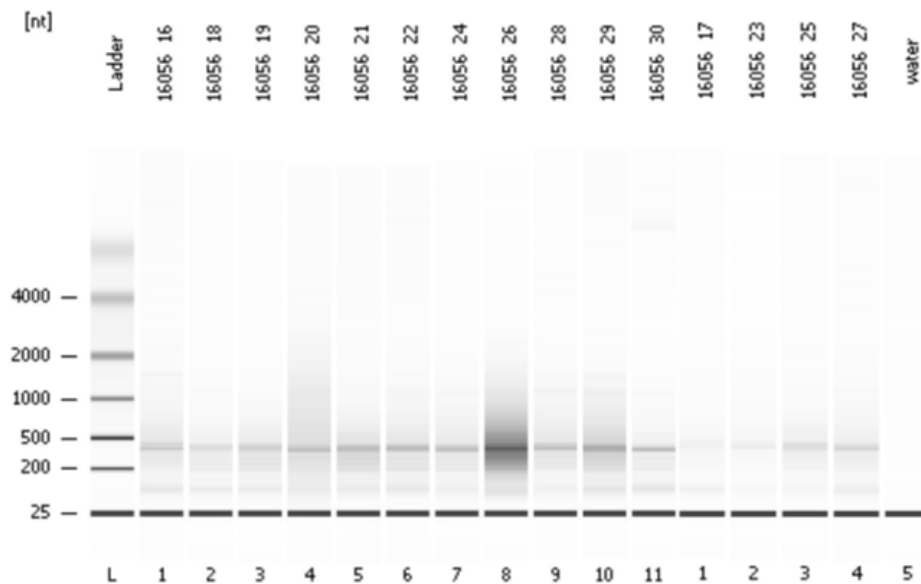


Figure 8.11 Agilent Bioanalyzer profiles of the RNA samples after using the RiboZero kit.

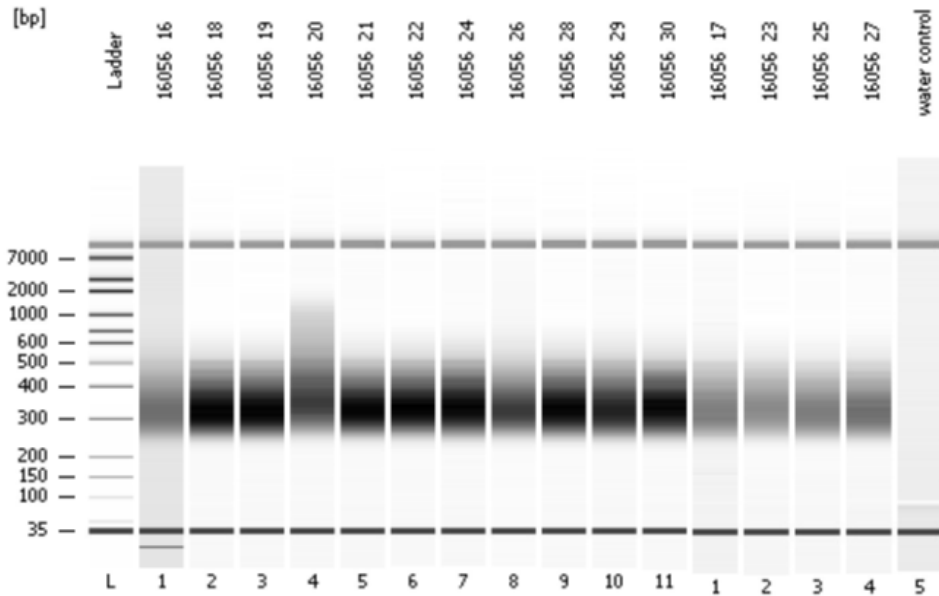


Figure 8.12 Agilent Bioanalyzer profiles of the RNaseq libraries

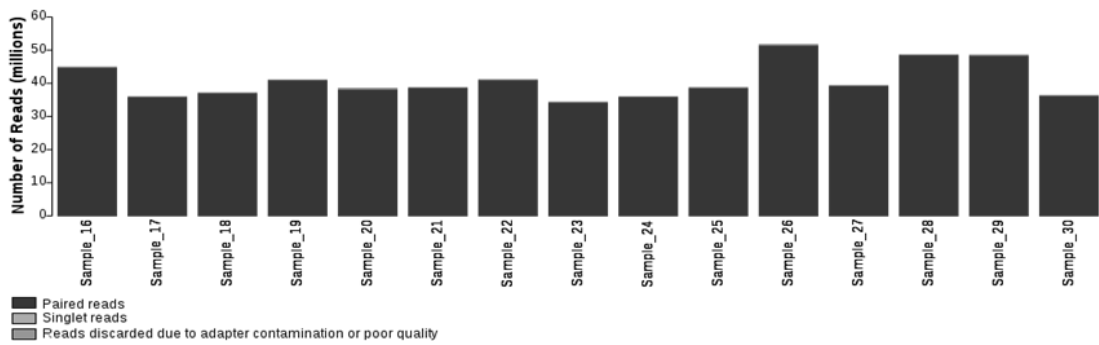


Figure 8.13 Post-processing statistics of RNA-sequencing data.

The figure is taken from http://cgr.liv.ac.uk/illum/LIMS16056_259a058713a41d9b/.

Table 8.3 RNA-Seq analysis statistics.

Fold coverage is calculated as the number of assigned paired reads, multiplied by the read length (150 bp) and divided by the genome size (2434428 bp).

| Sample | Repl- cate | Total number of paired reads | Overall alignment rate by Hisat2 | Aligned 1 time (Hisat2) | Assigned by Feature Count | Fold coverage |
|---------------|-----------------------|---|---|--|--|--------------------------|
| Mutant 0 Zn | 1 | 22167701 | 98.52% | 81.82% | 57.66% | 787.55 |
| Mutant 0 Zn | 2 | 17713760 | 97.59% | 89.40% | 62.00% | 676.70 |
| Mutant 0 Zn | 3 | 18310727 | 97.99% | 90.78% | 60.60% | 683.76 |
| Mutant 772 nM | 1 | 20242792 | 98.11% | 90.02% | 64.59% | 805.58 |
| Mutant 772 nM | 2 | 18872501 | 97.13% | 90.16% | 70.21% | 816.38 |
| Mutant 772 nM | 3 | 19113013 | 97.74% | 89.57% | 65.34% | 769.54 |
| WT 0 Zn | 1 | 20290010 | 98.60% | 87.09% | 62.60% | 782.67 |
| WT 0 Zn | 2 | 16902906 | 95.91% | 88.21% | 66.20% | 689.51 |
| WT 0 Zn | 3 | 17746878 | 98.71% | 87.28% | 63.48% | 694.14 |
| WT 772 nM Zn | 1 | 19085877 | 98.67% | 87.82% | 62.91% | 739.80 |
| WT 772 nM Zn | 2 | 25473625 | 73.19% | 69.83% | 61.15% | 959.73 |
| WT 772 nM Zn | 3 | 19418897 | 98.42% | 86.85% | 61.73% | 738.63 |
| WT 2.5 uM Zn | 1 | 24009262 | 98.91% | 89.52% | 66.24% | 979.93 |
| WT 2.5 uM Zn | 2 | 23942645 | 99.11% | 92.05% | 74.54% | 1099.71 |
| WT 2.5 uM Zn | 3 | 17919536 | 98.65% | 86.82% | 56.08% | 619.16 |

