UNIVERSITY OF SOUTHAMPTON

Faculty of Natural and Environmental Sciences
Ocean and Earth Sciences

THE ENVIRONMENTAL, ELEMENTAL AND PROTEOMIC PLASTICITY OF *TRICHODESMIUM* IN THE (SUB)-TROPICAL ATLANTIC

by

JOSEPH THOMAS SNOW

Thesis for the degree of Doctor of Philosophy

May 2014

Joseph Thomas Snow: *The environmental, elemental and proteomic plasticity of Trichodesmium in the (sub)-tropical Atlantic,*
Dr C. M. Moore
Dr T. S. Bibby, Prof. E. Achterberg, Dr C. Mahaffey

This thesis is dedicated to my parents, thank you for insisting I pay attention to the world around me.

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

Faculty of Natural and Environmental Sciences
Ocean and Earth Sciences

Doctor of Philosophy

THE ENVIRONMENTAL, ELEMENTAL AND PROTEOMIC PLASTICITY OF *TRICHODESMIUM* IN THE (SUB)-TROPICAL ATLANTIC

by

JOSEPH THOMAS SNOW

Phytoplankton dominate surface ocean biomass and have major roles in global carbon and nutrient cycling. Of these, the cyanobacterial *Trichodesmium* sp. are considered the dominant and most conspicuous nitrogen fixers in the tropical and subtropical oceans, regimes frequently characterised by low concentrations of fixed nitrogen.

Despite being arguably the best studied marine diazotroph, the factors controlling the distribution and growth of *Trichodesmium* remain a subject of debate, with sea surface temperature, the partial pressure of CO_2 and nutrients including iron and phosphorus, all suggested to be important. Dominant controls on the biogeography of other important marine diazotrophic organisms are even less clear. Synthesising data from 7 research cruises collectively spanning large temporal and spatial scales across the Atlantic Ocean, including 2 previously unreported studies crossing the largely under-sampled South Atlantic gyre, we assessed the relationship between proposed environmental drivers and both community nitrogen fixation rates and the distribution of *Trichodesmium*. Simple linear regression analysis of the combined data set would suggest no relationship between any of the sampled environmental variables and nitrogen fixation rates. However, considering the concentrations of iron and phosphorus together within a resource-ratio framework indicates the combined effects these nutrients have on *Trichodesmium* and broader diazotroph biogeography. The resource ratio framework is argued to be consistent with both the previously described North-South Atlantic contrast in *Trichodesmium* abundance, and the presence and consequence of a substantial non-*Trichodesmium* diazotrophic community in the euphotic zone of the western South Atlantic subtropical gyre.

Using high-throughput chemical and biological analyses we were able to observe significant plasticity in *Trichodesmium*'s elemental composition. The macro- and micro- elemental composition of environmental *Trichodesmium* showed enrichments in arsenic, vanadium and

molybdenum along an environmental phosphorus gradient, which we attributed to phosphorus- stress induced accidental uptake. Stoichiometric comparison with bulk phytoplankton revealed enrichments in iron (alongside nickel, copper and zinc) to which we predict 12-37% is attributed to the process of nitrogen fixation.

To date, understanding *Trichodesmium*'s propensity for growth in iron-deplete oceanic regimes has focused on traditional metrics such as physiological rates and targeted molecular studies. Here a label-free quantitative proteomics technique (MS^E) was employed to examine the full complement of *Trichodesmium erythraeum* IMS101's proteome when grown with different iron availability. Iron stress resulted in increased abundances in proteins involved in iron-stress acclimation and of proteins involved in iron-uptake. Also a systematic decrease in the iron-binding proteins involved in photosynthesis and nitrogen fixation is reported. Such changes reveal potentially novel iron uptake pathways but also that *Trichodesmium* reallocates resource away from nitrogen fixation and towards components of the photosynthetic apparatus under iron stress.

Finally, utilising a bioinformatic approach we are able to generate a predicted metallo-proteome for *Trichodesmium*, detailing the relative protein-bound concentrations of iron, vanadium, arsenic, molybdenum, zinc, nickel, copper, manganese, cadmium and cobalt. By comparing this metallo-proteome to the observed intracellular metallome we were able to synthesise our findings into the description of discreet *Trichodesmium* phenotypes observed in the (sub)-tropical Atlantic. In doing so, we present a comprehensive, observation-based explanation of the interactions between *Trichodesmium* and iron, linking small scale physiology to basin-scale biogeochemical variability.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

AMRT Accurate mass retention time

AQUA Absolute Quantification Assay

ATP Adenosine triphosphate

BATS Bermuda Atlantic Time Series

BCA Bicinchoninic acid

CCM Carbon concentrating mechanism

Chl-a Chlorophyll A

CID Collision induced dissociation

DDA Data dependent analysis

DFe Dissolved Iron

DIA Data independent analysis

DIC Dissolved inorganic carbon

DIN Dissolved inorganic nitrogen

DIP Dissolved inorganic phosphorus

DNA Deoxyribonucleic acid

DON Dissolved organic nitrogen

DOP Dissolved organic phosphorus

DTT Dithiothrietol

EA-IRMS Elemental Analyser Isotope Ratio Mass Sepctrometer

ESI Electrospray ionisation

FIA Flow injection analysis

FIRe Fluoroescence Induction and Relaxation fluorometer

FRRf Fast Repetition Rate fluorometer

GF/F Glassfibre filters

HDPE High density polyethylene

HPHFe High phosphorus, high iron region

HPLC High performance liquid chromatography

HPLFe High phosphorus, low iron region

IAA Iodoacetamide

ICP-MS Inductively Coupled Plasma mass Spectrometer

ITCZ Intertropical convergence zone

LDPE Low density polyethylene

LDS Lithium dodecyl sulfate

LPHFe Low phosphorus, high iron region

LPLFe Low phosphorus, low iron region

LTQ Linear trap quadrupole mass analyser

MALDI Matrix assisted laser desorption ionisation

MRM Multiple reaction monitoring

MS Mass spectrometry

MUF Methylumbelliferyl

MUF-P Methylumbelliferyl phosphate

NADPH Nicotinamide adenine dinucleotide phosphate

NASG North Atlantic Subtropical Gyre

NCEP National Centers for Environmental Prediction

NiSOD Nickel superoxide dismutase

OaTOF Orthogonally accelerated time of flight

OEC Oxygen evolving complex

OTE Ocean Test Equipment

PG Phosphatidylglycerol

PIC Particulate inorganic carbon

P_{IC} Intracellular phosphorus

POC Particulate organic carbon

polyP Polyphosphate

PON Particulate organic nitrogen

POP Particulate organic phosphorus

PSB Protein solubilizing buffer

PSI Photosytem I

PSII Photosytem II

PSI Photosytem I

PTFE Polytetrafluoroethylene

Q-TOF Quadrupole Time of Flight mass analyser

QS Quorum sensing

RNA Ribonucleic acid

ROS Reactive oxygen species

SASG South Atlantic Subtropical Gyre

SDS Sodium dodecyl sulfate

SOD Superoxide dismutase

SQDG Sulfoquinovosyldiacylglycerol

SRM Single reaction monitoring

SRM Standard reference material

SSI Sea surface irradiance

SSS Sea surface salinity

SST Sea surface temperature

SXRF Synchrotron X-ray fluorescence

TA Total Alkalinity

TIC Total ion chromatogram

TM Trace-metal

TOF Time of flight

XIC Extracted ion chromatogram

DECLARATION

I ***Joseph Thomas Snow*** declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

The environmental, elemental and proteomic plasticity of *Trichodesmium* in the (sub)-tropical Atlantic

I confirm that -

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

   Schlosser, C., Klar, J., Wake, B. D., **Snow, J.T**, Honey, D. J., & Woodward, E. M. S. (2013). *Seasonal ITCZ migration dynamically controls the location of the (sub-)tropical Atlantic biogeochemical divide.* Proceedings of the National Academy of Sciences of the United States of America.

Joseph Thomas Snow,
Date : _____

ACKNOWLEDGEMENTS

There are of course many people and organisations to thank for their help with this research and the production of this thesis. I would like to thank my supervisors, Mark Moore, Tom Bibby, Eric Achterberg and Claire Mahaffey for their willingness to take me on as a student. The trust they placed in me is greatly appreciated and I hope I was able to fulfil your expectations. The constant encouragement, guidance and enthusiasm each of you provided throughout my PhD studies and beyond will not be forgotten. A further thank you must go to Claire Mahaffey and George Wolff at the University of Liverpool for introducing me to ocean sciences and encouraging me to pursue a PhD, I'm sure their kind words played no small part in my success in applying for this project.

My sincere thanks go to Richard Geider, Metodi and Gergana Metodieva and María Huete-Ortega at the University of Essex for starting me off on the road to quantitative proteomics. On more than one occasion Maria and I shared in the pain of proteomics-gone-wrong and although we never got the AQUA method up and running everyone in Essex were consistently welcoming and supportive. Without their collective guidance and support I would never have been able to continue my proteomic research in Southampton. Paul Skipp and Dom Burg stepped in at one of the darkest stages of my PhD, without their contribution my thesis would have had a very different outcome. I'm extremely grateful to the open-armed welcome they gave me.

In addition to those listed above, many people have helped in the collection of samples and data for this thesis along with invaluable discussion and advice. To attempt to list them all would be foolish however special acknowledgements must go to Elizabeth Sargent, Christian Schlosser, David Honey, Chris Marsay, Tommy Ryan-Koegh, Sophie Richier, Despo Polyviou, Andy Hitchcock, Andreas Johansson, Anna Macey, Nicola Pratt and Nathan Chrismas. A special thanks go to the participants of D361 and AMT21, scientific and non-scientific, together you made those cruises unforgettable.

I would like to thank my mum and dad who taught me the most important thing to do in life is whatever makes you the happiest (although there was a slight academic bias let's be honest). To Sally, Gabba and Fergus I accept your thanks for my role in entering the 'family business' so that you didn't have to.

Most importantly I thank Sophie, my best friend and loving partner, who has supported and encouraged me for *nearly* a decade. I have little doubt that without her unwavering support I would not

be where I am today. She has stuck by me throughout the ups and downs of my PhD even when she (nor I) fully understood what it was I was doing or why. Thank you for your patience and thank you for making my life considerably less boring.

Finally, I sincerely apologise if I have left anyone out from this section, know that your assistance was greatly appreciated and will not be forgotten. Thank you.

1 INTRODUCTION

1.1 A SYSTEMS PERSPECTIVE

Any attempt at understanding the bidirectional interactions of an organism and its environment insists on the utilisation of a system-wide approach. The recently re-defined field of systems biology attempts to describe the functional processes of biological systems utilising knowledge on their molecular constituents. This involves the integration of large amounts of different levels of 'omics' data where the mode, patterns and properties of interactions can be elicited (Banci and Bertini, 2013) (Figure 1). In this study I borrow key components and principles from the field of systems biology; I expand upon these ideas and apply them to *Trichodesmium* sp., both on a cellular level but additionally in the global context within which this cyanobacteria plays a key biogeochemical role. In doing so the breadth of the study must attempt to process the linkages between small scale physiology through to the fundamental and unequivocal linkages to basin-scale biogeochemical variability.

The Earth's current redox state is an emergent consequence of microbial life on a planetary scale. The first half of the Earth's biogeochemical history was fundamentally different to the present day, the advent of microbial organisms and the evolution of photosynthesis are thought to have altered the chemical speciation of virtually every element on the planets surface (Falkowski et al., 2008). Of the 92 naturally occurring elements around 30 are known to have biological functions; of these 30, six major 'macro-nutrients' - hydrogen (H), carbon (C), oxygen (O), sulphur (S), nitrogen (N) and phosphorus (P) comprise the majority of cellular biomass whilst the major ions Mg, K and Ca along with 'micro-nutrients' such as strontium (Sr), iron (Fe), manganese (Mn), nickel (Ni), zinc (Zn), copper (Cu), molybdenum (Mo), cadmium (Cd) and cobalt (Co) (amongst others) are all essential in trace amounts. To a general approximation the macro-nutrients assume both functional and structural roles within a cell whilst the micro-nutrients, particularly the transition metals mostly assume functional roles. Protein complexes carry out almost all of the biochemical, signalling and functional processes in the cell, and as such they have a wide range of shapes, sizes and functions. It has been estimated that between one quarter and one third of these proteins have a metal constituent (Waldron and Robinson, 2009; Dupont et al., 2006). Consequently any study of global biogeochemical cycles is intricately linked to the study of the metallo-proteomic toolbox

wielded by the microbial engineers that drive these cycles. As such, research on the biology of metals in cells and their effects on the marine biosphere falls at the interface between bioinorganic chemistry, cell biology and ocean biogeochemistry (Figure 1).

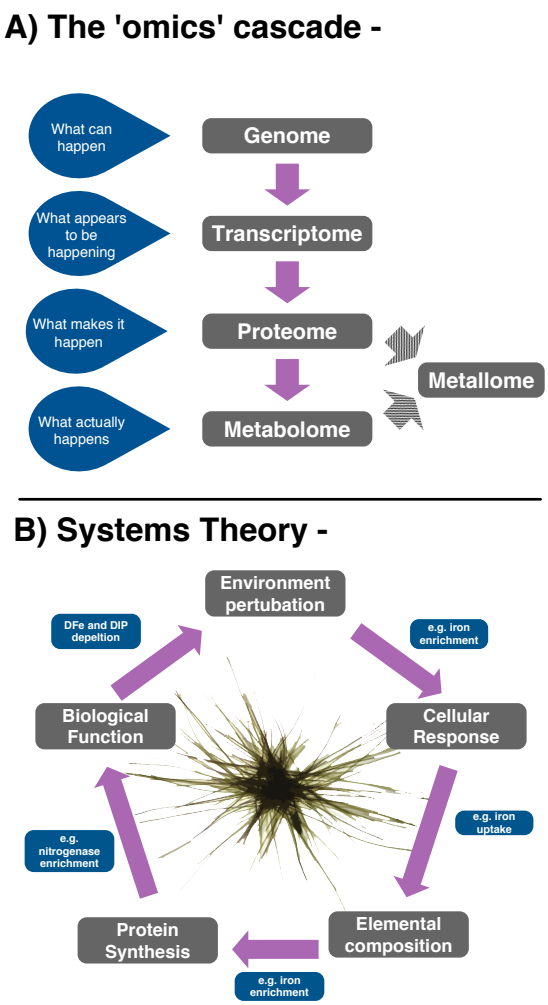

Figure 1: A) Diagrammatic representation of the 'omics' cascade, highlighting a few data types that contribute to the field of 'system's biology'. B) The system's perspective employed in this thesis, expanding the scope of 'systems biology' and applying these core concepts to the marine environment and global biogeochemical cycles.

1.2 MARINE PRIMARY PRODUCTION AND THE NITROGEN CYCLE

Marine primary producers are fundamental to the oceanic food chain, the overall biogeochemistry of the oceans and contribute 46.2% of the global primary production annually (Field, 1998). This biological con-

version of CO_2 to organic material by phytoplankton and its subsequent sinking to the deep oceans forms the basis of the 'biological carbon pump' (Longhurst and Harrison, 1989). Numerous parameters control the strength and efficiency of this organic matter export with microbial nitrogen fixation being a particularly important one as it contributes a supply of 'new' N (as opposed to 'regenerated' N from existing organic material) to the surface ocean (Dugdale and Goering, 1967; Eppley and Peterson, 1979) (Figure 2). Unlike 'regenerated' sources of N, which will be associated with an approximately stoichiometric equivalent CO_2 contributions, this diazotrophically derived new N along with any atmospherically deposited nitrogen or NO^3- enriched upwelling are the only source of N that can lead to a net sequestration of C (Figure 2) (Eppley and Peterson, 1979).

It is difficult to overstate the importance of marine nitrogen fixation to the global nitrogen cycle, and by extension the global biosphere and anthroposphere. There are two significant processes by which nitrogen fixation can occur, those being the Haber-Bosch processes and biological nitrogen fixation. Any secondary school chemistry student will attest the importance the Haber-Bosch processes plays in the industrial world, in ammonia and hence fertilizer production, and the significance of its discovery in 1908. Nitrogen fixing (diazotrophic) Cyanobacteria are responsible for more than double the nitrogen fixed per year as by the Haber-Bosch process (Galloway et al., 2004) and have been fixing nitrogen for the last 3.5 billion years (Falkowski, 1997; Boyd and Peters, 2013; Navarro-González et al., 2001). Achieved using a unique set of proteins, containing within them a unique combination of metal co-factors, diazotrophic cyanobacteria fix nitrogen at ambient temperatures and pressures. Of the marine diazotrophs *Trichodesmium* is the most well studied (Berman-Frank et al., 2001b; Carpenter et al., 2004; Capone et al., 2005; Van Mooy et al., 2012; Dyhrman et al., 2006). With *Trichodesmium* accounting for $\approx$ 40-80 % of annual 'new' nitrogen introduced to the Atlantic ocean (Capone et al., 2005), furthering our understanding of this enigmatic organism and its role in global biogeochemical cycling is of pivotal importance.

1.2.1 *Photosynthesis*

The key process driving marine primary production and in turn the biological carbon pump is that of photosynthesis. Oxygenic photosynthesis involves the conversion of sunlight into chemical energy through the production of oxygen and the assimilation of carbon dioxide into organic matter (Figure 3). The initial chemical reaction in photosynthesis involves the extraction of electrons from water; these electrons are then used to drive the photosynthetic electron-transport chain, located within the thylakoid membrane of photosynthetic organisms, and thus generate the reductant NADPH and the chemical

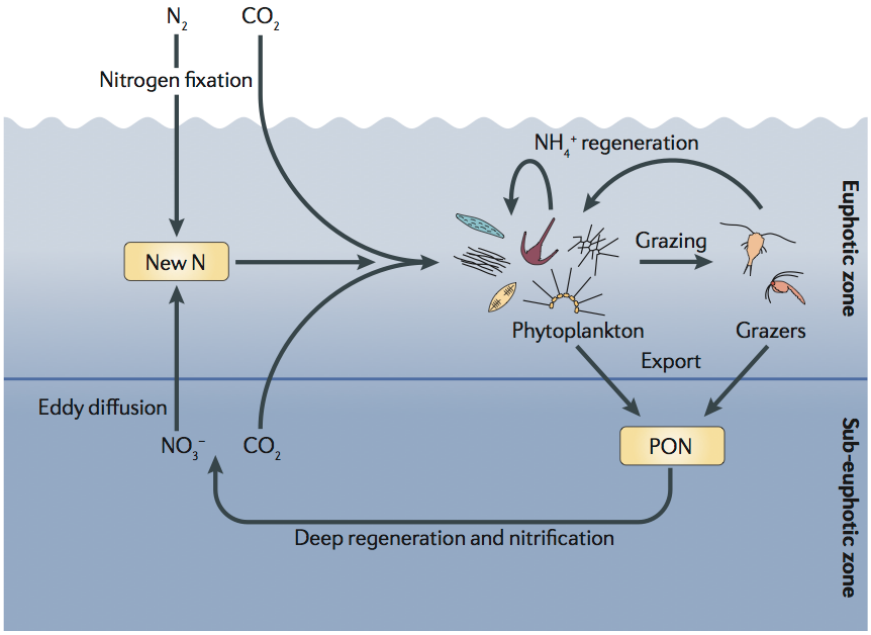

Figure 2: Overview of the importance of nitrogen fixation. Primary production fuelled by nitrogen fixation derived new N allows for a net sequestration of CO_2 whilst particulate organic nitrogen (PON) remineralisation at depth and its subsequent upwelling is a source of both N and C and therefore results in minimal net C sequestration (Sohm et al., 2011a).

energy carrier ATP. The photosynthetic electron-transport chain consists of two major, multi-subunit, protein complexes: photosystem II (PSII) responsible for light driven oxygen evolution and photosystem I (PSI) responsible for electron transfer from plastocyanin to ferredoxin and the subsequent generation of NADPH. Numerous metal-cofactors are utilised within the photosynthetic electron-transport chain including, amongst others, Fe, Mn, Cu and Zn (Yruela, 2013). Further details regarding the use of metal-cofactors in photosynthesis can be found in Section 1.7.1 whilst Chapter 4 and Chapter 6 discuss these metals and metalloproteins with regards to their environmental and biological distributions.

1.2.2 *Nitrogen Fixation*

Although N_2 is the most abundant gas in the Earth's atmosphere, its extreme lack of reactivity makes it inaccessible to all but a small number of organisms. N_2 fixation, performed by diazotrophs, is the biologically mediated process by which N_2 gas is reduced into two molecules of ammonia (Figure 3).

$$N_2 + 8H^+ + 16ATP + 8e^- \rightarrow 2NH_3 + H_2 + 16ADP + 16P_i \quad (1)$$

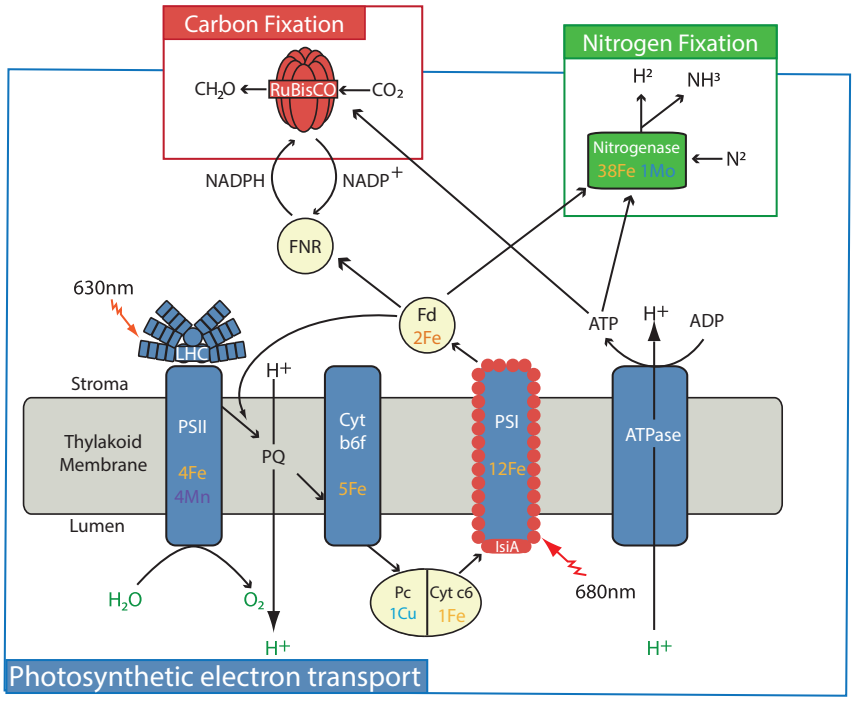

Figure 3: Simplified schematic model illustrating the flow of electrons through the principle components in photosynthesis, nitrogen fixation, respiration and carbon fixation in cyanobacteria. The metal composition of each complex is also shown. Abbreviations include: PSII - photosystem II, LHC - phycobilisome light harvesting complex, Cytb$_6$f - cytochrome b$_6$f complex, PSI - photosystem I, ATPase - ATP synthase, Pc - plastocyanin, Cyt c6 - cytochrome c6, PQ - plastoquinone, Fd - ferredoxin, FNR - ferredoxin:NADP$^+$ reductase, IsiA - iron stress induced protein, ATP - adenosine triphosphate, ADP - adenosine diphosphate, NADP$^+$/H - Nicotinamide adenine dinucleotide phosphate, RuBisCO - Ribulose-1,5-bisphosphate carboxylase/oxygenase.

N_2 fixation is an energetically costly process due to the strength of the N $\equiv$ N bond (941.4 kJ/mol) and as such the reaction is catalysed by the nitrogenase enzyme complex (Figure 4).

Three distinct nitrogenase metallotypes have been documented, MoFe-nitrogenase, VFe-nitrogenase and FeFe-nitrogenase (Eady, 2003). These metallotypes have unique protein subunits, each of which requires a different suite of metal-cofactors. The most ubiquitous of these metallotypes is the MoFe-nitrogenase. The MoFe nitrogenase enzyme system consists of three primary subunits, nitrogenase reductase (NifH) containing a single Fe_4S_4 cluster and the heterotetrameric dinitrogen reductase ($\alpha_2\beta_2$) containing an Fe_7MoS_7 (α-subunit, NifD) and Fe_8S_7 (β-subunit, NifK) (Rubio and Ludden, 2008). Biological nitrogen fixation, as achieved using one of the above

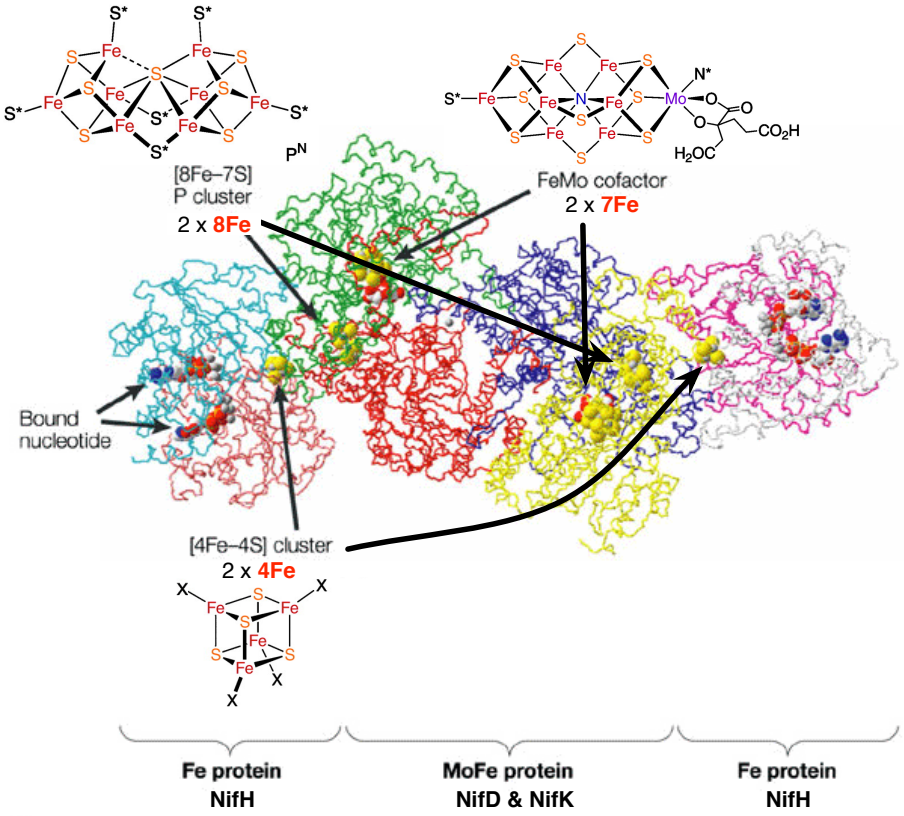

Figure 4: Structural representation of the complete MoFe-nitrogenase enzyme complex. Both Fe-protein and MoFe-protein are depicted with their associated Fe stoichiometries. Nitrogenase ribbon diagram modified from Dixon and Kahn (2004)

described metallo-enzyme complexes, is thought to be an ancient and perhaps even primordial process with estimates placing its evolution between ≈2.2-3.5 Ga (Falkowski, 1997; Navarro-González et al., 2001; Boyd and Peters, 2013). Little evidence exists as to the availability of ammonia or other fixed forms of nitrogen over the geological past however the geological imbalance in supply and demand for fixed nitrogen is thought to have exerted significant selective pressure favouring the emergence of the diazotrophs (Boyd and Peters, 2013). As a result of the kinetic isotope effect, biological nitrogen fixation preferentially introduces N with a $\delta PO^{15}N$ of ≈1‰ which is close to atmospheric N_2 (where $\delta^{15}N = ({}^{15}N/{}^{14}N)_{sample}/({}^{15}N/{}^{14}N)_{reference} - 1$ where the references is atmospheric N_2) (Montoya et al., 1996). Low $\delta^{15}N$ values from sediment cores have been used to describe the paleo-oceanographic history of nitrogen fixation versus denitrification (Straub et al., 2013) alongside serving as a tracer for the current day geographical N_{Fix} distribution.

1.3 DIAZOTROPHS

Of the extant marine diazotrophs the most conspicuous and well studied is *Trichodesmium* sp., a non-heterocystous, filamentous, cyanobacterium commonly found in 'puff 'or 'tuft' colonies (see Section 1.3.1). *Trichodesmium* alone are thought to contribute up to 50% of oceanic new nitrogen, with unicellular diazotrophs and diatom-diazotroph assemblages accounting for the remainder. Recently work has begun to re-evaluate the importance of unicellular diazotrophs such as *Crocosphaera watsonii*, a cyanobacterium exhibiting a unique iron thrift mechanism (Saito et al., 2011), UCYN-A, an often endosymbiotic cyanobacterium lacking PSII (Tripp et al., 2010; Zehr et al., 2008; Thompson et al., 2012), group B and the as yet unidentified UCYN-C (Foster et al., 2007).

Nitrogenase is irreversibly inhibited upon exposure to molecular oxygen. Cyanobacteria, being the only diazotrophs to actively produce oxygen as a by-product of photosynthesis have developed numerous evolutionary adaptations that protect nitrogenase from this molecular oxygen (Berman-Frank et al., 2003). One such oxygen-avoidance strategy is that of the *heterocystous* cyanobacteria. Heterocysts are specialised nitrogen-fixing cells containing an additional glycolipid cell wall that acts as a hydrophobic oxygen barrier and reduces the inward diffusion of oxygen. They are characterised by an increase in PSI activity which supplies ATP for nitrogen fixation but are free of PSII, the supercomplex responsible for molecular oxygen formation (Berman-Frank et al., 2003). This process of spatial segregation allows nitrogen fixation to occur during the photoperiod (Figure 5), however the heterocysts lose their ability to divide which is a costly trade-off (Adams, 2000; Berman-Frank et al., 2003).

Contrasting the spatial segregation performed by heterocystous cyanobacteria is a temporal oxygen-avoidance strategy performed by some unicellular and non-heterocystous cyanobacteria (e.g. *Crocosphaera watsonii*). Here photosynthesis is performed during the photoperiod whilst nitrogen fixation is conducted during the scotoperiod (Colón-López et al., 1997; Schneegurt and Tucker, 2000) (Figure 5). This allows for the nitrogenase enzyme systems to be present in all cells, however it relies on stored carbon accumulated during the day/-light period to fuel nitrogen fixation along with sufficient respiratory activity to minimise the cellular oxygen in order to reduce oxygenic inhibition (Berman-Frank et al., 2003).

Trichodesmium combat oxygenic inhibition of diazotrophy using a unique combination of both spatial and temporal segregation. Fine temporal control over both photosynthetic oxygen production and nitrogen fixation is observed during the photoperiod (Figure 5). This results in a period of lowered oxygen evolution around midday followed by an up-regulation of respiration and the Mehler reaction dur-

ing peak nitrogenase activity (Kana, 1993; Berman-Frank et al., 2001b; Küpper and Ferimazova, 2004; Milligan et al., 2007). This reduction in molecular oxygen formation allows for nitrogen fixation to occur uninhibited. *Trichodesmium* do not form heterocysts but they do sequester the majority of the trichomes nitrogenase enzyme complex into a small group of cells known as diazocytes (Berman-Frank et al., 2001b). Unlike heterocysts, diazocytes have no reinforced cell membrane and retain their ability to divide and return to non-fixing cells (Figure 6). These diazocytes are merely aggregations of nitrogenase rich cells that occur in regularly spaced groups along the trichome and their abundance varies through the photoperiod in line with the temporal cycle outlined above (Fredriksson and Bergman, 1995; Lin et al., 1998).

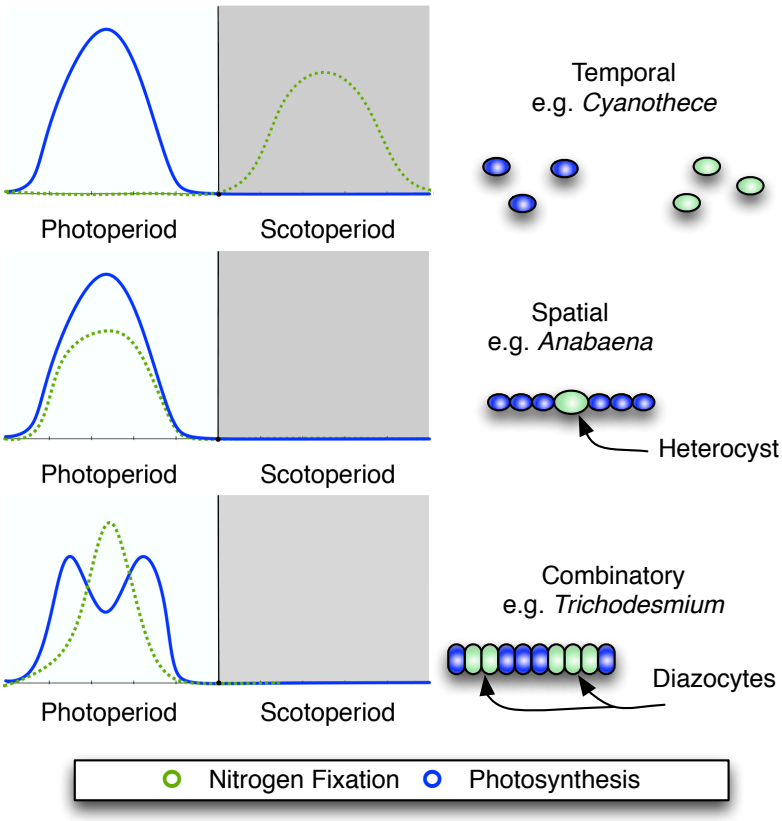

Figure 5: Differing spatial and temporal nitrogen fixation strategies employed by cyanobacteria as discussed in Berman-Frank et al. (2001b).

1.3.1 *Trichodesmium*

Trichodesmium sp. are a colonial, nitrogen fixing cyanobacteria responsible for 40-80% of the global new N input (Capone et al., 2005). *Tri-*

chodesmium is typically found in the oligotrophic tropical and (sub)-tropical oceans with an apparent preference for warm, well stratified water columns (Capone, 1997). Free floating *Trichodesmium* filaments, known as trichomes are frequently observed however the most readily identifiable characteristic of *Trichodesmium* are its two common colony morphotypes (Figure 6). These two morphologies are spherical 'puffs' where the trichomes are arranged in radially symmetrical patterns, and fusiform 'tufts', where the trichomes are arranged as parallel rafts often with a twisted orientation (Figure 6) (Janson et al., 1999; Hynes et al., 2012). It was once thought that these colony morphologies were a species specific trait but with six described species (T. *erythraeum*, T. *thiebautii*, T. *tenue*, T. *hildenbrandtii*, T. *contortum* and T. *spiralis*) each with overlapping morphological characteristics this practice has been superceded by genetic tools (Orcutt and Rasmussen, 2002; Lundgren and Janson, 2005; Hynes, 2009). The most well studied culture strain of *Trichodesmium* is that of *T. erythraeum* IMS101, originally isolated from approx 35°N, 76°W in 1991 (Prufert-Bebout et al., 1993); the genome for *T. erythraeum* IMS101 was published in 2003 (Joint Genome Institute, http://genome.jgi-psf.org/trier/trier.home.html), however by that time the strain had been in culture for 12 years. The genome of *Trichodesmium erythraeum* IMS101 is 7.21 Mbp large consisting of a single circular chromosome, of which only 60% is thought to code for the 4342 predicted proteins. Compared with the genomes of other marine microbes such as *Prochlorococcus, Synechoccocus or Crocosphaera watsonii* (Bench et al., 2013; Swan and Tupper, 2013), *Trichodesmium erythraeum* IMS101 has an unusually large genome and has a much larger fraction of non-coding DNA (Bergman et al., 2012; Larsson et al., 2011).

The benefits of *Trichodesmium*'s colonial growth are the subject of current debate with numerous competing and complimentary hypotheses being considered. This research is focussed both on *Trichodesmium* as a species but also *Trichodesmium* colonies as a consortia where the microbial communities growing within the *Trichodesmium* colonies are considered (Hewson et al., 2009; Rubin et al., 2011; Van Mooy et al., 2012). Rubin et al. (2011) has demonstrated *Trichodesmium* colonies actively acquiring dust particles and transporting them to the centre of their colonies. This is believed to be an elaborate iron uptake strategy whereby Fe-rich dust particles are transported to a potentially suboxic microenvironment near the colonies interior. Van Mooy et al. (2012) demonstrated a direct symbiotic relationship between *Trichodesmium* and its microbial consortia through the use of quorum sensing (QS), a cell-density dependent signalling system. It was shown that through QS *Trichodesmium* and its consortia were able to coordinate the acquisition of a not-readily bioavailable form of phosphorus. Recently work has been done to acquire the genome of an environmental *Trichodesmium* consortia isolated from the Bermuda

Atlantic time series (BATS) which will help in future studies of *Trichodesmium* consortia.

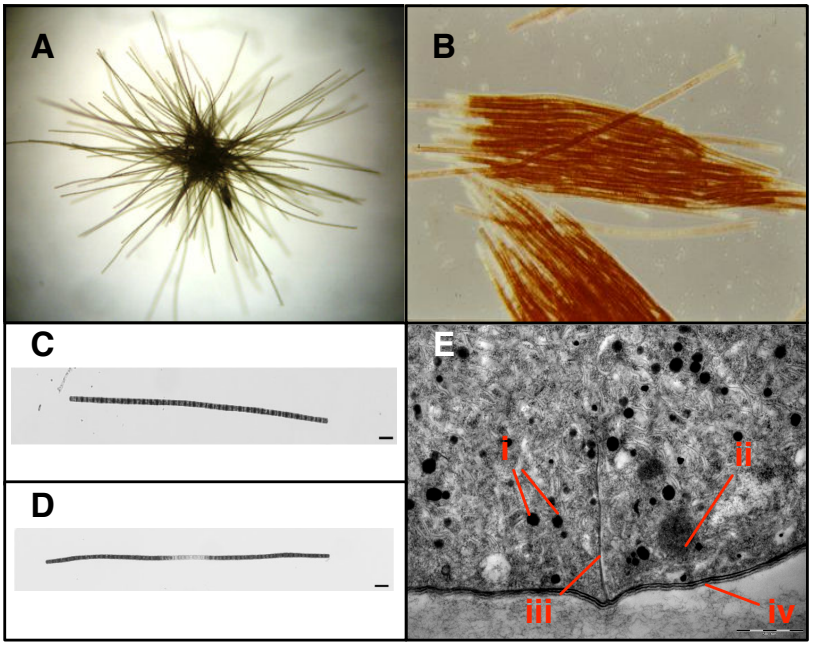

Figure 6: Microscopy images of *Trichodesmium* morphologies, trichomes and ultrastructure. A) a puff colony (image credit: Elizabeth Sargent), B) a tuft colony (image credit: Hans Paerl), C) a free trichome showing no diazocytes, D) a free trichome showing a lighter region indicative of diazocyte cells, scale bar = 20 µm (Images C & D: Sandh et al. (2011)) and finally E) a transmission electron micrograph of *Trichodesmium erythraeum* IMS101 showing various ultrastructure components - i) Cyanophycin granules, ii) carboxysomes, iii) intercellular space and iv) outer membrane, scale bar = 500 nm (modified from Sargent (2013)).

1.4 CONTROLS ON DIAZOTROPHIC PRODUCTION

The controls on diazotrophic production are the subject of much debate with numerous physical and chemical parameters being considered.

Arguably the use of *Trichodesmium* as a pseudo-model organism, representing marine diazotrophy as a whole, has skewed our understanding of the controls on diazotrophy. Physical parameters such as temperature and light have long been linked to nitrogen fixation (LaRoche and Breitbarth, 2005; Breitbarth et al., 2007) and it is undeniable that these parameters affect the growth and nitrogen fixing activity of diazotrophs, specifically *Trichodesmium*. However, as I argue in Chapter 3 for *Trichodesmium*, these controls are not obligatory

and are of little predicted importance in certain regions. Evidence of high latitude diazotrophy, at low temperatures and light levels, has been presented on numerous occasions (Staal et al., 2003; Stal, 2009; Pandey et al., 2004) and argues against the assumed diazotrophic requirement for high light and temperature. Moisander et al. (2010) support this concept by showing differential distributions of the diazotrophs *C.watsonii*, UCYN-A and *Trichdodesmium* at different latitudinal, depth and temperature ranges.

In a similar vein, the effects of elevated pCO_2 on diazotrophy and *Trichodesmium* growth has been well documented (Hutchins et al., 2007, 2013; Barcelos e Ramos et al., 2007). Experimental doubling of pCO_2 to projected end-of-century levels have been shown to increase CO_2 and N_2 fixation by 15-128% and 35-100% respectively (Hutchins et al., 2013). However similar studies on environmental *Trichodesmium* colonies have shown this affect to be inconsistent (Gradoville et al., 2014). Furthermore, current geographical variability in surface water pCO_2 is substantially less than these projected concentrations (Rödenbeck et al., 2013) as discussed further in Chapter 3.

Trichodesmium's cellular apparatus and unique photosynthetic/diazotrophic growth strategy (Berman-Frank et al., 2001b) result in the absolution for the need of a source of fixed nitrogen, but requires elevated Fe and has an absolute P requirement. The research regarding the nutrient controls of *Trichodesmium* have focused predominately on these two nutrients, Fe (Shi et al., 2007; Richier et al., 2012; Berman-Frank et al., 2007) and P (Orchard et al., 2009; Dyhrman et al., 2006). The meridional distribution of DFe and DIP in the sub-tropical Atlantic follows an antithetical gradient with orders of magnitude variability in the concentrations of these two nutrients (Chapter 3). This thesis will subsequently focus on the distribution of these nutrients along with *Trichodesmium*'s response to scarcity of either P or Fe.