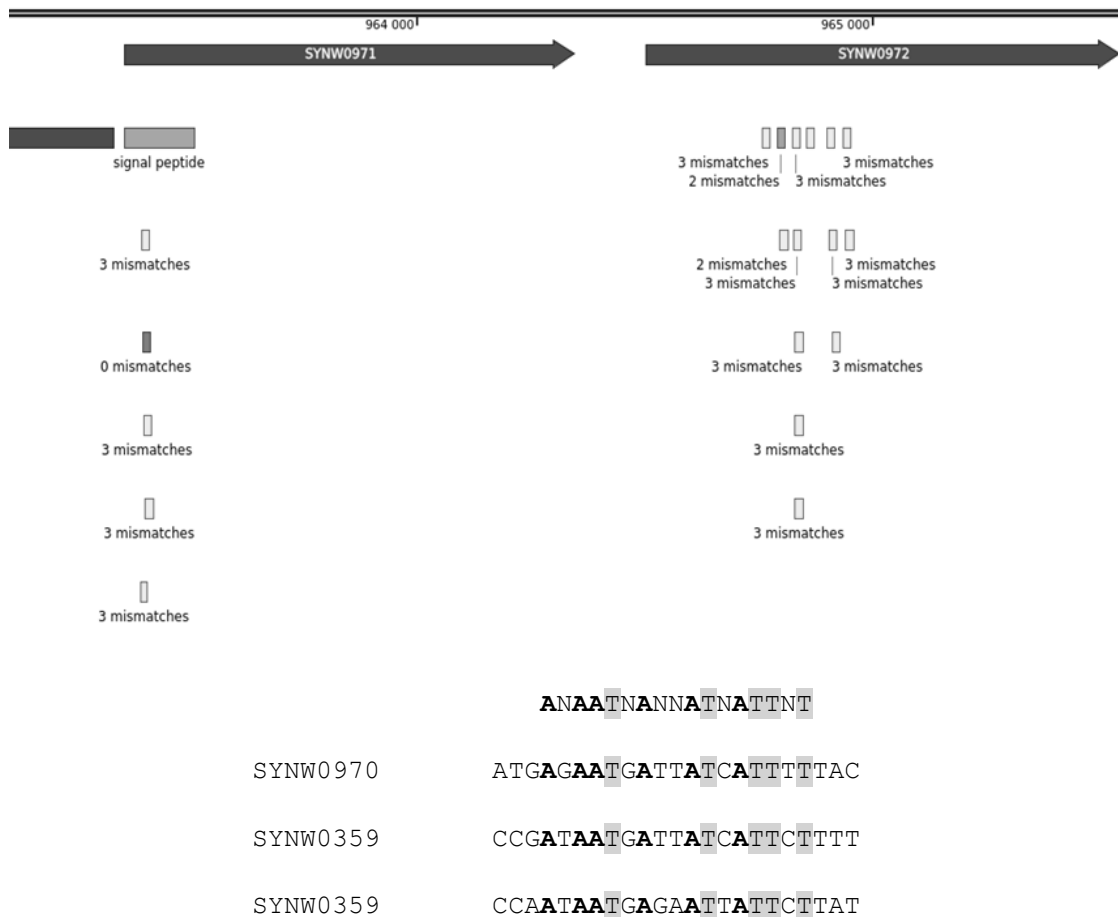


Figure 8.14 Identification of the predicted zur-boxes.

A. Predicted zur-boxes upstream of SYNW0971 (*znuA*) and in the coding region of SYNW0972.
B. A predicted known cyanobacterial Zur binding box was used to identify Zur-regulated genes in *Synechococcus* sp. WH8102 and its specific Zur binding box (ANAATGANNATNATTNT). Fully conserved A bases are shown in bold letters. Fully conserved T bases are highlighted with grey.

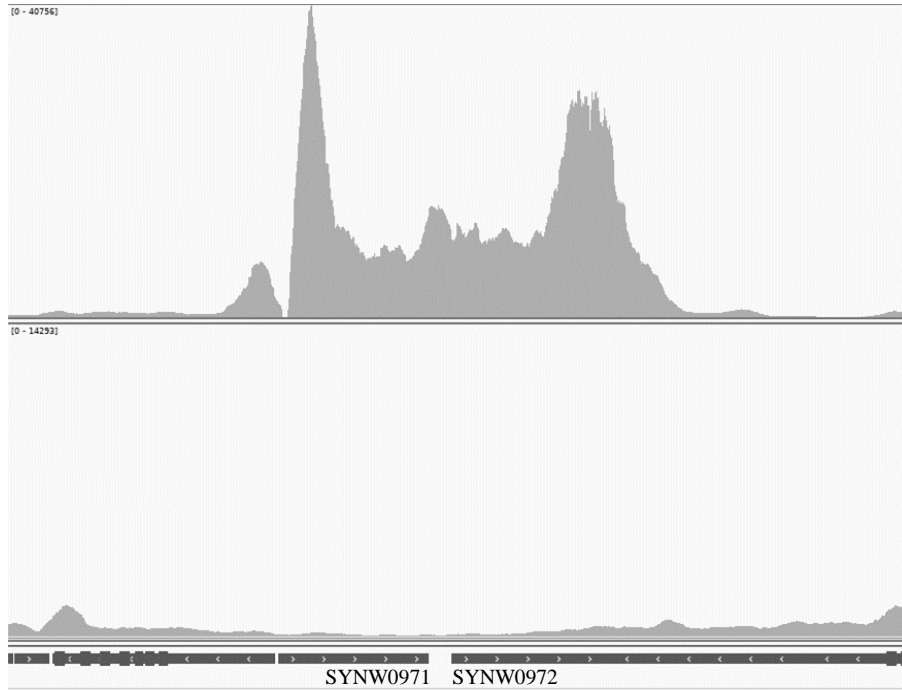


Figure 8.15 Transcript coverage of SYNW0971 and SYNW0972 in the *zur* mutant (top) and WT (bottom) checked by IGV software.

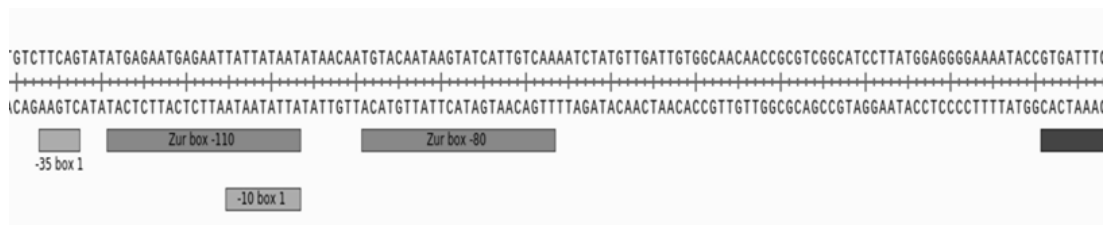


Figure 8.16 Location of zur-boxes and their position in the promoter elements of *Anabaena* sp. PCC 7120 *znuA* (*all0833*).

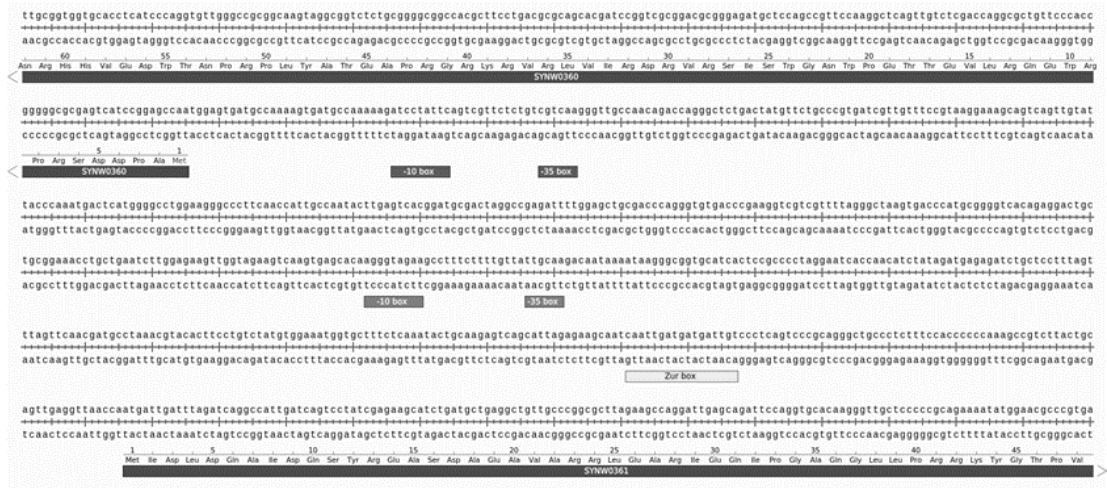


Figure 8.17 Location of a possible zur-box and predicted RNA-polymerase binding sites in the operator-promoter region of SYNW0360.

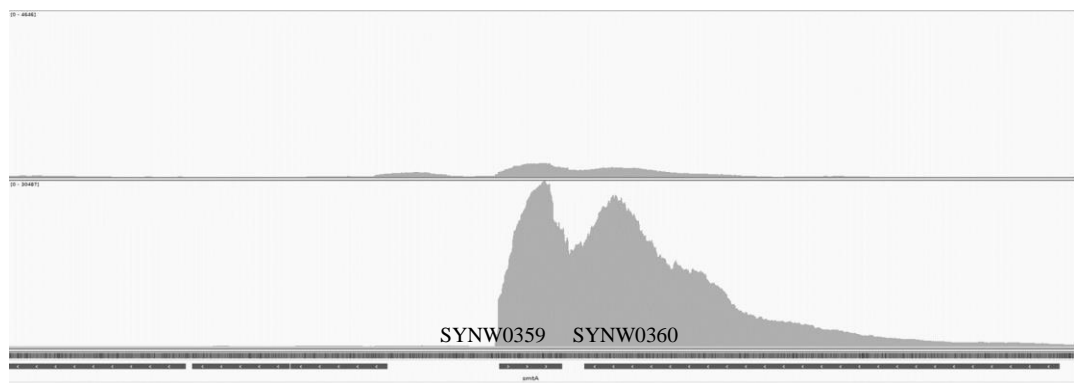


Figure 8.18 Transcript coverage of SYNW0359 and SYNW0360 in the mutant (top) and WT (bottom) checked by IGV software.