1.5 CONCEPTS OF NUTRIENT LIMITATION

The foundations of our understanding on biological nutrient limitation stems from the early agrochemical work of Justus Von Liebig 1855 and A.C. Blackman 1905. Von Liebig described a mechanism by which nutrient limitation affects biomass yield whilst Blackman, building upon Liebig's work, described the rate of photosynthesis, and by extension growth rate, being constrained by a rate limiting nutrient (Box 1.5).

Box 1.5

Liebig's Law of the Minimum

"When a given piece of land contains a certain amount of all the mineral

constituents in equal quantity in an available form, it becomes barren for any one kind of plant when, by a series of crops, one only of these constituents-as for example soluble silica-has been removed, that the remaining quantity is no longer sufficient for a crop"

- Liebig 1855 as cited by de Baar (1994).

Blackman's Law of Limiting Factors
"When a process is conditioned as to its rapidity by a number of separate factors the rate of the process is limited by the pace of the slowest factor"

- Blackman (1905).

In Liebig's definition 'limitation' is a consequence of disparate cellular nutrient stoichiometry and its environmental counterpart. Alfred Redfield expanded this principle by examining the macronutrient stoichiometry within phytoplankton and found an empirical relationship between the intracellular stoichiometry (C_{106}:N_{16}:P_1) and that of the major nutrients in the oceans interior (*16* NO_3^-: *1* PO_4^{3-}) (Redfield, 1934, 1958). Termed the Redfield Ratio, this conserved ratio has acted as a vital diagnostic tool in biogeochemical oceanography since its inception, whereby deviations to this ratio allow for determination of the deficient nutrient.

1.5.1 *Stress and Deficiency*

There exists an important difference between the concepts of nutrient *stress* and nutrient *deficiency*. Here I define nutrient stress as a physiological response to the scarcity of a required nutrient (See Section 1.5.4) whilst nutrient deficiency or limitation describes a stoichiometric lack of one element relative to others. Here the Redfield ratio concept is a diagnostic tool of nutrient deficiency such that an N:P value of >16:1 would signify phosphorus *deficiency*.

1.5.2 *Co-limitation*

In addition to single-nutrient stress or deficiency, three types of multi-element co-limitation have been defined (Saito et al., 2008):

- Type I - Independent Nutrient co-limitation - describes limitation by two elements which are biochemically independent but are *both* found in the environment at sufficiently low concentration so as to be potentially limiting.
- Type II - Biochemical Substitution co-limitation - involves two elements that can substitute for the same biochemical role, be

it a direct elemental substitution within the same enzyme (cambialistic) or synthesis of an alternative enzyme capable of the same function but a lower limiting substrate requirement. A well studied example of these substitutions is for super oxide dismutases (SODs), an important antioxidant defence in cells exposed to reactive oxygen species (ROS). A variety of SODs exist with Ni, Mn, Fe or both Cu and Zn forming their active sites (see Wolfe-Simon et al. (2005) and references therein). Cambialistic substitution involving Mn and Fe has also been documented within a single SOD enzyme (Saito et al., 2011; Sugio et al., 2000).

- Type III - Biochemically dependent limitation - here the uptake of one substrate is dependent on the sufficient nutrition with regards to a second. For example, cyanobacterial nitrogen fixation being directly dependent upon sufficient availability of Fe and Mo, the metal co-factors associated with the nitrogenase enzyme system.

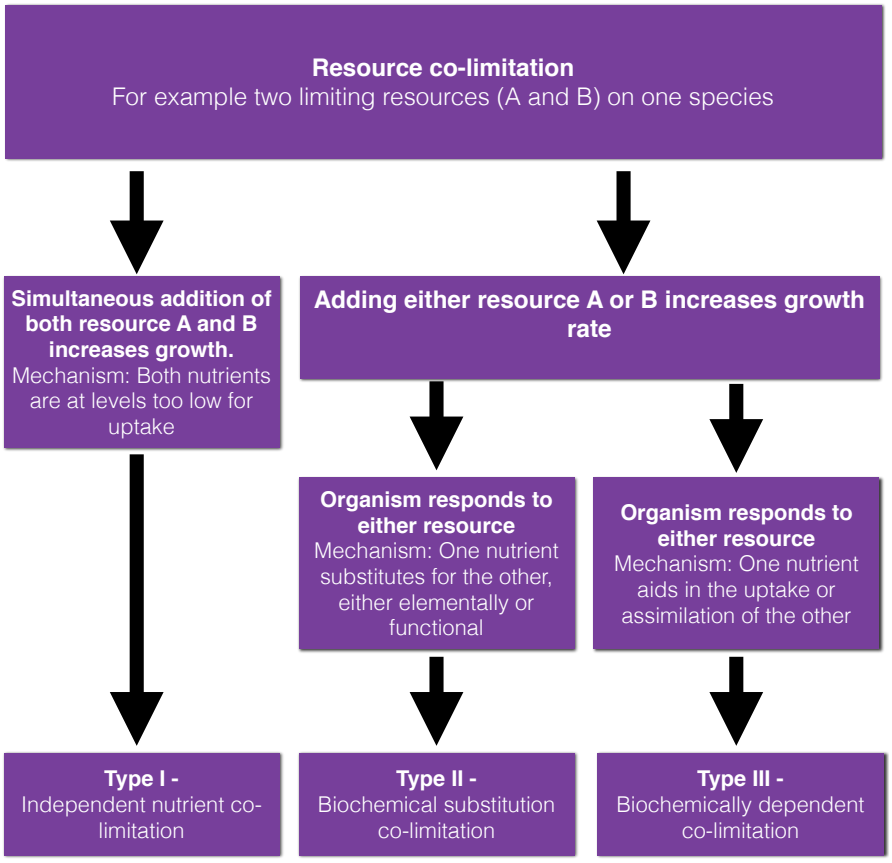

Figure 7: Flow chart for determining whether co-limitation experienced by a single organism is Type I, II or III.

These types of co-limitation can arise amongst mixed communities, whereby differing organisms have differing nutritional requirements (Type I) or within a single organism (Types I, II and III) (Figure 7).

1.5.3 *Resource Competition Theory*

Resource competition theory dictates that in an equilibrium environment the organism best able to utilise a given resource will deplete that resource down to a well defined minimum concentration, termed the R^* value (Tilman, 1980). In steady-state this framework supports a niche for as many species as there are competing resources, a state which could result in a type I co-limited community. The elevated Fe requirements and relatively low absolute growth rate of *Trichodesmium* make them extremely poor competitors for either Fe or P in regions of elevated fixed N (Figure 8A, B and C). However *Trichodesmium*'s diazotrophic absolution of a fixed N source, afforded to them by their elevated Fe and P requirements result in an effective $R^*_N = 0$. This allows *Trichodesmium* to avoid competitive exclusion in regions of fixed N oligotrophy and subsequently distribution can be described by the concentrations of Fe and P alone, assuming no other nutrient is limiting and other physical parameters (e.g. SST and pCO_2) are of secondary importance (Figure 8C and D).

The sub-tropical Atlantic is a region well documented as N deficient (Mills et al., 2004) where abundant populations of *Trichodesmium* have been observed (Carpenter et al., 2004; Capone et al., 2005; Moore et al., 2009a). Chapter 3 of this thesis aims to better explain the biogeography of *Trichodesmium* in the sub-tropical Atlantic, focusing on the concentrations of DFe and DIP, using resource competition theory as a basis.

1.5.4 *Response to Nutrient Scarcity*

The mechanisms by which any organism responds to nutrient stress or limitation can be grouped into 3 fundamental categories (Straus, 2004):

1. Acquisition - whereby the organism alters their nutrient acquisitional tactics in order to maximise its uptake of the limiting nutrient (this can include practices such as luxury uptake).

2. Compensation - involves a physiological change in an organism intended to reduce its requirement for the nutrient in question and increase its efficiency of use.

3. Retrenchment - finally if both acquisitional and compensatory techniques are still unable to supply the organism with suffi-

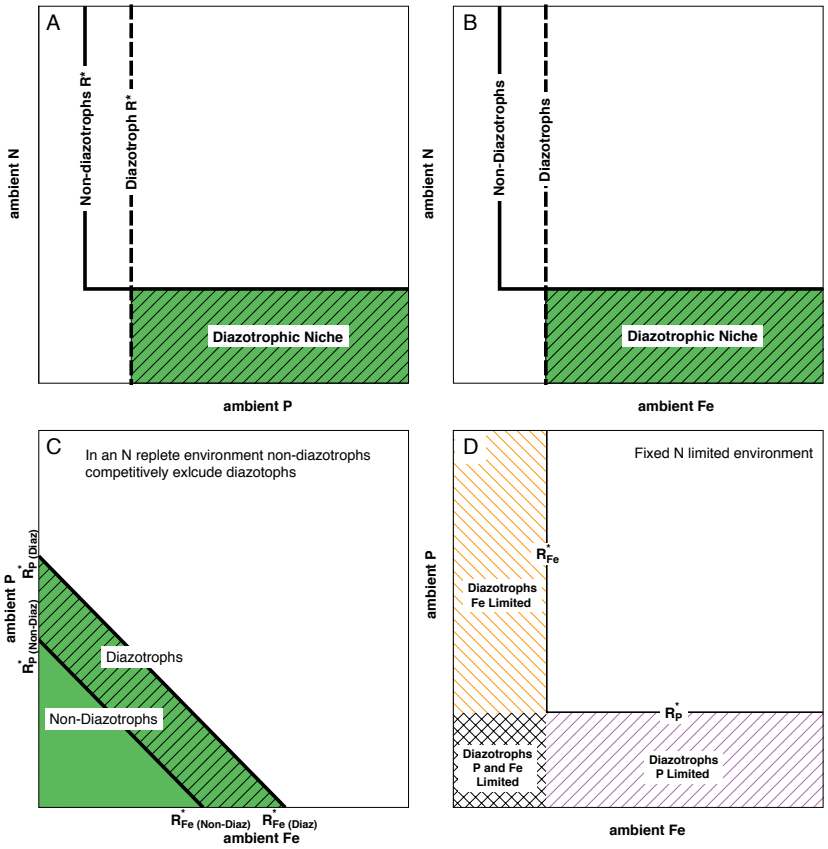

Figure 8: Graphical interpretations of differing resource competition scenarios concerning diazotrophs and non-diazotrophs. A) and B) Competition between diazotrophs and non-diazotrophs for N and P (A) or N and Fe (B), diazotrophic absolution of a fixed N source allows for a diazotrophic niche to exist at ambient N<$R^*_{N(non-diaz)}$. C) Under N replete conditions non-diazotrophs are able to outcompete diazotrophs due to $R^*_{P(non-diaz)} < R^*_{P(diaz)}$ and $R^*_{Fe(non-diaz)} < R^*_{Fe(diaz)}$. D) Resultant nutrient controls on *Trichodesmium* under N deplete conditions, here P and Fe alone can be used to describe the distribution.

cient nutrients, the organism will reduce its cellular processes and ultimately will become unviable.

An acquisitional or compensatory response strives to maintain the organisms cellular processes, thus I term such a response as nutrient *stress*; conversely a retrenchment response would signify the organism is nutrient *limited* (Straus, 2004).

A wide range of nutrient elements can potentially become scarce in the environment. Of significant interest in the sub-tropical Atlantic are the scarcity of DFe and DIP. Below I will outline the current understanding of *Trichodesmium*'s response to scarcity of either nutrient. Meanwhile in Chapter 4 and Chapter 5 of this thesis I consider

the metallomic and proteomic response of *Trichodesmium* to Fe or P scarcity.

1.6 PHOSPHORUS

Phosphorus is an essential component of all living organisms where it is used in both structural and functional molecules. Lipids are a group of P containing molecules acting as structural components of cell membranes, energy storage and signalling molecules; approximately 50% of the intracellular P in marine cyanobacteria can be attributed to phospholipids (Bertilsson et al., 2003). Other substantial intracellular P pools include deoxyribonucleic acids (DNA), ribonucleic acid (RNA), adenosine triphosphate (ATP) and polyphosphate molecules (see Section 1.6.2). Phosphorus in the ocean exists as two operationally defined pools, dissolved (<0.2 μm) and particulate (>0.2 μm). The dissolved pool can be further differentiated into dissolved inorganic phosphorus (DIP) generally in the form of orthophosphate (PO_4^{3-}) whilst dissolved organic phosphorus (DOP) includes P-esters, P-diesters and phosphonates. DIP is more readily bioavailable than DOP and as such is generally observed at lower concentrations (Roy et al., 1982; Dyhrman et al., 2006; Orchard et al., 2009). The residence time of oceanic P in the surface oceans is approximately 1-3 years during which time it will transition between these pools (Minster, 1983), deep water P turnover is much slower at around 69,000 years (Bruland, 1983; Byrne et al., 1988). The primary sources of 'new' Phosphorus to the oceans is via riverine transport ($\approx$ 17.7-30.4 Tg.P.yr^{-1}) and a small atmospheric source ($\approx$ 0.31 Tg.P.yr^{-1}) (Compton et al., 2000; Graham and Duce, 1979). Given the small size of the atmospheric P source along with the minimal effect of riverine inputs on open ocean environments, internal processes such as lateral advection are of greater importance to regions such as the (sub)-tropical Atlantic (Mather et al., 2008; Moore et al., 2009a).

1.6.1 *Phosphorus Acquisition*

Cyanobacteria such as *Trichodesmium* have three key pathways that mediate P acquisition and metabolism, the first and second involve extracellular P acquisition whilst; the third pathway involves degradation of previously synthesised P molecules (see Section 1.6.2):

1. Inorganic phosphate uptake, generally considered as preferential due to DIP being directly available for growth (Romanova et al., 2002). DIP uptake in *Trichodesmium* occurs via the P specific transport system (Pst) which comprises two high affinity phosphate binding proteins SphX (Tery_3534) and PstS (Tery_3537).

2. Dissolved organic phosphorus (DOP) hydrolysis. The enzyme alkaline phosphatase hydrolyses the RO–P bond of P-esters providing *Trichodesmium* access to this alternative source of P. *Trichodesmium* have two putative alkaline phosphatases genes *phoA* (Tery_3467) and *phoX* (Tery_3845) both of which are regulated by P supply in cultured *Trichodesmium* (Orchard et al., 2009). Phosphonates, where P is directly bound to the parent molecule with a C–P bond, are less bioavailabe than DIP and P-esters. *Trichodemsium* access the phosphonate pool via a C–P lyase pathway protein, PhnD (Tery_3868), which is capable of hydrolysing a broad range of substrates (Dyhrman et al., 2006).

3. Polyphosphate (polyP) metabolism. If P uptake is in excess of cellular requirements, inorganic phosphate can be stored as polyphosphate granules (Kornberg, 1995; Kornberg et al., 1999; Romans et al., 1994). Polyphosphate synthesis and catabolism in *Trichodesmium* is controlled by the *ppk* (Tery_1114) gene encoding for a polyP kinase which reversibly adds phosphate to the end of the polyP chain (Tzeng and Kornberg, 1998). Accumulation of polyP may occur as a luxury uptake strategy where P is in excess of Fe or as an overplus response to a short-term P enrichment event (Jacobson and Halmann, 1982; Bolier et al., 1992; Orchard et al., 2010b).

1.6.2 *Phosphorus Compensation*

Phosphorus related compensatory strategies in *Trichodesmium* involve cellular adaptation to either P depletion or repletion conditions (relative to other required nutrients). The reversible formation of polyP granules in *Trichodesmium* during periods of P enrichment creates a reservoir of intracellular P. Under periods of P limitation it is thought that the polyP synthesis pathway is reversed so as to liberate phosphate molecules for cellular processes (Kornberg, 1995; Kornberg et al., 1999; Romans et al., 1994). Bertilsson et al. (2003) attributed 50% of intracellular P to being phospholipid bound. However more recently this phospholipid pool has been shown to be flexible under P limiting conditions where phosphorus containing membrane lipids can be substituted with non-phosphorus containing counterparts (Van Mooy et al., 2006, 2009). Specifically, under P stress, *Trichodesmium* decrease the production of phosphatidylglycerol (PG) and increase production of sulfoquinovosyldiacylglycerol (SQDG), a sulphur containing, phosphorus free functional equivalent (Table 1). This was demonstrated by Van Mooy et al. (2006) who saw an SQDG/PG ratio change from 7.8 during P replete conditions to 18.5 under P deplete conditions.

Group	Phospholipid	Substitute Lipid	Δ Atomic Ratio (P:C)
Eukaryotes			
	Phosphatidylcholine [PC]	Betaine Lipids [BL]	-0.025 ±0.002
	Phosphatidylethanolamine [PE]	Betaine Lipids [BL]	-0.025 ±0.002
Cyanobacteria			
	Phosphatidylglycerol [PC]	Sulfoquinosyldiacylglycerol [SQDG]	-0.025 ±0.002

Table 1: Lipid substitution as a phosphorus stress compensatory response (Van Mooy et al., 2006, 2009).

1.6.3 *Phosphorus Retrenchment*

Retrenchment simply refers to a reduction in an organisms physiological activities. Phytoplankton are known to alter their cellular N:P ratio under P stress conditions, this reduction in the intracellular P pool (relative to N) is largely considered a retrenchment response (Redfield, 1958; Bertilsson et al., 2003). Using Redfield-like ratios as an indicator of P retrenchment in *Trichodesmium* is difficult given the observed flexibility in cellular stoichiometry (Krauk et al., 2006; White et al., 2006). The P compensatory strategies outlined above will actively alter a *Trichodesmium* cell's N:P stoichiometry meaning an altered N:P serves as a poor diagnostic of nutrient limitation.

1.7 IRON

Despite being the fourth most abundant element in the Earth's crust iron has been widely demonstrated to limit primary production in the marine environment (Boyd et al., 2007). Iron exhibits two environmentally relevant oxidation states Fe(II) and Fe(III) in aqueous solutions. It is widely understood that the iron-rich photosynthetic electron transport chain and the nitrogenase enzyme complex, are ancient and highly conserved, evolving between 2.3 and 2.2 10^9 years ago when the more bioavailable Fe(II) was abundant (Falkowski, 2006; Kranzler et al., 2013) due to the low oxidation state of the oceans prior to the great oxidation event (Holland, 2002, 2006; Lyons et al., 2014). The evolution of oxygenic photosynthesis drove an accumulation of atmospheric oxygen and led to oxygenation of the marine environment. This resulted in the lower valance Fe(II) undergoing rapid oxidation to its more thermodynamically stable form, Fe(III). Ferric iron (III) is substantially less soluble than its ferrous counterpart and as a result is less bioavailable. The primary source of Fe to the (sub)-tropical Atlantic ocean is from atmospheric deposition of iron-rich Saharan dust (Baker, 2003). Subsequently this Fe is rapidly removed by aggregation and sinking and results in Fe in this region having a residence time of just 6-62 days (Croot, 2004).

1.7.1 *Primary Iron Requirements*

The total iron requirement of the photosynthetic apparatus is 23-24 Fe atoms per complete photosystem. PSII contains 4 Fe atoms, one in cytochrome c_{550}, 2 in cytochrome b_{559} and one Fe bound to the reaction centre protein D2. The cytochrome b_6/f complex which bridges the two photosystems contains 5 Fe atoms, 2 associated with cytochrome b_6, one associated with cytochrome f and the remaining 2 making up the Rieske iron-sulfur protein centre (Fe_2S_2). The 9th iron is dependent on whether plastocyanin (no Fe) or cytochrome c553 is used (1 Fe) between the Cyt b_6/f complex and PSI. The 3 Fe_4S_4 complexes (F_X, F_A and F_B) associated with PSI comprise 12 Fe and the final 2 Fe are found in the Fe_2S_2 centre of ferredoxin (Ferreira and Straus, 1994) (Table 2 and Figure 3).

As previously mentioned, in *Trichodesmium* nitrogenase is a heterodimeric enzyme consisting of two proteins. The Fe-protein NifH, (Tery_4136) consists of a single redox-active Fe_4S_4 cluster which acts to hydrolyse MgATP and uses the required energy to supply the MoFe-protein with electrons. The MoFe-protein comprises two clusters; the first being the P cluster NifK, (Tery_4138) consisting of an Fe_8S_7 centre that accepts the electrons from the Fe-proteins' Fe_4S_4 cluster. The second cluster is the FeMo-co-centre NifD, (Tery_4137) comprising an Fe_7MoS_9 centre which contains the active site of nitrogen fixation responsible for the $N \equiv N$ reduction. Assuming a 5:1 stoichiometry of Fe-protein monomer:FeMo-protein homodimer this results in each complete (heterodimeric) nitrogenase system requiring a minimum of 50 Fe atoms.

1.7.2 *Iron Acquisition*

Two Fe uptake strategies have been demonstrated by cyanobacteria, the first being a transport system designed for specific Fe compounds, such as Fe-citrate, Fe-siderophores and hemes (Hopkinson and Morel, 2009); and the second being reductive iron uptake. Whilst many cyanobacteria are known to synthesise siderophores, including the diazotrophic *Anabaena* sp. (Goldman et al., 1983), this iron acquisition strategy has not been demonstrated by *Trichodesmium* sp. directly (Goldman et al., 1983; Hopkinson and Morel, 2009). *Trichodesmium* do however encode homologues of TonB (Tery_1560) and exbBD (Tery_4448/4449) which may be potential components of siderophore uptake systems. The colonial nature of *Trichodesmium* and its exosymbiotic community, some of which are capable of siderophore synthesis, may exist in a commensalistic relationship whereby *Trichodesmium* acquire symbiont synthesised siderophores. The relationship between *Trichodesmium* and their hosted heterotrophic bacteria may thus play a key role in *Trichodesmium*'s iron uptake (Rubin et al., 2011).

Complex	Fe-containing cofactor	No. of Fe atoms
Photosynthetic Apparatus		23-24
PSII	1 D1/D2 non-haem iron	1
	1 Cyt c_{550} haem	1
	1 Cyt b_{559} haem	2
Cytochrome b_6f	3 Cyt b_6 haems	2
	1 Cyt *f* haem	1
	1 Rieske Protein Fe_2-S_2	2
PSI	1 F_X [Fe_4S_4]	4
	1 F_A [Fe_4S_4]	4
	1 F_B [Fe_4S_4]	4
Cytochrome c_{553}	1 haem	1
Ferredoxin	1 [Fe_2-S_2] centre	2
N_{Fix} Apparatus		38
Nitrogenase reductase	2 [Fe_4S_4]	8
Dinitrogen reductase	2 P-cluster [Fe_8S_7]	16
	2 FeMo-co-centre [Fe_7MoS_9]	14

Table 2: Iron components of the electron transport chains of the photosynthetic apparatus (Shi et al., 2007; Michel and Pistorius, 2004) and the nitrogenase complex of the nitrogen fixation apparatus (Raven, 1988).

Luxury Fe uptake has been theorised in *Trichodesmium* due to the presence of ferritin Dps_{Tery}, (Tery_2787), an iron storage protein (Sandh et al., 2011). Under Fe replete conditions it is possible for a cells Fe uptake to exceed its current physiological requirement, in which case this excess Fe must be stored in a way so as to prevent interaction with various redox active components (Keren et al., 2004). Sandh et al. (2011) observed a 50-fold increase in the presence of *Trichodesmium*'s Dps_{Tery} when grown diazotrophically compared to growth on a source of fixed N. They attributed this increase to the greater iron burden associated with the nitrogenase enzyme.

1.7.3 *Iron Compensation*

Perhaps the most well studied iron stress response in marine cyanobacteria is that of the IsiA protein (Tery_1667) (Bibby et al., 2001; Ryan-Keogh et al., 2012). This iron stress induced protein acts to increase the organisms light harvesting ability by up to 7 times (Yeremenko et al., 2004) by expanding PSI's effective antennae size by 296% (Chauhan et al., 2011). By increasing PSI's light harvesting ability, IsiA allows for a reduction in the abundance of the iron-rich PSI complexes relative to PSII complexes. This permits a reduction in the cellular iron requirement whilst attempting to maintain cellu-

lar processes. An elevated IsiA response is well documented at both the transcriptional and protein level in *Trichodesmium* subject to iron stress conditions (Shi et al., 2007; Richier et al., 2012). The substitution of the iron containing protein ferredoxin with an Fe-free flavodoxin protein (IsiB, Tery_1666) has also been observed in *Trichodesmium* under iron stress (Chappell and Webb, 2010; LaRoche et al., 1996). Such substitution allows *Trichodesmium* to maintain cellular processes but at a higher iron use efficiency.

Other iron compensation mechanisms are less well studied, however one novel strategy performed by the unicellular diazotroph *Crocosphaera watsonii* is that of iron 'hot-bunking' (Saito et al., 2011). Here *Crocosphaera watsonii* selectively re-purpose the Fe contained in their metalloproteins across the diel cycle in order to suit their current cellular function. A re-purposing of photosynthetic Fe to nitrogenase Fe was observed and allowed for a substantial increase in both nitrogen fixation and photosynthesis given their disparate temporal activity (Figure 5) (Saito et al., 2011). The presence or absence of such an iron-thrift mechanism has not yet been investigated in *Trichodesmium*.

1.7.4 *Iron Retrenchment*

As with a P retrenchment response, care must be taken when using Fe:N or Fe:P ratios as a diagnostic of Fe stress due to *Trichodesmium*'s use of the iron storage protein ferritin along with its various Fe compensatory responses detailed above.

Iron requiring cellular processes are routinely down-regulated under severe iron stress. Nitrogen fixation is reduced substantially under Fe stress, this decrease has been observed both as a reduction in N_{Fix} rates, as a reduction in nitrogenase protein abundance and finally as a change in nitrogenase gene expression (Shi et al., 2007; Richier et al., 2012; Berman-Frank et al., 2007). Another Fe retrenchment strategy is the reduction in the PSI:PSII ratio (Richier et al., 2012; Shi et al., 2007; Berman-Frank et al., 2007). Whilst initially this response maybe classified as compensatory due to the activity of IsiA in increasing PSI's efficiency, under severe Fe stress this becomes a retrenchment response where the efficiency of PSI can longer be increased and so the process of photosynthesis is reduced (Richier et al., 2012; Shi et al., 2007).

1.8 TRANSITION METALS AS MICRO-NUTRIENTS

Whilst the importance of nutrient Fe and P on *Trichodesmium* is the primary focus of this thesis, there remains a wealth of other micronutrients essential to the healthy proliferation of this cyanobacterium. With an estimated 1 in 3 proteins containing a metal co-factor (Waldron and Robinson, 2009) the wider study of transition metals in

the marine environment is increasingly being pursued. The intracellular metal complement, the metallome, of an organism can provide key insights into its nutritional requirements and cellular function. This information, when combined with water-column concentrations of the metals of interest, can inform and identify previously unexplored avenues of nutrient limitation and species distribution. Studies have shown that *Trichodesmium* are stoichiometrically enriched in a number of trace metals relative to bulk phytoplankton (Quigg et al., 2011; Nuester et al., 2012; Tovar-Sanchez and Sañudo Wilhemy, 2006; Sañudo Wilhelmy et al., 2001). Substantial iron enrichment has consistently been observed and indirectly attributed to the nitrogen fixation apparatus (Nuester et al., 2012; Tovar-Sanchez and Sañudo Wilhemy, 2006; Sañudo Wilhelmy et al., 2001; Nuester et al., 2014). Other trace-metal enrichments are less consistently observed in the literature with Nuester et al. (2012) showing elevated V, Mo, Ni and Zn whilst Tovar-Sanchez and Sañudo Wilhemy (2006) demonstrate a substantially smaller Mo enrichment and a distinct lack of V enrichment. A consistent hurdle in interpreting the results of these studies is a lack of supporting biological data, this results in any discussion being somewhat speculative in nature. In Chapter 4 of this study I present metallomic data for environmental *Trichodesmium*, collected along an antithetical surface water Fe and P gradient of the (sub)-tropical Atlantic. The intracellular data, collected using trace-metal clean practices and analysed by acid-digestion and inductively coupled plasma mass spectrometry (ICP-MS), is supported by high-resolution multi-trace-metal water column data along with being used to help inform and support Chapter 5 and Chapter 6.

1.9 PROTEOMIC MASS SPECTROMETRY

Proteins are the primary functional building blocks of all cells and as such researchers have investigated their function and diversity for many decades. Initially these investigations were focused, small-scale experiments. However, with the advent of new instrumentation and methodologies alongside major advances in bioinformatics, it has become possible to investigate a cells full protein complement, the proteome. Proteomic analysis allows for investigation into cellular processes on an unprecedented scale; when applied to microbial organisms the technique theoretically allows for the simultaneous observation of all proteins.

1.9.1 *MS Instrumentation and Theory*

In short biological mass spectrometry instrumentation relies upon three essential components:

- An ionization source, which converts the sample solution, containing the molecules of interest, into charged gas-phase ions.
- One or more mass analysers dependent on the level of structural information required. Used to separate the gas phase ions according to their mass-to-charge ratio.
- An ion detector for counting the ions as they emerge from the last mass analyser.

Recent technological advances in ion source and mass analysers, along with computational improvements in data analysis have ushered in the so called 'third generation' of proteomics (Lamond et al., 2012). These advances have fundamentally changed the proteomic landscape, moving it, beyond simple high sensitivity protein identification towards providing whole proteome quantification allowing for the characterisation of proteome dynamics and establishing itself as an indispensable tool for cell biology.

1.9.1.1 *Ion Source*

One of the most important advances in biological mass spectrometry was the introduction of so called 'soft ionisation' technologies that allow for the efficient ionisation of biomolecules (peptides in this instance) with minimal sample degradation. Two primary ionisation techniques are used in proteomic MS analyses; the matrix-assisted laser desorption ionisation (MALDI) technique is a pulsed ion source and so is typically linked to time-of-flight (TOF) mass analysers (Harvey, 2003; Karas and Hillenkamp, 1988) whilst electrospray ionisation (ESI) is a continuous ion source and is better suited to quadrupole or ion-trap mass analysers (Fenn et al., 1989). Both techniques are ideally suited to the ionization of large biomolecules, however ESI techniques are better suited for coupling to upstream liquid chromatography units.

1.9.1.2 *Mass Analysers*

Following sample ionisation by either MALDI or ESI the ions need to be separated according to their mass to charge ratio (m/z). This is achieved using either time (e.g. TOF) or an electric or magnetic field (e.g. quadrupole or ion-trap). There are many types of mass analysers, often used in hybrid instruments which combine two mass analysers in one platform so as to leverage the strengths of each type. Only the three used during this study will be covered: Time-Of-Flight analysers (TOF), Linear Trap Quadrupole (LTQ), and Orbitrap analysers.

TIME-OF-FLIGHT ANALYSERS –

Conceptually the simplest type of mass analysers, TOF systems determine an ions m/z value via a time measurement. An electric

field is used to accelerate ions of a given charge state to a constant kinetic energy. This constant kinetic energy, coupled with each ions unique m/z values, results in each ion travelling at a specific velocity and subsequently arriving at the detector at a distinct time. This ion specific flight-time is then used to calculate the ions m/z value (Stephens, 1946). Given the requirement of a known start point for the flight-time, initially TOF instruments were ideally suited to a pulsed ion source such as MALDI. However introduction of the orthogonal accelerated Time-of-Flight analyser (OaTOF) allowed for continuous ion sources (such as ESI) to be utilised (Guilhaus et al., 2000). Here an accelerating potential is applied at right angles to the ion beam. This accelerating potential is pulsed so as to chop the continuous ion beam into distinct pulses of ions which are then recorded in the ToF analyser. Since a OaToF analyser measures all ions m/z values on the micro-second time scale the duty-cycle is much shorter than that of a scanning instrument (e.g. LTQ or Orbitrap) hence the sensitivity and amount of data acquired is higher.

LINEAR TRAP QUADRUPOLE ANALYSERS –

The LTQ is a derivation of a quadrupole mass analyser consisting of 4 parallel rods to which constant oscillating voltages are applied. This generates an electric field that allows ions of a particular mass-to-charge ratio to pass between the rods and so reach an ion detector. Tuning this electric field allows for ion mass filtering where the LTQ selects for ion of different m/z values. Ions with unfavourable m/z values will have unstable trajectories and will discharge themselves against the rods and not reach the ion detector.

ORBITRAP ANALYSERS –

First made commercially available in 2005, the Orbitrap is an entirely novel mass analyser which offers high resolution and accurate mass detection (Scigelova and Makarov, 2006; Makarov, 2000). Similar to the LTQ the Orbitrap is an ion trap type detector however the ions are trapped in an electrostatic field. The ions are electrostatically attracted to the central electrode and compensated by a centrifugal force arising from their initial tangential velocity. The electrostatic field moves the ions in complex spiral patterns, the axial component of which is detected as an image current on the two halves of an electrode encapsulating the orbitrap (Figure 9). This axial component is independent of initial energy, angles and position and is used to obtain oscillating frequencies for ions with different masses, resulting in the accurate determination of the ions m/z (Makarov, 2000; Scigelova and Makarov, 2006).

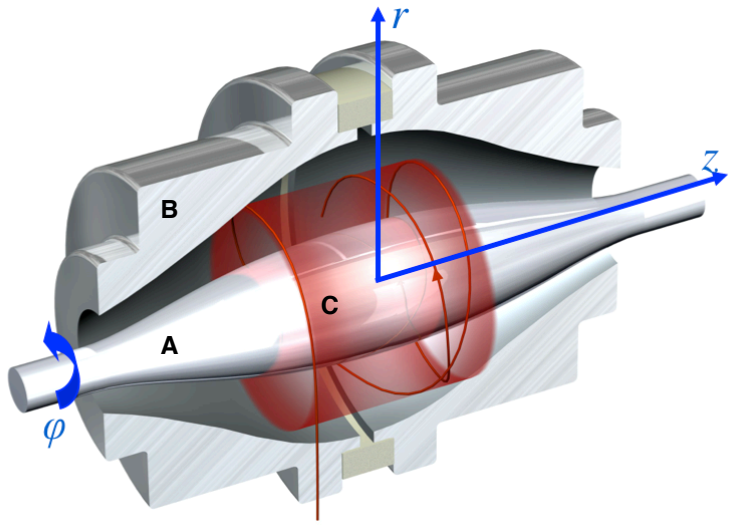

Figure 9: A cut-away model of the orbitrap mass analyser. A) Central electrode, B) An outer electrode, split in half by an insulating ceramic ring, C) Complex orbital path of an ion. Image modified from Scigelova and Makarov (2006)

1.9.2 *Tandem Mass spectroscopy and collision induced dissociation*

Tandem mass spectroscopy such as that used in proteomics relies on 'soft ionisation' producing minimal fragmentation of pre-cursor ions. These precursor ions are then isolated in the first mass analyser before undergoing fragmentation to yield product or transition ions which can then be measured in a second spectrometer. This results in the collection of both precursor and product ion m/z values. Fragmentation is achieved using collision induced dissociation (CID) whereby a precursor ion undergoes constant collision with a collision gas in order to bring the ion into an excited state sufficient to overcome the fragmentation threshold. The precursor ion then undergoes unimolecular decomposition into product ions. This two-stage analysis can be done spatially, using two distinct instruments, or temporally using the same instrument for both precursor and product ion with an intermediary ion storage device. In this study both spatial and temporal analysis were performed for shotgun and single reaction monitoring (SRM) or multiple reaction monitoring (MRM) data acquisition respectively.

1.9.3 *Peptide mass fingerprinting*

For peptides this fragmentation of precursor to product ion, when conducted at low-energies results in structurally significant fragmen-

tation along the αC-C, C-N amide linkage or N-αC bond (Figure 10). If the charge is retained on the N-terminal fragment these fragments are termed a- b- and c-type ions respectively; alternatively if the charge is retained on the C-terminal fragment they are termed x- y- and z-type ions (Figure 11B). These reproducible structural fragmentation patterns along with the *m/z* of both precursor and product ions allow for the derivation of the peptides amino acid structure followed by comparison to predicted proteolytic peptides generated *in-silico* from relevant genomic libraries. This allows for the translation of an MS spectra to an amino-acid sequence followed by its proteotypic assignation to a protein described in the genomic library.

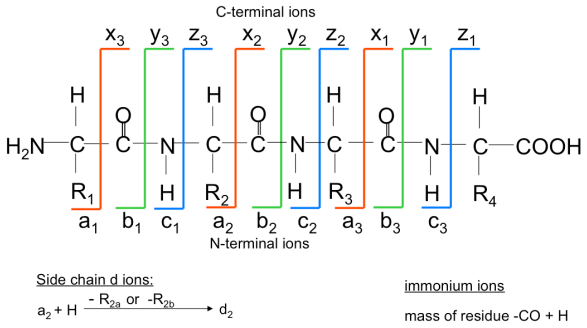

Figure 10: The standard nomenclature for ions derived from backbone fragmentation of a peptide by collision induced dissociation (CID) (Biemann, 1988).

1.9.4 *Shotgun Proteomics*

Deriving it's name from the quasi-random firing pattern of a shotgun, shotgun proteomics is intrinsically a bottom-up technique for identifying relative protein abundances in complex mixtures. In our case the complex mixtures are whole colony digests and as such they are in effect a pseudo-mixed community lysate. Shotgun proteomics on the LTQ Orbitrap Velos was performed in a data dependent acquisition method whereby a predetermined number of precursor ions whose m/z values were selected from a survey scan were subject to a second stage of mass selection in an MS/MS analysis. A data-dependent acquisition (DDA) strategy limits the number of observable peptides due to instrument time constraints. As such many less-abundant peptides in the sample are excluded from the precursor selection and hence produce no data.

1.9.5 *Absolute QUantification Assay (AQUA)*

Where shotgun proteomics provides relative quantification of entire proteome inventories, targeted quantitative proteomics using multiple reaction monitoring mass spectrometry (MRM) allows for the absolute quantification of select peptides (Gerber et al., 2003). The selection of proteotypic peptides, for which identifiable and reproducible spectra are observed, is made possible by leveraging acquired shotgun proteomic data. From these selected peptide targets isotopically labelled synthetic counterparts were produced. End-terminus Arganine (R) or Lysine (K) labelling provides a mass discrepancy of 8 or 10 Daltons respectively. Isotopic labelling preserves the chemical characteristics of the native peptide whilst allowing spectral differentiation of native and labelled peptide. I am then able to calculate an absolute concentration of the native peptide given the known concentration of labelled standard (Figure 11). Data acquisition is performed in a similar manner to the DDA method described for shotgun proteomics however the precursor ion m/z selection is pre-selected as the native and labelled target m/z values allowing for an increased probability of detection.

1.9.6 *Label-free*

The label-free method, Hi3 or MS^E (Waters Corporation), is a data independent analysis (DIA) method whereby no precursor ion selection is performed. A Q-TOF mass detector rapidly alternates between low energy and elevated energy (MS^E) modes of acquisition without any mass selection (Figure 12). During low energy acquisition *all* ions are transmitted from the source, through the collision cell (set to low collision energy so no fragmentation occurs) to the mass analyser. The data collected during low energy acquisition is the precursor ion spectrum. During elevated energy acquisition *all* ions are transmitted from the ion source, through the collision cell (high collision energy so as to generate maximum ion fragmentation) to the mass analyser and recorded as a fragment ion spectrum (Waters Corporation, 2011). This results in the generation of the complete MS dataset where the instrument records 'all the data all the time' (Waters Corporation, 2011). Using accurate mass and retention time pairs (AMRT) the fragment ion spectra are assigned to their associated precursor ion peaks (by matching their retention times) after which peptide mass fingerprinting and database searching can occur (Silva et al., 2005, 2006b).

Quantification relies on the relationship between MS signal response and protein concentration, Silva et al. (2006b) discovered that the average MS signal response for the three most intense tryptic peptides per mole of protein is constant within a coefficient of variance of $\pm$ 10%. Utilising an internal standard of known concentration this

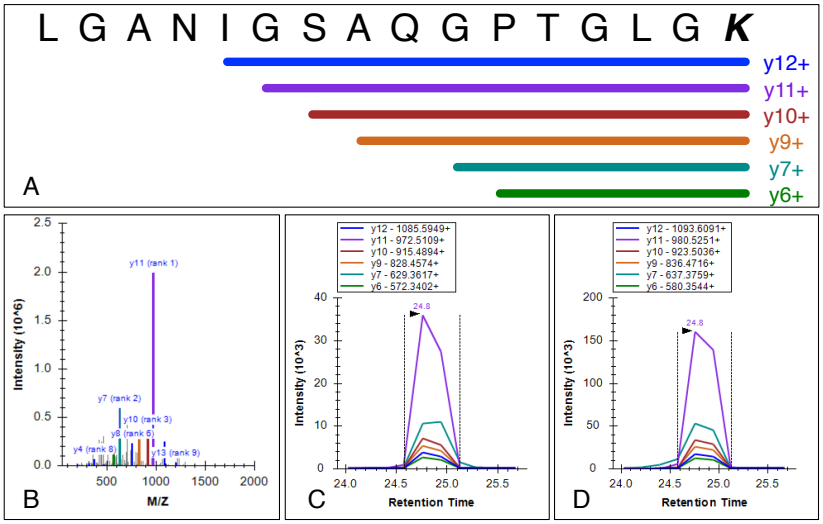

Figure 11: MRM-triggered MS/MS experiment for the validation of a PsbC/CP43 peptide (LGANIGSAQGPTGLGK) acquired on a Thermo Scientific LTQ Orbitrap Velos. A) Peptide amino acid showing the six selected transitions. B) Full MS^2 spectra for the protein showing relative abundance of differing transitions. C) MRM traces for the native peptide showing consistent transition abundance across 4 sequential MS^2 spectra. D) MRM traces for the isotopically labelled peptide standard showing similar transition abundance across 4 spectra along with chromatographic co-elution with the native peptide, transition m/z is offset by 5 daltons due to the isotopic labelling of the end-terminus Lysine.

relationship is used to calculate a universal signal response factor (counts.mol^{-1}) and thus allow the absolute quantification of sample peptides.