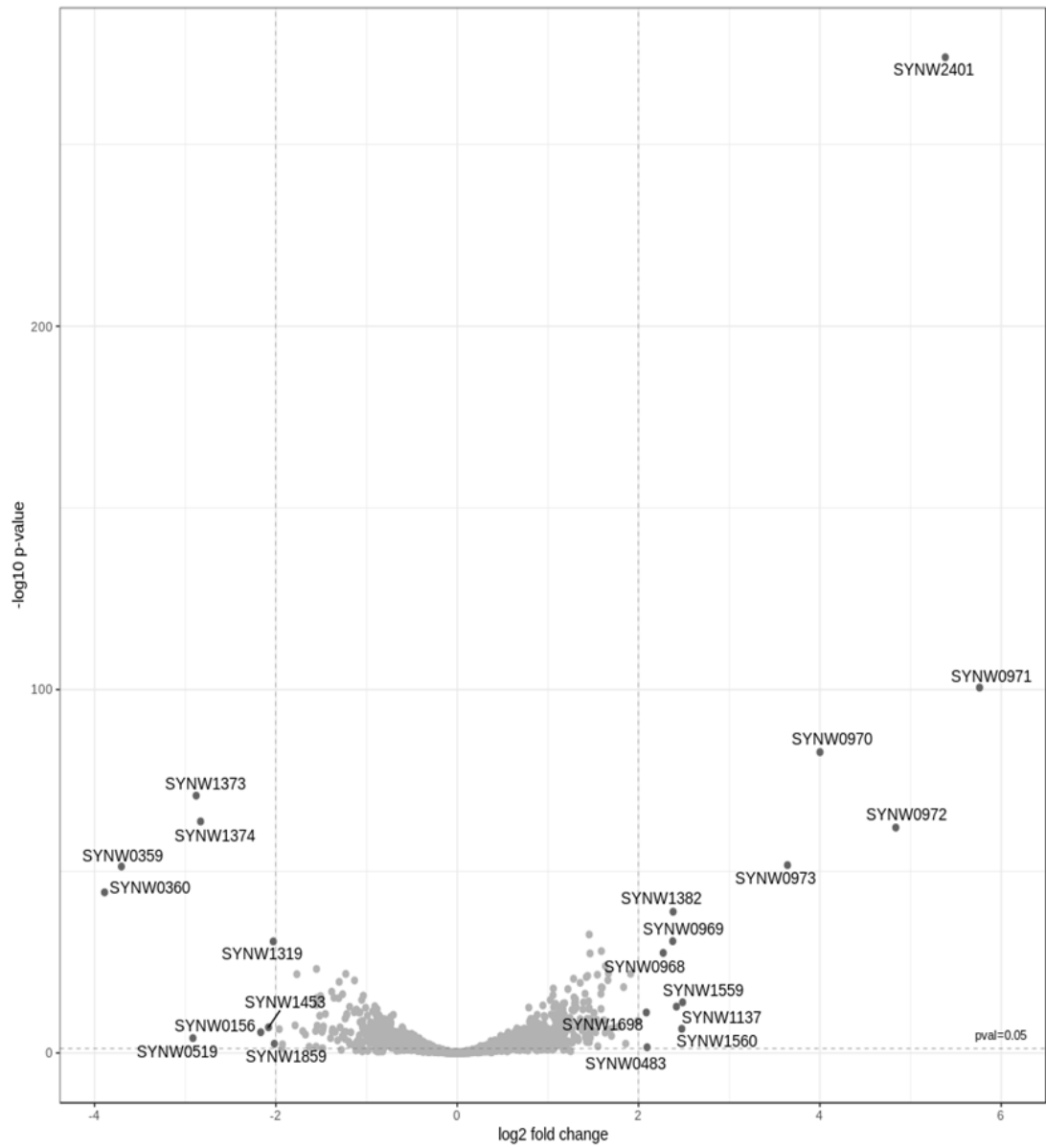


Figure 8.19 Genes differentially expressed in the *zur* knockout mutant of *Synechococcus* sp. WH8102 compared to the wild type at 0 added zinc.

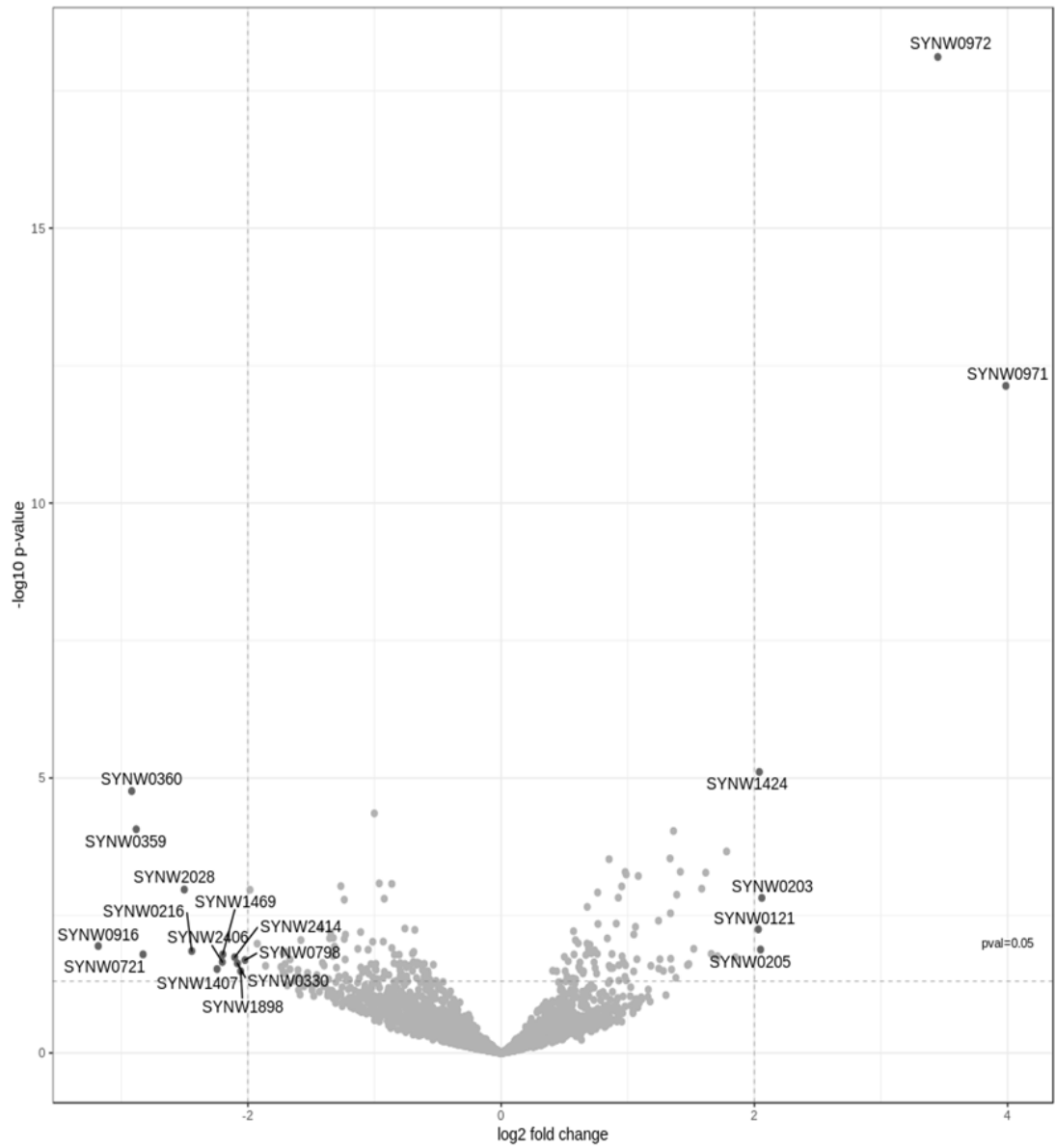


Figure 8.20 Genes differentially expressed in *Synechococcus* sp. WH8102 at 0 added zinc compared to 772 nM zinc.

Table 8.4 List of genes common to the different comparisons found from Venn diagrams.

| Down-regulated genes | | |
|--|--|--|
| Common for all three comparisons | Common for M0/WT0 and M772/WT772 | Common for WT0/WT772 and M772/WT772 |
| SYNW0360 SYNW0359 | SYNW0153 SYNW0298 SYNW0519 SYNW0600 SYNW0832 SYNW0876 SYNW0953 SYNW1070 SYNW1071 SYNW1281 SYNW1319 SYNW1374 SYNW1622 SYNW1682 SYNW1684 SYNW1826 SYNW2063 SYNW2324 | SYNW0095 SYNW0216 SYNW0330 SYNW0696 SYNW0774 SYNW0935 SYNW1469 SYNW1752 SYNW1804 SYNW1821 SYNW2043 SYNW2380 |
| Up-regulated genes | | |
| Common for all three comparisons | Common for M0/WT0 and M772/WT772 | Common for WT0/WT772 and M772/WT772 |
| SYNW0973 SYNW0970 SYNW0972 SYNW0971 | SYNW0145 SYNW0350 SYNW0879 SYNW0968 SYNW0969 SYNW1035 SYNW1103 SYNW1187 SYNW1188 SYNW1257 SYNW1368 SYNW1381 SYNW1382 SYNW1476 SYNW1569 SYNW1576 SYNW2176 SYNW2235 SYNW2400 SYNW2401 | SYNW0203 SYNW0750 SYNW0762 SYNW1417 SYNW1423 SYNW1424 SYNW1425 SYNW2068 SYNW2154 SYNW2205 SYNW2390 SYNW2427 |

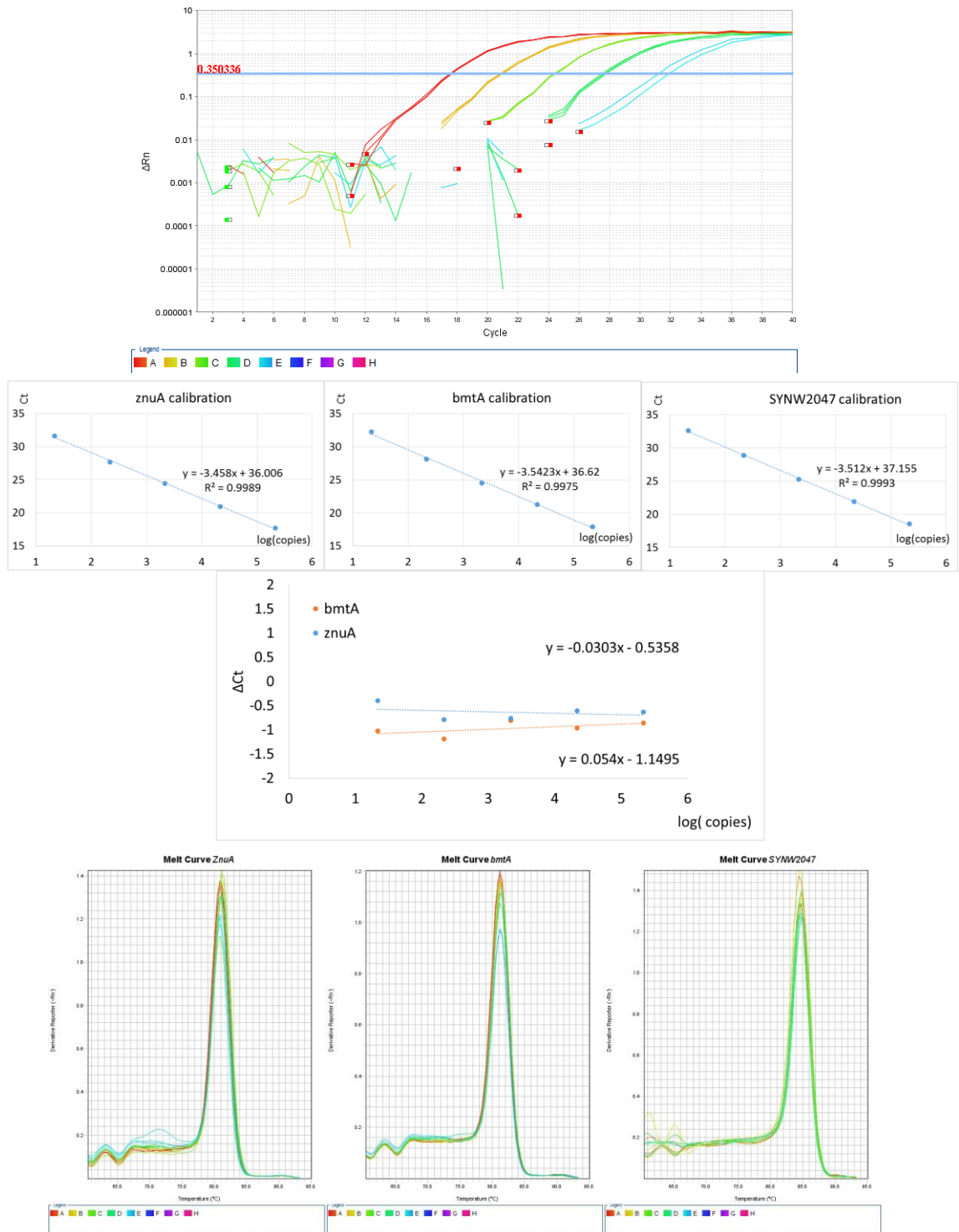


Figure 8.21 qPCR amplification using a known concentration of genomic DNA as a template and corresponding primers.

A. A typical qPCR amplification curve (*znuA*, primer concentration 400 nM, template 0.54 ng – 54 fg). B. Amplification efficiency curves. C. Graph of ΔCt against log(copies) showing relative amplification efficiency (normalised to SYNW2047). D. Melting curves showing formation of only one product.