1.10 THESIS AIMS AND OUTLINE

My aim throughout this thesis was to explore the influence nutrients Fe and P have on the distribution and cellular characteristics of *Trichodesmium*. I hypothesis that neither nutrient alone (Fe or P) can fully explain the observed distribution of *Trichodesmium* in the (sub)-tropical Atlantic and instead suggest a framework of Fe and P co-limitation using the resource-ratio theory, proposed by Tilman (1980), is required. To further explore this proposed co-limitation hypothesis I aimed to directly compared the intracellular composition of *Trichodesmium* with a select set of specific Fe and P stress proteins. The hypothesis here being that varying degrees of Fe or P stress would be evident in variations in *Trichodesmium*'s intracellular composition. Further to this I suggest the nutrient stress biomarkers should correlate with other markers of nutrient stress (e.g. water column nu-

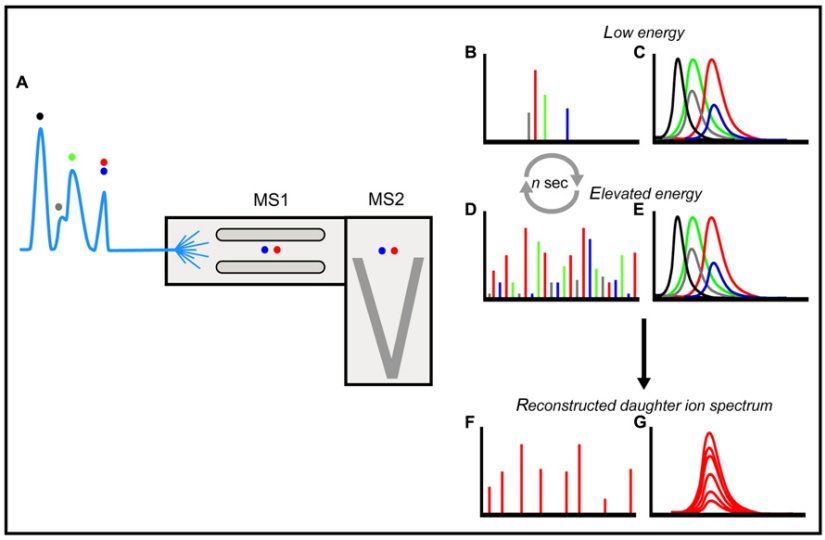

Figure 12: Separation of peptides (coloured dots) via liquid chromatography and Quadrupole-OaTOF in a data independent (DIA) mode (A). The continuous ion beam provided by the orthogonal acceleration allows for the generation of all parent (B) and product ions (D). Ion-accounting algorithms allow for the deconvolution of all parent (C) and product (E) ions based on accurate mass and retention time pairs (AMRT). This allows for the reconstruction of matched parent and product ion spectra (F) and allows for correct identification of peptide sequence. Figure from Kramer et al. (2012).

trient distribution and *Trichodesmium* intracellular composition) with the potential that these biomarker protein could be used as a direct assessment of nutrient limitation in future studies.

Data and samples presented in this thesis were collected during one culture study and field studies in the central (sub)-tropical Atlantic. Chapter 2 contains descriptions of the methods used for sample collection, processing and analysis along with some preliminary proteomic method development work. The ocean scale distribution of *Trichodesmium* is presented and discussed in Chapter 3. This includes data from 7 oceanic cruise spanning 2002-2012 and focuses on the nutrient requirements of *Trichodesmium* and diazotrophs as a whole. Chapter 4 presents intracellular metal concentration data collected during 2 oceanic cruises from contrasting seasons, here I discuss the oligotrophic conditions of the study region as a means to interpret *Trichodesmium*'s unique biogeochemical composition. The results of a iron limitation culture experiment are presented in Chapter 5, here I show a high resolution, quantified proteomic dataset and explore a number of protein level responses to iron-limitation. Chapter 6 aims to synthesise the preceding chapters introducing *Trichodesmium*'s metallo-proteome, derived from a bioinformatical anal-

ysis of the results presented in Chapter 5. I combine these data with the metallomic data presented in Chapter 4 aiming to establish a cohesive understanding of the nutritional requirements of this globally important cyanobacterium.

2

MATERIALS AND METHODS

The new data presented in this thesis were collected during 2 meridional cruises in the central Atlantic and through growing *Trichodesmium* culture's in the laboratory at the National Oceanography Centre, Southampton. The first cruise, D361, took place between February 7th and March 19th 2011 as part of the GEOTRACES UK project whilst the second cruise, AMT21, took place between October 30th and November 19th 2011 and served as the Atlantic Meridional Transect project's 21st voyage (Figure 13). Both cruises were aboard the RRS Discovery.

2.1 FIELD SAMPLING

2.1.1 *Trichodesmium methods*

2.1.1.1 *Drift net deployments*

Drift net deployments were performed using a 50 cm aperture, 100 μm mesh size drift net. Deployments were performed during the predawn and solar noon stations of both cruises whereby nets were deployed to approximately 10-15 m depth for approximately 10 minutes per deployment. Two nets were used during each deployment, the first used for collecting non-trace metal *Trichodesmium* samples; the second net, consisting of a nylon string bridle in place of the usual metal bridle and acid cleaned cod-ends (10% HCl), was used for collection of trace-metal (TM) clean *Trichodesmium* samples. Drift net cod-ends were sealed immediately after retrieval and the TM nets cod-end was zip-lock bagged and transferred to a laminar flow hood. Drift net samples were transferred to a TM clean laboratory, emptied into acid-cleaned sample buckets where *Trichodesmium* colonies were individually picked using plastic inoculating loops or plastic pasteur pipettes.

2.1.1.2 *Chlorophyll*-a

Chl-*a* was measured in order to provide both community and *Trichodesmium* specific information on the concentration of this photosynthetic pigment and as an overall index of community phytoplankton biomass but also *Trichodesmium* specific biomass. Community Chl-*a* was assessed by filtering 200 ml of seawater collected from Ocean Test Equipment (OTE) bottles through 25 mm glass fibre filters (Whatman GF/F or Fisherbrand MF300) whilst *Trichodesmium* specific Chl-*a*

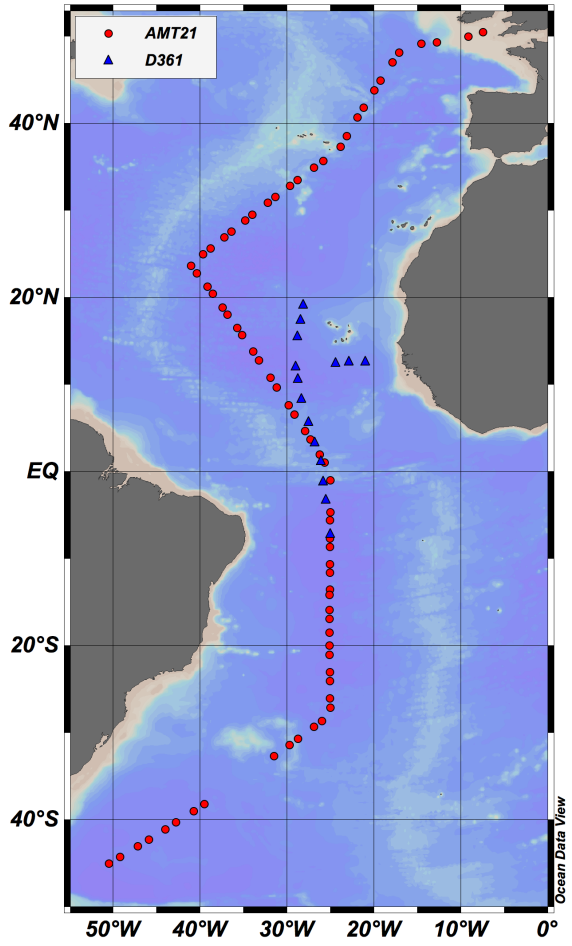

Figure 13: Map of the Atlantic ocean showing the station locations for D361 (Blue triangles) and AMT21 (Red circles). A number of stations planned for AMT21 in the South Atlantic were cancelled due to inclement weather, these are evident in the apparent gap in the AMT21 transect between ≈32-37°S.

was assessed by picking 5-10 colonies, suspending in 0.22 µm filtered seawater and filtered onto glass-fibre filters. Extraction was done in 90% acetone for 24 hours in the dark at 4°C or -20°C for D361 or AMT21 respectively. Fluorometric measurements were undertaken on a Turner Designs TD-700 (D361) or a Turner Designs Trilogy (AMT21) following the method described in Welschmeyer (1994).

2.1.1.3 *Trichodesmium abundance*

Trichodesmium abundance samples were collected in order to give an indication of the geographical distribution of *Trichodesmium* throughout both cruises. An entire 20 L OTE bottle, fired at the surface (≈2 m), was then gravity filtered (directly from the OTE bottle) onto Millipore Isopore 10 µm pore size, 47 mm diameter filters, colonies and

free trichomes were gently agitated to remove from the surface of the filter before being preserved in 2% Lugol's iodide and stored in the dark, filters were visually inspected for complete re-suspension of *Trichodesmium* colonies and free trichomes. Colony and free trichome abundance, colony morphology and approximate colony size were enumerated using light microscopy and a bogorov tray. Colonial cell abundances were estimated assuming 100 trichomes per colony (Carpenter et al., 2004).

2.1.1.4 *PON, POC and Nitrogen fixation rates*

Whole community N_{Fix} - Following the method described by Montoya et al. (1996), 4.5 L clear Nalgene bottles were filled, ensuring no air bubbles remained, and sealed with screw cap seals with silicone septa. Each bottle was spiked with 4 ml $^{15}N_2$ (Cambridge Isotope Laboratory, UK) using a gas-tight syringe and incubated for 24 hours at sea surface temperature (SST), achieved using the ships under-way flow system, and 100%, 69%, 20% and 0.1% sea surface irradiance (SSI), achieved using 0.15 neutral density LEE™ optical filters. Following incubation over 24 hours the contents of each bottle were filtered onto pre-ashed glassfibre filters (450°C for 12 hours, Whatman GF/F or Fisherbrand MF300), folded into 1.5 ml tubes (Eppendorf) before being dried for 24 hours at 40°C. Upon return to the lab samples were fumed for 24 hours in a vacuum desiccator containing a small volume of sulphurous acid so as to remove any particulate inorganic carbon. D361 samples were pelleted into aluminium capsules and analysed for organic nitrogen (PON) and $^{15}N/^{14}N$ ratio with an elemental analyser isotope ratio mass spectrometer (EA-IRMS) (Thermo Finnigan EA 1112 Series Flash Elemental Analyser) at the National Oceanography Center Southampton. AMT21 samples were encapsulated in aluminium capsules and sent to the Stable Isotope Facility, UC Davis for analysis on an Elementar Vario EL Cube or Micro Cube elemental analyser (Elementar Analysensysteme GmbH, Hanau, Germany) interfaced to a PDZ Europa 20-20 isotope ratio mass spectrometer (Sercon Ltd., Cheshire, UK). Nitrogen fixation rates alongside PON and POC concentration were obtained for AMT21 samples.

***Trichodesmium* specific N_{Fix} rates** - 50 colonies were picked using plastic inoculation loops and placed into 125 ml Nalgene bottles filled with 0.2 μm filtered trace-metal clean surface water (see Section 2.1.2.3). Bottles were filled entirely, sealed with screw-top, septum containing caps and injected with 0.5 ml $^{15}N_2$ using a gas tight syringe. Bottles were incubated for 12 hours over the diazotrophic period at SST and SSI before being filtered onto pre-ashed glass-fibre filters, dried and stored until analysis alongside the whole community N_{Fix} samples.

A systematic ≈2 fold underestimation of the absolute N_2 fixation rates is likely observed due to the method utilised (Groß kopf et al.,

2012; Mohr et al., 2010; Wilson et al., 2012), however given that observed community rates and *Trichodesmium* specific rates spanned >2 and >3 orders of magnitude respectively this underestimation will have little bearing on an qualitative conclusions I draw from the data.

2.1.1.5 *Photosynthetic efficiency*

The apparent PSII photosynthetic efficiency, F_v/F_m, was assessed during D361 using a Chelsea Scientific Instruments Fasttracka MkII Fast Repetition Rate fluorometer (FRRf) integrated with a FastAct Laboratory System (Chelsea Technologies Group Ltd., West Mosley, Surrey, UK) (Kolber et al., 1998) and during AMT 21 using a SATLANTIC Fluorescence Induction and Relaxation fluorometer (FIRe) (Gorbunov and Falkowski, 2004). Approximately 5 *Trichodesmium* colonies were suspended in 0.22 μm filtered seawater and dark acclimated for 15 minutes, measurements were blank corrected using 0.22 μm filtered seawater.

2.1.1.6 *Trichodesmium protein samples*

Due to the ease of potential protein degradation, collection of *Trichodesmium* protein samples took priority over most other sampling. As such it was ensured that sampling took no longer than 30 minutes between drift net collection and flash freezing of protein samples. 50-100 colonies were picked, re-suspended in 0.22 μm filtered seawater, and filtered onto 25 mm diameter glass-fibre filters (D361) or 25 mm, 2.0 μm Millipore Isopore polycarbonate filters (AMT21), placed into cryovials and flash frozen in liquid N_2.

2.1.1.7 ^{55}*Fe Uptake*

Trichodesmium colonies collected from the trace-metal clean tow net were picked using an acid washed inoculating loop. Colonies were rinsed 3 times using 0.22 μm filtered tow-fish water and then finally picked into 125 ml of a further clean sample of tow-fish seawater. Incubations were spiked with ≈1.3 kBq of carrier free ^{55}Fe added as weakly acidic Fe(III)Cl. Spiking was performed ≈1 hour post net haul and always before dawn. Spiked bioassays were incubated for 10-12 hours at 500 μmol photons $.m^{-2}.s^{-1}$. This ^{55}Fe addition of <1.5 kBq corresponds to an iron enrichment of <3 pM a minimal perturbation to the total iron pool. At the end of the incubation period samples were filtered onto 1.2 μm polycarbonate filters under <100 mBar vacuum. Filters and samples were rinsed using a buffered Ti-EDTA-citrate solution so to differentiate between uptake into cells relative to apparent uptake due to adsorption of ^{55}Fe to external cell surfaces (Hudson and Morel, 1989). Filtered and rinsed samples were then placed in 5 ml Ultima Gold (Perkin Elmer, MA, USA) before

being counted on a liquid scintillation counter (Perkin Elmer TriCarb 3100TR).

2.1.1.8 ^{33}P *Uptake and Alkaline Phosphatase Activity*

During D361, at stations where *Trichodesmium* were found, 20 colonies were placed in 50 ml 0.22 μm filtered seawater in a 125 ml, acid-washed, polycarbonate bottle. These samples were spiked with 3 kBq ^{33}P resulting in a minimal P addition of ≈0.01 nM. Samples were then placed in an on-deck incubator at 55% surface irradiance for 3 hours. At the end of the incubation samples were filtered onto 1.2 μm polycarbonate filters. Filters were placed in 7 ml scintillation vials, 5 ml Ultima Gold (Perkin Elmer, MA, USA) was added before being counted on a liquid counter (Perkin Elmer TriCarb 3100TR).

Alkaline phosphatase activity was determined using methods described by Ammerman (1993). Specifically, ≈20 *Trichodesmium* colonies were collected from net tows and resuspended in 50 ml 0.22 μm filtered seawater. 800 nM of 4-methyumbellyferyl phosphate (MUF-P) was added to each sample and a filtered seawater blank. Upon addition of MUF-P, 3 ml of spiked samples was removed immediately, 1 ml Borate buffer was added, fluorescence measured and T_{Zero} fluorescence data collected. Samples and blanks were incubated for 12 hours in a SST, 55% sea-surface irradiance on-deck incubator. Fluorescence was determined using a Turner 10AU field fluorometer fitted with 365 nm excitation and 455 nm emission optical filters (Optical kit #10-302R). An MUF standard curve (200, 400, 800 nM), blanks and killed controls (boiled seawater plus 200 or 400 nM MUF-P) were incubated each day alongside sample incubations. Rates of alkaline phosphatase hydrolysis was determined by calculating the change in fluorescence observed during the incubation time and dividing by the fluorescence of the appropriate MUF standard. Rates were normalised to number of *Trichodesmium* colonies.

2.1.2 *Biogeochemical Sampling*

2.1.2.1 *Salinity and Temperature*

During both D361 and AMT21 both sea surface temperature (SST) and sea surface salinity (SSS) data were obtained from the RRS Discovery's non-toxic under-way supply system using a ThermoSalinoGraph.

2.1.2.2 *Nutrients*

During AMT21 micro-molar nutrient samples were collected into acid-washed 60 ml Nalgene HDPE sample bottles. Analysis was carried out using a 4 channel (nitrate, nitrite, phosphate and silicate) Bran and Luebbe AAIII segmented flow, colorimetric, auto-analyser

(Woodward and Rees, 2001). For D361 nanomolar NO_3^-, NO_2^- and PO_4^{3-} were determined using colorimetric methodologies as with standard segmented flow analysis however 2 m Liquid waveguide capillary cells were used as flow cells (Patey et al., 2008).

2.1.2.3 *Dissolved Fe*

Water samples were collected using a near-surface water tow-fish pumped into the TM clean laboratory container using a Teflon diaphragm pump (D361) or a peristaltic pump (AMT21). Dissolved iron (DFe) samples were filtered through Supor Acropak (D361) or Sartobran P300 (AMT21) 0.2 μm filter units and stored in 125 ml, acid cleaned, nalgene LDPE bottles (Soaked for 1 week in Decon, 1 week in 6 M HCl, 1 week in 6 M HNO_3, rinsed 5 times and stored sealed with a small amount of DI water). Aboard D361, samples were acidified to pH ≈2 using HCl (Romil UpA) whilst for AMT21 samples were kept at -20°C until returning to the laboratory where they were acidified in the same manner. Analysis was performed using an online flow injection analysis (FIA) system with Toyopearl column via chemiluminescence detection (Klunder et al., 2011). Although DFe data for D361 and AMT21 was acquired by both FIA and ICP-MS methods (see Section 2.1.2.4) the DFe data presented throughout this thesis are from FIA only.

2.1.2.4 *Multi-element ICP-MS analysis of dissolved metal concentrations*

WATER SAMPLING AND FIRST TREATMENT - Surface seawater was pumped into the trace metal clean laboratory using a Teflon diaphragm pump connected by acid-washed braided PVC tubing to a towed 'fish' positioned at 3 – 4 m. Samples were filtered in-line through 0.8/0.2 μm cartridge filter (AcroPak1000™) into acid washed low density polyethylene (LDPE) bottles. Water column samples were collected using trace metal clean OTE bottles mounted to a titanium frame rosette deployed on a Kevlar wire. After recovery, OTE bottles were transferred into the clean sampling container where all sample handling was performed. Dissolved samples were filtered in-line through 0.8/0.2 μm cartridge filter (AcroPak500™) into acid washed LDPE bottles, while total dissolvable samples were dispensed unfiltered into acid washed LDPE bottles. Dissolved and total dissolvable seawater samples were acidified within 24 hours of collection with concentrated trace metal grade hydrochloric acid (HCl, UpA, Romil) to pH 1.7 (≈20 mmol H^+ L^{-1}). Acidified seawater samples for multi element analysis were shipped to the National Oceanography Centre Southampton and analysed by isotope dilution and standard addition inductively coupled plasma - mass spectrometry (ICP-MS), respectively (Milne et al., 2010).

TRACE METAL ANALYSIS The pre-concentration and ICP-MS analysis was adapted from the method outlined by Milne et al. (2010). Approximately one year after collection, 12 ml of acidified seawater was transferred into 30 ml fluorinated ethylene propylene (FEP) bottles and spiked with a spike solution containing mainly the artificially enriched isotope of iron (^{57}Fe), zinc (^{68}Zn), cadmium (^{111}Cd), lead (^{207}Pb), nickel (^{62}Ni), and copper (^{65}Cu). For the analysis of manganese (^{55}Mn), aluminium (^{27}Al), and cobalt (^{59}Co) a series of three standard additions were performed on every fifth sample. To obtain equimolar conditions between the spike and the natural seawater concentration, larger amounts of spike was added to the total dissolvable seawater samples. All samples were irradiated with strong ultraviolet (UV) light for a minimum of 3.5 hours. Subsequently, the sample solution was buffered to pH 6.4 using a 2 M ammonium acetate solution (pH 9.2, Fisher Optima grade ammonia and acetic acid, glacial). Immediately after buffer addition the solution was pre-concentrated using an automated system (Preplab, PS Analytical) that was equipped with a metal chelating resin (WACO) (Kagaya et al., 2009). Any remaining seawater salts were rinsed off using deionized water (> 18 MΩ cm, MilliQ, Millipore). The metals retained on the resin were eluted using 1 ml of a 1 M sub-boiled HNO_3 solution, which was collected in acid cleaned 4 ml polypropylene (PP) vials. The collected vials were placed into the Elemental, Service, & Instruments (ESI) auto-sampler of the ICP-MS (Element XR, Thermo). Certified seawater standards (SAFe and GEOTRACES) were pre-concentrated and analysed with each batch of samples, in order to validate our sample concentration (Table 3). Values obtained were in good agreement with reported census values for the GEOTRACES and the SAFe standard seawater (Table 3). The analytical limit of detection (LOD) and the method blank for each trace metal analysed are listed at the right side of the Table 3. More detailed information about the used standard seawaters can be retrieved from the website: http://es.ucsc.edu/ kbruland/GeotracesSaFe/kwbGeotracesSaFe.html.

2.1.2.5 pCO_2

During D361 and AMT21 dissolved inorganic carbon (DIC) and total alkalinity (TA) were sampled from near-surface fired OTE bottles. DIC analysis was done using the standard colourimetric technique described by (Dickson et al., 2007) whilst TA was analysed using standard titrimetric techniques (Dickson et al., 2007). The partial pressure of CO_2 was derived from these parameters following equations outlined in Dickson et al. (2007).

Metal	SAFe S	SAFe D1	SAFe D2	GS	GD	LOD	Blank
Cd (pmol kg^{-1})	7.2 ± 1.6	967 ± 49	949 ± 66	8.1 ± 1.6	263 ± 4	1	0.4
Cd census	1.1 ± 0.3	991 ± 31	986 ± 23	2.1 ± 0.6	271 ± 6		
Pb (pmol kg^{-1})	48.4 ± 6.1	23.9 ± 3.1	28.9 ± 4.4	26.9 ± 3.2	45.3 ± 0.3	0.6	2.8
Pb census	48.0 ± 2.2	27.7 ± 2.6	27.7 ± 1.5	28.6 ± 1.0	42.7 ± 1.5		
Ni (nmol kg^{-1})	2.56 ± 0.55	8.59 ± 0.62	8.60 ± 0.36	2.54 ± 0.30	4.16 ± 0.12	0.1	0.15
Ni census	2.28 ± 0.09	8.58 ± 0.26	8.63 ± 0.25	2.08 ± 0.06	4.00 ± 0.10		
Cu (nmol kg^{-1})	0.55 ± 0.06	2.02 ± 0.07	2.15 ± 0.16	0.86 ± 0.05	1.65 ± 0.06	0.01	0.05
Cu census	0.52 ± 0.05	2.27 ± 0.11	2.28 ± 0.15	0.84 ± 0.06	1.62 ± 0.07		
Fe (nmol kg^{-1})	0.087 ± 0.025	0.60 ± 0.07	0.90 ± 0.10	0.56 ± 0.05	1.28 ± 0.15	0.032	0.056
Fe census	0.093 ± 0.008	0.67 ± 0.04	0.93 ± 0.02	0.55 ± 0.05	1.00 ± 0.10		
Zn (nmol kg^{-1})	0.07 ± 0.06	6.90 ± 0.56	7.29 ± 0.27	0.07 ± 0.10	1.71 ± 0.10	0.129	0.1
Zn census	0.069 ± 0.010	7.40 ± 0.35	7.43 ± 0.25	0.04 ± 0.01	1.71 ± 0.12		
Mn (nmol kg^{-1})	0.75 ± 0.05	0.36 ± 0.04	0.40 ± 0.05	1.40 ± 0.13	0.24 ± 0.02	0.001	0.004
Mn census	0.79 ± 0.06	0.35 ± 0.05	0.35 ± 0.05	1.50 ± 0.11	0.21 ± 0.03		
Co (pmol kg^{-1})	2.9 ± 0.8	38.2 ± 2.3	40.1 ± 3.9	26.3 ± 2.2	57.2 ± 5.0	0.07	0.01
Co census	4.8 ± 1.2	45.4 ± 4.7	45.7 ± 2.9	31.8 ± 1.1	65.2 ± 1.2		
n =	25	5	19	6	3		

Table 3: Observed and census values for certified seawater standards (SAFe and GEOTRACES) for multi-element ICP-MS seawater analysis.

2.1.3 *Sample collection summary*

Table 4 lists all samples collected during both D361 and AMT21 and details the person(s) responsible for their collection.
Table 5 lists all the data sources used in Chapter 3, these include data from cruises M55, M60, AMT17, D326 and JC32 which were not generated as part of this study.

	Sample Type	No. collected (D361)	No. collected (AMT21)	Person Responsible
Biogeochemical	N_{Fix} Community	91	272	*D Honey* (D361), *JT Snow* (AMT21)
	Underway Dissolved Fe	250	56	*C Schlosser* (AMT21 & D361), *B Wake* (D361), JT Snow (AMT21)
	Underway Trace-metals	250	56	*C Schlosser* (AMT21 & D361), JT Snow (AMT21)
	Underway Nanonutrients	231	56	*M Woodward* (AMT21 & D361), *FE Legiret* (D361), JT Snow (AMT21)
	Underway PON/POP	•	39	JT Snow (AMT21)
Trichodemsium	N_{Fix}	11	11	JT Snow (D361 and AMT21)
	Protein	28	22	JT Snow (D361 and AMT21)
	Intracellular Metals	19	22	JT Snow (D361 and AMT21)
	CHN	24	21	JT Snow (D361 and AMT21)
	POP	23	19	JT Snow (D361 and AMT21)
	FRRF	16	22	JT Snow (D361 and AMT21)
	Chl a	13	22	JT Snow (D361 and AMT21)
	Abundance	16	15	JT Snow (D361 and AMT21)
	Underway sampling	34	•	JT Snow (D361)
	^{55}Fe uptake	11	•	*CM Moore* (D361)
	^{33}P Uptake	34	•	*S Reynolds* & *C Mahaffey* (D361)

Table 4: Summary of all samples collected during D361 and AMT21 along with the person responsible for sample collection and/or analysis (italics).

Cruise	SST	pCO_2	DIN and DIP	DFe	Trichodesmium Abundance	Surface N_{Fix}
M55	Underway	A Kortzinger[6]	P Fritsche[3]	P. Croot[8]	NA	NA
M60	Underway	H Luger[3]	F. Malien[3]	NA	NA	NA
AMT17	Underway	G Moore[2] and D Bakker[5]	M Woodward[2]	S Ussher[2]	CM Moore[1]	CM Moore[1]
D326	Underway	C Dumousseaud[1]	M Patey[1]	M Rijkenberg[6]	CM Moore[1] & C Mahaffey[4]	CM Moore[1] & C Mahaffey[4]
JC32	Underway	U Shuster[7]	S Torres-Valdes[1]	NA	CM Moore[1]	CM Moore[1]
D361	Underway	E Tynan[1]	M Woodward[2] & FE Legiret[1]	C Schlosser[1,3]	JT Snow[1]	JT Snow[1] & DJ Honey[1]
AMT21	Underway	V Kitidis[2]	JT Snow[1] & M Woodward[2]	C Schlosser[1,3]	JT Snow[1]	JT Snow[1]

Table 5: Summary of data presented in Chapter 3 for cruises M55, M60, AMT17, D326, JC32, D361 and AMT21 along with the person responsible for sample collection. Affiliations: [1] National Oceanography Centre Southampton, University of Southampton; [2]Plymouth Marine Laboratory, UK; [3]GEOMAR Helmholtz-Zentrum fur Ozeanforschung, Germany; [4]School of Environmental Sciences, University of Liverpool, [5]School of Environmental Sciences, University of East Anglia, [6]Department of Marine Chemistry and Geology, Royal Netherlands Institute for Sea Research, [7]College of Life and Environmental Sciences, University of Exeter, [8]National University of Ireland, Galway

2.2 TRICHODESMIUM CULTURING

Trichodesmium erythraeum IMS101 was grown in filter sterilised, modified YBC-II growth media (Chen et al., 1996). Trace metal concentrations were modified in line with the methods reported by Shi and Kranz (2012) to result in Cu - 8 nM, Zn - 20 nM, Co - 8 nM, Mn - 18 nM, Mo - 100 nM, Ni - 20 nM, Se - 10 nM and EDTA - 20 μM. Batch cultures were grown in 25 cm^3, 0.2 μm vented sterile polystyrene flasks (Corning Inc., NY, USA). Growth conditions were maintained at 27°C, 12:12 hour light:dark cycle at 100-160 μmol photons.m^{-2}.s^{-1} and subject to gentle agitation using an orbital shaker (Orbitron, INFORS-HT). Culture inoculations were performed using late exponential phase cells. Inoculant was pre-concentrated onto 5 μm track etched polycarbonate membranes (Nucleopore, Whatman, Kent, UK), washed with iron-free modified YBC-II growth media. Cells were resuspended in a small volume of iron-free growth media before being used to inoculate new flasks. Cultures were maintained at discreet iron concentrations (0, 4, 40, 120 and 400 nM).

2.3 TRICHODESMIUM INTRACELLULAR METAL SAMPLES

Sample were collected using drift nets as previously described (Section 2.1.1). Between 50-100 colonies were picked under class 100 laminar air flow, suspended in 0.2 μm filtered seawater and filtered onto acid-cleaned (24-48 hours, 10% HCl) 25 mm diameter 4.0 μm pore size polycarbonate filters. In order to remove metals bound extracellularly the collected colonies and filters were washed using the Oxalate-EDTA solution (100 mM $Na_2C_2O_4$, 50 mM Na_2EDTA, 0.3 M NaCl, 0.01 M KCl) described in Tang and Morel (2006). The extensive pre-cleaning of a citrate containing Oxalte-EDTA wash solution described in Tovar-Sanchez et al. (2003) is thought to reduce the solution Fe content by ≈30 fold. However this cleaning was deemed impractical given the duration of both research cruises and the solutions comparatively short shelf life, as a result the citrate free solution described in Tang and Morel (2006) was used. Tang and Morel (2006) showed the removal of citrate from this solution effectively reduced the Fe contamination by half. They then went on to show that extensive rinsing of the sample and filter with trace-metal clean water was an effective way of eliminating contamination for all elements aside from Ba. As such my samples were extensively rinsed with 0.2 μm filtered seawater in order to minimise contamination from the wash solution itself (Tang and Morel, 2006). Consequently two 10 minute, 10 ml Oxalate-EDTA washes were followed by 15 rinses of low trace-metal, 0.2 μm filtered seawater. At select stations parallel samples were collected without undergoing an oxalate wash, these were used to assess the efficiency of the oxalate reagent.

2.3.1 *Acid Digestion*

Intracellular metal samples were analysed in a class 100 clean laboratory under laminar flow. Frozen samples were defrosted and filters unfurled onto the sides of 25 ml PTFE digestion vessels (Savillex). Samples were digested with 0.5 ml concentrated HNO_3 (sub-boiling distilled, Romil SpA grade) and 1.5 ml concentrated HCl (sub-boiling distilled, Fisher scientific, PrimarPlus grade) for 4 hours on a 150 °C hotplate. Following digestion, remaining filters were removed from the digestion vessels before the digested samples were evaporated and re-suspended in 2% HNO_3. A certified reference material, NIST 1573a, was digested alongside the samples providing data on the efficiency of the extraction.

2.3.2 *HR-ICP-MS analysis*

Samples were run alongside blanks (2 % HNO_3 containing internal standards ^{115}In and ^{9}Be) and a series of multi-element standards which were prepared to using Romil PrimAg-xtra and Inorganic Ventures single element reference solutions and gravimetrically diluted as required using the same 2 % HNO_3. Standard concentrations were chosen so as to closely match the expected intracellular composition of *Trichodesmium* based of previously published data including Nuester et al. (2012), (Sañudo Wilhelmy et al., 2001) and Tovar-Sanchez and Sañudo Wilhelmy (2011).

Analysis was performed on a Thermo Scientific ELEMENT 2XR high resolution inductively coupled plasma mass spectrometer (HR-ICP-MS). The instrument had a 60 second uptake time for each sample/standard followed by a four minute wash time (2 % HNO_3). Samples were introduced via an Elemental Scientific Inc. autosampler (ESI SC2; Omaha, USA) using a pumped micro nebuliser (ESI PFA-ST) into a peltier cooled PFA cyclonic spray chamber (ESI PC_3). Low-resolution (LRM, $m/\Delta m = 300$), medium-resolution (MRM, $m/\Delta m = 3000$) and high-resolution (HRM, $m/\Delta m = 7500$) modes were utilised to avoid spectral interferences such as those between ^{111}Cd & ^{98}Mo and ^{56}Fe & ArO. Additional argon (Ar) "add" gas was ported into the spray chamber to improve sensitivity and signal stability. The instrument used the standard torch, sapphire injector and guard electrode, along with adjusting ion lens settings to maximise sensitivity, and was tuned to achieve sensitivity of around 1.4 x 106 counts per second in low resolution mode for 1 $ng.g^{-1}$ ^{115}In. Oxide formation was monitored using the $^{238}U^{16}O^{+}/^{238}U^{+}$ ratio, which was minimised as part of the tuning procedure. Further settings are given in Table 6. Table 7 lists the isotopes measured alongside the instrument resolution used for each isotope. Internal standards of 2 ppb ^{115}In and 1 ppb ^{9}Be were used. Details on blanks, digestion efficiency and

Parameter	Value
RF Power (W	1200
Sample uptake Rate (ml .min^{-1})	0.15, pumped
Argon gas flow rates (L min^{-1}	
Cooling	15
Auxillary	0.95
Nebuliser	0.7 (via nebuliser)
Add gas	0.3 (ported into spray chamber)
Acquisition mode	E-scan, Mode 1, 3 runs x 3 passes per resolution
Mass window (%)	20 (LRM), 125 (MRM and HRM)
Search window (%)	0 (LRM), 60 (MRM and HRM)
Integration window (%)	20 (LRM), 60 (MRM and HRM)
No. of samples per peak	100 (LRM), 25 (MRM and HRM)

Table 6: Operating parameters used during ICP-MS analysis.

the efficiency of the Oxalate-EDTA wash procedure are presented in Chapter 4 and Appendix A.

2.4 SHOTGUN AND AQUA PROTEOMIC METHOD DEVELOPMENT

2.4.1 *Protein Preparation and Separation*

2.4.1.1 *Extraction optimisation*

The extraction protocol described in Richier et al. (2012) was modified in a number of ways in order to determine the optimal extraction method for environmental *Trichodesmium* samples. Two alternate extraction buffers were trialled, Novex Tris-Glycine SDS Sample Buffer (SDS), an ionic detergent and Triton X-114 Detergent Solution (X-114), a non-ionic detergent. An additional protein precipitation step was introduced using MeOH and Chloroform (Wessel and Flügge, 1984) in order to remove possible contamination from disintegrated glass-fibre filter remnants, salts and detergents. The precipitation stage also allowed for a 5-10 fold concentration of the extracted protein sample. Total ion chromatograms (TIC) for the differing extraction buffers show a greater concentration of extracted protein was achieved using the SDS buffer than when using the X-114, the TIC also illustrates the level of pre-concentration of the samples using the MeOH/$CHCl_3$ precipitation step (Figure 14). The MeOH/$CHCl_3$ precipitation was superseded by a -20° C Acetone precipitation which allowed for the assessment of chlorophyll concentration in the supernatant whilst maintaining the 5-10 fold concentration of the extracted protein. The Acetone precipitation method resulted in marginally

Isotope	Instrument Method	Instrument Resolution (m/Δm)
^{111}Cd	LRM	300
^{115}In	LRM	300
^{118}Sn	LRM	300
^{98}Mo	LRM	300
^{9}Be	LRM	300
^{9}Be	MRM	3000
^{27}Al	MRM	3000
^{31}P	MRM	3000
^{51}V	MRM	3000
^{55}Mn	MRM	3000
^{56}Fe	MRM	3000
^{59}Co	MRM	3000
^{60}Ni	MRM	3000
^{65}Cu	MRM	3000
^{66}Zn	MRM	3000
^{115}In	MRM	3000
^{75}As	HRM	7500

Table 7: [Summary of analytes observed and methods utilised for ICP-MS analysis.] Summary of analytes observed and methods utilised for ICP-MS analysis.

lower protein recovery however it was deemed acceptable given it also provides chlorophyll concentration. The use of the acetone precipitation allowed segregation and concentration of the protein from the extracted sample. As such the non-protein components of the extracted sample (e.g. lipids and chlorophyll) were conserved in the acetone-based supernatant. Fluorometric analysis of this chlorophyll-containing supernatant allowed for effective calculation of extraction efficiency whereby % extraction is equal to chlorophyll concentration in supernatant / in-situ chlorophyll concentration (for environmental) or the Chl concentration of non-extracted replicate sample (for culture samples). This optimised extraction procedure routinely resulted in an extraction efficiency of ≈77% (±6%).

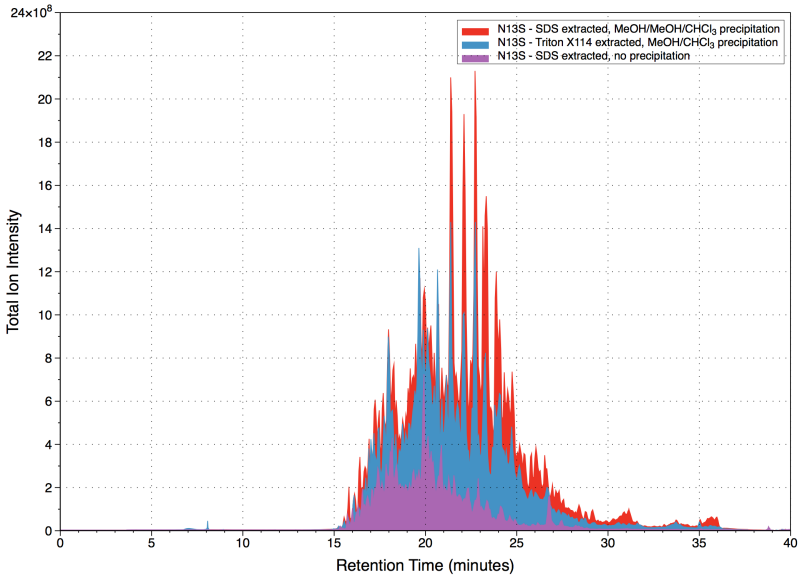

Figure 14: Total ion chromatogram (TIC) showing SDS extracted samples pre-precipitation (purple), precipitation via MeOH/Chloroform (Red) and Triton X114 extracted MeOH/Chloroform precipitated samples (Blue).

2.4.1.2 *Total Protein Quantification*

Total protein concentration was determined using the Pierce BCA Protein Assay Kit (No. 23225) based on the bicinchoninic acid assay first described by Smith et al. (1985). Analysis was performed in 96-well plates using bovine serum albumin (BSA) as a comparative protein standard allowing for a standard range of 0-2000 µg ml^{-1}. Absorbance was measured at 544 nm using a FUKOstar Optima plate reader (BMG Labtech). In order to avoid fluorometric interference from chlorophyll and other photosynthetic pigments the assay was performed on samples post-precipitation. Due to low protein sample concentration and the relatively high biomass requirement of the

BCA assay, total protein concentration was only assessed at this stage of the sample preparation. Qualitative assessment of protein concentration throughout further stages was done by spotting 1 µL of extract onto nitrocellulose membrane followed by amido-black staining.

2.4.1.3 *Protein isolation*

Following protein reduction (10 mM dithiothreitol, 80°C for 5 minutes) and alkylation (50 mM iodoacetamide, RT, dark for 30 minutes) the proteins were isolated from any remaining contaminants and extraction buffers. Isolation was either by gel electrophoresis using Novex 10% non-graduated polyacrylamide gels or by nitrocellulose membrane spotting. Similar to the method described in Alldridge and Metodieva (2008) protein electrophoresis was halted upon separation of protein sample from SDS buffer at the interface between polyacrylamide gel loading and resolving densities (4% and 10% respectively). The protein containing gel band was then excised and cut into 1 mm^3 cubes. Nitrocellulose membrane spotting involved the direct spotting of extracted proteins onto a fragment of nitrocellulose membrane. Excised membrane spots were then gently washed in 5x Coomassie destain solution and 5 x MilliQ water before digestion.

2.4.1.4 *Proteolysis*

Proteolysis was performed using the serine protease Trypsin (Promega, V5113). Trypsin cleaves peptide chains at the carboxyl side of either amino acid lysine (Lys or K) or arginine (Arg or R) resulting in shorter peptide chains more suited for mass spectrometry analysis. Proteolysis was performed on either excised polyacrylamide cubes or nitrocellulose membrane spots using 25 µL 40:1 modified Trypsin at 30°C for 12-18 hours. Digested peptides were then re-suspended in 25 mM Ammonium bicarbonate, 0.1% Formic acid solution ready for analysis. For AQUA analysis this running solution was prepared with 20 $fmol.\mu l^{-1}$ of each chosen AQUA target peptide.

2.4.1.5 *Data Acquisition*

Shotgun Analysis – The peptide mixtures generated by membrane spotted digestion were analysed as described in Greenwood et al. (2012). In brief, analysis was performed using a hybrid LTQ/Orbitrap Velos instrument (Thermo Fisher) interfaced to a split-less nano scale HPLC (Ultimate 3000, Dionex). The peptides were injected by autosampler, desalted and concentrated online at a flow of 1 $\mu L.min^{-1}$ on a 2 cm long, 0.1 mm i.d. trap column packed with 5 µm C18 particles (Dionex). Following concentration/desalting the peptides were separated along a 90-min gradient of 2-30% (v/v) acetonitrile in 0.1% (v/v) formic acid with a flow rate of 0.3 $\mu L.min^{-1}$. The separation column was a 15 cm, long 0.1 mm i.d. pulled tip packed with 5 µm C18

particles (Nikkyo Technos Co., Tokyo, Japan). The eluting peptides were electrosprayed directly from the packed tip into the LTQ/Orbitrap Velos mass spectrometer by applying a voltage of 1.75kV via a liquid junction interface. The LTQ/Orbitrap Velos was operated in the Top20 data-dependent mode, here two high resolutions scans of 30,000 (at 400 m/z) were followed by MS^2 scans for each of the 20 most abundant charged peptide ions. During the high-resolution MS^1 scans the orbitrap analyser was set to accumulate 10^6 ions for the maximum 0.5 s. During MS^2 scans the LTQ was set to accumulate 5000 precursor ions for the maximum 0.1 s. The normalised collision energy was set to 30; minimum signal intensity required was 500, activation time of 10 (ms) and activation Q to 0.250. A dynamic exclusion was implemented to avoid repetitive analysis of abundant peptide ions: after a peptide ion has been analysed once its m/z was put in the exclusion list for 30 s. The mass calibration was internal by means of lock mass. The ambient ion of 445.12 m/z was used for this purpose.

AQUA Analysis – Data acquisition for the AQUA method was similar to that described above with the following alterations. Following concentration and desalting the peptides were run along a shorter 40-min gradient. Heavy and native peptide m/z values were set as filter windows in place of the high-resolution scan for the determination of the top 20 most abundant peptide ions. Given the set filter windows no exclusion list was applied.

2.4.2 *Data Processing and Database searching*

2.4.2.1 *ESI-LC-MS/MS - Shotgun Analysis*

Preliminary MS/MS data (used for the picking of heavy labelled peptide targets) were analysed using the X!Tandem (Craig and Beavis (2004)) and Mascot (Cottrell and London (1999)) search engines. Database searching was performed using a protein translation of the *Trichodesmium erythraeum* IMS101 genome downloaded from Uniprot in June 2012. X!Tandem and Mascot search parameters were as follows: Enzyme specificity was set to trypic; fixed modifications included cysteine alkylation to carbamidomethyl; variable modifications included N-terminal acetylation and methionine oxidation. Subsequent MS/MS data were analysed using the MaxQuant software package (version 1.3.0.5). Database searching was done using the Andromeda search engine (Cox and Mann (2008)). Again, the most recent Uniprot proteome for *Trichodesmium erythraeum* IMS101 was used. Andromeda search parameters were: Enzyme specificity was set to tryptic, fixed modifications included cysteine alkylation to carbamidomethyl, variable modifications included N-terminal acetylation and methionine oxidation, and a maximum 2 missed cleavages were permitted.

2.4.2.2 *ESI-LC-MS/MS - AQUA Analysis*

Data processing for AQUA quantification was done using Skyline (Version 1.4.0.4222). Peptide settings included: Enzyme - Trypsin [KR|P], Missed cleavages - 0, Structural Modification - Carbomidomethyl Cysteine, Isotope label type - C-term K [$^{13}C(6)^{15}N(2)$] and C-term R [$^{13}C(6)^{15}N(4)$]. A *Trichodesmium* spectral library was built using previous X!Tandem, Mascot and Andromeda search results in order to aid in peptide searching. Transition settings remained as defaults aside from: Precursor mass analyser - QIT, Acquisition method - Targeted, Product mass analyser - QIT. Finally the 'Integrate All' option was enabled.

2.4.3 *Labelled Peptide Selection*

A prerequisite of the AQUA method is the use of isotopically labelled synthetic peptides for use as internal standards (henceforth known as AQUA targets). Candidate peptides for my proteins of interest were selected from preliminary shotgun data. Given *Trichodesmium*'s unique iron requirement detailed in Chapter 1 we selected two proteins from each of the major iron containing complexes. These included PsaC and PsaB for PSI, PsbA and PsbC for PSII, NifK and NifH for nitrogenase and PetB for the cytochrome b_6f complex (Table 8). Alongside these proteins were a number of proteins involved in Fe or P acquisition or compensation. These included the iron stress protein IsiA, the iron storage protein Ferritin (DpsTery), an alkaline phosphatase protein PhoA, a phosphonate uptake protein PhnD and two lipid synthesis proteins, one representing phospholipid synthesis (CDP) and the other sulfolipid synthesis (SqdB) (Table 8). In addition RplC, a protein subunit of the 50S Ribosomal multi-subunit complex was used as a house keeping proteins.

The amino-acid composition of the selected peptide along with its fragmentation spectra were carefully considered based on the following criteria from Lange et al. (2008):

- Peptide must be proteotypic to the protein of interest.
- For each peptide multiple fragmentation spectra were assessed for a clear, reproducible fragmentation pattern. Certain amino-acids, such as proline (P), are more readily fragmented by CID.
- Peptide mass was constrained between 800-2400 Da.
- Cysteine (C, Cys) containing peptides were avoided as cysteine is readily alkylated during sample preparation.
- Peptides containing methionine (M, Met) or tryptophan (W, Trp) were avoided as these amino-acids are prone to oxidation.

- Glutamine (G, Gln) and asparagine (N, Asn) were avoided as these amino-acids readily convert to glutamate (Q, Gln) and aspartate (D, Asp).
- Peptide sequences featuring neighbouring basic amino acids as either end member of the chain were avoided (e.g. VYLTGGF**KK**).
- The two peptides for a given protein were selected from distant regions of the protein sequence whenever possible. In doing so I aimed to better quantify the protein of interest by minimising any inconsistencies in reduction, alkylation or digestion resulting from the tertiary structure of the protein.

These criteria were adhered to as best as was possible, however due to the low abundance of some proteins of interest coupled with the inherently stochastic nature of shotgun proteomics, it was not always possible to select peptide targets that met *all* the criteria. As such a few target peptides included cysteine residues Table 8, as such these standards required alkylation prior to use so as to ensure complete conversion to carbamidomethyl cysteine. Synthetic peptides were ordered from ThermoFisher Scientific (Ulm, Germany) as either HeavyPeptide™ AQUA Basic or AQUA Crude. Given the digestion enzyme chosen was trypsin, which cleaved at at either arginine (R) or lysine (K), isotopic labelling was either on the end terminus R (^{13}C6; ^{15}N4, + 10 Da) or the end terminus K (^{13}C6; ^{15}N2, + 8 Da) (Table 8) of each peptide sequence.

Peptide Key	Protein	Protein Description	Sequence	Type	MW
PsaC_01	Tery_0454	PSI Iron-Sulfur centre	VYLGAETT[Arg(13C6;15N4)]	Crude	1019.86
PsaC_02	Tery_0454	PSI Iron-Sulfur centre	CETACPTDFLSV[Arg(13C6;15N4)]	Crude	1452.39
PsbA_01	Tery_0182/3	PSII reaction centre protein D1	EWELSY[Arg(13C6;15N4)]	Crude	992.79
PsbA_02	Tery_0182/3	PSII reaction centre protein D1	VINTWADIIN[Arg(13C6;15N4)]	Crude	1325.25
PetB_01	Tery_1137	Cytochrome b_6	LEVQALADDISS[Lys(13C6; 15N4)]	Crude	1396.55
PetB_02	Tery_1137	Cytochrome b_6	GSTSVGQATLT[Arg(13C6;15N4)]	Crude	1188.01
RplC_01	Tery_3012	50S Ribosomal Protein	LPGSTGPGTTPG[Arg(13C6;15N4)]	Crude	1208.05
RplC_02	Tery_3012	50S Ribosomal Protein	SGTLISIAPANIVGQ[Lys(13C6; 15N2)]	Crude	1576.84
NifK_01	Tery_4138	Nitrogenase MoFe beta subunit	GILIALTDG[Lys(13C6; 15N2)]	Crude	1008.21
NifK_02	Tery_4138	Nitrogenase MoFe beta subunit	IGYPIFD[Arg(13C6;15N4)]	Crude	990.86
PsbC_02	Tery_0513	PSII antenna protein, CP43	SPTGEIIFGGETM[Arg(13C6;15N4)]	Crude	1505.42
PsaB_01	Tery_0454	PSI P700 apoprotein A2	DKPVALSIVQA[Arg(13C6;15N4)]	Crude	1307.3
PsaB_02	Tery_0454	PSI P700 apoprotein A2	DFGYAFPCDGPG[Arg(13C6;15N4)]	Crude	1412.3
IsiA2_01	Tery_1667	Iron stress induced protein, CP43′	LGITPYFADSV[Lys(13C6; 15N2)]	Crude	1318.53
IsiA2_02	Tery_1667	Iron stress induced protein, CP43′	AMYFDGLYDATTQTV[Arg(13C6;15N4)]	Crude	1862.8
DpsTery_01	Tery_4282	Iron storage protein, Ferritin	AEQDVIKLL[Arg(13C6;15N4)]	Crude	1195.13
DpsTery_02	Tery_4282	Iron storage protein, Ferritin	LAAQAESLGD[Arg(13C6;15N4)]	Crude	1140.95
PhoA_01	Tery_3467	Alkaline phosphatase protein	IFGPETNAA[Lys(13C6; 15N2)]	Crude	1055.2
PhoA_02	Tery_3467	Alkaline phosphatase protein	DYDNFSEEV[Arg(13C6;15N4)]	Crude	1284.06
SqdB_01	Tery_0398	Sulfolipid synthesis protein	LDYDGVFGTALN[Arg(13C6;15N4)]	Crude	1451.31

Peptide Key	Protein	Protein Description	Sequence	Type	MW
SqdB_02	Tery_0398	Sulfolipid synthesis protein	LGTMGEYGTPNIDIEEGYI[Lys(13C6; 15N2)]	Crude	2208.49
CDP_01	Tery_0298	Phospholipid synthesis protein	LLGIPIILYTL[Lys(13C6; 15N2)]	Crude	1364.77
CDP_02	Tery_0298	Phospholipid synthesis protein	LLILAPLLALIELG[Lys(13C6; 15N2)]	Crude	1598.08
PhnD_01	Tery_4993	Phosphonate metabolism protein	ILLESEPIPGSPLAY[Arg(13C6;15N4)]	Crude	1765.79
PhnD_02	Tery_4993	Phosphonate metabolism protein	SELIMGLIPAESNEEMVE[Lys(13C6; 15N2)]	Crude	2127.47
NifH_01	Tery_4136	Nitrogenase Fe protein	LGGLICNS[[Arg(13C6;15N4)]	Basic	943.1
NifH_02	Tery_4136	Nitrogenase Fe protein	ETELIENLAA[Arg(13C6;15N4)]	Basic	1269.4
PsbC_01	Tery_0513	PSII antenna protein, CP43	LGANIGSAQGPTGLG[Lys(13C6; 15N2)]	Basic	1449.4

Table 8: Isotopically heavy labelled peptides sequences used as standards during AQUA analysis.

2.4.4 *Preliminary AQUA results*

The selected AQUA targets were trialled using extracted and digested *Trichodesmium erythraeum* IMS101 culture samples as a protein background. An upper limit for the number of MS^2 scans was set at 24 MS^2 scans per MS^1 cycle. This allowed for a maximum of 12 AQUA target and native pairs (12 AQUA standards + 12 Native peptides = 24). Given the requirement for 2 peptides per protein, a total of 6 proteins were quantifiable per instrument run.

Acquired MS^2 spectra were searched for a given peptide spectral signature in Skyline, whereby a previously acquired spectra for the peptide of interest was used as a template for searching and allowed for the calculation of a dot product similarity score to aid in correct identification. A minimum of 3 matching product ions were required per precursor ion and a minimum of 2 concurrent MS^2 spectra were required before being considered a correct identification (Figure 15). A ratio was calculated from the intensity of the target and native product ion peaks (H/L ratio) and converted to a concentration given the known AQUA target concentration of 100 fmol on column (5 µl injection of a 20 $fmol.\mu l^{-1}$ solution).

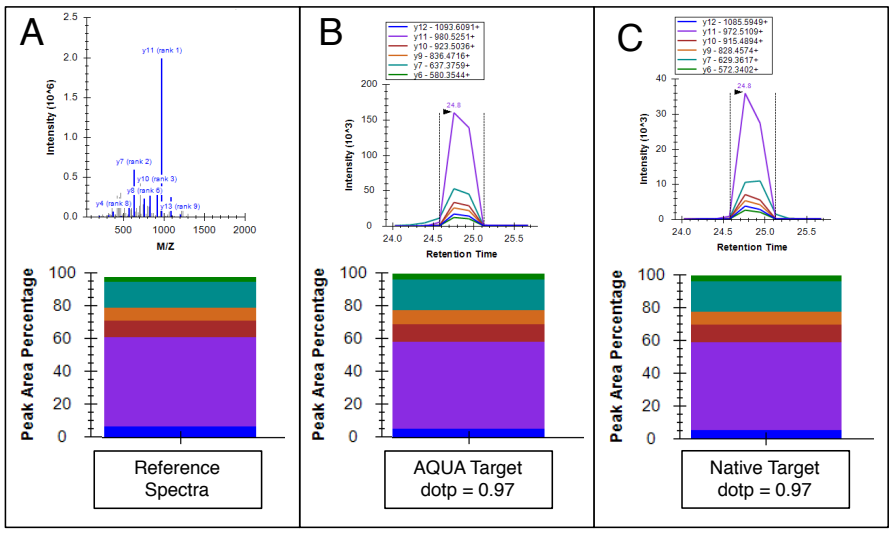

Figure 15: Example of confident AQUA peptide quantification - A) Reference spectra for my protein of interest (PsbC, amino-acid sequence - LGANIGSAQGPTGLGK), previously acquired during shotgun analysis. B) and C) correctly assigned AQUA target (B) and native target (C) spectra showing 6 identified product ion peaks across 4 concurrent MS^2 spectra. Dot product similarity score compared to reference spectra are 0.97 for both AQUA and native targets.

Each AQUA target was assessed in the manner described above, a dot product score of >0.85 was used as an indicator of a good target whilst a score <0.85 was an indicator of a poor peptide (Table 9).

The majority of AQUA targets were shown to be very inconsistent between runs with targets showing very high quality spectral matches in some analyses whilst being entirely absent in others (Table 9). An apparent distinction between peptide types is observed with the AQUA Basic peptides performing well in the standards only run whilst the AQUA Crude standards perform well when a background protein complex is used.

In the few instances where both AQUA and native targets for both peptides of a given protein (e.g. RplC_01 and RplC_02) were confidently observed the calculated H/L ratios for the two peptide pairs were not conserved. Given equal confidence in either peptides AQUA/native pair it became impossible to select one H/L ratios as being more accurate (Table 10).

The discrepancy between peptide pairs may result from numerous factors, a few of which include:

- Inefficient alkylation may inconsistently degrade a protein's tertiary structure. Here certain regions of a protein may be better exposed to the trypsin digest and so digest more efficiently.
- Non-comparable AQUA target concentrations, given the low quality of the AQUA targets used (AQUA crude) it is possible that their individual concentrations may not have been comparable. This could have resulted in the loaded '100 fmol' of each AQUA target being vastly different to one another. That being said the higher quality AQUA Basic targets, NifH_ 01 and NifH_ 02, showed equally poor reproducibility in H/L ratio.
- Differential recovery of peptides from the nitrocellulose membrane may cause discrepancies in the observed concentrations of native peptides.

Whilst AQUA quantification remains the 'gold standard' in quantitative proteomics my preliminary results demonstrated a need for further, exhaustive method development in order to get a reliable and reproducible set of AQUA targets. Given this required further method development along with the time constrains associated with the project I diverted my attention to a label-free quantification technique. Label-free quantification has in very recent years reached a level of near comparability with the AQUA method (Neilson et al., 2011). The data-independent acquisition method of MS^E eliminates the need for specific target identification and offers quantification of all observed peptides (Silva et al., 2005, 2006a).

Peptide Key	Type	Standards Only	Culture Background
PsaC_01	Crude	NA	0.93
PsaC_02	Crude	NA	n.d.
PsbA_01	Crude	NA	0.67
PsbA_02	Crude	NA	0.23
PetB_01	Crude	NA	0.94
PetB_02	Crude	NA	0.98
RplC_01	Crude	0.00	0.96
RplC_02	Crude	0.39	0.99
NifK_01	Crude	0.44	0.96
NifK_02	Crude	0.22	0.89
PsbC_01	Basic	0.97	n.d.
PsbC_02	Crude	0.65	0.96
PsaB_01	Crude	0.41	0.80
PsaB_02	Crude	0.07	0.93
IsiA_01	Crude	0.09	0.99
IsiA_02	Crude	0.61	0.83*
DpsTerry_01	Crude	NA	n.d.
DpsTerry_02	Crude	NA	0.97
PhoA_01	Crude	0.31	n.d.
PhoA_02	Crude	0.2	0.93
SqdB_01	Crude	NA	0.97
SqdB_02	Crude	NA	0.93
CDP_01	Crude	NA	n.d.
CDP_02	Crude	NA	n.d.
PhnD_01	Crude	NA	0.97
PhnD_02	Crude	NA	0.98
NifH_01	Basic	0.96	0.49
NifH_02	Basic	0.96	0.60

Table 9: 'Best match' dot product scores for the AQUA targets during either a standards only analysis or a *Trichodesmium erythraeum* IMS101 protein background. Dot product scores > 0.85 (green) indicate good peptide identification. Scores <0.85 (Red) indicate poor peptide identification. NA - not analysed, n.d. - no peptide match found. *IsiA_02, with a dot product of 0.83 is due to a noisy spectra however it is a confident identification.

Peptide Key	Native dot product	AQUA dot product	H/L Ratio
IsiA_01	0.99	0.99	0.39
IsiA_02	0.97	0.83*	2.15
RplC_01	0.95	0.96	6.4
RplC_02	0.99	0.99	4.96
NifH_01	0.94	0.92	1.18
NifH_02	0.96	0.96	0.09

Table 10: Results where both AQUA and native targets are confidently identified for both peptides of a given protein. This allows for the calculation of two independent H/L ratios for the protein of interest.

2.5 MS^E PROTEOMICS

2.5.1 *Protein Preparation and separation*

2.5.1.1 *Protein Extraction, Precipitation, Isolation and Proteolysis*

Preparation of samples for MS^E analysis involved the same protein extraction and precipitation methods as those described previously for shotgun or AQUA analysis consisting of PSB 1x extraction buffer, 3 x freeze-sonicate cycles followed by an acetone precipitation procedure. Membrane spot isolation was not attempted for MS^E samples. Protein isolation was instead only attempted by polyacrylamide gel separation as described above. Protein digestion of the excised gel cubes was achieved using 25 µL 40:1 modified Trypsin incubated at 30 °C for 12-18 hours.

2.5.1.2 *Total Protein Quantification*

The Direct Detect (Merck Millipore) system was used to determine total protein and peptide concentration at multiple stages of sample preparation. The Direct Detect system uses infra-red spectroscopy to excite a protein or peptide's amide bonds. Amide bonds are an intrinsic component of every protein/peptide thus offering a quantification method independent of amino-acid composition. The Direct-Detect measures the intensity of the Amide I peak of the IR spectrum from a 2 µL protein/peptide sample and subtracts any buffer interference. This small sample volume and compatibility with multiple buffer solutions including SDS/LDS based buffers allowed for multiple protein quantification measurements throughout sample extraction/precipitation and crucially post-digestion.

2.5.2 *Data Acquisition*

1D-UPLC-MS^E - Separations were performed using a nanoAcquity UPLC system (Waters). 1.5 µl of the prepared protein lysates (500 ng on column) containing 100 fmol of the internal standard digests Phosphorylase B and Enolase digest were injected onto a Symmetry C18, 180 µm x 20 mm trapping cartridge (Waters). After 5 min washing of the trap column, peptides were separated using a 75 µm ID x 200 mm,1.7µm BEH130 C18, column (Waters) using a linear gradient of 5 to 40% B (buffer A = 0.1% formic acid in water, buffer B = 0.1% formic acid in acetonitrile) over 90 min with a wash to 85% B at a flow rate of 300 nl/min. All separations were automated and performed on-line to a mass spectrometer.

All mass spectrometry was performed using a Waters G2-S HDMS mass spectrometer operating in MS^E mode. Data was acquired from 50 to 2000 m/z using alternate low and high collision energy (CE) scans. Low CE was 5V and elevated, ramped from 20-40V. The lock mass Glu-fibrinopeptide, $(M+2H)^{+2}$, m/z = 785.8426) at a concentration of 500 fmol/µl was infused at 250 nl/min and acquired every 13 seconds.

2.5.3 *Data Processing and Database Searching*

LC-MS^E data were processed using Protein Lynx Global Server (PLGS) version 2.5 for submission to the IDENTITY^E search engine (Waters corporation, Milford, MA). Briefly, LC-MS^E spectra were lock-mass corrected, centroided, deisotoped and charge state reduced and intensity measurements reduced (Geromanos et al., 2009). Precursor and fragment ion for each detected compound were expressed as an AMRT pair. These AMRT pairs were then submitted to IDENTITY^E for database searching. Processed LC-MS^E data was submitted to the IDENTITY^E database search algorithm version 2.5. Details of the search strategy can be found in Geromanos et al. (2009) and Li et al. (2009). Each processed file was searched against a protein translation of the *Trichodesmium erythraeum* IMS101 genome sequence acquired from Uniprot on 20/11/2012.

Search parameters were as follows: Automatic precursor and product ion tolerances were set. Enzyme specificity was set to tryptic; fixed modifications included cysteine alkylation to carbamidomethyl; variable modifications included N-terminal acetylation and methionine oxidation. A maximum of 2 missed cleavages were allowed. A false discovery rate of 4% was applied. Each peptide was assigned with $\geqslant 3$ fragment ions and each protein was identified with $\geqslant 3$ different peptides.

2.5.4 *Data Manipulation and Statistical Analysis*

Individual protein concentration was normalised between samples by dividing the observed concentration of the protein in question with the sum of all protein concentrations observed. This mitigated any variability between sample replicates and allowed for more accurate comparison. For each protein a mean concentration values from the three technical replicates was calculated, statistical significance was then determined using this mean value from each of the three biological replicates analysed per experimental condition. A two-way analysis of variance was conducted employing a Bonferroni correction so as to minimise the problem of multiple comparisons. Whilst offering a simple means to control family-wise error rate the Bonferroni correction is considered conservative and hence the total suite of differentially abundant proteins may have been an underestimate.

2.5.5 *MS^E Results*

The MS^E analysis method was employed on an iron-stressed culture experiment detailed in Chapter 5. During this experiment 4 discreet sampling points were undertaken both for low iron and high iron (designated by '+' or '-') but also for 2 and 6 hours after the onset of the photoperiod (designated by PS or NF respectively). Sampling was performed in biological triplicate each being harvested from a discreet culture flask. Analysis for each biological replicate was performed in triplicate technical replicates. Figure 16 shows the consistence observed between technical replicates for one biological replicate from each sampling point. Sequence coverage for confidently identified proteins ranged from 0.5-95.6% with a mean of $\approx$ 24% (Figure 17).

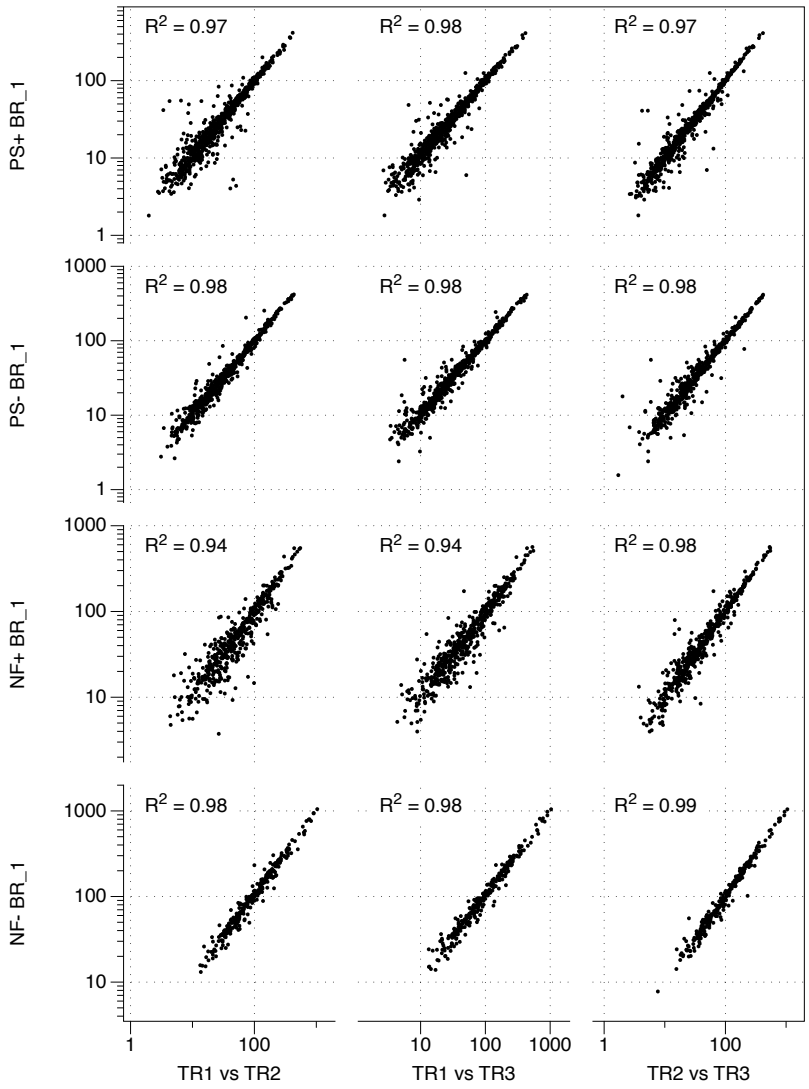

Figure 16: Log-log scatter plots showing reproducibility of technical replicates. Plots are shown only for the first biological replicate from each sampling point (PS+ BR_ 1, PS- BR_ 1, NF+ BR_ 1, NF- BR_ 1.). Correlation coefficients are displayed on each plot individually.

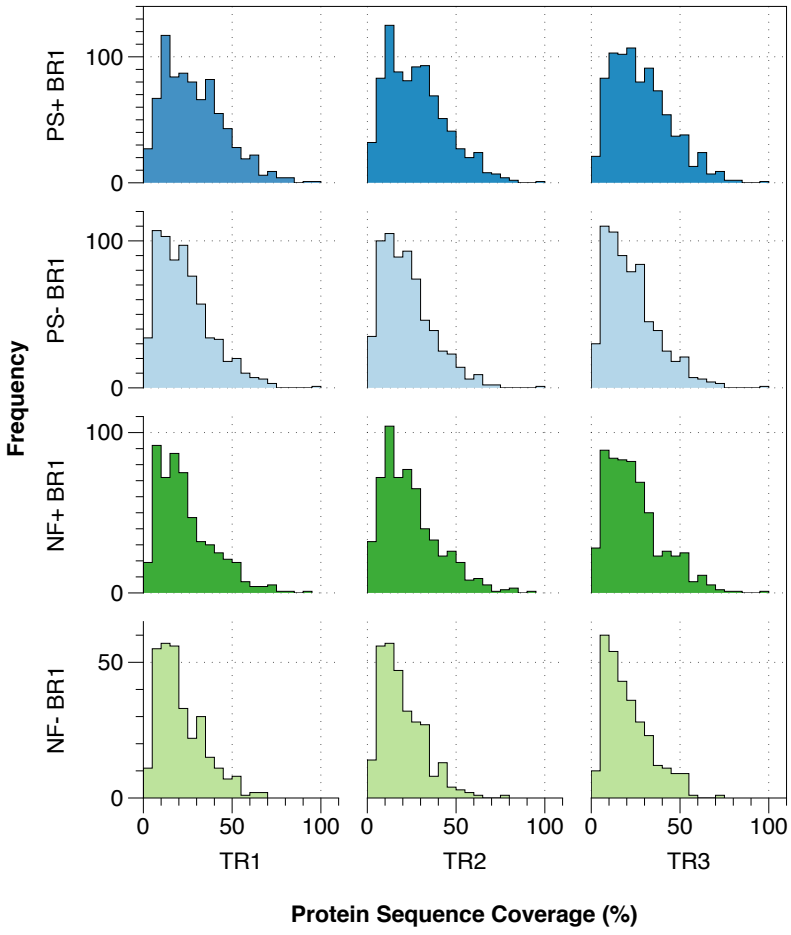

Figure 17: Sequence coverage (%) histograms for MS^E technical replicates PS+ BR_ 1, PS- BR_ 1, NF+ BR_ 1, NF- BR_ 1.

3

A RESOURCE-RATIO FRAMEWORK FOR INTERPRETING THE DIAZOTROPHIC BIOGEOGRAPHY OF THE (SUB)-TROPICAL ATLANTIC

3.1 INTRODUCTION

Trichodesmium, a colonial, diazotrophic, cyanobacterium has been a major focus of research in recent years due to the importance of this organism in influencing the nutrient biogeochemistry of the tropical and subtropical oceans and its role as a significant contributor to global biological nitrogen fixation (Sohm et al., 2011b). The balance of different environmental controls on the growth and biogeography of *Trichodesmium* currently remain a matter of debate. Effects of sea surface temperature (SST) , pCO_2 and nutrient concentrations on the growth and nitrogen fixation (N_{Fix}) capabilities of *Trichodesmium* have all been suggested to be important (LaRoche and Breitbarth, 2005; Kranz and Sültemeyer, 2009; Kranz et al., 2010; Hutchins et al., 2013; Fu et al., 2014).

It is frequently assumed that natural populations of *Trichodesmium* may be physiologically confined to a SST niche of ≈20-34°C due in part to correlative abundance observations (LaRoche and Breitbarth, 2005). The numerous culture strains of *Trichodesmium erythraeum* also displays optimal growth between 24-30°C (Breitbarth et al., 2007; Fu et al., 2014). However, phytoplankton are expected to retain temperature optima reflecting that of their isolation environment (Thomas et al., 2012). Additionally other diazotrophs, including heterocystous cyanobacteria, are prevalent in much colder environments suggesting that elevated temperatures are not a *de facto* prerequisite for diazotrophic growth (Staal et al., 2003; Stal, 2009; Pandey et al., 2004; Blais et al., 2012). Moreover theoretical studies have begun questioning the assumption that elevated temperatures are a primary driver of diazotroph biogeography arguing the importance of co-varying parameters such as oligotrophy and water column stratification. The implication being that the observed correlation between diazotroph abundance and temperature is not directly causal but reflects a secondary adaptation to conditions found within the realised niche (Monteiro et al., 2011; Ward et al., 2013; Dutkiewicz et al., 2012, 2014).

Numerous *Trichodesmium* isolates have shown elevated rates of N_2 and carbon fixation and increased cellular C:N ratios when cultured at projected end-of-century pCO_2 levels (750-1000ppm) (Hutchins et al., 2007, 2013; Barcelos e Ramos et al., 2007). Experimental dou-

bling in pCO_2 has been shown to result in CO_2 fixation rates increasing 15-128% whilst N_2 fixation increased by 35-100% (Hutchins et al., 2007, 2013; Barcelos e Ramos et al., 2007). The physiological effects of elevated pCO_2 have been attributed to a down-regulation of *Trichodesmium*'s active carbon concentrating mechanism (CCM) resulting in an increase in the availability of cellular energy resources for other processes such as CO_2 and N_2 fixation (Kranz and Sültemeyer, 2009; Kranz et al., 2010).

Investigation into the nutritional requirements of *Trichodesmium* has led to the discovery of a number of novel nutrient acquisition and utilisation strategies for iron and phosphorus. Along with all organisms *Trichodesmium* require phosphorus as an essential component of both structural and functional molecules (e.g. lipids, deoxyribonucleic acid (DNA), ribonucleic acid (RNA), adenosine triphosphate (ATP)). *Trichodesmium* sp. are capable of accessing dissolved inorganic phosphorus (DIP) as PO_4^{3-}, alongside dissolved organic phosphorus (DOP), in the form of phosphonates or phosphomonoesters, which are typically less bioavailable to microorganisms (Roy et al., 1982; Orchard et al., 2009; Dyhrman et al., 2006). Recent work has also shown the structural requirement of phosphorus for lipid formation is adaptable to conditions of low phosphorus availability whereby phosphorus containing lipids are substituted for phosphorus free alternatives in both prokaryotes (e.g. *Trichodesmium erythraeum*) and eukaryotes (e.g. *Thalassiosira psuedonana*) (Van Mooy et al., 2006, 2009).
Trichodesmium has a high absolute iron requirement as iron is an essential component for both the photosynthetic and nitrogen fixing apparatus (Shi et al., 2007; Kustka, 2003a). Within the photosynthetic apparatus photosystem I accounts 12 Fe atoms whilst the remaining photosynthetic apparatus accounts for a further 11-12 Fe atoms (Ferreira and Straus, 1994). Nitrogenase, the heterodimeric enzyme system responsible for nitrogen fixation has a stoichiometric equivalent of 50 iron atoms making it one of the most iron rich enzyme systems in nature (Raven, 1988). The iron requirement is compounded by *Trichodesmium*'s unique ability to simultaneously fix CO_2 and N_2 during the day, within the same cell, achieved by a complex combination of temporal and intercellular-spatial segregation (Berman-Frank et al., 2001b; Sandh et al., 2009). The co-occurrence of two iron-rich metabolic processes suggests *Trichodesmium* may be uniquely susceptible to low iron availability (Richier et al., 2012; Berman-Frank et al., 2007).

The interface between the north Atlantic subtropical gyre (NASG) and the south Atlantic subtropical gyre (SASG) has repeatedly been noted as an environment abundant in *Trichodesmium* which can dominate the region's diazotrophic community (Capone et al., 2005; Carpenter et al., 2004; Moore et al., 2009a; Tyrrell and Marañón, 2003; Groß kopf et al., 2012). Net down-welling within the subtropical gyres

prevents the re-supply of inorganic nutrients from deep waters, maintaining the system in a state of oligotrophy (Williams and Follows, 2003). Surface water observations have shown that whilst dissolved inorganic nitrogen (DIN) concentrations are persistently low (≈ 2-60 nM) in both gyres the concentration of DIP is significantly lower in the NASG (≈ 25 nM) when compared to the SASG (≈ 157 nM), as is evident from the SASG's elevated P* (where P* = DIP-DIN/ Redfield n/p [16]) (Deutsch et al., 2007). In contrast surface water dissolved iron (DFe) concentrations are higher in the NASG (≈0.39 nM) than in the SASG (≈ 0.12 nM) owing to the atmospheric deposition of iron-rich Saharan dust into the NASG (Jickells et al., 2005; Bowie et al., 2002; Measures and Vink, 1999; Baker et al., 2013). A significant contributor to this atmospheric deposition of iron is thought to be associated with the inter-tropical convergence zone (ITCZ). The ITCZ is the area in which the north-east and south-east trade winds meet producing increased rainfall and thus increased wet deposition of iron-rich continentally-derived dust (Schlosser et al., 2014; Jickells et al., 2005). The relative location of the ITCZ is affected by sea surface temperature (SST) and as such the ITCZ shows an asymmetrical seasonal migration (Mitchell and Wallace, 1992). This seasonally influenced iron flux is thought to play a dominant role in the differing biogeochemical environment between the two gyres (Moore et al., 2009a; Schlosser et al., 2014). The observed high rates of nitrogen fixation are thought to be fuelled by this influx of DFe which in turn depletes both DIP and DOP pools of the NASG and so drive these concentrations to minimal values, limiting further diazotrophic growth. Consequently the resupply of DIP/DOP to the NASG via lateral advection of P enriched water may potentially become the limiting factor to diazotrophy in this system (Mather et al., 2008; Moore et al., 2009a; Palter et al., 2010; Straub et al., 2013).

It is thought that nitrogen fixation may be coupled to long-term natural climate cycles whereby increases in the amount of CO_2 they sequester may reduce greenhouse forcing and subsequently cool the earth (Hutchins et al., 2007; Falkowski, 1997). As such understanding the potential environmental controls on diazotrophy is particularly important in the context of global change. Certain increases in SST and pCO_2 alongside the less-well understood changes in dust-driven iron fertilisation and physical excess P (re)-supply have all been hypothesised to influence rates of N_{Fix} resulting in a possible shift in the balance between local and global of N_{Fix} and denitrification (Hutchins et al., 2013; Moore et al., 2009a; Breitbarth et al., 2007; Falkowski, 1997; Straub et al., 2013).

In this study I assembled a data set of diazotrophic abundances and rates alongside hypothesised environmental drivers collected across 7 oceanic cruises using consistent methods in the period 2002-2011. Within these environmental data sets, observations of *Trichodesmium*

abundance and N_{Fix} rates were collected during two contrasting seasons, the N.hemisphere Spring (Feb-Mar) and Autumn (Oct-Nov). Using this comprehensive data set I aim to assess the hypothesised importance of temperature, pCO_2, iron and phosphorus variability on the distribution of *Trichodesmium* and bulk community, making particular use of well established resource-ratio theory in the interpretation of observed nutrient (Fe, P) patterns.

3.2 METHODS

3.2.1 *Sampling and hydrography*

Sampling was conducted during 7 oceanic cruises, M55 (Oct-Nov 2002, *R/V Meteor*) (Mills et al., 2004), M60 (Mar-Apr 2004, *R/V Meteor*) (Moore et al., 2006, 2008), AMT17 (Oct-Nov 2005, *RRS Discovery*) (Moore et al., 2009a), D326 (Jan-Feb 2008, *RRS Discovery*) (Richier et al., 2012), JC32 (Mar-May 2009, *RRS James Cook*), D361 (Feb-Mar 2011, *RRS Discovery*) (Schlosser et al., 2014) and AMT21 (Oct-Nov 2011, *RRS Discovery*) Figure 18. AMT17, D361 and AMT21 followed a north-south transect approximately associated with the 25°W meridian, as such these 3 cruises will be discussed in a meridional context. M55, M60, D326 and JC32 cover a broader longitudinal area therefore they will be excluded from longitudinal analysis. Precipitation data sourced from the NCEP Reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, (http://www.esrl.noaa.gov/psd/) was used as a proxy for seasonality of the ITCZ. I have defined the ITCZ's location during D361 (representative of N.hemisphere Spring) and AMT21 (representative of N.hemisphere Autumn) as the area of surface precipitation $> 2\times10^{-5}Kg.m^{-2}.s^{-1}$ (Figure 19).