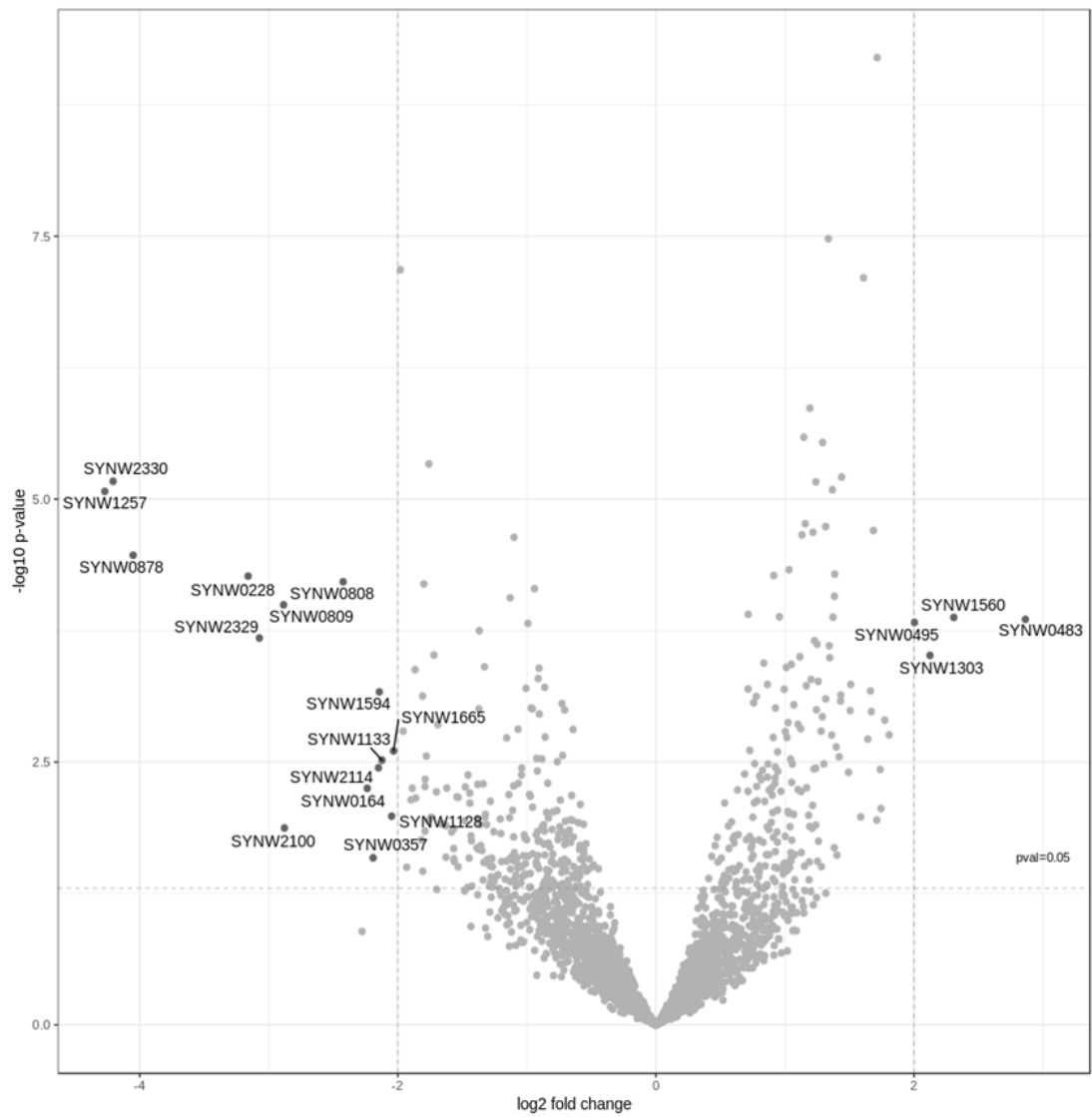


Figure 8.24 Genes differentially expressed in the *Synechococcus* sp. WH8102 *zur-* mutant at 0 added zinc compared to 772 nM zinc.

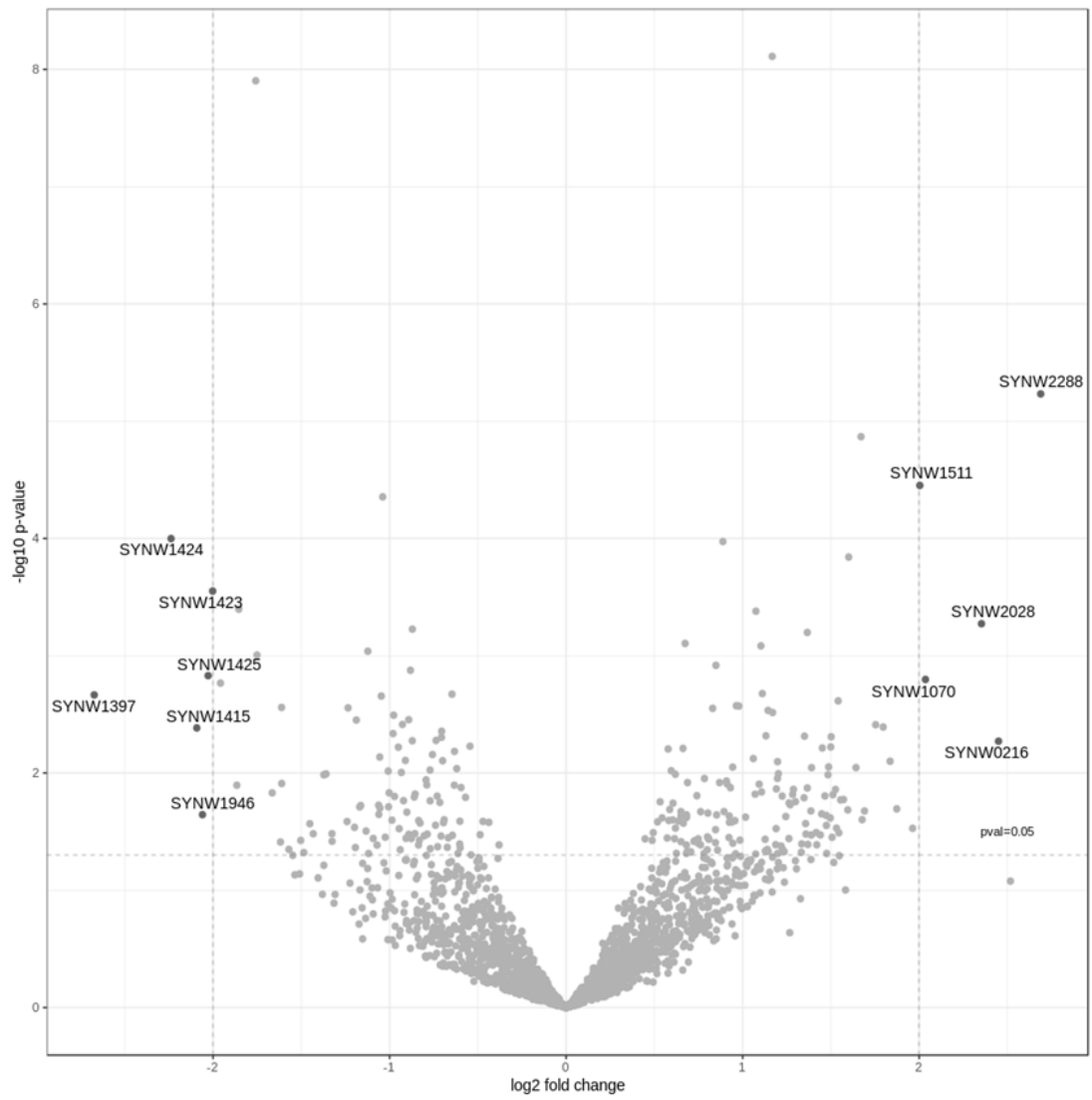


Figure 8.25 Genes differentially expressed in the *Synechococcus* sp. WH8102 *zur*-mutant at 772 nM zinc compared to 2.5 μ M zinc.