3.2.2 *DIN, DIP and DFe*

For AMT21 macro-molar nitrate (NO_3^-), nitrite (NO_2^-) and phosphate (PO_4^{3-}) were measured using a Bran and Luebbe AAIII segmented flow, colorimetric, auto-analyser using standard colorimetric techniques (Brewer and Riley, 1965; Grasshoff, 1976). For cruises D326, D361 and AMT17 nanomolar NO_3^-, NO_2^- and PO_4^{3-} were determined using colorimetric methodologies as with standard segmented flow analysis however 2 metre Liquid waveguide capillary cells were used as flow cells (Patey et al., 2008). Dissolved iron concentrations were determined from surface-water samples collected using a trace-metal clean tow fish. Samples were filtered via 0.2µm cartridge filter (Acropak/Sartobran P300) and acidified with ultraclean HCl. Analysis was performed using an online flow injection analysis system with a Toyopearl column via luminol chemilumines-

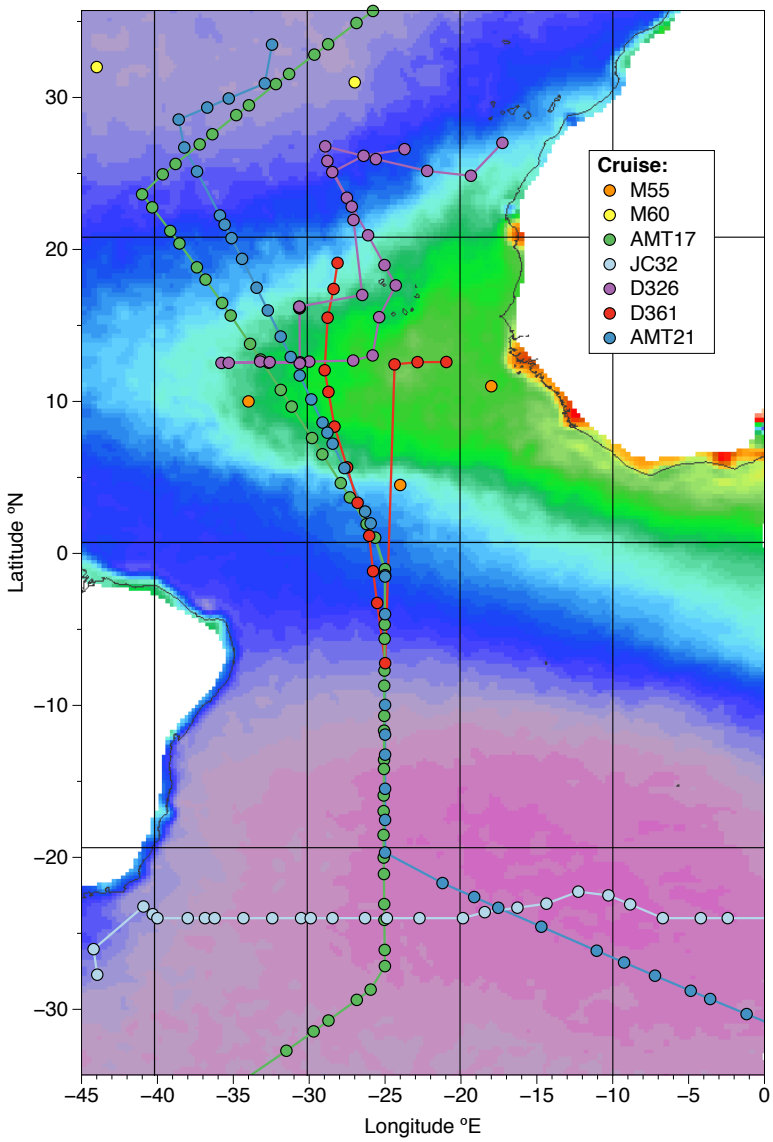

Figure 18: Cruise transects for the 7 cruises included in this study including M55, M60, AMT17, D326, JC32, D361 and AMT21. Transects are overlaid on the 2002-2011 time-averaged SeaWIFS 550nm Aerosol Optical Depth.

cence (Klunder et al., 2011) or via isotope dilution method detailed in (Chapter 2).

3.2.3 *pCO_2 and SST*

Near surface temperature (henceforth assumed to be SST) was measured throughout all 7 cruises using a Seabird 38 or similar temperature sensor mounted to the underside of the vessel. During each of the 7 cruises dissolved inorganic carbon (DIC) and total alkalinity (TA) were sampled from near-surface fired Ocean Test Equipment (OTE) bottles. DIC analysis was done using the standard coulometric technique described by (Dickson et al., 2007) whilst TA was analysed using standard titrimetric techniques (Dickson et al., 2007). The partial pressure of CO_2 was derived from these parameters following equations outlined in Dickson et al. (2007).

3.2.4 *Trichodesmium sp. abundance*

Entire 20 L OTE bottles fired near the surface (≈2 m) were gravity filtered through polycarbonate filters (Millpore Isopore, 10 μm pore size, 47 mm diameter). Colonies and free trichomes were gently agitated to remove from the surface of the filter before being preserved in 2% Lugol's iodine and stored in the dark. Filters were visually inspected for complete re-suspension of *Trichodesmium* colonies and free trichomes. Colony and free trichome abundance, colony morphology and approximate colony size were calculated and enumerated using light microscopy. The limit of detection for *Trichodesmium* abundance was equivalent to 0.05 Trichomes.L^{-1}.

3.2.5 *N_2 fixation rates*

Whole community N_{Fix} - Nitrogen fixation rate measurements were conducted on surface water samples collected from a pre-dawn CTD cast, incubations were performed following the method described by Montoya et al. (1996). For the whole community, 4.5 L polycarbonate bottles (Nalgene) were filled, ensuring no air bubbles remained, and sealed with rubber septa containing screw-caps. Each bottle was spiked with 4 ml $^{15}N_2$ (Cambridge Isotope Laboratory, UK) gas and incubated for 24 hours at sea surface temperature (SST) and ≈97 % sea surface irradiance (SSI) achieved using the ship's underway seawater supply system and optical filters (0.15 neutral density, LEE). Following incubation each sample was filtered onto a pre-ashed GF/F filters (450 °C for 12 hours, Whatman GF/F or Fisherbrand MF300), folded into a 1.5 ml tube (Eppendorf) before being dried for 24 hours at 40 °C. Upon return to the lab samples were fumed for 24 hours in a vacuum desiccator containing a small volume of sulphurous acid so

as to remove any particulate inorganic carbon. Each sample and filter was encapsulated in a tin disc and analysed for organic nitrogen and $^{15}N/^{14}N$ ratio with an Elemental Analyser Isotope Ratio Mass Spectrometer (Thermo Finnigan EA 1112 Series Flash Elemental Analyser). For *Trichodesmium* specific N_{Fix} rates, 50 colonies were isolated using plastic inoculation loops and placed into 125 ml polycarbonate bottles (Nalgene) filled with 0.22µm filtered trace-metal clean surface water. Bottles were filled entirely, sealed with rubber septa containing screw-caps and spiked with 0.5 ml $^{15}N_2$. Bottles were incubated for 12 hours at SST and sea surface irradiance before being filtered onto pre-ashed GF/F filters, dried and stored until analysis according to the method described for the whole community. A systematic $\approx$2 fold underestimation of the absolute N_2 fixation rates is likely to have been observed due to the method utilised (Groß kopf et al., 2012; Mohr et al., 2010; Wilson et al., 2012), however given that observed community rates and *Trichodesmium* specific rates spanned >2 and >3 orders of magnitude respectively such underestimation has no bearing on our conclusions which are based on the observed interrelationship between variables within the data set.

3.3 RESULTS

3.3.1 *Hydrography and the ITCZ*

I observed the region of maximal precipitation during the D361 between 4 °N and 8 °S and between 11.5 °N and 2 °N during AMT21 (Figure 19). The ITCZ is at its Northern and Southern most position in July and January respectively (Philander et al., 1996; Xie and Saito, 2001). Given the dates of D361 and AMT21 are mid-phase in relation to this seasonality it is likely that there is a greater latitudinal migration than our observations suggest.

3.3.2 *pCO_2 and SST*

Sea surface temperatures ranged from 19.6 - 28.7 °C with one JC32 station located in the upwelling region to the west of Africa having a lower temperature of 16.4 °C (23.7 °S, 13.7 °E). All observed temperatures are within previously published *Trichodesmium erythraeum* IMS101 tolerances (20-34°C) (Breitbarth et al., 2007; LaRoche and Breitbarth, 2005). pCO_2 varied across the basin from 332 - 464 µAtm (mean $\pm$SD = 387 $\pm$26, n = 91).

3.3.3 *DIN, DIP and DFe*

Throughout the sample region surface nitrate concentrations remained consistently low with a mean concentration of 12$\pm$10 nM

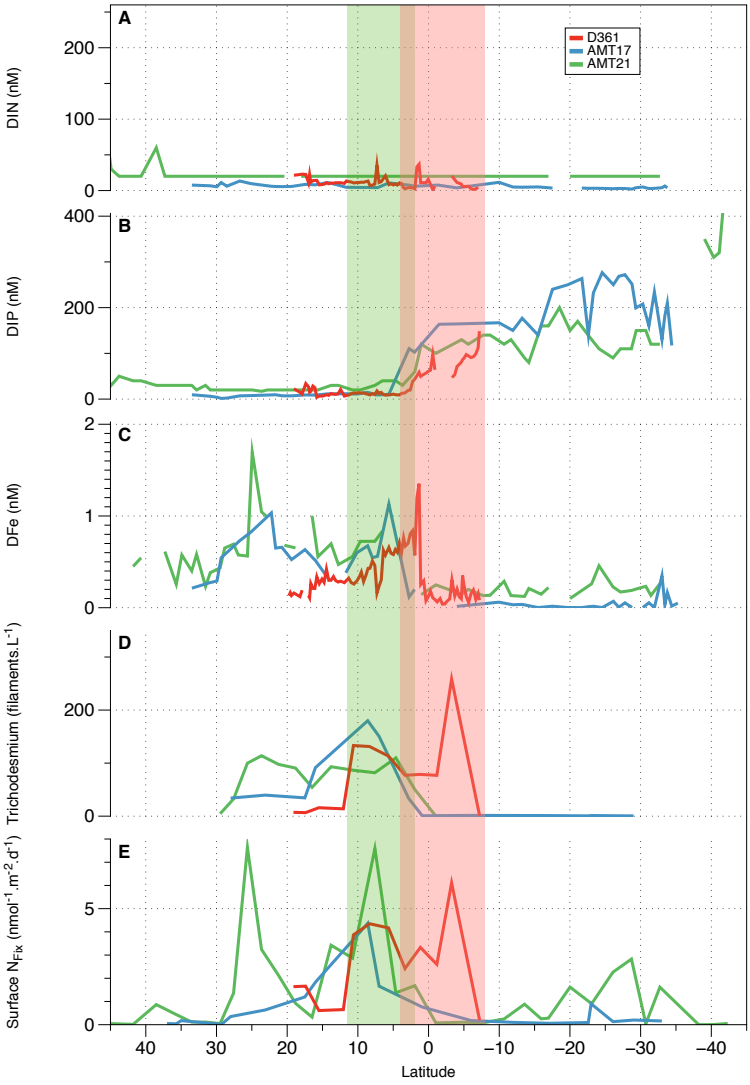

Figure 19: Data collected from 3 cruises along an Atlantic meridional transect along ≈25°W. Cruises shown are D361 (red), AMT17 (blue) and AMT21 (green). A) Dissolved inorganic nitrogen, DIN (nM), B) Dissolved inorganic phosphorus, DIP (nM), C) Dissolved Iron, DFe (nM), D) *Trichodesmium* abundance (Trichomes.L^{-1}) E) Whole community surface N_{Fix} rates (nmol.m^{-2}.d^{-1}). Shaded regions represent >2x10^{-5} Kg.m^{-2}.s^{-1} surface precipitation rate derived from NCEP reanalysis daily average surface flux data and approximately describe the location of the ITCZ during D361 (7th February - 19th March 2011, red) and AMT21 (29th September - 14th November 2011, green).

(mean ± SD) (Figure 19A). The NASG when compared with the SASG consistently displayed greatly diminished DIP concentrations (22 ± 19 nM and 126 ± 60 nM respectively) (Figure 20A and Figure 19B). Conversely the distribution of DFe revealed consistently elevated concentrations in the Saharan-dust influenced NASG (0.42 ± 0.33 nM) when compared with the SASG (0.13 ± 0.11 nM) where little dust-deposition is evident (Jickells, 2006) (Figure 20B & Figure 19C). During both N.hemisphere spring (D361) and autumn (AMT17 & 21) I observed a seasonal consistency in DIP concentrations along this latitudinal range. The distribution of DFe is observed to migrate north between spring and autumn, thought to be caused by the seasonal migration of the ITCZ (Schlosser et al., 2014) (Figure 19). AMT 21 and 17 show elevated DFe concentrations in the central NASG around 20 - 5 °N whilst D361 (red) shows this peak shifted south between 5 °N and 0 °N.

3.3.4 *Trichodesmium Biomass*

Trichodesmium colonies and free trichomes were observed during all 7 cruises. Near-surface *Trichodesmium* abundance data shows the densest populations residing in the southern region of the NASG and the equatorial region (≈15 °N - 7 °S) (Figure 20). *Trichodesmium* was consistently abundant throughout the NASG before sharply declining to undetectable concentrations at ≈0 °N for both AMT 21 & 17 whilst the observed population maxima during D361 is located further south with *Trichodesmium* seen between 19 °N and ≈7 °S (Figure 19). *Trichodesmium* was not observed in the central SASG during any of these cruises.

3.3.5 *Nitrogen Fixation*

Surface nitrogen fixation rates observed across large regions of the North and South (sub-)tropical Atlantic correlated well with *Trichodesmium* biomass (R^2=0.55) particularly at higher N_{Fix} rates / *Trichodesmium* abundances (Figure 21). At lower *Trichodesmium* abundance this relationship decouples, likely representing the contribution to measured N_{Fix} rates attributed to the non-*Trichodesmium* diazotrophic community representing a higher proportion of the total community rate. Mean *Trichodesmium* specific nitrogen fixation rates assessed during both D361 and AMT21 (n= 24) highlight the increased contribution of the non-*Trichodesmium* diazotrophic community at lower N_{Fix} rates, however our data show elevated whole community N_{Fix} rates are a strong indicator of *Trichodesmium* specific N_{Fix}.

Surface community N_{Fix} rates show a similar north-south divide as observed for *Trichodesmium* abundance. The bulk of diazotrophic

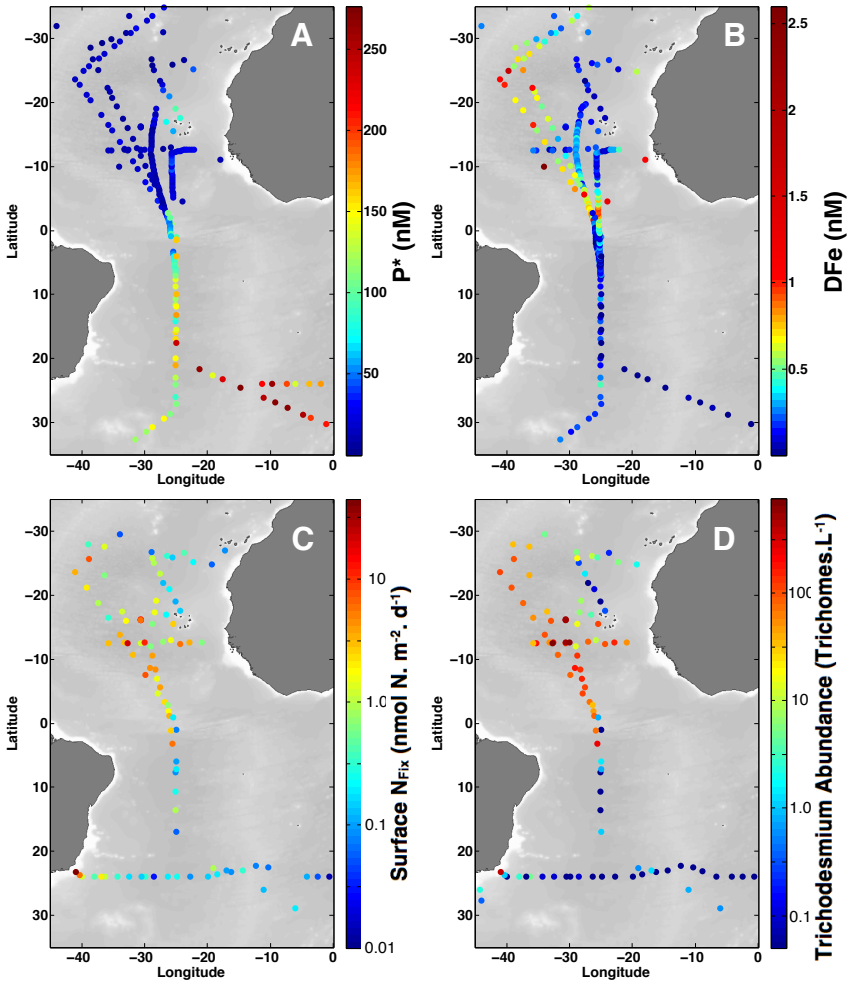

Figure 20: Data collected from cruises M55 (Oct-Nov 2002, R/V Meteor), M60 (Nov-Dec 2003, R/V Meteor), AMT17 (Oct-Nov 2005, RRS Discovery), D326 (Jan-Feb 2008, RRS James Cook), JC32 (Mar-May 2009, RRS James Cook), D361 (Feb-Mar 2011, RRS Discovery) and AMT21 (Oct-Nov 2011, RRS Discovery). A) P* = $DIP - DIN/r_{N/P}$), B) Dissolved Iron (nM), C) Whole community surface nitrogen fixation (nmol.m^{-2}.d^{-1}), D) *Trichodesmium* abundance (Trichomes.L^{-1}).

activity is observed in the southern flanks of the NASG and the equatorial region (Figure 20). A strong seasonality is observed between spring (D361, red) and autumn (AMT17 & 21, green) cruises where peak surface N_{Fix} rates follow the pattern of DFe and reside further south during D361 when compared with either AMT cruise (Figure 19).

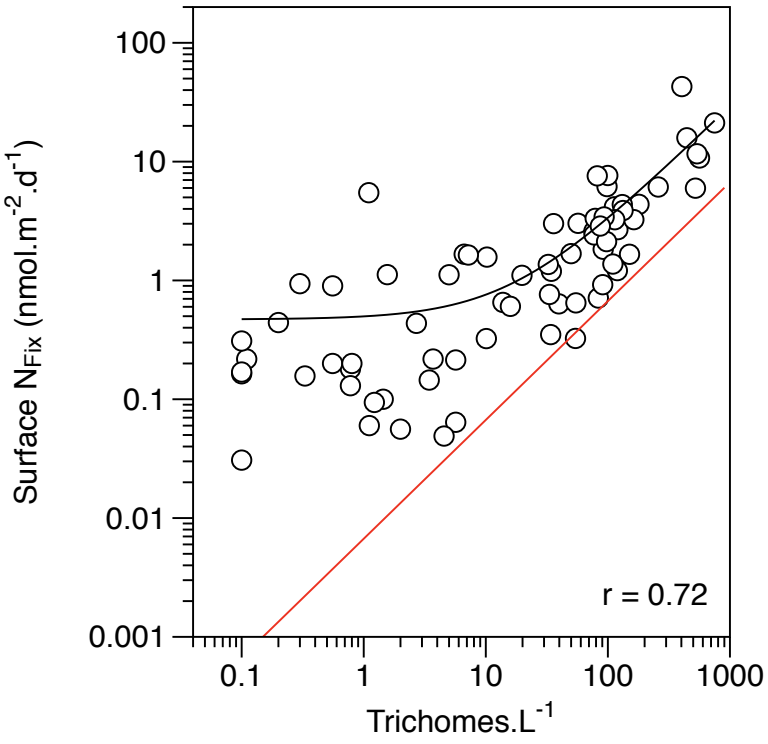

Figure 21: Correlation between *Trichodesmium* abundance (Trichomes.L^{-1}) and whole community N_{Fix} rates (nmol.m^{-2}.d^{-1}). Linear regression is shown as a black line, Pearson correlation coefficient (r) = 0.72. Independently calculated mean nitrogen fixation per Trichome (nmol N. Trichome^{-1}.d^{-1}) is shown in red.

In addition to the *Trichodesmium* attributed diazotrophic activity of the NASG and equatorial Atlantic, our observation show diazotrophy in the western-SASG (Figure 20 and Figure 22). Depth integrated comparison of NASG and SASG N_{Fix} along the AMT21 transect show similar rates between the two gyres with ≈1.1 and ≈0.90 mol N.$m^{-1}d^{-1}$ fixed in the NASG and SASG leg of this transect respectively (Figure 23). Furthermore the natural abundance of ^{15}N relative to ^{14}N observed during JC32 show a strong longitudinal shift with a lower $\delta^{15}N$ PON values observed in the western Atlantic than in the eastern Atlantic (Figure 23). This reduction in $^{15}N/^{14}N$ is a clear indicator of diazotrophic activity (Montoya et al., 1996).

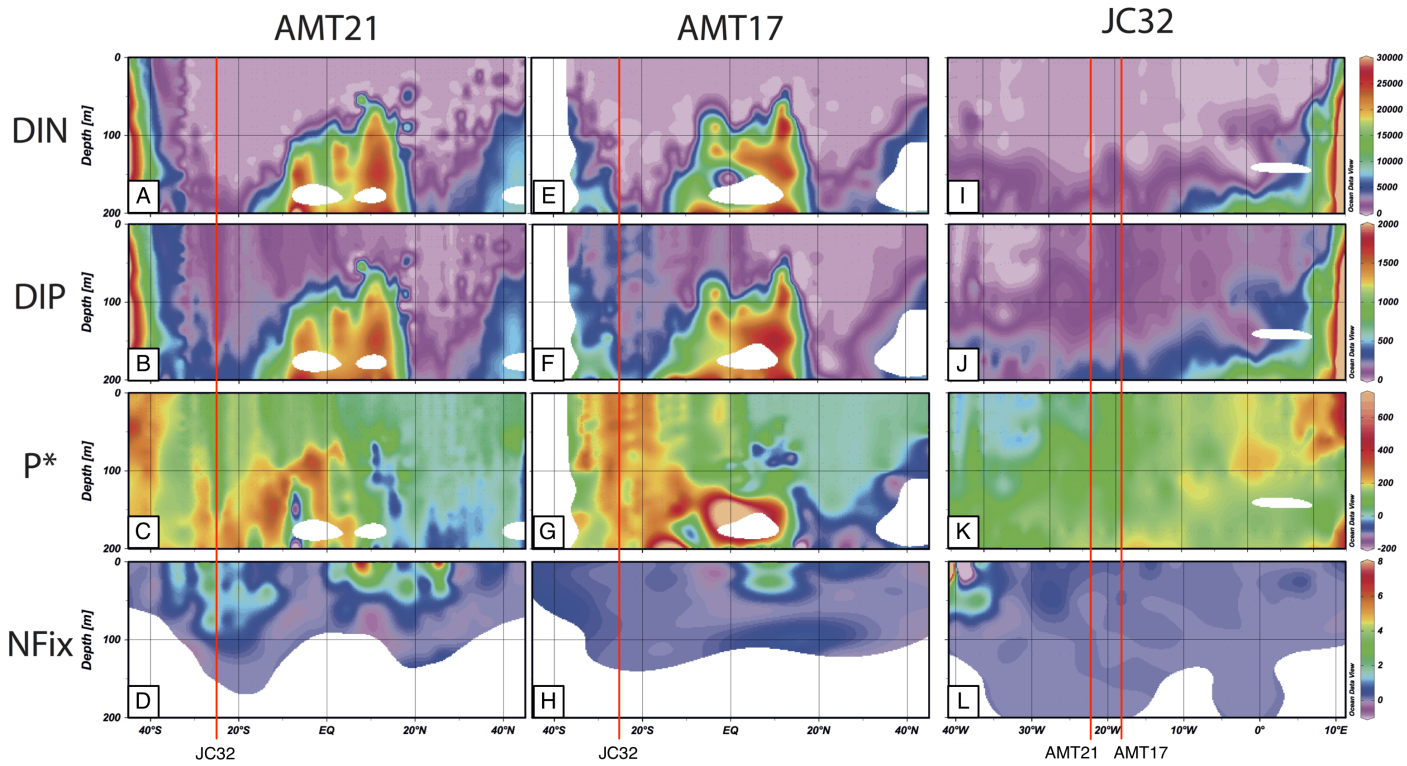

Figure 22: Contour plots showing DIN (nM), DIP (nM), P* (nM) and N_{Fix} ($nmol^{-1}.m^{-2}.d^{-1}$) within the euphotic zone for AMT21 (A,B,C and D respectively), AMT17 (E, F, G and H respectively) and JC32 (I, J, K and L respectively). Red lines indicate points of cruise intersection with JC32 crossing the AMT21 transect at ≈25 °E, crossing AMT17 at ≈17.25 °E. AMT17 and AMT21 both cross the JC32 transect at ≈24-25 °N

3.4 DISCUSSION

3.4.1 *Seasonality and the ITCZ*

The seasonal synchronicity between precipitation and dissolved iron distributions in the equatorial Atlantic as described in Schlosser et al. (2014) is further supported in this study with the inclusions of the previously unpublished N_{Fix} and *Trichodesmium* abundance data from AMT21 (Figure 19). Schlosser et al. (2014) show that this synchronicity extends to an observed seasonal migration of *Trichodesmium* populations and their attributed nitrogen fixation. The NASG, with its consistent atmospheric resupply of DFe supports a low abundance yet persistent diazotrophic community who remove the DIP from the surface water thus preserving the state of DIP oligotrophy. Conversely the SASG lacks a consistent DFe supply, as such diazotrophy is unsupported and DIP can accumulate (Moore et al., 2009a; Schlosser et al., 2014). Between these two extremes resides the optimal region of *Trichodesmium* growth whereby high biomass is observed in a presumably favourable environment for DIP/DOP resupply (Mather et al., 2008). Our observations support a mechanism analogous to that of seasonality in higher latitudes whereby winter resupply of nutrients through convective mixing and favourable light conditions allow for the formation of a spring bloom (Honjo and Manganini, 1993). Here I observe spatial changes in DFe and DIP supporting a region of elevated *Trichodesmium* abundance responsible for an increase in N_{Fix}.

3.4.2 *Linear Regression Analyses*

Relationships between environmental and diazotroph related variables were initially investigated using simple linear regression analysis. A weak yet significant correlation was observed between surface N_2 fixation and DIP however no other significant correlation was found between SST, pCO_2 DIP or DFe and *Trichodesmium* abundance or SST, pCO_2 or DFe and surface N_2 fixation rates (Figure 24 and Table 11). This is in contrast to the findings of Luo et al. (2013) who observed correlations between temperature and N_{Fix} across a broader temperature and geographical range. Such poor correlation suggest SST or pCO_2 are not important controls over the study region given that neither potential driver can be strongly reciprocally influenced by diazotrophy. It is apparent that variations in SST and pCO_2, across the observed environmental ranges, are of limited predictive value on size and location of *Trichodesmium* populations or their nitrogen fixation activity in the (sub)-tropical Atlantic (Figure 24). Further, I agree with the conclusions of LaRoche and Breitbarth (2005) and Ward et al. (2013) that the correlation of observed SST and *Trichodesmium* populations is unlikely an indication of a direct relation-

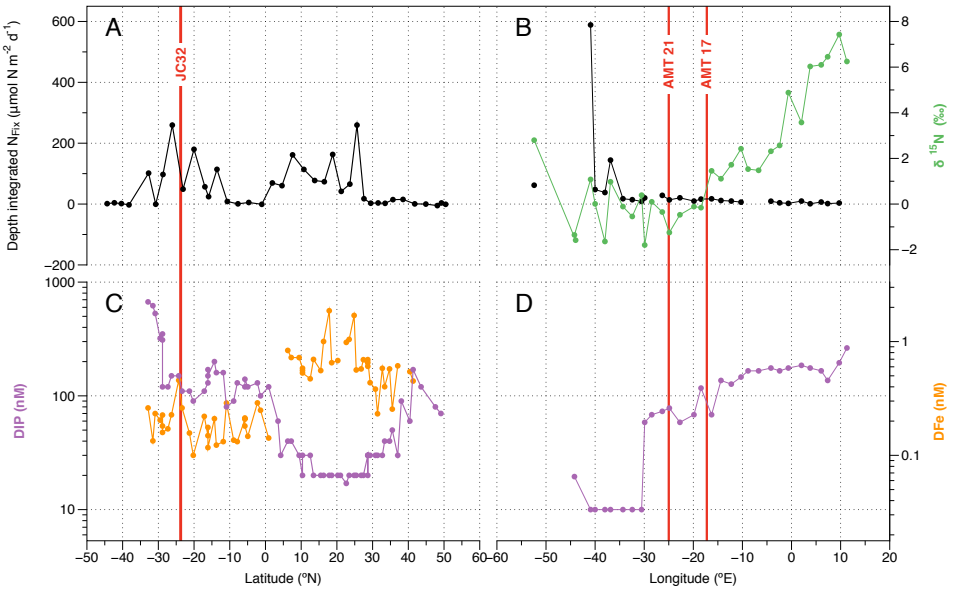

Figure 23: Depth integrated N_2 fixation rates and $\delta^{15}N$ natural abundance data observed during AMT21 (A, N_{Fix} only) and JC32 (B). DFe (nM, orange) and DIP (nM, purple) surface water concentrations observed during AMT21 (C) and JC32 (D, DIP only). Red lines indicate points of cruise intersection with JC32 crossing the AMT21 transect at ≈25 °E, crossing AMT17 at ≈17.25 °E. AMT17 and AMT21 both cross the JC32 transect at ≈24-25 °N

ship but is rather a result of factors known to co-vary with temperature such as low-nutrients and increased stratification. Moreover I suggest this lack of a hypothesised correlation may also reflect the overriding importance of nutrient availability. However considering DIP and DFe, individually, these variables also provided poor predictive capability (Figure 24 and Table 11) with equally poor correlations as those described above for pCO_2 and SST (Figure 24). In this instance a lack of direct correlation might be expected due to the reciprocal feedbacks between nutrients and diazotrophs which, unlike SST and pCO_2, maybe strongly influenced by the activity of diazotrophs including *Trichodesmium*. Moreover I regard phosphorus and iron as essential nutrients (Tilman, 1980), with *Trichodesmium* requiring both for growth, their effects on *Trichodesmium* abundance and N_{Fix} must thus be considered together.

3.4.3 *Resource competition framework*

Resource competition theory dictates that in an equilibrium environment the species best able to utilise a given resource will deplete that resource down to a well defined minimum concentration, termed the R* value (Tilman, 1980). At steady-state this framework supports a niche for as many competing species as their are different potential

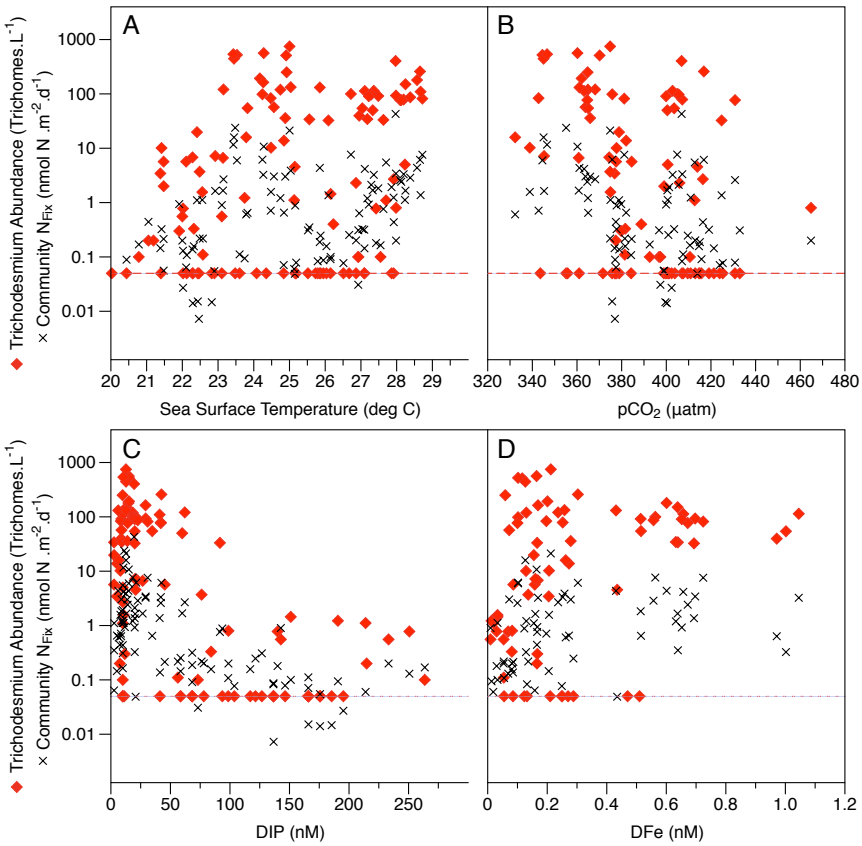

Figure 24: Linear regression analysis of DIP, DFe, SST and pCO_2 on *Trichodesmium* abundance and whole community N_{Fix} observed in all cruises and seasons. *Trichodesmium* abundance (Trichomes.L^{-1}) (black crosses) and whole community N_{Fix} ($nmol^{-1}.m^{-2}.d^{-1}$) (red diamonds) as a function of either A) SST (°C), B) pCO_2 (µAtm), C) DIP (nM) or D) DFe (nM).

	DIP	**DFe**	**pCO_2**	**SST**
***Trichodesmium* Abundance**	0.07 (n=74)	0.00 (n=62)	0.08 (n=59)	0.01 (n=79)
Surface N_{Fix} rate	**0.06 (n=105)**	0.04 (n=69)	0.03 (n=84)	0.03 (n=107)

Table 11: Coefficient of determination (R^2) for either community N_2 fixation or *Trichodesmium* abundance compared with SST, pCO_2, DFe and DIP. Sample size is shown in brackets, significant correlation as determined by Student's t-test (P<0.01) are highlighted in bold font.

limiting resources. Due to their elevated P and Fe requirements diazotrophs such as *Trichodesmium* are only able to avoid competitive exclusion in certain regions due to N limitation of the non-diazotrophic community (Dutkiewicz et al., 2012; Monteiro et al., 2011; Ward et al., 2013). In the (sub-)tropical Atlantic this steady-state approximation is maintained with regards to N limitation where diazotroph such as *Trichodesmium* can flourish. Given the stability of this N limitation niche the abundance and activity of the diazotrophic community within regions of N limitation are hypothesised to be controlled by the availability of P and Fe (Ward et al., 2013).

To first order, *Trichodesmium* can be assumed to consume Fe and P at a given ratio which can be used to define a consumption vector which dictates the change in resource availability (Figure 25), deviations away from this assumed ratio are discussed in Chapter 4 and Chapter 5. This Fe or P utilisation drives the individual resource down to its R^* concentration, here the consumption is balanced by the supply and so a stable population is supported but zero net growth is achievable, the absolute size of this population being dependent on the nutrients supply rate (Dutkiewicz et al., 2012). For any given supply of Fe or P an L-shaped relationship between these two resources is expected. This L-shape defines the transition between either Fe or P limitation with populations located along the R^*_{Fe} isocline being Fe limited, populations along the R^*_P isocline can be considered P limited and the apex of the L representing the ideal supply of Fe:P and so signifying Fe and P co-limitation. Plotting Fe against P from our data set I find *Trichodesmium* abundance and community N_{Fix} are both at maximal values in regions observed to be deplete in both DFe and DIP. As such I suggest this concomitant depletion of DIP and DFe is a result of the activity of *Trichodemsium* populations rather than *Trichodesmium* preferentially favouring DFe and DIP deplete regions.

I can conceptually model the Fe and P requirements of *Trichodesmium* using Equation 2 & Equation 3 (Tilman, 1977, 1980). Where k_{Fe} and k_P represent the growth half saturation coefficients for Fe and P respectively whilst μ and m refers to *Trichodesmium*'s maximum growth rate and mortality rate respectively.

$$R_{Fe^*} = \frac{k_{Fe}}{\frac{\mu}{m} - 1} \tag{2}$$

$$R_{P^*} = \frac{k_P}{\frac{\mu}{m} - 1} \tag{3}$$

As previously discussed (Ward et al., 2013) these R^* concentration allow us to define 4 distinct, *Trichodesmium* specific, oceanic environments (Figure 25A):

1. Low P, High Fe (LPHFe) – (where $DFe > R^*_{Fe}$,$DIP < R^*_P$) Characterised by surplus Fe, non-diazotrophic growth is constrained by DIN limitation whilst diazotrophic growth is constrained by P supply. Typical of the NASG (Sohm et al., 2011b; Mather et al., 2008).

2. High P, Low Fe (HPLFe) – (Where $DFe < R^*_{Fe}$,$DIP > R^*_P$) Surplus phosphorus, non-diazotrophic community is constrained by N limitation whilst diazotrophic growth is constrained by Fe supply. Typical of the SASG (Sohm et al., 2011b; Moore et al., 2009a).

3. Low P, Low Fe (LPLFe) – (Where $DFe < R^*_{Fe}$,$DIP < R^*_P$) Region in which both Fe and P concentrations are below the *Trichodesmium* R^*_P and R^*_{Fe}, here non-diazotrophic growth is limited by N supply whilst diazotrophic growth is limited by both P and Fe supply and diazotrophs are effectively excluded. I suggest that this province does not exist in the sampled (sub)-tropical Atlantic and potentially may not be generated as no mechanism exists to reduce P and Fe below R^*_P and R^*_{Fe}.

4. High P, High Fe (HPHFe) – (Where $DFe > R^*_{Fe}$,$DIP > R^*_P$) Region in which both Fe and P concentration exceed growth limiting concentrations. Non-diazotrophs are competitively excluded due to N limitation, diazotrophs prosper and draw both P and Fe down to their R^* concentrations. This is an inherently transient environment as nutrient consumption will actively deplete Fe and P down to either R^*_{Fe} or R^*_P values. This region is thought to be typical of the ITCZ influenced central equatorial region where variable resupply of DFe and DIP may temporarily support concentrations above their R^* values.

Utilising the data collected in the equatorial Atlantic I am able to approximate the R^* concentrations for Fe and P and thus define the transition concentrations for the 4 provinces. Observations suggest an $R_{Fe^*<0.1}$ nM and an $R_{P^*<5}$ nM resulting in a *Trichodesmium* $k_{Fe} : k_P$ ratio of $\approx 2x10^{-2} : 1$ or 20 $mmol.mol^{-1}$ (Figure 25). Interestingly this ratio shows strong agreement with previously published intracellular Fe:P ratios where Nuester et al. (2012) observed Fe:P ranging from 13 to 63 $mmol.mol^{-1}$ whilst Sañudo Wilhelmy et al. (2001) observed Fe:P ratios of 5.4-39 $mmol.mol^{-1}$.

I suggest that the predominant state of the (sub)-tropical Atlantic is one of nutrient limitation for both non-diazotrophs and diazotrophs alike, given the consumption vectors (synonymous to *Trichodesmium* growth) act to drive the two resources to their respective R^* concentrations (Figure 25A). With DFe and DIP concentrations residing at or close to their R^* concentrations. The conformity of our environmental observations to the theoretical expectations strongly supports

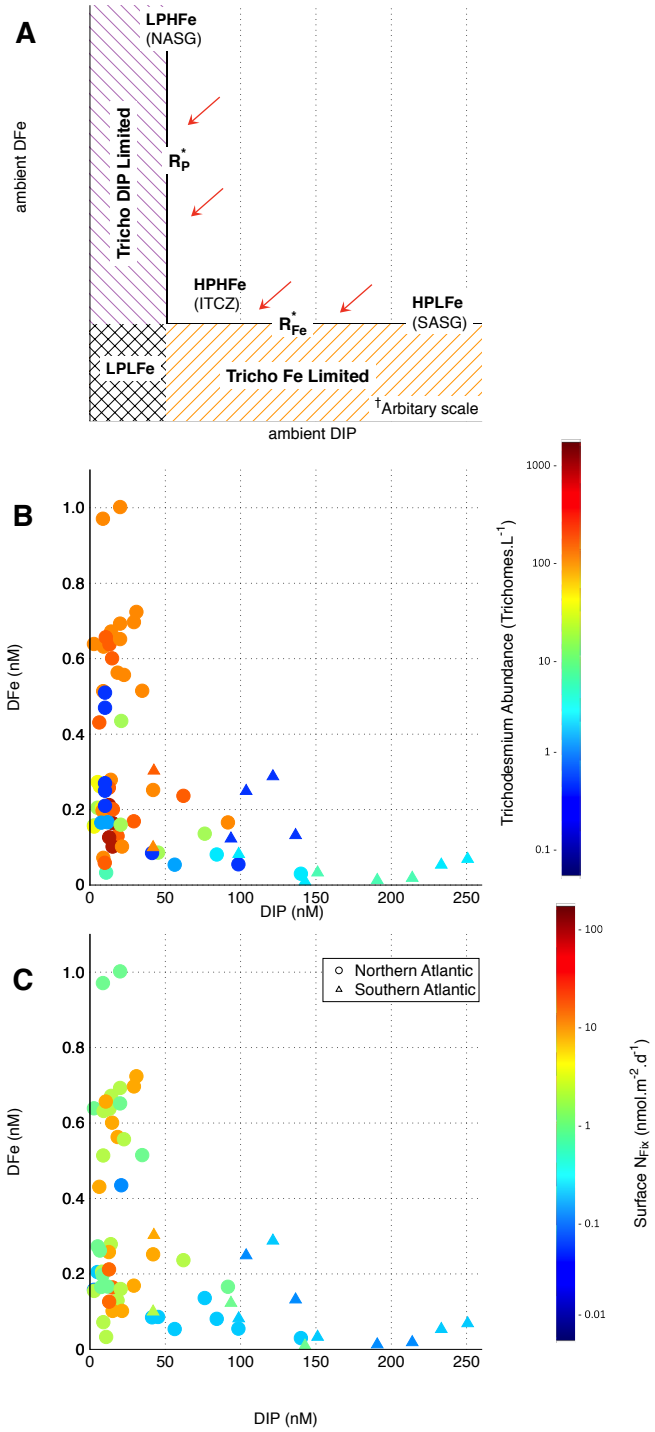

Figure 25: A) Graphical representation of the resource competition theory conceptual model. Red arrows indicate *Trichodesmium* consumption vectors. B) *Trichodesmium* abundance (Trichomes.L^{-1}) as a function of DIP (nM) and DFe (nM) and C) Surface nitrogen fixation (nmol.m^{-2}.d^{-1}) as a function of DFe (nM) and DIP(nM). Observations taken from the northern hemisphere are signified by circles, observations from the southern hemisphere are signified by triangles.

an argument for persistent contrasting regions of Fe and P limitation characterising the sampled region of the Atlantic ocean whilst at the same time questioning the first order importance of other proposed control mechanisms such as SST and pCO_2 when applied to the Atlantic Ocean (Hutchins et al., 2013; Breitbarth et al., 2007).