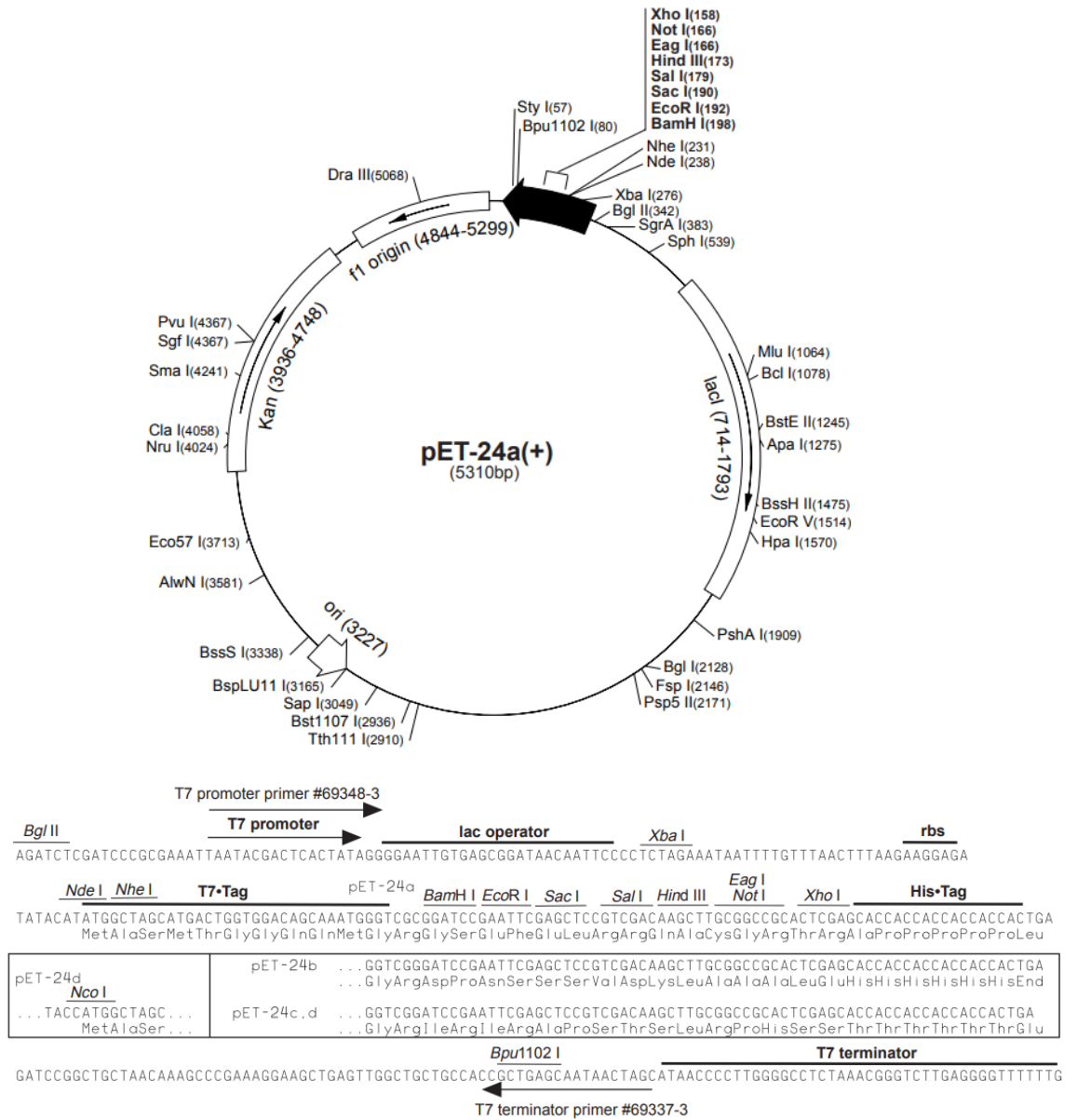


Figure 8.26 pET-24a(+) vector map.

The figure is taken from the Novagen brochure, TB070 12/98.

agag**gagctc**GTGTCTAGAGCAGTTGGGAGAGGTTGCAAGGATTGAATAAATCTAT
 GCCGATAATGATTATCATTCTTTTGTGCCAATAATGAGAATTATTCTTATACTTTA
 GGCCGCTTATTGGTCTAGTGTGTGGCTGAAAATTGCCTGG**tctaga**ctct

Figure 8.27 Sequence of *pbmtA* used in this study.

Zur-boxes are highlighted with grey and the part of the sequence belonging to the promoter is shown with capital letters.


```

-7.....GCGCCAT.....0
1..ATGTCTACAGCAATCAAATGTGCATGCCCTAAGTGCACCTGTATGGTCGC.....50
51..AGAAGAGAGCGCAATTGTTCTTCAAGGTAAGTTCTTTTGCTCGACGTCCT.....100
101.GTTCAACAGGGCATGCCAATAATGAGCCCTGTCATGGCGAAGGCTCTTGT.....150
151.GGATGTAAGTGCGGAGAGTAAGAGCTCGTCG.....181

```

Figure 8.28 Sequence of the SYNW0359 PCR product. *bmtA* is underlined.

Primers are shown in bold (for the reverse primer the reverse-complement sequence is shown). The restricted sites are highlighted in grey.

Table 8.5 DynaFit scripts used for calculation of SynZur-*pbmtA* K_d.

| Model 1: Zur ₄ + <i>pbmtA</i> | Model 2: 4Zur+ <i>pbmtA</i> | Model 3: 2Zur ₂ + <i>pbmtA</i> |
|---|--|---|
| <pre> [task] task = fit data = equilibria [mechanism] Zur4 + DNA <==> Zur4DNA : Kd dissoc [constants] ; units: nM Kd = 6.4 ? [concentrations] ; units: nM DNA = 4.96 [responses] ; units: DNA %/ DNA nM DNA = 20.16 [data] variable Zur4 set titration [set:titration] Zur4 total, nM DNA,% 0 100 1.25 100 2.5 93.795 12.5 7.236 25 0 37.5 0 50 0 125 0 250 0 0 100 0.25 100 0.5 100 1.5 100 2 100 2.5 100 3 97.255 4 91.887 5 90.578 10 37.265 25 0 [output] directory ./Zur/output/ [end] </pre> | <pre> [task] task = fit data = equilibria [mechanism] Zur + Zur + Zur + Zur + DNA <==> ZurZurZurZurDNA : Kd dissoc [constants] ; units: 10-36 M or nM4 Kd = 3.95e+006 ? [concentrations] ; units: nM DNA = 4.96 [responses] ; units: DNA %/ DNA nM DNA = 20.16 [data] variable Zur set titration [set:titration] Zur total, nM DNA,% 0 100 5 100 10 93.795 50 7.236 100 0 150 0 200 0 500 0 1000 0 0 100 1 100 2 100 6 100 8 100 10 100 12 97.255 16 91.887 20 90.578 40 37.265 100 0 [output] directory ./Zur/output/ [end] </pre> | <pre> [task] task = fit data = equilibria [mechanism] ZurZur + ZurZur + DNA <==> ZurZurZurZurDNA : Kd dissoc [constants] ; units: 10-18 M or nM2 Kd = 162 ? [concentrations] ; units: nM DNA = 4.96 [responses] ; units: 100 %/DNA nM DNA = 20.16 [data] variable ZurZur set titration [set:titration] ZurZur total, nM DNA,% 0 100 2.5 100 5 93.795 25 7.236 50 0 75 0 100 0 250 0 500 0 0 100 0.5 100 1 100 3 100 4 100 5 100 6 97.255 8 91.887 10 90.578 20 37.265 50 0 [output] directory ./Zur/output/ [end] </pre> |

```

[task]
    task = fit
    data = equilibria
[mechanism]
    Zn + Quin <==> ZnQuin : Kd1 dissociation
    Zn + BmtA <==> ZnBmtA : Kd2 dissociation
[constants] ; units: uM, %
    Kd1 = .0000037
    Kd2 = 1.03e-8 ?
[concentrations] ; units: uM
    Quin = 33.67
    BmtA = 49.59
[responses] ; units: % Quin/uM Quin
    Quin = 2.97
[data]
    variable Zn
    offset auto ?
    set titration
[set:titration]
    Zn, uM      Quin, %
    0           100
    5.496183206 100.2527168
    10.90909091 99.45523269
    16.2406015  100.255012
    21.49253731 99.92201391
    26.66666667 99.94453659
    31.76470588 99.89809767
    36.78832117 98.84210472
    41.73913043 98.6202962
    46.61870504 95.87221435
    51.42857143 88.89948527
    56.17021277 80.52214412
    60.84507042 68.56152325
    65.45454545 56.85180696
    70           42.98986183
    74.48275862 26.20754093
    78.90410959 10.89529116
    83.26530612 0
    87.56756757 0.216993156
    91.81208054 1.454962928
    96           2.91837529
[output]
    directory ./Zur/output/
[end]

```

Figure 8.29 DynaFit script used for calculation of SynBmtA-Zn K_D .

$Kd1$: The Quin-2-Zn dissociation constant, μM . $Kd2$: The SynBmtA-Zn dissociation constant, μM . This value is fitted by the software to match the experimental curve (parameters marked by a question mark are being fitted). Quin-2: The total Quin-2 concentration, μM ; BmtA: The total SynBmtA concentration, μM ; R: The Quin-2 response coefficient = $100\%/\text{Quin-2}$, μM ; The Free Quin-2 fraction calculated at 261 nm spectrophotometrically.

Table 8.6 Side chain signals of Asn, Gln and un-assigned residues of SynBmtA.

| Residue | N | H_N | H_α | H_β | H_γ |
|----------------|----------|----------------------|----------------------|----------------------|----------------------|
| GlyI | 108.7 | 8.28 | 4.352, 4.168 | | |
| GlyII | 110.7 | 7.803 | 4.214, 3.668 | | |
| GlyIII | 109.8 | 8.224 | 3.249, 2.936 | | |
| GlyIV | 108.2 | 7.909 | 4.546, 3.562 | | |
| Gly/Ser1 | 120.1 | 8.095 | 3.985 | | |
| Type?1 | 132.6 | 7.289 | | | |
| TypeA/T1 | 129 | 7.968 | 4.126 | 1.253 | |
| TypeJ1 | 120.6 | 5.47 | 4.182 | 3.775 | |
| H2O | | 4.762 | | | |
| Type?2 | 132.5 | 12.77 | | | |
| TypeA/T2 | 129.2 | 7.895 | 4.055 | 1.266 | |
| TypeJ2 | 122.4 | 8.369 | 4.462 | 3.667 | |
| Type?3 | 115.1 | 7.881 | 4.268 | - | |
| TypeJ3 | 122.7 | 8.917 | 3.696 | 2.782 | |
| TypeU3 | 122.6 | 7.848 | 4.105 | 2.279 | 1.88 |
| Type?4 | 122.7 | 7.268 | 1.297 | - | |
| TypeJ4 | 125.9 | 8.851 | 4.365 | 3.325 | |
| TypeU4 | 122.9 | 7.942 | 4.119 | 2.247 | 1.928 |
| TypeJ5 | 121.7 | 9.037 | 3.962 | - | |
| TypeU6 | 125.4 | 8.384 | 4.319 | - | |
| TypeJ7 | 122.8 | 8.743 | 4.427 | 3.584 | |
| TypeU7 | 130.8 | 7.899 | 4.105 | 1.715 | 1.399 |
| TypeJ8 | 114.9 | 8.343 | 4.476 | 4.022 | |
| TypeJ9 | 116.1 | 9.162 | 2.864 | 2.478 | |
| TypeU9 | 125.6 | 7.841 | 4.072 | 1.714 | 1.349 |
| TypeU10 | 131.1 | 8.777 | 4.438 | 2.117 | 1.529 |
| TypeU11 | 124 | 7.188 | 3.939 | 2.144 | 1.879 |
| TypeU12 | 114.6 | 8.08 | | | 1.569 |
| Gln24sc1 | 111.7 | 7.456 | | | |
| Gln24sc2 | 111.6 | 6.82 | | | |
| Asn39sc2 | 111.3 | 6.745 | | | |
| Asn39sc1 | 111.2 | 7.685 | | | |
| Asn40sc1 | 114 | 7.589 | | | |
| Asn40sc2 | 114.1 | 6.867 | | | |