The observations of elevated *Trichodesmium* abundance and N_{Fix} at low DIP but variable DFe concentrations results in a potential LPHFe niche within which an organism with a R^*_P lower than that of *Trichodesmium* could be capable of utilising the excess DFe. However these observations could also potentially be explained by *Trichodesmium*'s phosphorus compensatory strategies and the periodic and spatially variable supply of DFe and DIP (Figure 25C & D). *Trichodesmium* are capable of accessing DOP, either as P-esters via the alkaline phosphatase pathway or as phosphonates via the CP lyase pathway, *Trichodesmium* also have a flexibility in cellular P requirement due to lipid substitution strategies (Dyhrman et al., 2006; Orchard et al., 2009; Van Mooy et al., 2009). These strategies, where DIP is substitutable for DOP, result in DIP no longer representing the sole available pool of essential P according to the definition set by Tilman (1980) and highlight the importance of considering re-supply rates of DIP and DOP together. This adaptability to both P source allows for DIP concentrations at R^*_P values to support additional biomass through DOP uptake. The described flexibility in P requirement will result in changes to *Trichodesmium*'s Fe:P consumption vector potentially introducing a curved vector due to the disparate substitutability of Fe and P, this is discussed further in Chapter 4 and Chapter 5. Neither strategy however allows for depletion of DIP below the R^*_P value. The observations of elevated DFe could be an artefact of the disparity between the resupply rate of DFe from atmospheric deposition and the comparatively slow resupply of DIP/DOP through advection from elsewhere. This prevalence of DFe$>R^*_{Fe}$ in regions supporting *Trichodesmium* populations serves as a strong indicator of potential phosphorus stress possibly combated with the above mentioned compensatory strategies.

Given the uniquely high iron requirement of *Trichodesmium* there exists within our theoretical framework a niche for diazotrophs with a greater iron use efficiency than that of *Trichodesmium*. *Crocosphaera watsonii* for example employ a mechanism of iron 'hot-bunking' whereby temporal segregation of N_{Fix} and photosynthesis allows for an intracellular recycling of metalloprotein bound iron (Saito et al., 2011). This is thought to allow for a ≈40% reduction in metalloenzyme iron requirement relative to a non-hot-bunking cell. I hypothesis this iron conservation strategy is unlikely in *Trichodesmium* given the shorter time period between peak N_{Fix} and photosynthesis. Whilst segregation of N_{Fix} and photosynthesis either spatially or temporally and iron hot-bunking both allow for a reduction in a cells

absolute iron requirement it does not necessitate a reduction in the R^*_{Fe} concentration. Unicellular diazotrophs such as *C. watsonii* are subject to increase grazing pressure due to their size. I speculate that this grazing pressure may elevate the mortality parameter of Equation 2 for *C. watsonii* and results in a higher R^*_{Fe} concentration and thus they fail to out-compete *Trichodesmium* in all situations.

3.4.4 *Non-Trichodesmium diazotrophy*

As has been previously observed, diazotrophy in the NASG is very much confined to the surface waters (Figure 22). In contrast our observation of SASG diazotrophy shows a more homogenous distribution across the euphotic zone (Figure 22). Considering the depth integrated rates of N_2 fixation observed during AMT21 I see, along the AMT21 transect at least, that N_2 fixation between the 2 gyres is approximately equal in magnitude (Figure 23). The SASG is grossly under sampled with regards to N_2 fixation rates, few published data exist but of that data the majority is surface-water only rates (Luo et al., 2012). The presence of a SASG diazotrophic community has previously been inferred from H_2 supersaturation where the cellular process of N_2 fixation results in the production of a molar equivalent of 2 H_2 molecules which is then used as a tracer for N_2 fixation (Fraser, 2012; Moore et al., 2009b). The fact that colonial diazotrophs such as *Trichodesmium* are rare in the SASG (Tyrrell and Marañón, 2003; Moore et al., 2009a), together with reported negative N* values (Gruber and Sarmiento, 1997) would suggest a lack of N_2 fixation in the region. However our observations along the AMT21 transect (Figure 20 and Figure 22) along with those of Sohm et al. (2011a) in the eastern-south Atlantic would appear to contradict these assumptions. Montoya et al. (2004) highlighted that unicellular diazotrophs are more homogeneously distributed throughout the euphotic zone when compared with *Trichodesmium*, that tend to aggregate in surface waters such as is seen in the NASG. Further Moisander et al. (2010) observed a UCYN-A and *C.watsonii* preference for deeper waters than *Trichodesmium*. I speculate that this SASG diazotrophic signature is unicellular in origin although without supporting data it is impossible to confirm.

Considering this potentially non-*Trichodesmium* diazotrophic community within the confines of our conceptual model I observe that all observations of N_{Fix} in the South Atlantic fit within the HPLFe regime whereby DFe concentrations are close to our observed *Trichodesmium* R^*_{Fe} concentrations (Figure 25). This observation of elevated N_{Fix} in a region of apparent Fe limitation further suggests a non-*Trichodesmium* source of this diazotrophy. *Trichodesmium*'s uniquely high Fe requirement, caused by its concurrent N_{Fix} and pho-

tosynthetic activity, could potentially result in competitive exclusion by a hypothesised unicellular diazotrophic community.

3.5 CONCLUSION

Using previously unpublished *Trichodesmium* abundance and nitrogen fixation data I support the ITCZ driven seasonal migration of *Trichodesmium* outlined by Schlosser et al. (2014). Previous studies have utilised Fe:N ratios as an indicator of potential diazotrophic niches (Dutkiewicz et al., 2012; Ward et al., 2013), here I utilise Fe:P ratios in order to describe the finer scale distribution of diazotrophs in the (sub)-tropical Atlantic. In doing so I describe 3 distinct biogeochemical provinces along a meridional transect of the (sub)-tropical Atlantic. With the use of resource-ratio theory I demonstrate that these biogeochemical provinces do not simply dictate the abundance or diazotrophic capability of the *Trichodesmium* community but rather are a reciprocal result of the unique nutrient requirements of this globally important diazotroph. In doing so I further our understanding of the controls on *Trichodesmium*'s distribution highlighting a distinct lack of correlation between any one parameter, be it SST, pCO_2, DIP or DFe. I conclude, that whilst SST and pCO_2 may be capable of influencing *Trichodesmium*'s growth and nitrogen fixation rates, in the (sub)-tropical Atlantic these effects are overridden by nutrient availability. These findings suggest changes in nutrient availability due to global change will have a greater effect on the diazotrophic community than the more confidently predicted changes in SST and pCO_2. As such a better understanding of of these potential changes is required in order to confidently predict the future of low-latitude diazotrophy.

4

EXPLORING THE PLASTICITY OF *TRICHODESMIUM*'S INTRACELLULAR ELEMENTAL COMPOSITION ALONG THE ANTITHETICAL FE AND P GRADIENTS OF THE (SUB)-TROPICAL ATLANTIC

4.1 INTRODUCTION

The biogeochemical cycling for many of the trace nutrients is in large part controlled by the plankton community. The incorporation of bioactive trace metals into plankton biomass in the surface waters and subsequent re-mineralisation at depth results in distinctive vertical profiles analogous to those of fixed nitrogen and inorganic phosphorus (Sarmiento and Gruber, 2006). The extrapolation of this bidirectional feedback beyond the initial macronutrient C, N, Si and P proposed by Redfield (1958) towards the inclusion of bio-active trace metals is complicated significantly by community composition (Ho et al., 2003; Moore et al., 2013). As has been shown for diazotrophs, different organisms and community composition often depart from the canonical Redfield ratios of C_{106}:N_{16}:P_1 (Carpenter et al., 2004; Krauk et al., 2006). Such a departure is arguably more prevalent for trace-metals given their environmental scarcity and broad range of lithogenic inputs, redox transformations and scavenging behaviours (Morel et al., 2008). Vertical distribution profiles suggest the bioactive metals Fe, Cd, Co, Cu, Mn, Ni and Zn play critical roles in oceanic primary productivity and biogeochemistry (Donat and Bruland, 1995; Saito and Moffett, 2002). The nitrogenase proteins, used to fixed nitrogen, results in diazotrophs having an elevated requirement for Mo and potentially V (Burris, 1991; Nuester et al., 2012), both of which are trace-metals previously observed at elevated, vertically homogeneous, concentrations.

Here I present the intracellular composition (C, N, P, Fe, Ni, Cd, Co, Cu, Mn, Zn and As) of the non-heterocystous cyanobacteria *Trichodesmium*. Collected data is from two cruises in the central Atlantic covering an antithetical Fe and P gradient thought to have a controlling influence on the distribution and activities of this globally important nitrogen fixer (Chapter 3). This data is presented alongside surface water dissolved P, Fe, Mn, Zn, Ni, Cu, Cd and Co concentrations. Correlations between cell quotas and ambient surface water concentrations are interrogated along these two latitudinal transects. Here I infer the inadvertent bio-accumulation of As, V and Mo as a result of potential P stress and aim to describe the unique biogeochem-

ical composition in relation to *Trichodesmium*'s diazotrophic growth strategies.

4.2 METHOD

4.2.1 *Biomass Collection*

Sampling was conducted during two oceanic cruises, D361 consisted of 19 triplicate samples being collected across 14 stations whilst AMT21 consisted of 22 triplicate samples collected from 22 stations. Samples were collected using drift net deployments as described in Chapter 2. Between 50 (D361) or 100 (AMT21) colonies were picked under Class-100 laminar flow, suspended in 0.2 μm filtered seawater and filtered onto acid-cleaned (24-48 hours, 10% HCl) 25 mm diameter 4.0 μm pore size polycarbonate filters (Millipore Nucleopore). In order to remove metals bound extracellularly the collected colonies and filters were washed using the Oxalate-EDTA solution (100 mM $Na_2C_2O_4$, 50 mM Na_2EDTA, 0.3 M NaCl, 0.01 M KCl) described in Tang and Morel (2006). The extensive pre-cleaning of a citrate containing Oxalte-EDTA wash solution described in Tovar-Sanchez et al. (2003) is thought to reduce the solution Fe content by ≈30 fold. However this cleaning was deemed impractical given the duration of both research cruises and the solutions comparatively short shelf life, as a result the citrate free solution described in Tang and Morel (2006) was used. Tang and Morel (2006) showed the removal of citrate from this solution effectively reduced the Fe contamination by half. They then went on to show that extensive rinsing of the sample and filter with trace-metal clean water was an effective way of eliminating contamination for all elements aside from Ba. Consequently two 10 minute, 10 ml Oxalate-EDTA washes were followed by 15 rinses of low trace-metal, 0.2 μm filtered seawater. Samples were subsequently flash frozen in liquid nitrogen and stored at -80°C until further analysis. Sample blanks were collected at every station, these involved every stage of sample collection and processing (i.e. filters and washing procedure) aside from actually having sample loaded onto them.

4.2.2 *Acid Digestion*

In order to avoid sample contamination all laboratory work was carried out in a class 100 clean laboratory under laminar flow. Samples were allowed to defrost for a few minutes prior to being transferred to 25 ml, conical based, PTFE digestion vessels (Savillex) using acid-cleaned PTFE tweezers. Sample filters were unfurled along the wall of the digestion vessel so as to avoid direct contact with the base (Marsay, 2012). Digestion was achieved with 0.5 ml HNO_3 (sub-boiling distilled, Romil SpA grade) and 1.5 ml HCl (sub-boiling dis-

tilled, Fisher scientific, PrimarPlus grade) (Similar to Tovar-Sanchez and Sañudo Wilhemy (2006) with HF omitted). Digestion vessels were sealed and transferred to a 150°C hotplate for 4 hours. The acidic vapours resulting from this continuous boiling acted to digest any biological matter on the filters whilst reducing the degradation of the filters themselves. Following digestion, remaining biomass-free filters were removed from the digestion vessels. The digested vessels were then returned open to the hotplate to allow the evaporation of the digestion solution. Following this the analyte containing precipitate was fully re-dissolved in 2% ultra-clean HNO_3 ready for analysis. Sample blanks were handled in the same manner described above for samples. In addition triplicate analytical blanks were performed at this stage consisting of a reagent only digestion. A standard reference material (SRM), SRM-1573a (National Institute of Standards and Technology, U.S Department of Commerce) was digested in triplicate alongside the samples.

4.2.3 *Inductively coupled plasma mass spectrometric (ICP-MS) Analysis*

The solutions resulting from the above described digestion were analysed by high resolution inductively coupled plasma mass spectrometry (Thermo Fisher Scientific Element 2 XR HR-ICP-MS). Analytes measured included Cd and Mo using the Low resolution method (LRM, $m/\delta m = 300$), Al, P, V, Mn, Fe, Co, Ni, Cu and Zn using the medium resolution method (MRM, $m/\delta m = 3000$), and finally As using the high resolution method (HRM, $m/\delta m = 7500$) (Table 7). Internals standards included In and Be (measured at both LRM and MRM). Samples were run along side a series of multi-element standards prepared from certified single-element reference solutions with 2% HNO_3 as the diluent (Romil PrimAg-xtra or Inorganic Ventures). This multi-element cocktail of standards was prepared to match previously published *Trichodesmium* elemental composition so as to best match our sample composition (Tovar-Sanchez and Sañudo Wilhelmy, 2011).

The solutions resulting from the above described digestion were analysed by high resolution inductively coupled plasma mass spectrometry (Thermo Fisher Scientific Element 2 XR HR-ICP-MS). Analytes measured included Fe, Ni, P, Cd, Co, Cu, Mn, Mo, V, Zn, and As, along with In and Be used as internal standards. Samples were run along side a series of multi-element standards prepared from certified single-element reference solutions with 2% HNO_3 as the diluent (Romil PrimAg-xtra or Inorganic Ventures). This multi-element cocktail of standards was prepared to match previously published *Trichodesmium* elemental composition so as to best match our sample composition (Tovar-Sanchez and Sañudo Wilhelmy, 2011).

4.2.4 *Method Validation*

Sample blanks that took into account possible metal contributions from filter apparatus, filters, Oxalate-EDTA wash or subsequent seawater washes were measured. For each triplicate set of samples, data for the corresponding blanks were averaged then subtracted from the raw signal intensity values for each analyte prior to any further data processing. Expressed as an average percentage of total analyte ± SE, blanks for D361 and AMT21 can be found in Table 12. Blanks for some analytes were elevated however absolute concentrations for individual analytes show good consistency amongst both intra- and inter-cruise blank comparisons (see Table A1, Table A2, Table A3). This suggests the observation from D361, where blank values represent a higher percentage of observed sample values are a result of reduced sample biomass as discussed later. Analytical blanks (no filter, biomass or washing procedure) were <10 % for the total observed concentrations of all analytes except for Al, Zn and Cu (D361 only). Furthermore these analytical blanks accounted for >40 % of the total observed sample blank concentrations for Cd, Mo, Al, P, V, Mn (D361 only), Fe (AMT21 only) Co, Cu, Zn and As. This suggests a significant proportion of our blank measurement is derived from our choice of digestion reagents rather than filter or sampling contamination.

Analyte	D361 Blank (%)	AMT21 Blank (%)
Cd	11.07 ± 1.49	3.94 ± 0.41
Mo	20.94 ± 1.52	7.29 ± 1.01
P	18.03 ± 1.27	2.55 ± 0.21
V	6.59 ± 0.56	1.96 ± 0.38
Mn	23.24 ± 1.69	18.76 ± 2
Fe	32.91 ± 2.88	35.85 ± 6.98
Co	31.77 ± 2.16	29.05 ± 4.81
Ni	39.97 ± 3.2	32.93 ± 7.67
Cu	57.06 ± 4.71	36.31 ± 5.33
Zn	67.65 ± 5.24	40.28 ± 3.91
As	8.05 ± 1.16	1.89 ± 0.35

Table 12: Blank values observed during D361 and AMT21 expressed as mean percentages of observed samples (i.e. ([Absolute Blank (ppb)]/[Absolute Sample [ppb]) x 100 = %. Further blank details along with absolute values (ppb) for instrument, reagent and filter blanks can be found in Table A1, Table A2 and Table A3.

Data regarding the efficiency of digestion of the SRM helped ascertain the accuracy of analysis for all analytes considered. The vari-

ability in recoveries between individual digestions was observed to be very low with most analytes showing a standard error of <5% for both D361 (n=15) and AMT21 (n=24) (Table 13). Zinc showed the largest variability in recovery with D361 showing a 20.1% SE between digests. Recovery of Mo during analysis of AMT21 samples was substantially lower than that of D361 samples. The only change in analytical technique between D361 and AMT21 digestions was a different batch of sub-boiled HCl which should not have affected absolute recovery rates, subsequently the reasons for these differences in recoveries are unclear.

Analyte	Certified Value (ppm)	NIST1573a (D361) (ppm)	NIST1573a (AMT21) (ppm)	NIST1573a (D361) (%)	NIST1573a (AMT21) (%)
Cadmium	1.5	0.97 ±0.04	1.08 ±0.02	63.96 ±2.59	70.77 ±1.6
Molybdenum	0.5	0.22 ±0.02	0.19 ±0.08	46.98 ±3.97	25.25 ±4.14
Aluminium	598.0	146.92 ±7.39	157.25 ±9.65	24.57 ±1.24	26.3 ±1.61
Phosphorus	2160.0	1729.36 ±84.74	2136.08 ±90.5	80.06 ±3.92	98.89 ±4.19
Vanadium	0.8	0.32 ±0.01	0.33 ±0.01	38.74 ±1.8	39.27 ±1.71
Manganese	246.0	195.88 ±9.23	223.09 ±7.47	79.63 ±3.75	90.69 ±3.04
Iron	368.0	208 ±10.75	237.01 ±7.17	56.52 ±2.92	64.4 ±1.95
Cobalt	0.6	0.45 ±0.02	0.49 ±0.02	78.09 ±2.94	85.8 ±2.95
Nickel	1.6	1.23 ±0.06	0.98 ±0.3	77.35 ±3.48	80.11 ±3.78
Copper	4.7	3.66 ±0.25	3.88 ±0.26	77.89 ±5.27	82.5 ±5.47
Zinc	30.9	20.98 ±6.22	24.66 ±2.76	67.9 ±20.12	79.8 ±8.93

Table 13: Recovery efficiencies for SRM NIST 1573a. Showing certified values (ppm) alongside observed values (mean ± SE) and percentage recoveries (mean ± SE) split between cruises D361 (n = 15) and AMT21 (n =24) so as to highlight any discrepancies in recoveries.

4.3 RESULTS AND DISCUSSION

4.3.1 *Removal of adsorbed extracellular metals*

An assessment on the removal of extracellularly bound metals achieved by the Oxalate-EDTA reagent was conducted by repeat sampling of washed and un-washed samples. This assessment was performed on environmentally collected *Trichodesmium* colonies at 3 distinct locations along the AMT21 transect, labelled as TN005, TN011 and TN018 (Table 14). In all three assessments, data showed significant removal of extracellular Mo, P, V, Mn whilst Cu was only significantly removed for samples TN005 and TN011 (Student's t-test $P<0.05$, $n=3$) (Table 14). Meanwhile analysis showed statistically insignificant removal of extracellularly bound Cd, Al, Fe, Co nor Ni (in one test Ni showed a small yet statistically significant increase in the washed treatment, this anomaly is attributed to an inherent sample variability when working with *Trichodesmium* colonies) (Table 14). These data suggest either poor removal efficiency or low extracellular concentrations of each metal. The observed lack of significant Fe removal is interesting given *Trichodesmium*'s ability to produce Fe-rich transparent exopolymeric particles (TEP) during auto-catalytic programmed cell death (PCD) (Berman-Frank and Bidle, 2004). *Trichodesmium* PCD can be triggered during periods of Fe or P stress Berman-Frank and Bidle (2004), conditions demonstrated in Chapter 3. Thus we might predict significant removal of TEP bound Fe in from our *Trichodesmium* samples although this appear to not be the case.

Analyte	TN005 (% removal)	TN011 (% removal)	TN018 (% removal)	Mean $\pm\sigma$
Cadmium	ns	ns	ns	na
Molybdenum	30.87	44.17	49.42	41.49 ± 9.56
Phosphorus	61.31	64.46	43.66	56.48 ± 11.21
Vanadium	93.04	85.85	52.84	77.24 ± 21.44
Manganese	78.51	74.57	51.47	68.18 ± 14.61
Iron	ns	ns	ns	na
Cobalt	ns	ns	ns	na
Nickel	ns	ns	ns	na
Copper	54.48	47.01	ns	50.75 ± 5.29
Zinc	ns	46.34	ns	na
Arsenic	97.86	98.42	61.90	86.06 ± 20.93

Table 14: Percentage removal of adsorbed extracellular metals assessed at three distinct locations/stations - TN005, TN011 and TN018. Significant removal of Mo, P, V, Mn, Cu and As was consistently observed. Statistical significance was assessed using a students t-test ($p<0.05$, $n=3$). Non-significant removal of extracellular metals were observed for Cd, Fe, Co and Ni in all 3 sample sets. Mean and standard deviation values for all three sample sets are shown for analytes showing statistically significant removal.

Using the same Oxalate-EDTA wash reagent and performing analysis via Synchrotron X-ray Fluorescence (SXRF) Nuester et al. (2012) observed minimal extracellular Fe (similar to our data) alongside, Mn and V to which I observed substantial extracellular removal. Removal of extracellular P (47% lower in washed samples) was similar to our observations (33.7-64.5% lower in washed samples) (Nuester et al., 2012). Assessing the intracellular composition of the Baltic diazotrophs *Aphanizomenon* sp. and *Nodularia spumigena*, Walve et al. (2014) observed minimal removal of extracellularly bound Fe, Ni and V yet significant removal of P, Mn, Cu Mo and Co.

4.3.2 *Colony Composition*

4.3.2.1 *PON and P_{IC}*

Per colony particulate organic nitrogen (PON) and intracellular phosphorus (P_{IC}) concentrations were fairly consistent during D361 with little variability. PON ranged from 6.7-23.8 nmol.col^{-1} (mean ± SD - 13.9 ±4.3 nmol.col^{-1}) whilst P_{IC} ranged from 0.12-0.46 nmol.col^{-1} (mean ± SD - 0.25 ±0.11 nmol.col^{-1}) (Figure 26). Colony specific N:P ratios varied between 27-138 (mean ± SD - 76 ±39). Meanwhile colony Chl-*a* concentrations ranged between 2.10 - 23.11 ng Chl.col^{-1} (mean ± SD - 6.03 ±5.76 ng Chl.col^{-1}).

Comparing observations from AMT 21 to those of D361 I see slightly elevated PON concentrations ranging from 9.0-36.1 nmol.col^{-1} (mean ± SD - 18.7 ±8.9 nmol.col^{-1}) but substantially higher colony P_{IC} concentrations ranging from 0.14-1.43 nmol.col^{-1} (mean ± SD - 0.57 ±0.33 nmol.col^{-1}) (Figure 26C & D). Particulate organic carbon concentrations observed during AMT21 ranged from 46.6-197.4 nmol.col^{-1} (mean ± SD - 107.3 ±50.9 nmol.col^{-1}) and showed a similar latitudinal gradient to both PON and P_{IC} (Figure 26B). A clear latitudinal gradient in POC, PON, P_{IC} and therefore inferred colony size is evident during AMT 21 with far larger colonies observed to the north (Figure 26). AMT21's *Trichodesmium* specific N:P ratios (N:P - 17-63, mean ± SD - 31 ±14) were far lower than D361's equivalent and do not show the same latitudinal gradient as either PON and P_{IC}. Near Redfield stoichiometries were observed in *Trichodesmium* C:N ratios from AMT21, ranging from 5.2-6.4:1 (mean ± SD - 5.72 ±0.41) with little latitudinal variability (Figure 26E). Chlorophyll concentrations observed during AMT21 ranged from 2.28 - 58.17 ng Chl.col^{-1} (mean ± SD - 17.09 ±15.74 ng Chl.col^{-1})

Per colony concentrations of PON, intracellular P (P_{IC}) and Chl-a demonstrate both a substantial inter and intra-cruise variability in colony size. AMT21 colonies are generally larger than D361 and a clearer latitudinal gradient is evident in the data that is absent in D361's data. This substantial variability, whilst corrobo-

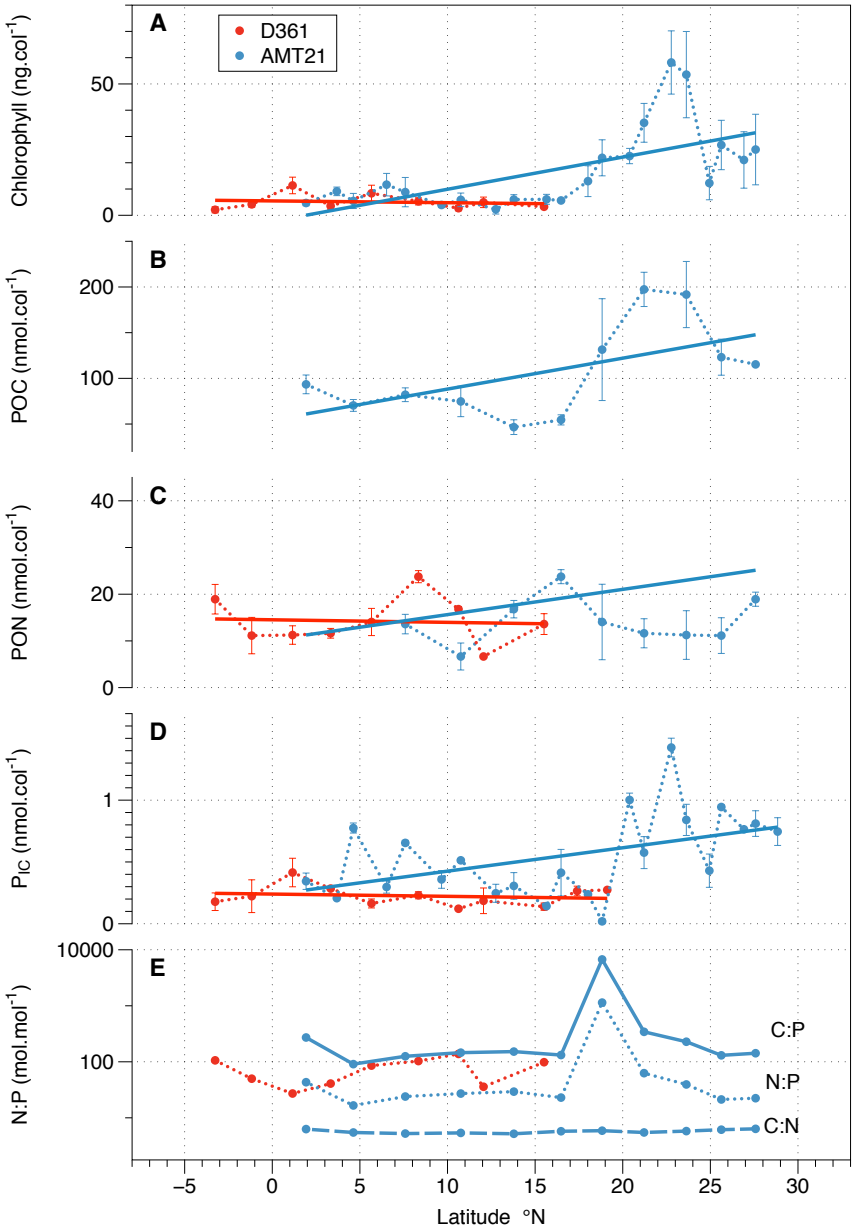

Figure 26: Per colony chlorophyll concentration (A), Particulate inorganic carbon (POC, B), Particulate organic nitrogen (PON, C), intracellular phosphorus (P_{IC}, D) concentrations and C:P, C:N and N:P ratios (E) of isolated *Trichodesmium* colonies from either D361 (red) or AMT21 (blue). Linear fit curves are displayed in order to highlight any latitudinal gradient.

rated by literature ranges (Carpenter et al., 2004; Nuester et al., 2012; Sañudo Wilhelmy et al., 2001), will influence the absolute intracellular metal concentrations with the expectation of larger colonies having elevated metal concentrations. Given this observed variability, biomass normalised concentrations ($Metal_{Intracellular}$

(M_{IC}):phosphorus$_{Intracellular}$(P_{IC})) will be used. POC concentrations offer the most direct proxy for cellular biomass, however POC was not analysed using the same method as P_{IC} therefore it will not be used.

N:P ratios agree with previous observations whereby diazotrophic organisms, and *Trichodesmium* specifically, have N:P values consistently above the canonical Redfield value of 16:1 (Figure 26) (Carpenter et al., 2004; Nuester et al., 2012; Sañudo Wilhelmy et al., 2001; Krauk et al., 2006). A substantial difference in P_{IC} concentrations were observed between D361 and AMT21, this is thought to have elevated the N:P concentration observed during D361 and could potentially signify an enhanced degree of P stress during this cruise when compared with AMT21.

4.3.2.2 *Intracellular Metallome*

As was evidenced in PON, P_{IC} and Chl-a concentration, D361 colonies were again characterised by lower concentrations of all intracellular metals analysed. Observed metal and P_{IC} (pmol.col^{-1}) concentrations show poor correlations with little of the observed metal variability being explainable by P_{IC} alone (Figure 27). Of the observed trace metals, Fe was most abundant with a mean concentrations of ≈34.76 pmol.col^{-1} then in descending order - Zn - 2.94, Ni - 1.95, V - 1.21, As - 0.98, Mn - 0.65, Cu - 0.65, Mo - 0.54, Cd - 0.06 and Co - 0.03 pmol.col^{-1} (Figure 27).

In contrast to this are the significantly higher concentrations and strong metal to P_{IC} correlations observed during AMT21 with all analytes showing a general trend of increased metal concentration with a P_{IC} inferred biomass increase (Figure 28). Metal concentrations were systematically elevated relative to values observed during D361, average values were - Fe 37.43, V - 13.41, Zn - 9.58, As - 8.0, Ni 3.99, Mo - 2.98, Cu - 1.26, Mn - 1.15, Cd - 0.06, Co - 0.04 pmol.col^{-1}. Of particular note are the strong correlations between P and Mn, V, Cu, As, Ni and Mo all of which show R^2 values >0.30.

I attribute the poor correlations observed during D361 in part to the substantial size difference between AMT21 and D361 colonies. The majority of colonies observed during D361 show a P_{IC} < 300 pmol.col^{-1} whilst a large number of samples collected during AMT21 show a P_{IC} > 300 pmol.col^{-1}. The latitudinal consistency in colony P_{IC} concentrations observed during D361 along with the comparatively small P_{IC} range observed supports our intracellular metal data whereby minimal variations in metal concentrations are seen (Figure 28). The correlation between intracellular Mn and P with an $R^2 \approx 0.84$ suggesting a clear relationship between Mn and P concentrations or Mn and colony size. Manganese plays a critical component in the photosynthetic apparatus where the Mn_4OxCa cluster catalyses the water oxidising complex (Najafpour et al., 2012). Similar cor-

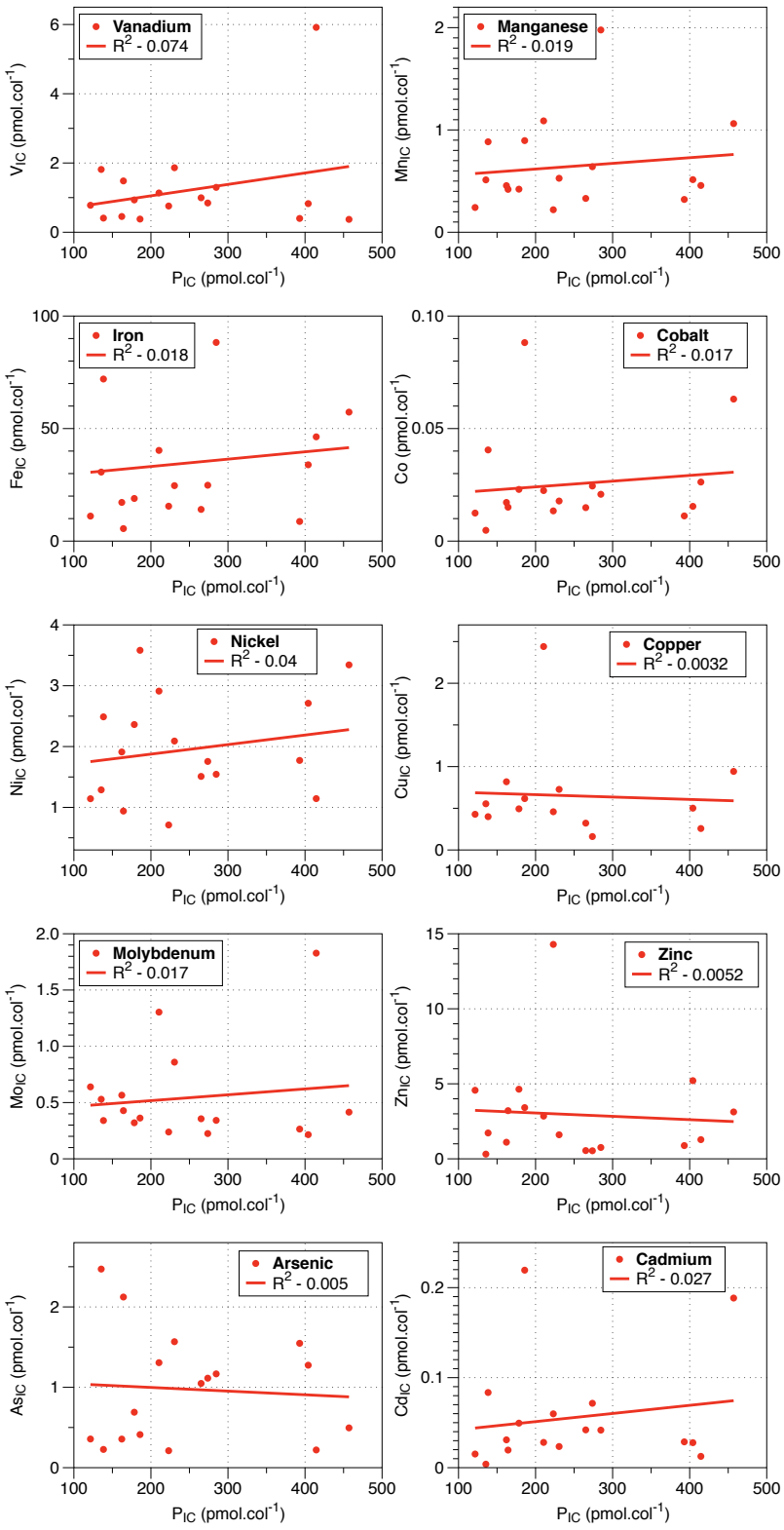

Figure 27: *Trichodesmium*'s intracellular concentrations of V, Mn, Fe, Co, Ni, Cu, Mo, Zn, As and Cd as observed during D361 when compared with intracellular P concentration, used here as a proxy for colony size. R^2 values show the coefficient of determination for each relationship.

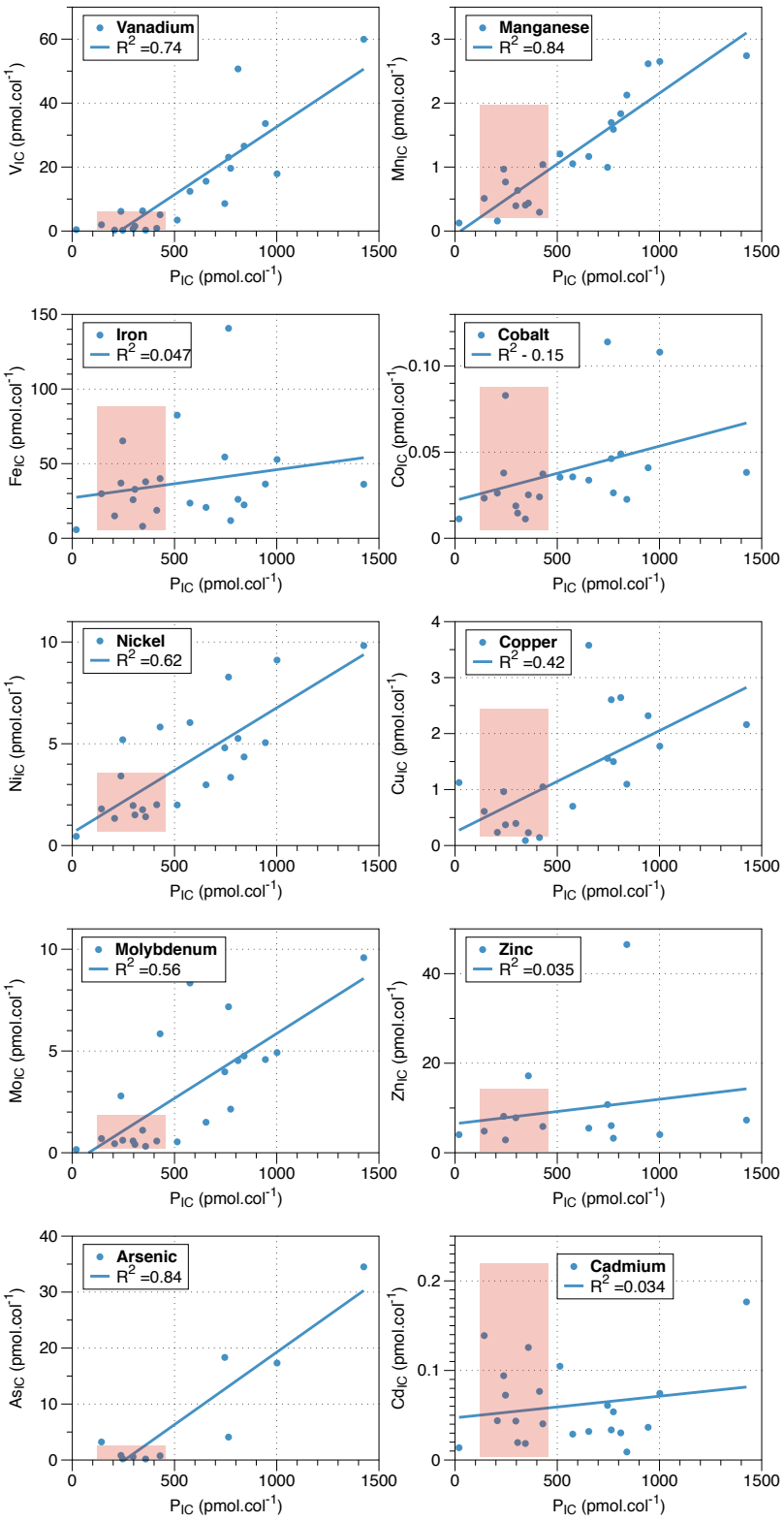

Figure 28: *Trichodesmium*'s intracellular concentrations of V, Mn, Fe, Co, Ni, Cu, Mo, Zn, As and Cd as observed during AMT21 when compared with intracellular P concentration, used here as a proxy for colony size. The ranges of intracellular concentrations observed during D361 are shown by the shaded red region on each plot. R^2 values show the coefficient of determination for each relationship.

relations observed between P and As or V are less easily explained as discussed below.

4.3.3 *Latitudinal gradients in intracellular composition*

In order to account for the observed gradient in colony size (inferred from PON and P_{IC}, Figure 26) P-normalised metal stoichiometries were used to ascertain any latitudinal gradient in intracellular metal composition (Figure 29). Whilst an environmental DIP gradient does exist across our study region, conducting the same analysis on N-normalised metal stoichiometries provides the same results as follows. Whilst this biomass normalised data is generally quite variable a clear gradient is seen along the AMT21 transect with all metals, except Fe, showing elevated stoichiometries in the north and depressed stoichiometries in the south (Figure 29). Given the similar gradient in P_{IC} concentrations it appears that large colonies have a higher *stoichiometric* metal quota relative to small colonies for all these metals. Considering that small colonies will have a higher surface area to volume ratio, these stoichiometric enrichments appear to be located within the colonies rather than being adsorbed onto the cell surface. Moreover the use of the oxalate-EDTA wash, demonstrated earlier to provide substantial removal of adsorbed Mo, P, V, Mn, Cu and As, suggests these enrichments to be truly intracellular.

Considering the antithetical Fe and P gradients prevalent in the central Atlantic (Chapter 3, Figure 19), we transition from a LPHFe region to a HPLFe region and so observe a gradient in DFe:DIP (Chapter 3). Such a pronounced gradient is not readily observed in intracellular Fe:P. A degree of intracellular plasticity is well documented in *Trichodesmium* (Van Mooy et al., 2012; Nuester et al., 2012; Krauk et al., 2006), however as discussed in Chapter 3 the concentrations of DFe and DIP observed in the far north and far south of the tropical Atlantic are not conducive to *Trichodesmium* growth and I therefore suggest these regions to have DFe and DIP concentrations beyond *Trichodesmium*'s intracellular flexibility.

A pronounced gradient in V:P, Mo:P and As:P representing greatly increased values in the northern sub-tropical gyre was observed (Figure 29 and Figure 30). These metals are present in the marine environment as vanadate (HVO_4^{-2}), molybdate (MoO_4^{-2}) and arsenate (AsO_4^{-3}) respectively. A tetrahedral structure and similar electronic properties are conserved amongst them and the dominant DIP anion, phosphate (PO_4^{-3}).

Despite vanadate being known to inhibit phosphate metabolising systems including ATPase (Cantley and Josephson, 1977) and alkaline phosphatase (Seargeant and Stinson, 1979) no vanadium specific transport system had been identified until 2006 (Pratte and Thiel, 2006). Pratte and Thiel (2006) identified a high-affinity V-uptake sys-

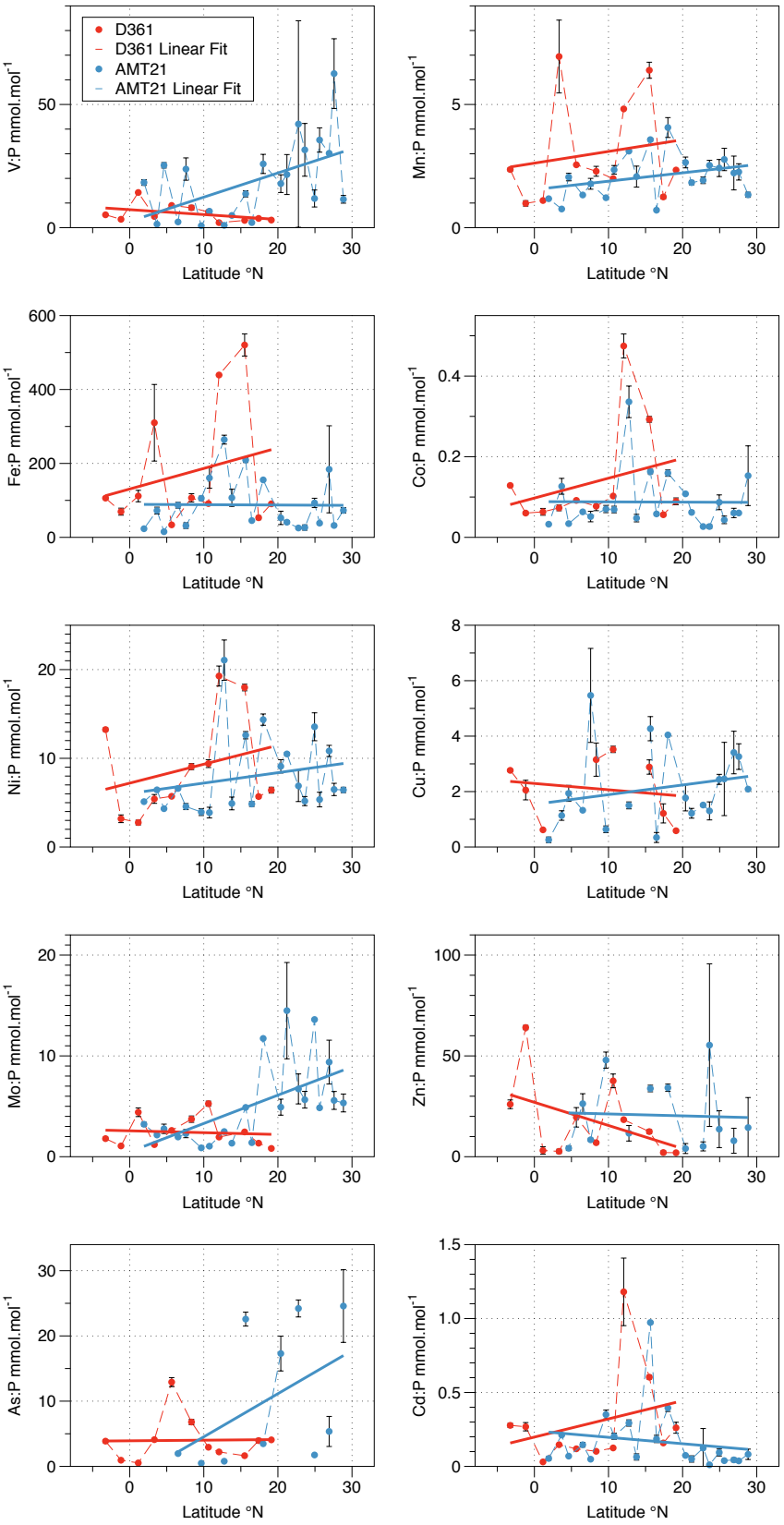

Figure 29: Latitudinal variation in intracellular M:P stoichiometry observed during D361 (red) and AMT21 (blue). Error bars show standard deviations from triplicate measurements.

tem, *vupABC*, in the model diazotroph *Anabaena variabilis* alongside an alternative low-affinity V transport system of a *vupB* mutant. No *vupABC* homologs are found in the *Trichodesmium erythraeum* IMS101 genome (Nuester et al., 2012), this is perhaps unsurprising given a lack of identified vanadium containing proteins within the genome (see Chapter 5 and Chapter 6).

Molybdate uptake in cyanobacteria is mediated by a high-affinity ABC-type transport system, *modABC* (*modA*, Tery_4147 and *modBC* Tery_4148). Tery_4147 and Tery_4148 in *Trichodesmium erythraeum* IMS101's genome encode for *modA* and *modBC* respectively. Zahalak et al. (2004) demonstrated an additional low-affinity molybdate uptake pathway induced in a *modBC* mutant of *Anabaena variabilis* ATCC 29413 when grown in sulphate deplete media. Here it was inferred that the ABC-type sulphate transport system may function as a low affinity molybdate transport system. As discussed previously (Chapter 1) molybdenum is a required co-factor in the nitrogenase enzyme system, however this does not fully account for our observed enrichments (see Section 4.3.6.1).

Mistaken uptake, reduction and subsequent excretion of arsenate is well documented in marine phytoplankton when subject to severe P depletion (Oremland and Stolz, 2003; Karadjova et al., 2008; Hellwegerl and Farley, 2003; Wurl et al., 2013) whereby the P transporters acts as a low affinity As uptake pathway. Whilst no direct evidence is available for a similar occurrence of vanadate or molybdate uptake in marine organisms, Bowman (1983) observed elevated vanadate uptake via a phosphate uptake pathway in the fungi *Neurospora crassa* when grown, in DIP deplete conditions, on an organic P source. The chemical similarities of arsenate, vanadate and molybdate to the anion phosphate, the substantial evidence for low-affinity transport systems for each of these anions alongside our observed environmental gradients suggests a potential mistaken uptake in the P deplete NASG (Figure 30). Moreover a substantial proportion of the total As was removed by the Oxalate-EDTA wash (Table 14), the extracellular abundance of this toxic As is in-keeping with the described mechanisms of rapid uptake of arsenate [As(V)], reduction to arsenite [As(III)] and subsequent excretion (Hellwegerl and Farley, 2003) implying preferential localisation of As at or near the cell wall. A similar V enrichment was observed in *Trichodesmium* colonies from the P deplete Sargasso Sea (Nuester et al., 2012). The V enrichment observed by (Nuester et al., 2012) was again hypothesised to be mistaken uptake, however they did not observe a similar enrichment in Mo as observed in this study nor did they obtain any As data so it is unknown if these V enriched colonies were also enriched in As.

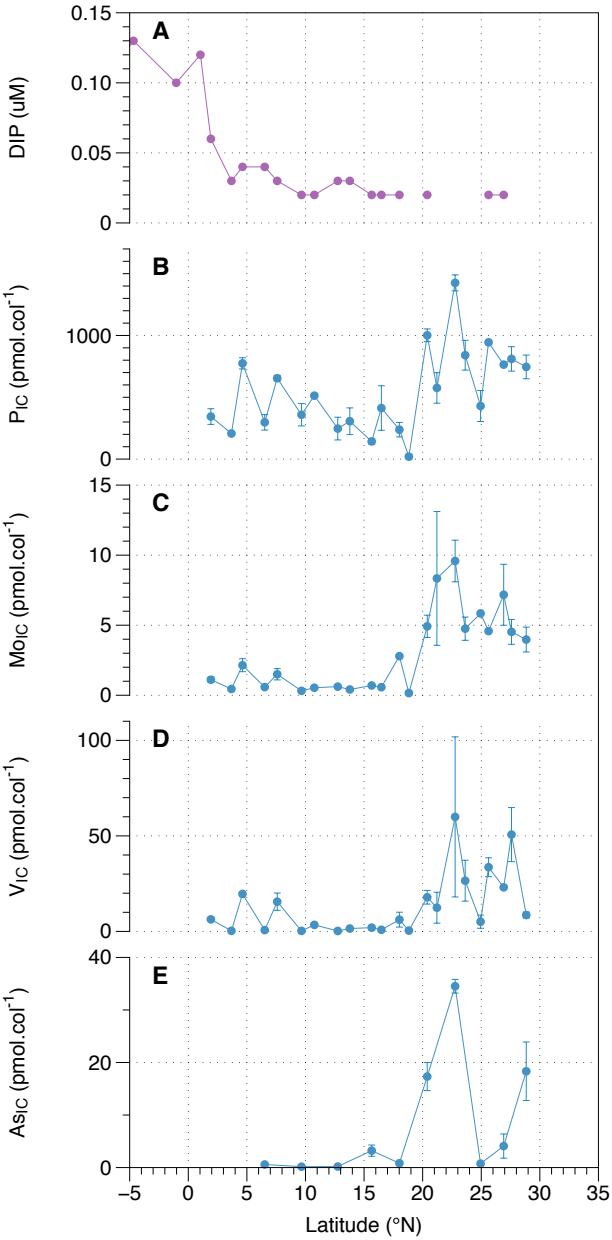

Figure 30: Latitudinal gradient in surface water dissolved inorganic phosphorus concentration (A) alongside intracellular phosphorus (P_{IC}, B), Molybdenum (Mo_{IC}, C), Vanadium (V_{IC}, D) and arsenic (As_{IC}, E) collected during AMT21.

4.3.4 *Dissolved and Particulate comparison*

No clear relationship is observed between intracellular concentrations and surface water dissolved concentration for any element (Figure 31). This is unsurprising given the complexities of metal bioavailability along with the bidirectional feedback of an organisms intracellular stoichiometry and its nutrient source. Moreover Tovar-Sanchez and Sañudo Wilhelmy (2011) observe a similarly weak relationship between *Trichodesmium*'s intracellular concentration and that of the water column.

If however, I partition the Atlantic into 3 distinct regions, as described in Schlosser et al. (2014) and Chapter 3, I observe a similar partitioning with regards to intracellular and dissolved fraction for some other metals. No evidence of *Trichodesmium* was observed in the HPLFe region of the SASG (≈<0°N) therefore I will focus our discussion instead between the LPHFe region of the NASG (≈>11 °N) and the transitional HPHFe region of the ITCZ (0-11 °N). The intracellular concentrations of Mo, Mn, V, As and Ni are all elevated in the NASG region, relative to the ITCZ region, however only DIP, DFe and DCo show significantly different concentrations in the dissolved fraction with the NASG being deplete in both (Figure 32). Intracellular Mo, V and As and dissolved P show a partitioning in-line with our previous discussion on accidental uptake.

4.3.5 *Stoichiometric comparison*

Cellular stoichiometry has long been used as a prognostic indicator of an organism or communities elemental requirements. The efficacy of this technique is reliant on an accurate baseline stoichiometry to which repletion or depletion of a given element can be gauged. Fortunately numerous studies have assessed the elemental stoichiometry of marine organisms (Ho et al., 2003; Quigg et al., 2011; Twining et al., 2011). Similarly numerous studies have assessed the elemental stoichiometry of isolated *Trichodesmium* (Nuester et al., 2012; Sañudo Wilhelmy et al., 2001; Tovar-Sanchez and Sañudo Wilhelmy, 2011; Tovar-Sanchez and Sañudo Wilhemy, 2006).

Considering the mean stoichiometric ratios of *Trichodesmium* samples analysed from D361 (n=51) I see the same general trends in relative abundance seen across most other phytoplankton with N > P > Fe > Zn≈Mn≈V≈Ni≈Cu≈Mo > Cd≈Co (Ho et al., 2003; Quigg et al., 2011). Samples analysed from AMT21 (n = 66) however show a different trend in-tune with the previously discussed V, Mo and As discrepancies C > N > P > Fe > V≈Zn≈As > Ni≈Mo≈Cu≈Mn > Cd≈Co. A substantial flexibility in *Trichodesmium*'s stoichiometric composition is observed when comparing our data to the environmental observations of Tovar-Sanchez

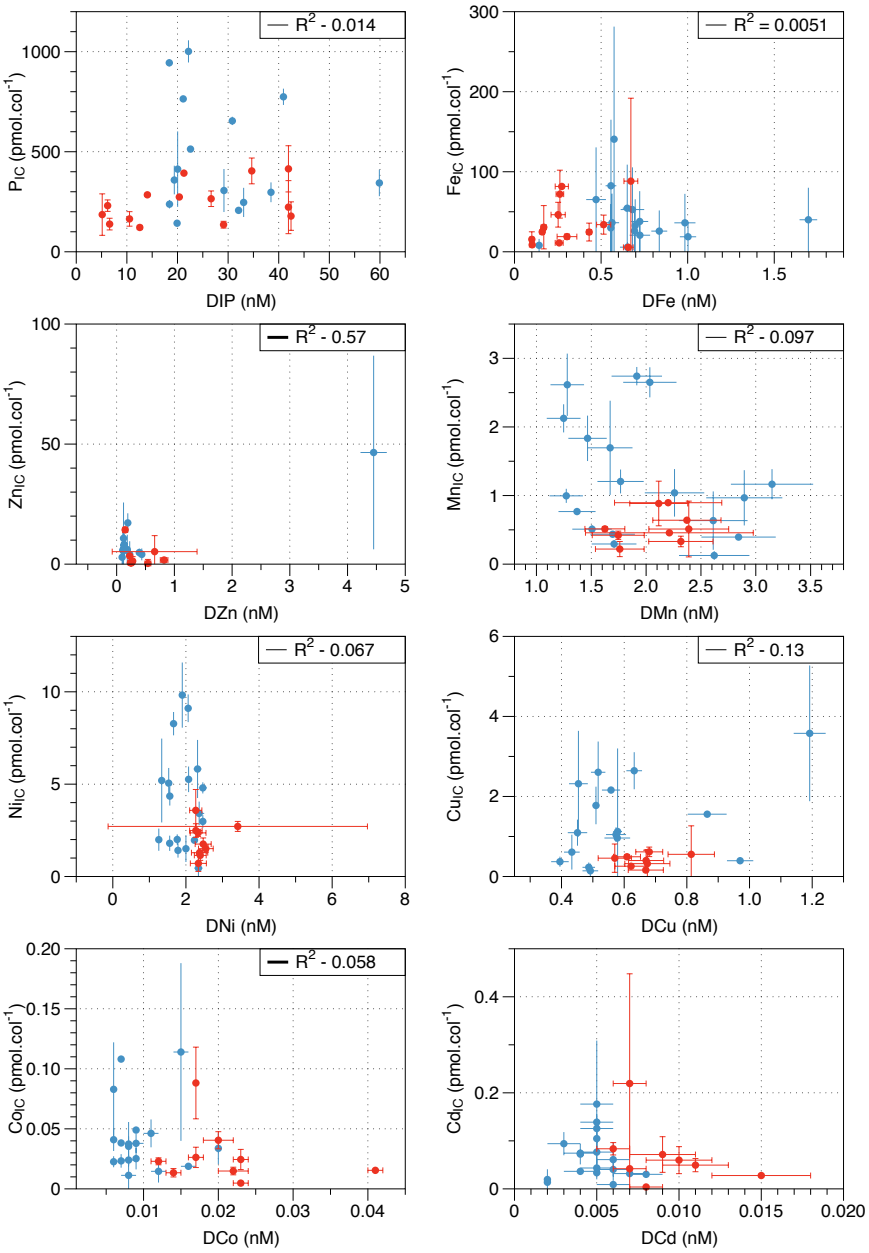

Figure 31: Comparison of surface water dissolved fraction and *Trichodesmium* particulate fraction for select metals. Data displayed is from D361 (red) and AMT21 (blue), error bars represent 1 s.d. of triplicate measurements.

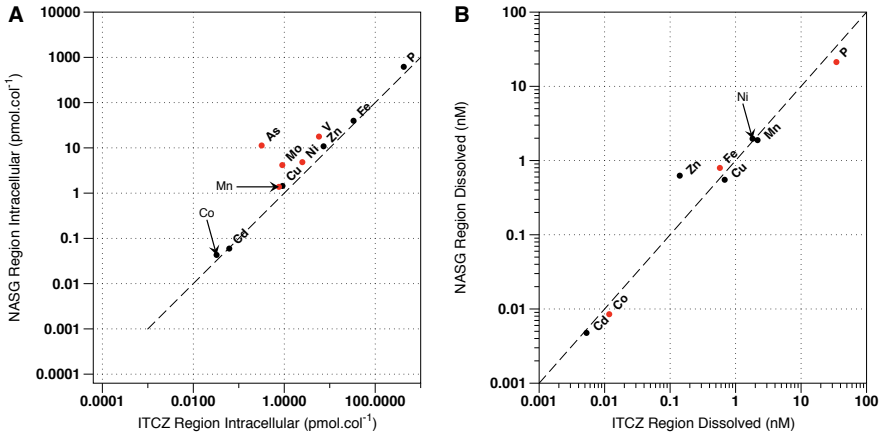

Figure 32: Mean *Trichodesmium* intracellular metal concentrations (A) and surface water dissolved metal concentrations (B) for the LPHFe NASG region (>11°N) and the transient HPHFe region of the ITCZ (0-11°N). Statistically different values are highlighted in red as determined by a student's t-test (P<0.05).

and Sañudo Wilhemy (2006),Tovar-Sanchez and Sañudo Wilhelmy (2011) and Nuester et al. (2012) but also the *Trichodesmium erythraeum* IMS101 culture observation of Quigg et al. (2011) (Figure 33). Our intracellular Cd, Co and Cu data are all shown to be marginally yet consistently enriched in *Trichodesmium* colonies isolated from the central Atlantic (Figure 33), the reason for this enrichment is unclear. Meanwhile V, Mo, Zn, Ni, Mn and Fe all show variable patterns of enrichment or depletion relative to the 4 literature data sets in question.

For both AMT21 and D361 our Fe_{IC} concentrations are very similar to those observed in cultured *Trichodesmium erythraeum* IMS101 (Quigg et al., 2011) however when compared to environmental *Trichodesmium* our Fe_{IC} appears marginally elevated (Nuester et al., 2012; Sañudo Wilhelmy et al., 2001; Tovar-Sanchez and Sañudo Wilhelmy, 2011; Tovar-Sanchez and Sañudo Wilhemy, 2006). The (sub)-tropical Atlantic is well documented as an area of elevated DFe input (Schlosser et al., 2014; Jickells et al., 2005) where average DFe concentrations for the NASG and ITCZ influenced region during AMT21 were ≈0.79 nM and ≈0.57 nM respectively (Figure 32). The Fe_{IC} concentrations observed by Nuester et al. (2012) in the Sargasso Sea and Tovar-Sanchez and Sañudo Wilhelmy (2011) in the Amazon River plume were also recorded in regions of episodic DFe input. The Sargasso sea has previously been observed to have a DFe gradient ranging from 0.3-1.6 nM (Shelley et al., 2012) whilst Tovar-Sanchez and Sañudo Wilhelmy (2011) observed highly variable DFe concentration along a pronounced salinity gradient (0.41 - 8.70 nM). As noted in

Section 4.3.4 little correlation is observed between the dissolved and intracellular fraction both with regards to data presented in this study or data presented in Tovar-Sanchez and Sañudo Wilhelmy (2011). As such I conclude that, above the R^*_{Fe} value outlined in Chapter 3, DFe concentration has minimal affect on *Trichodesmium*'s Fe_{IC} concentration. As discussed previously (Section 4.3.3) V_{IC} and Mo_{IC} observed during D361 are substantially lower than those observed during AMT21. D361 V_{IC} and Mo_{IC} are similar to concentrations observed by Tovar-Sanchez and Sañudo Wilhelmy (2011) whilst AMT21 concentrations more closely match those observed in the P deplete Sargasso Sea (Nuester et al., 2012).

Metal	Protein	Reference
Fe	Cytochromes Ferredoxin Other Fe-S proteins Chelatase Nitrogenase Catalase Peroxide Superoxide Dismutase	(Berman-Frank et al., 2003; Kustka, 2003a)
Zn	Carbonic Anhydrase Alkaline Phosphatase RNA Polymerase tRNA synthetase Reverse Transcriptase Carboxypeptidase Superoxide Dismutase	(Yentsch et al., 1972; Mulholland et al., 2002) (Carpenter et al., 1992)
Ni	Urease Superoxide Dismutase	(Dupont et al., 2008; Ho, 2013; Ho et al., 2013)
Cu	Plastocyanin Cytochrome oxidase Ascorbate oxidase Superoxide dismutase Multicopper ferroxidase	(Ferreira and Straus, 1994) (Carpenter et al., 1992)
Mo	Nitrate Reductase Nitrogenase	(Berman-Frank et al., 2003; Kustka, 2003a)

Table 15: Common Fe, Zn, Ni, Cu and Mo metalloproteins found in marine phytoplankton. Green text signifies a metalloprotein of significant cyanobacterial or *Trichodesmium* importance. Red text signifies proteins known not to be present in cyanobacteria or *Trichodesmium* specifically. Table is modified from Twining and Baines (2013).

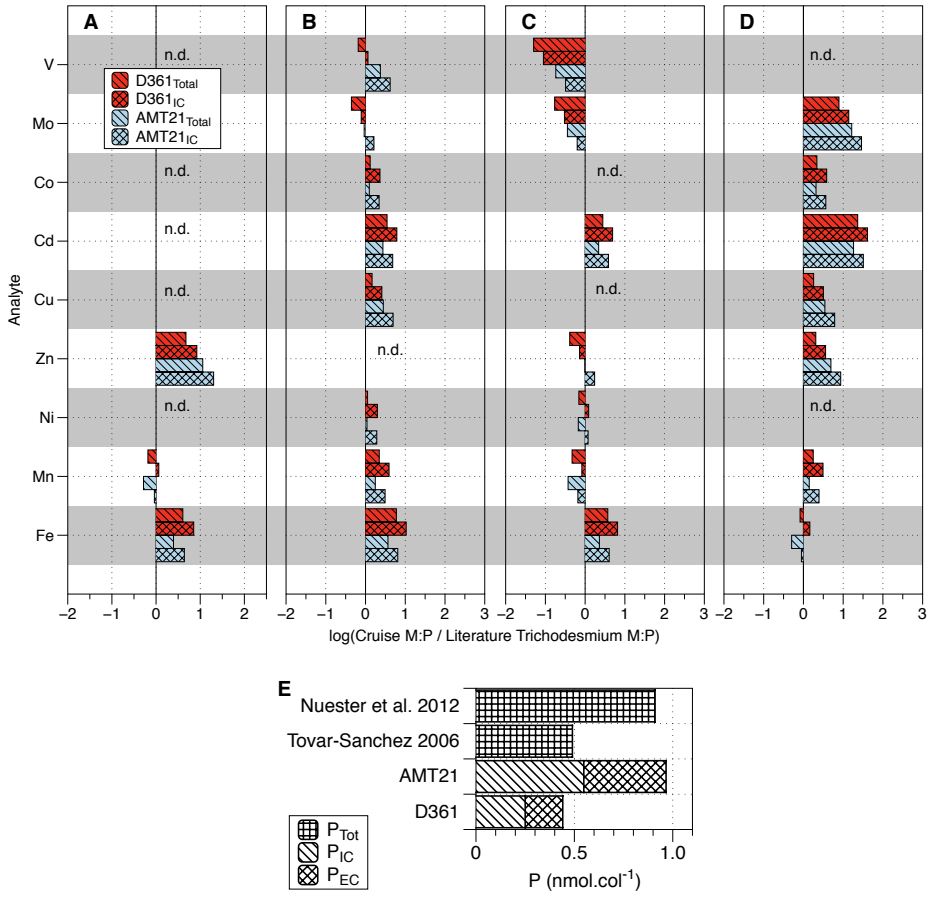

Figure 33: Mean *Trichodesmium* elemental stoichiometry from D361 and AMT21 compared that published in the literature for culture and environmental *Trichodesmium*. Data are displayed as log transformed enrichment factors relative to literature derived *Trichodesmium* stoichiometries (=log(Sample M:P / Literature M:P). *Trichodesmium* specific comparisons for A)Tovar-Sanchez and Sañudo Wilhemy (2006), B) Tovar-Sanchez and Sañudo Wilhelmy (2011), C) Nuester et al. (2012) and D) Quigg et al. (2011). Given the majority of literature sources did not implement a washing procedure to correct for extracellular bound metals, M:P enrichment factors are displayed for both total P and intracellular only quotas when considering direct *Trichodesmium* comparisons. Plot F) shows direct P_{IC} washed and unwashed comparisons with (Nuester et al., 2012; Tovar-Sanchez and Sañudo Wilhemy, 2006).

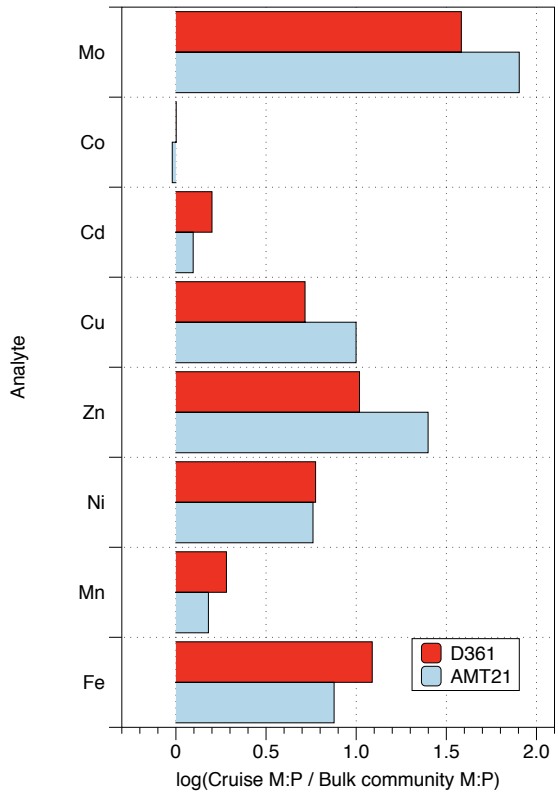

Figure 34: Mean *Trichodesmium* elemental stoichiometry from D361 and AMT21 compared to the elemental stoichiometry of bulk phytoplankton published in the literature. Data are displayed as log transformed enrichment factors relative to literature derived bulk community stoichiometry (=log(Sample M:P / Literature M:P). Literature sources include Ho et al. (2003); Quigg et al. (2011); Twining et al. (2011).

Comparing these data to those published for bulk phytoplankton I see a distinct pattern of elemental enrichments (Figure 34). Ratios of Mo:P, Cu:P, Zn:P, Ni:P and Fe:P all show substantial enrichment in *Trichodesmium* relative to the bulk phytoplankton community (Ho et al., 2003; Quigg et al., 2011; Twining et al., 2011). Understanding the diazotrophic strategies utilised by *Trichodesmium* provides insight into these enrichments (Table 15).

As outlined in Chapter 1, *Trichodesmium* have a uniquely high iron requirement due to a reliance on their photosynthetic and diazotrophic apparatus. This elevated iron requirement is further compounded by *Trichodesmium*'s unique spatio-temporal strategy for both C_{Fix} and N_{Fix} limiting any potential for intracellular iron recycling (Berman-Frank et al., 2001b; Saito et al., 2011). There are three documented metallotypes of nitrogenase, each of which requires a specific nitrogenase reductase, NifH supporting MoFe-nitrogenase, VnfH supporting VFe-nitrogenase, and AnfH supporting FeFe-nitrogenase (Bothe et al., 2010). Of these three metallotypes the *Trichodesmium erythraeum* IMS101 genome contains all required subunits for NifDK (Tery_4137 and Tery_4138 encoding for MoFe-nitrogenase) but lacks required subunits for both VFe-nitrogenase and FeFe-nitrogenase (Eady, 2003). Further to this VnfDGK expression in *Azotobacter vinelandii* occurs only under reduced Mo conditions or at low temperatures. Given the consistently high Mo concentration in seawater (≈100 nM)(Collier, 1985) and *Trichodesmium*'s absence of required subunits for VFe-nitrogenase it seems highly likely that the MoFe-nitrogenase enzyme is responsible for both Fe and Mo enrichments relative to the bulk community (Figure 34).

Alkaline phosphatase, the enzyme responsible for enabling access to phospho-ester DIP is well documented in *Trichodesmium* (Yentsch et al., 1972; Mulholland et al., 2002; Orchard et al., 2010a) where it has been found to account for upwards of 25% of the *Trichodesmium*'s P uptake (Orchard et al., 2010a). Of the three putative alkaline phosphatases encoded by *Trichodesmium*, *phoA* and *phoX* are both pho box regulated and, whilst currently unconfirmed in *Trichodesmium*, *phoA* is known to require a zinc containing co-factor (Wu et al., 2007). Whilst many marine organisms are thought capable of P-ester hydrolysis by alkaline phosphatase, studies have shown *Trichodesmium* colonies to be better competitors for P-ester uptake than for inorganic P uptake (Orchard et al., 2010a; Sohm and Capone, 2006). Our understanding of the biological role of zinc in cyanobacteria such as *Trichodesmium* is poorly explored compared with that of Fe or Mo. However this propensity for P-ester hydrolysis could result in an elevated cellular zinc requirements as observed across both D361 and AMT21.

Reactive oxygen species (ROS) produced during photosynthesis irreversibly inactivates the nitrogenase Fe-protein (Gallon, 1981). The

spatio-temporal segregation of nitrogen fixation and photosynthesis is one strategy by which *Trichodesmium* combat this oxygenic inhibition of nitrogenase (Berman-Frank et al., 2003). Additionally the cellular antioxidant enzyme, superoxide dismutase (SOD), is thought to play an important role in the disproportionation of superoxide radicals produced during photosynthesis. Whilst numerous different SOD species have been documented, the *Trichodesmium erythraeum* IMS101 genome shows only nickel-based and manganese-based SODs (NiSOD and MnSOD respectively) (Dupont et al., 2008; Ho, 2013; Ho et al., 2013). Culture studies have demonstrated an upregulation in NiSOD activity and Ni uptake rates both when *Trichodesmium* is grown diazotrophically versus growing on a source of fixed N, but also at increasing light levels (Ho, 2013; Ho et al., 2013). These observations strongly imply an important role of NiSOD in combating oxygenic inhibition of *Trichodemsium*'s nitrogen fixation. *Trichodesmium*'s intracellular Ni enrichment is likely a result of an elevated NiSOD requirement compared with non-diazotrophic phytoplankton.

Twining and Baines (2013) list 5 common copper containing metalloproteins as being present in marine phytoplankton; plastocyanin, cytochrome oxidase, ascorbate oxidase, superoxide dismutase and multicopper ferroxidase. *Trichodesmium* lack any copper containing SODs and show no evidence of elevated cytochrome oxidase nor ascorbate oxidase. An abundance of the Cu-containing enzyme plastocyanin, a functional replacement for the Fe-containing cytochrome c_{553} of PSII could attribute *Trichodesmium* with an elevated stoichiometric requirement for copper. Although the substitution of cytochrome c_6 for plastocyanin would reduce *Trichodesmium*'s cellular Fe burden, it is currently believed that plastocyanin in cyanobacteria is regulated by Cu rather than Fe supply (Peers and Price, 2006). Nonetheless the use of plastocyanin in cyanobacteria including *Trichodesmium* could result in a cellular Cu enrichment relative to a non-cyanobacteria bulk community. Multicopper ferroxidase is an oxidoreductase enzyme that catalyses the oxidation of Fe(II) to Fe(III). *Trichodesmium* encode two proteins (Tery_2878 and Tery_2879), homologous to the ferroxidase proteins FeoA and FeoB present in *Synechocystis*, however their metal centres are currently unknown.

4.3.6 *Metal Use Efficiency*

4.3.6.1 *Iron and Molybdenum*

Using in-situ *Trichodesmium* nitrogen fixation rates, alongside literature sourced values for both the stoichiometry and catalytic rate of the nitrogenase enzyme complex, it is possible to asses what proportion of observed intracellular iron that would be predicted to be asso-

ciated with nitrogen fixation. A number of assumptions and literature values have been used in this analysis, these include:

- Here I assume the nitrogenase complex to be operating at its maximal reaction rate, therefore a NifH:NifD/K stoichiometry of 5:2 (2.5:1) is assumed (Kustka, 2003a) (discussed further in Chapter 5). This results in an Fe:N_2ase stoichiometry of 50:1 and a Mo:N_2ase of 2:1 given each NifH protein contains 4 Fe and each NifDK pair contains 15 Fe
- Nitrogenase's maximal catalytic rate as observed in *Azotobacter vinelandii* is 1.33 mol N. mol N_2ase.S^{-1} (Vichitphan, 2001). Given the similarities between *Azotobacter vinelandii* and *Trichodesmium*'s nitrogenase this maximal reaction rate will be used for *Trichodesmium*.
- Our observed N_{Fix} rates, measured over the entire 12 hour photoperiod and previously represented in units of nmol $N_{Fix}.col^{-1}.d^{-1}$, are converted to nmol $N_{Fix}.col^{-1}.s^{-1}$ assuming a 6 hour active diazotrophic period as described in Berman-Frank et al. (2007).

Using the literature values for NifH:NifDK stoichiometry (Kustka, 2003a) and N_2ase's maximal catalytic rate (Vichitphan, 2001) I can calculate a theoretical NifH maximal catalytic rate :

$$\mathrm{NifH\,rate\ (mol\,N.mol\,NifH^{-1}.s^{-1})} = \frac{\mathrm{N_2ase\,rate}}{\mathrm{NifH:NifDK}}$$

$$\frac{1.33\ \mathrm{mol\,N.mol\,N_2ase^{-1}.s^{-1}}}{2.5} = 0.532\ \mathrm{(mol\,N.mol\,NifH^{-1}.s^{-1})} \quad (4)$$

This NifH specific maximal reaction rate is calculated in the same manner as in Kustka (2003a) although they assumed a NifH:NifDK ratio of 4:2 or 6:2 to produce a NifH maximal reaction rate of 0.44-0.67 mol N.mol N_2ase.s^{-1}. Utilising our NifH maximal reaction rate of 0.532 mol N.mol N_2ase.s^{-1} and my in situ nitrogen fixation observations I can calculate a theoretical concentration of the NifH protein:

$$\mathrm{[NifH]\ mol.col^{-1}} = \frac{\mathrm{N_{Fix}\,Rate\ (fmol.col^{-1}.s^{-1}}}{\mathrm{NifH\,rate\ (mol\,N.mol\,NifH^{-1}.s^{-1})}} \quad (5)$$

Leading on from this I am able to calculate the amount of iron required to enable our observed nitrogen fixation rates:

$$\mathrm{Nitrogenase\ Fe(pmol.col^{-1})} = \frac{(\mathrm{[NifH]} \times 4\,\mathrm{Fe} + \mathrm{[NifDK]} \times 15\,\mathrm{Fe})}{1000}$$

$$\text{Nitrogenase Fe}(\text{pmol.col}^{-1}) = \frac{([\text{NifH}] \times 4\,\text{Fe} + \frac{[\text{NifH}]}{2.5} \times 15\,\text{Fe})}{1000} \quad (6)$$

Expressing this nitrogenase iron as a percentage of observed intracellular iron allows us to speculate as to the true iron burden of nitrogen fixation (Table 16).

CRUISE	FE_{IC} [a]	MO_{IC} [a]	N_{Fix} [b]	NIFH REACTION RATE[c]	% OF ACTUAL IRON	% OF OBSERVED MO_{IC}
D361	34.01	0.50	11.42	0.532 [1]	1.23 ± 0.50 %	1.57 ± 0.46 %
				0.056 [2]	11.67 ± 4.74 %	14.78 ± 4.40 %
AMT21	26.24	2.60	27.49	0.532 [1]	3.92 ± 1.86 %	3.70 ± 2.83 %
				0.056 [2]	37.19 ± 17.66 %	35.21 ± 26.88 %

Table 16: [a] Mean intracellular Fe/Mo ($pmol.col^{-1}$).
[b] Measured *Trichodesmium* specific N_{Fix} rates ($fmol.col^{-1}.s^{-1}$).
[c] Nitrogenase catalytic activity (mol N. mol $N_2ase^{-1}.s^{-1}$) as either [1] the theoretical maximal reaction rate (Vichitphan, 2001) or [2]assessed *Trichodesmium erythraeum* IMS101 cultures (Levitan et al., 2010).
[d] % of intracellular Fe/Mo allocated to nitrogen fixation.

Using nitrogenase's theoretical maximal reaction rate calculated above allows us to account for an average of only 1.23 ± 0.50 % or 3.92 ± 1.86% of the observed intracellular iron during D361 or AMT21 respectively (Table 16). Levitan et al. (2010) directly assessed *Trichodesmium*'s NifH reaction rate under typical environmental conditions (25°C and 400 μAtm pCO_2), with no nutrient stressors. Here *Trichodesmium erythraeum* IMS101 was shown to have a NifH reaction rate ranging from ≈0.027-0.095 mol N.mol N_2ase.s^{-1} (mean ≈0.056 mol N.mol N_2ase.s^{-1}). Utilising this observed mean NifH reaction rate I am then capable of accounting for 11.67 ± 4.74 % and 37.19 ± 17.66 % of the observed intracellular iron during D361 or AMT21 respectively.

Whilst initially our calculation appear to attribute less of the observed intracellular iron to the nitrogen fixation apparatus than the ≈80-90% observed by Richier et al. (2012) We must consider additional factors that result in our estimations being the minimum possible. The NifH reaction rate observed by Levitan et al. (2010) was for cultures grown under no nutrient, temperature, pH or light limitation and so represent fairly idealised growth conditions. Meanwhile our nitrogen fixation rates were observed from *Trichodesmium* populations thought to be either Fe or P limited (Chapter 3); it is therefore reasonable to assume these populations may not be fixing nitrogen at a comparable efficiency as the well controlled culture experiments. Furthermore the recent re-evaluation of the $^{15}N_2$ bubble method for the determination of nitrogen fixation suggesting an ≈2-fold systematic underestimation of gross N_{Fix} rates (Groß kopf et al., 2012). A hypothetical doubling of our N_{Fix} rates would in-fact double our percentage of Fe attributable to nitrogenase to 23.34 % and 74.39 % for D361 and AMT21 respectively. This substantial iron burden, attributed to nitrogen fixation, is in addition to that of the photosynthetic apparatus present in non-diazotrophic phytoplankton and supports our observed stoichiometric enrichment of Fe relative to the bulk community (Figure 34).

An abundance of literature on Fe-containing metalloenzymes allows for the above cellular budgeting approach to provide a reasonable approximation as to *Trichodesmium*'s Fe_{1C} allocation. Our calculations provide potential minimum values for the percentage of iron attributed to nitrogen fixation and as such appear as under-estimations when compared with previous estimates of 38-90 % (Kustka, 2003a; Richier et al., 2012). Despite this our data still highlight the substantial iron burden associated with a diazotrophic life style. A similar approach is not currently possible for the stoichiometric enrichments of Zn, Cu and Ni. To our knowledge, no rate or concentration data exists for NiSOD, plastocyanin or any other Zn or Cu containing metalloproteins in *Trichodesmium*.

4.4 CONCLUSION

In this chapter I have presented data for the metallome of environmental *Trichodesmium* populations sampled during two oceanic cruises along a transient nutrient gradient in the central Atlantic. Redfield theory, whereby dissolved and particulate fractions show similar stoichiometry, does not appear extend to *Trichodesmium*'s intracellular trace metal composition. Instead I see poor correlation in all dissolved metal and particulate metal comparisons. Although *Trichodesmium*'s intracellular composition showed marked variability along our environmental DIP/DFe gradient numerous elements showed a clear metal enrichment in the DFe rich, dust-deposition influenced, NASG when compared with the equatorial and SASG regions. The concentration of DFe is clearly of significant importance to *Trichodesmium*. However, aside from the above mentioned geographic split, the absolute concentration of DFe explains little of our observed Fe_{IC} variability. I suggest that *Trichodesmium* have an inherently flexible intracellular composition within regions where DFe > R^{*}_{Fe} (Chapter 3). Here cellular adaptations such as the compensatory strategies of altering photosystem and nitrogenase stoichiometries allow *Trichodesmium* to maintain niche dominance by altering their intracellular composition.

Of note were the AMT21 observations of substantial elevations in V:P, Mo:P and As:P in the P deplete NASG potentially indicative of a phosphorus stressed population. I suggest, as others have done previously for vanadium (Nuester et al., 2012), that these enrichments are a mistaken uptake of vanadate, molybdate and arsenate via a phosphate uptake pathway in this DIP deplete environment. Moreover our molybdenum use efficiencies fail to account for upwards of 90% of the observed intracellular Mo when the largest Mo-based metalloenzyme (NifD) is accounted for.

Our data attributes *Trichodesmium* with a unique metallomic signature in-keeping with the organisms previously understood cellular processes. Enrichments in Fe and Mo are typical of diazotrophs whilst Zn, Cu and Ni enrichments associate more broadly with cyanobacterial processes. *In situ* nitrogen fixation rates allow us to attribute upwards of ≈12-38% of our observed intracellular iron concentration to *Trichodesmium*'s diazotrophic growth strategy. This is in support of the findings of Richier et al. (2012), Kustka (2003a) and many others. *Trichodesmium*'s observed stoichiometric composition re-iterates the Fe burden associated with diazotrophs whilst highlighting the potential for Ni, Cu or Zn limitation. In order to fully elucidate the function roles of these metal enrichments careful study of *Trichodesmium*'s proteome under a variety of environmental conditions is warranted.

5

LABEL-FREE QUANTITATIVE PROTEOMIC ANALYSIS OF *TRICHODESMIUM ERYTHRAEUM* IMS101 IN RESPONSE TO IRON-LIMITATION

5.1 INTRODUCTION

To stay alive and remain competitive all organisms need to adapt to changes in their environment. Parameters such as temperature, light, pH and nutrient availability vary both spatially and temporally in the global oceans (Bunt, 1995; Arrigo, 2005; Fabry et al., 2008). Chapter 3 and Chapter 4 established the precedence that the ecologically important cyanobacteria, *Trichodesmium*, readily adapt their elemental composition in response to environmental variability and by doing so are able to span large variations in available Fe and P. Between 30-65% of an algal cells biomass is attributed to the proteome (Laws, 1991; Geider and Roche, 2002), these protein complexes carry out almost all of the biochemical, signalling and functional processes within a cell. Changes in the abundance and activity of these proteins play a large part in facilitating an organism's responses to environmental perturbations.

Trichodesmium sp. are a globally important diazotrophic cyanobacteria (Carpenter et al., 2004; Capone et al., 2005), their wide distribution across much of the tropical oceans is in spite of their potentially uniquely elevated cellular iron requirement (Kustka, 2003a,b). This iron burden is caused by their strategy of near concurrent oxygen-producing photosynthesis and oxygen sensitive nitrogen fixation (Berman-Frank et al., 2001b), two antagonistic cellular processes both requiring iron-rich enzyme systems. These antagonistic processes are both conducted during the photoperiod without specific complete segregation of the nitrogenase enzymes into discreet cells as performed by heterocystous cyanobacteria (Kana, 1993; Berman-Frank et al., 2001b; Küpper and Ferimazova, 2004; Milligan et al., 2007). Numerous proteomic adaptations for coping with this iron burden under frequent conditions of iron scarcity have been identified in *Trichodesmium*. These include the iron stress induced proteins IsiA (Shi et al., 2007; Richier et al., 2012) and IsiB (Chappell and Webb, 2010), a reworking of the photosynthetic and nitrogenase apparatus (Shi et al., 2007; Richier et al., 2012) and changes in iron uptake and storage processes (Sandh et al., 2011; Chappell and Webb, 2010). These proteomic changes have been probed through targeted research and as such have been assessed in relative isolation to one another.

Recent advances in proteomic technologies and their associated data processing software offer us the ability to elucidate an organisms mechanistic adaptations to environmental change. Theoretically these methodologies, when working with microbial organisms, offer us the ability to gain a snap-shot of the abundance of all cellular proteins simultaneously. Here I present data from a controlled iron-limitation study performed on *Trichodesmium erythraeum* IMS101 cultures. An iron stress and time-of-day change in proteomic composition was assessed using a label-free quantitative proteomic approach termed MS^E, this novel approach improves upon the qualitative nature of shotgun studies and offers greater coverage than the selective AQUA approach (Gerber et al., 2003).

I aim to describe the various mechanisms by which *Trichodesmium* are able to adapt to iron-deficiency whilst offering a revised and expanded understanding of the unique spatio-temporal strategy used to achieve both photosynthesis and nitrogen fixation concurrently.

5.2 METHODS

5.2.1 *Culture Conditions*

Trichodesmium erythraeum strain IMS101 was grown in filter sterilised, modified YBC-II growth media (Chen et al., 1996) for a total duration of 9 days. Trace metal concentrations were modified in line with those of Shi and Kranz (2012) to result in final concentrations of Cu - 8 nM, Zn - 20 nM, Co - 8 nM, Mn - 18 nM, Mo - 100 nM, Ni - 20 nM, Se - 10 nM and EDTA - 20 μM. Batch cultures were grown in 25 cm^3, 0.2 μm vented sterile polystyrene flasks (Corning Inc., NY, USA). Growth conditions were maintained at 27°C, 12:12 hour light:dark cycle at 100-160 $\mu mol.photons.m^{-2}.s^{-1}$ and subject to gentle agitation using an orbital shaker. Culture inoculations were performed using late exponential phase cells grown at 160 nM Fe. Inoculant was pre-concentrated onto 5 μm track etched polycarbonate membranes (Nucelopore, Whatman, Kent, UK), washed with iron-free modified YBC-II growth media. Cells were re-suspended in a small volume of iron-free growth media before being used to inoculate experimental flasks. Iron replete (Fe+, 400 nM), iron deplete (Fe-, 0 nM) and iron re-fed (initially 0 nM, later 400 nM) cultures were established. A set of triplicate cultures from each treatment were used for daily physiological assessment to establish iron-limitation whilst cultures for proteomic analysis remained sealed throughout the experiment in order to minimise potential iron contamination.

5.2.2 *Photosynthetic Physiology*

Daily assessment of each treatments photosynthetic physiology were performed 2.5 hours after the onset of the photo-period using a FASTtracka MkII Fast Repetition Rate fluorometer (FRRf) integrated with a FastAct™ Laboratory system (Chelsea Technologies Group Ltd, Surrey, UK). Cells were dark acclimated for 20 minutes upon sampling and spent 5 minutes at 21 μmol.photons.m^{-2}.s^{-1} immediately prior to analysis (Richier et al., 2012). F_V/F_M was used as an estimate of the apparent photosystem II (PSII) photochemical quantum efficiency whilst σ_{PSII} indicated the functional cross section of PSII (Kolber et al., 1998). Analyses were performed in triplicate with absolute values of F_V/F_M and σ_{PSII} being gain and blank corrected using cell-free growth media. Data shown are mean values of these triplicates with statistical significance determined using a student's t-test.

5.2.3 *Cell counts*

Cell counts were performed using a Sedgewick rafter counting chamber. Photographs were taken using a GX CAM-1.3 camera (GT Vision Ltd, Suffolk, UK) on an L1000A biological microscope (GT Vision Ltd, Suffolk, UK). Cell and filament lengths were identified using GX capture software (GX Optical, Suffolk, UK) and the number of cells per ml was calculated by dividing the total filament length by the average cell length. Cell counts were performed for in triplicate for each treatment every 2 days. Statistical differences was determined using a student's t-test.

5.2.4 *Statistical Analysis of Culture Data*

Data collected for cell counts, F_V/F_M, σ_{PSII} and relative chlorophyll concentration were subject to outlier identification and removal following Grubbs' test for outliers. A student's t-test was used to determining statistical significance ($n=3$, $P<0.05$).

5.2.5 *Sampling*

Sampling was conducted on day 7 of the experiment with cultures being harvested ($\approx$2 hours after onset of photoperiod and $\approx$6 hours after onset of photoperiod, these sampling points were denoted as PS for photosynthesis and NF for nitrogen fixation respectively according to the metabolic processes expected to dominant at the corresponding times of day (Berman-Frank et al., 2001a) (Figure 35). Cultures were gently filtered onto glass-fibre filters (Whatman) using an acid washed filtration rig. Samples were placed in 1.5 ml tubes (Eppendorf), snap-frozen in liquid nitrogen and stored at -80°C until

analysis. In total 4 treatments were assessed during this study, high iron 'photosynthesis' (PS+), high iron 'nitrogen fixation' (NF+), low iron 'photosynthesis' (PS-) and low iron 'nitrogen fixation' (NF-).

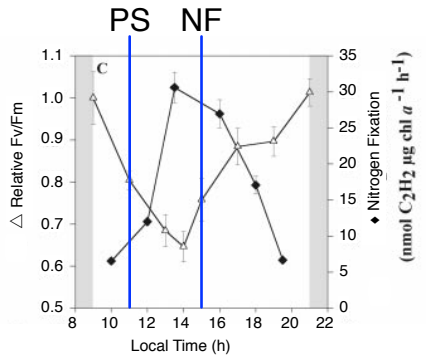

Figure 35: *Trichodesmium*'s diel cycle as published by Berman-Frank et al. (2001a). Observed variations in relative F_V/F_M and nitrogen fixation rates across the photoperiod are shown. Sampling time points for this study are overlaid for reference only. Time point "PS" was at 2 hours past the onset of the photoperiod whilst time point "NF" was at 6 hours. Figure modified from Berman-Frank et al. (2001a).

5.3 PROTEIN PREPARATION AND DIGESTION

Protein extraction was performed in 1x protein extraction buffer (1 x PSB) (140 mM Tris base, 105 mM Tris-HCl, 0.5 mM EDTA, 2% lithium dodecyl sulfate and 10% v/v glycerol). To each sample containing filter 500 µL of 1 x PSB was added in a 1.5 ml centrifuge tube before being frozen in liquid nitrogen. For extraction, frozen samples underwent 3 cycles of freeze-sonication (microtip attachment, duty cycle 35%) using a VibraCell sonicator (Sonics and Materials, Inc) until just thawed. Extracted protein and macerated filters were separated using 30 µm pore size polyethylene centrifuge filters (Pierce no. 89868, Thermo Scientific).

A 200 µL aliquot of extracted protein was acetone precipitated at 4:1 acetone:sample and -20°C for 1 hour. After precipitation samples were separated from acetone supernatant by centrifugation (10 minutes at 13,000 g), an aliquot of the supernatant was preserved for determination of chlorophyll concentration whilst the protein pellet was allowed to dry before being re-suspended in 40 µL 1 x PSB. This provided a 5x concentration of protein samples.

Precipitated protein was isolated using a Novex NuPAGE 10% polyacrylamide gel. Sample protein was allowed to migrate through the polyacrylamide gel to the 4% loading and 10% resolving division.

This change in gel density allowed for the sample protein to stack in a narrow band above the 10% resolving region. An ≈1 mm wide gel band containing all sample protein was excised and cut into ≈1 mm^3 cubes. Excised gel cubes were reduced using 10 mM dithiothrietol for 10 minutes at 80°C followed by alkylation with 50 mM iodoacetamide at room temperature in the dark for 30 minutes.

Reduced and alkylated gel cubes were then washed twice with 25 mM ammonium bicarbonate in 50% acetonitrile then allowed to dehydrate for 15 minutes in 100% acetonitrile. Dehydrated gel cubes were rehydrated with 10 µL activated trypsin solution (10 ng/µL Trypsin). Samples were then placed at 30°C to digest overnight. Upon complete digestion the remaining trypsin was quenched using 5 µL 0.1% Formic Acid, digested peptides were removed from the gel cubes by acetonitrile dehydration and finally vacuum centrifuged to dryness before being stored at -80 °C.

Prior to MS analysis peptide biomass was assessed using a Millipore Direct Detect. Peptide concentration suggested a protein recovery from digestion of ≈20-43%. Samples were biomass normalised and spiked with reference peptides (Yeast Enolase P00924, Rabbit Glycogen Phosphorylase P00489) so as to result in 100 fmol of each standard being introduced on column.

5.3.1 *1D-UPLC-MS^E*

Separations were performed using a nanoAcquity UPLC system (Waters). 1.5 µl of the prepared protein lysates (500 ng on column) containing 100 fmol of the internal standard digests Phosphorylase B and Enolase digest were injected onto a Symmetry C18, 180 µm x 20 mm trapping cartridge (Waters). After 5 min washing of the trap column, peptides were separated using a 75 µm ID x 200 mm, 1.7 µm BEH130 C18, column (Waters) using a linear gradient of 5 to 40% B (buffer A = 0.1% formic acid in water, buffer B = 0.1% formic acid in acetonitrile) over 90 min with a wash to 85% B at a flow rate of 300 nl/min. All separations were automated and performed on-line to a mass spectrometer.

All mass spectrometry was performed using a Waters G2-S HDMS mass spectrometer operating in MS^E mode. Data was acquired from 50 to 2000 m/z using alternate low and high collision energy (CE) scans. Low CE was 5V and elevated, ramped from 20-40V. The lock mass Glu-fibrinopeptide, $(M+2H)^{+2}$, m/z = 785.8426) at a concentration of 500 fmol/µl was infused at 250 nl/min and acquired every 13 seconds.

5.4 DATA PROCESSING AND STATISTICAL ANALYSIS

LC-MS^E data were processed using Protein Lynx Global Server (PLGS) version 2.5 for submission to the IDENTITY^E search engine (Waters corporation, Milford, MA). Briefly, LC-MS^E spectra were lock-mass corrected, centroided, deisotoped and charge state reduced and intensity measurements reduced (Geromanos et al., 2009). Precursor and fragment ion for each detected compound were expressed as an accurate mass retention time (AMRT) pair. These AMRT pairs were then submitted to IDENTITY^E for database searching. Processed LC-MS^E data was submitted to the IDENTITY^E database search algorithm version 2.5. Details of the search strategy can be found in Geromanos et al. (2009) and Li et al. (2009). Each processed file was searched against a protein translation of the *Trichodesmium erythraeum* IMS101 genome sequence acquired from Uniprot on 20/11/2012.

Search parameters were as follows: Automatic precursor and product ion tolerances were set. Enzyme specificity was set to tryptic; fixed modifications included cysteine alkylation to carbamidomethyl; variable modifications included N-terminal acetylation and methionine oxidation. A maximum of 2 missed cleavages were allowed. A false discovery rate of 4% was applied. Each peptide was assigned with $\geqslant 3$ fragment ions and each protein was identified with $\geqslant 3$ different peptides (Silva et al., 2006b).

Individual protein concentrations were normalised between samples by dividing the observed concentration of the protein in question with the sum of all protein concentrations observed. This mitigated any variability between sample replicates and allowed for more accurate comparison. For each protein mean concentration values from the three technical replicates was calculated, statistical significance was then determined using this mean value from each of the three biological replicates analysed per experimental condition. For proteins to which no data were collected at all for a given treatment (i.e.an absence in all biological replicates) no statistical analysis was performed and a presence/absence response recorded. For treatments where no data was observed for 1 or 2 of the biological replicates analysis was done using the method outlined in Glantz and Slinker (1990) where ANOVA analysis is replaced with multiple regression analysis. Biological replicates showed strong reproducibility with good correlations observed for all comparisons (Figure 36) however one biological replicate from each the NF- (NF- 1) and PS- (PS- 3) treatments was omitted from statistical analysis. Biological replicate 'PS- 3' showed a consistently reduced protein abundance when compared with PS- 1 and PS- 2 as is shown by the deviation away from the 1:1 line in Figure 36. Biological replicate NF- 1 showed poor correlation to both NF- 2 and NF- 3 ($R^2 < 0.95$ in both comparisons) (Figure 36).

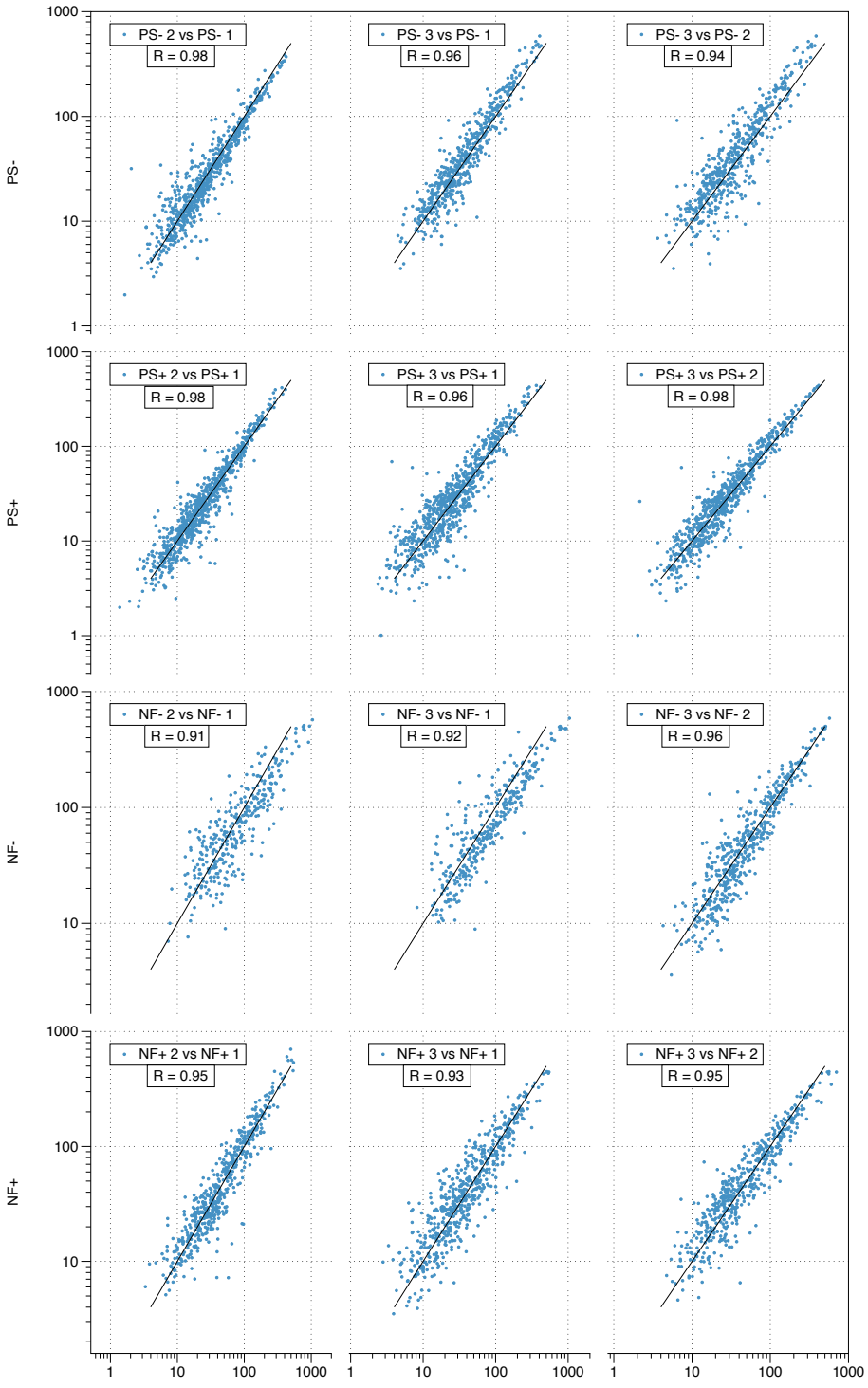

Figure 36: Log-log scatter plots showing reproducibility of biological replicates. Correlation coefficients are displayed on each plot individually. A black line is superimposed onto each plot to indicate a 1:1 relationship. Poor correlation and reduced protein abundance resulted in biological replicates NF- 1 and PS- 3 not being used for further statistical analysis.

A two-way analysis of variance was conducted employing a Bonferroni correction so as to minimise the problem of multiple comparisons. Whilst offering a simple means to control family-wise error rate the Bonferroni correction is considered conservative and hence the total suite of differentially abundant proteins may have been an underestimate. Statistical analysis was performed on both iron stress comparisons PS- vs PS+ and NF- vs NF+, along with diel variability comparisons PS- vs NF- and PS+ and NF+. Changes in protein concentrations were deemed significant if the multi-comparison adjusted P-value was <0.05 (Figure 37). Proteins observed to be statistically different between iron replete and deplete samples are listed in Table A5 in the appendix whilst proteins statistically different between time points are listed in Table A4.

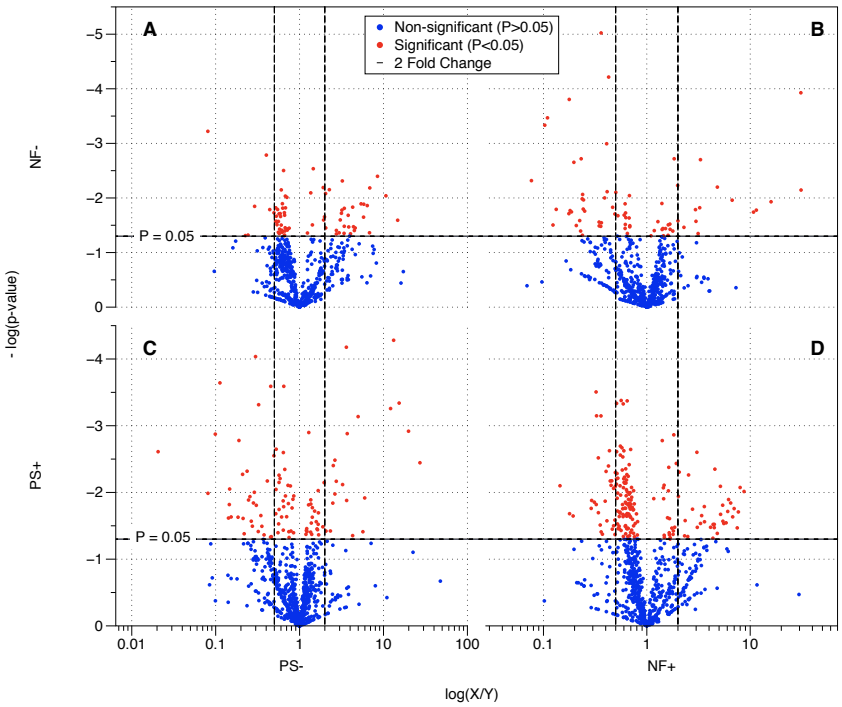

Figure 37: Volcano plots comparing mean protein concentrations (log(X/Y)) across our 4 different treatments against statistical significance (-log(p-value)). Proteins observed with a statistically significant (P<0.05) change in concentration are indicated with red points, non-significant changes as blue points. Solid black line shows P=0.05 whilst dashed lines indicates a 2-fold change. Plots show A) NF- vs PS-, B) NF- vs NF+, C) PS+ vs PS- and D) PS+ vs NF+.

5.5 RESULTS AND DISCUSSION

5.5.1 *Physiological and growth conditions*

After a 7 to 9 day incubation period, Fe+ *Trichodesmium* cultures showed a divergence in cell counts away from those observed in the Fe- treatments, suggesting Fe deficiency induced growth limitation in the Fe- cultures (Figure 38). Daily photosynthetic physiological measurements showed a significant divergence in F_v/F_m values after day 2 of the experiment where Fe+ and Fe- cultures showed F_v/F_m values ranging from ≈0.41-0.54 and ≈0.41-0.31 respectively. A difference in σ_{PSII} was evident from day 8 onwards with Fe+ cultures showing a small decrease in the cross-sectional area of PSII whilst the Fe- cultures demonstrated a consistent increase in σ_{PSII} values throughout the experimental period (Figure 38). Refed cultures were supplemented with 400 nM Fe on day 7 (when proteomic sampling was performed), whereupon a subsequent recovery of all parameters towards Fe+ culture values was observed (Figure 38). F_v/F_m data show this recovery to be quite rapid with the Refed culture diverging from Fe- cultures after just one day and being statistically indistinguishable from the Fe+ culture after 4 days (Figure 38).

The divergence of cell abundance, F_v/F_m, and σ_{PSII} are all physiological signs of stress. The observed decrease in photosynthetic competency (F_v/F_m) has previously been demonstrated as a signature of iron stress in *Trichodesmium* (Richier et al., 2012; Shi et al., 2007; Küpper et al., 2008), whilst previous studies differ regarding observed changes in σ_{PSII} in response to iron-stress. A similar σ-type photoacclimation (Six et al., 2008) was observed by Küpper et al. (2008) where increase in the cross sectional area of PSII under iron limitation were attributed to an elevated coupling of pigment complexes to PSII. This strategy is thought to increase the light absorbing efficiency of each photosynthetic unit, allow for a reduction in their number and subsequently a reduction in cellular iron requirement. The rebound in photosynthetic competency of refed cultures from day 7 onwards acts as confirmation of our intended iron stress conditions.

5.5.2 *Proteomic Overview*

Of the 4,342 protein encoding genes in the *Trichodesmium erythraeum* IMS101 genome a total of 1104 proteins were observed cumulatively across all samples. Of these, 573 were observed across all treatments (Figure 39 and Figure 40), 272 were present only in cultures sampled 2 hours after the onset of the photoperiod whilst 37 proteins were exclusively present in cultures sampled 6 hours after the onset of the photoperiod. Proteins exclusively present in iron stressed culture

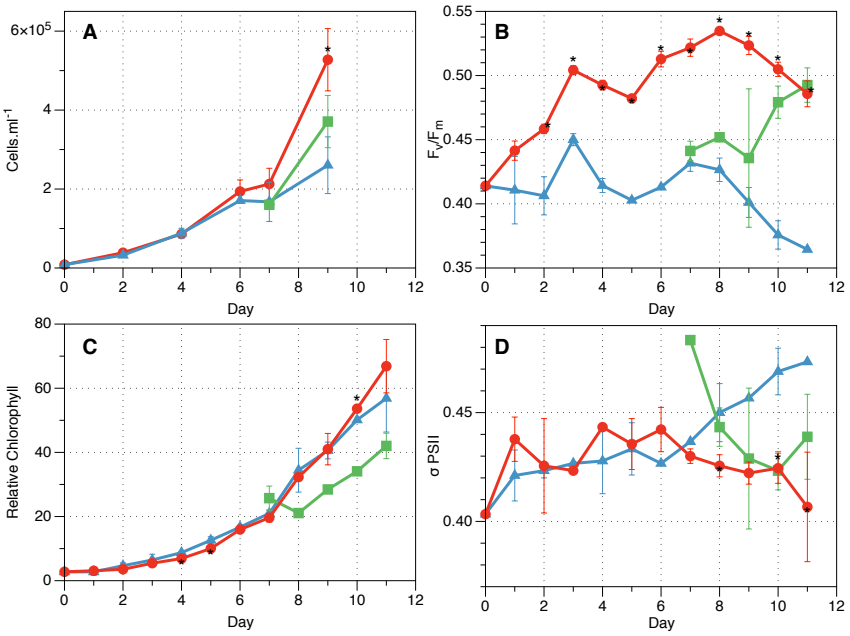

Figure 38: Physiological growth parameters observed throughout the duration of the culture experiment showing iron replete (red), iron deplete (blue) and iron refed (green). A) Cell counts (cells. ml^{-1}), B) photosynthetic efficiency (F_v/F_m), C) relative chlorophyll, D) Cross-sectional area of PSII (σPSII). Sampling was conducted on day 7, upon which assessment of refed cultures began. Error bars indicate standard deviation from triplicate measurements. Time points where data show a statistically significant difference (student's t-test, n=3, P<0.05) between Fe replete and Fe deplete cultures are indicated by an asterisk along the Fe replete trace.

totalled 47 and those exclusively observed in iron replete cultures totalled 202.

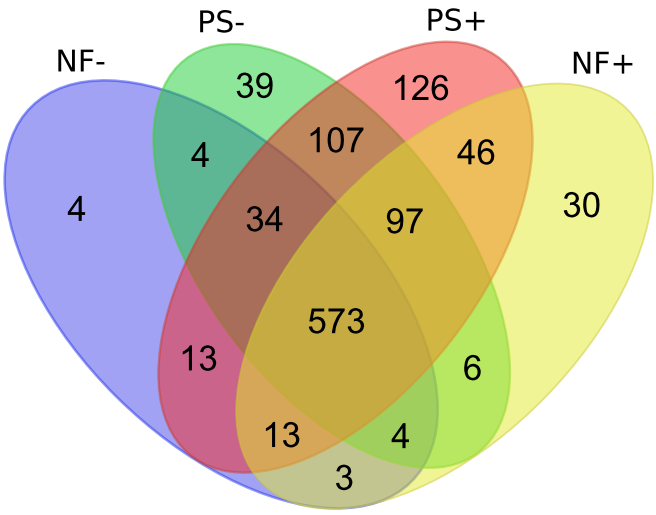

Figure 39: Venn diagram depicting all 1104 observed proteins and their distribution amongst our 4 sample treatments (PS+, PS-, NF+ and NF-).

5.5.3 *Multi-subunit Protein Complexes*

In order to facilitate data validation the literature values for stoichiometric ratios of known multi-protein complexes were compared to our observations. This approach is conceptually similar to that described Rogowska-Wrzesinska (2014) whereby a 'heteromer score' is calculated from the observed abundance of multi-subunit proteins and used as an internal data validation method.

Here I present data concerning the multi-subunit complexes PSII, Cytochrome b_6f, PSI, F_0F_1 type ATP-synthase, Nitrogenase and RuBisCO (Figure 41).

PSII - Cyanobacterial photosystem II comprises 20 subunits of which 12 were observed in this study. PSII contains 7 intrinsic protein, PsbA - D1, PsbB - CP47, PsbC - CP43, PsbD - D2, PsbE - α-subunit of cytochrome b_{559}, PsbF - β-subunit of cytochrome b_{559} and finally PsbI (Bricker et al., 2012) (Figure 42). Of these intrinsic proteins all but PsbF and PsbI are observed near to the 1:1 ratio described in the literature (Biesiadka et al., 2004). Of the extrinsic protein PsbO, PsbP, PsbU and PsbV only PsbO and PsbP are observed.

Of note are the two distinct isoforms of PsbA at near equimolar concentration and conserved PsbA1/2 : PSII stoichiometry (where

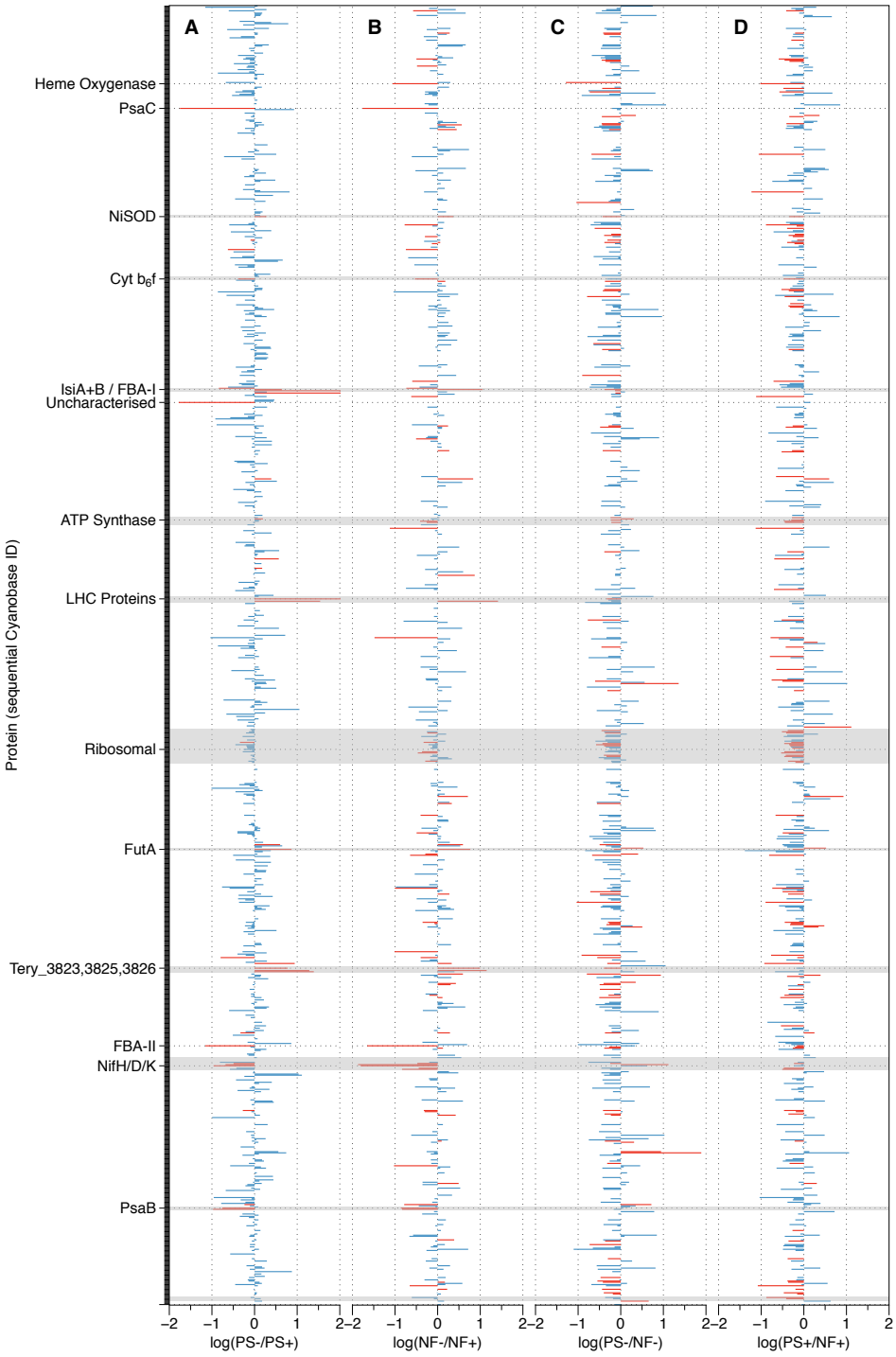

Figure 40: Overview depicting between treatment changes in protein abundance, plots A (PS-/PS+) and B (NF-/NF+) depict an iron induced response, plots C (PS-/NF-) and D (PS+/NF+) depict a time-of-day induced response. Significant and non-significant changes in protein abundance are coloured in red and blue respectively. The observed 1104 proteins are ordered sequentially using their KEGG/Cyanobase ID numbers.

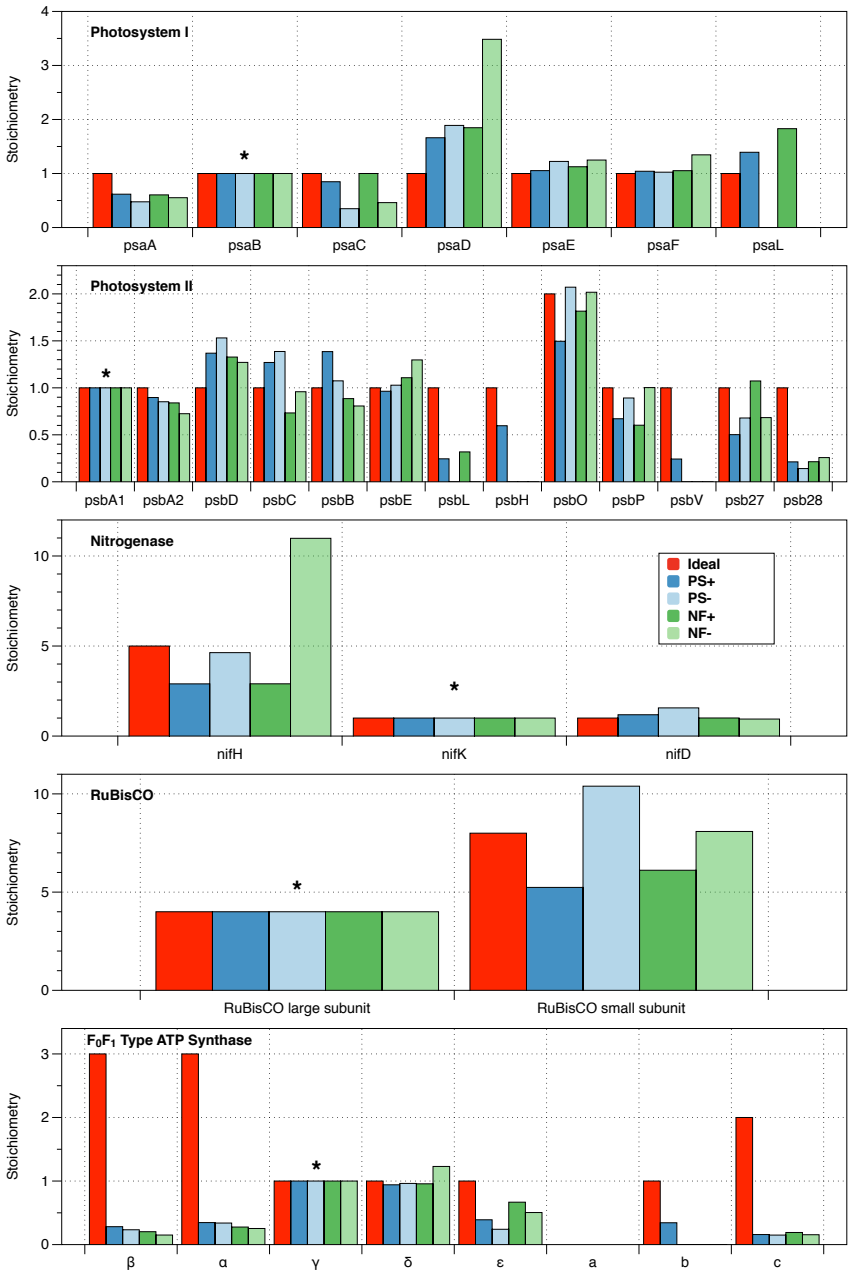

Figure 41: Relative stoichiometric ratios of select multi-protein complexes compared with literature values. Stoichiometries are calculated from the concentration of a given protein divided by the average concentration of the complete multi-protein complex. Literature derived 'ideal' stoichiometric ratios are shown in red. Bold label fonts denote membrane intrinsic proteins, non-bold font denotes membrane extrinsic protein and italic fonts denote non-membrane proteins.

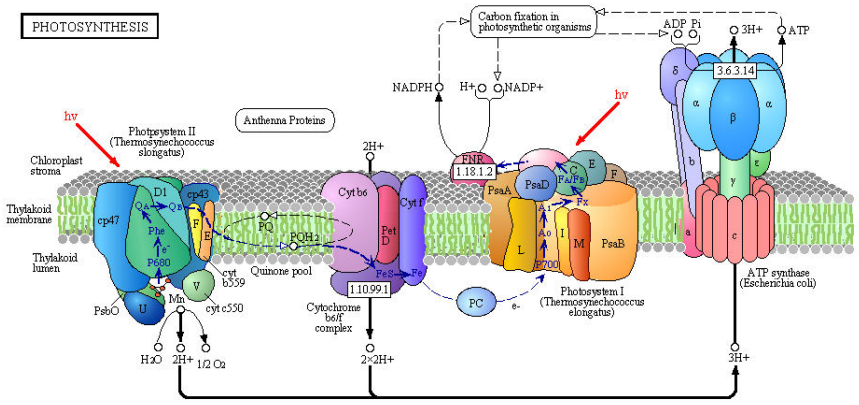

Figure 42: KEGG photosynthesis reference pathway depicting the structures and linkages of PSII, the Cytochrome b_6F complex, PSI and ATP synthase.

the concentration of PSII is the average concentration of all observed PSII proteins). The Mn containing oxygen evolving complex (OEC) of PSII is located on the lumenal surface of the PsbA/D1 protein and is stabilised by PsbO (Figure 42). Our data show PsbO to be present at a 2:1 PsbO:PSII stoichiometry. Although the stoichiometry of PsbO:PSII in crystal structure has been observed to be 1:1 there is currently debate as to the accuracy of these findings with experimental evidence suggesting a 2:1 stoichiometry in higher plants (Popelkova and Yocum, 2011; Xu and Bricker, 1992; Leuschner and Bricker, 1996). Given our observations of two PsbA isoforms, PsbA1 and PsbA2, along with a stoichiometric equivalent of 2 PsbO proteins to each complete PSII complex, it is possible that diazotrophic PSII includes two interchangeable PsbA-OEC-PsbO complexes. Both PsbA1 and 2 show no significant time-of-day variability suggesting both proteins are conserved across the diel cycle. Whilst one PsbA-OEC-PsbO complex will be fully functional and readily evolve oxygen, the second may be a non-oxygen evolving isoform utilised during the diazotrophic or within diazocyte cells period and used to maintain the structural integrity of the PSII super-complex and resist oxidative damage as suggested in Murray (2012). Murray (2012) determined that no such "rogue-D1" was present in the *Trichodesmium erythraeum* IMS101 genome, instead concluding that the structural differences between PsbA1 and 2 were not sufficient to attribute them with differing functions. Our observations of a ≈1:1 ratios for the two isoforms perhaps suggests an alternative strategy than that proposed by Murray (2012) but achieving a similar function.

CYTOCHROME B_6F - Cytochrome b_6f is a dimeric protein complex forming a major component of the photosynthetic electron trans-

port chain and aiding in the transfer of soluble electron carriers between PSII and PSI (Figure 42). Cytochrome b_6f contains 8 tightly bound polypeptide subunits (Baniulis et al., 2009) all present in a 1:1 stoichiometry. Of these 8 subunits Cytochrome f (PetA) and the Rieske iron-sulfur protein (PetC) are both membrane extrinsic. All membrane-intrinsic and extrinsic proteins are observed with the exception of PetF and PetG. Cytochrome f (1 Fe, PetA), cytochrome b6 (2 Fe, PetB) and the rieske protein (1 Fe, PetC) all show relative reductions in stoichiometry during iron-deplete conditions (Table A5 and Figure 41) however the absolute concentration of cytochrome b6 (PetB) shows unexplained elevation relative to all other observed subunits (Figure 43).

PSI - Cyanobacterial PSI contains 7 intrinsic proteins (PsaA, PsaB, PsaF, PsaI, PsaJ, PsaK and PsaL) and a further 3 extrinsic proteins (PsaC, PsaD and PsaE) (KEGG:ter00195) (Golbeck, 1992; Xia et al., 1998). The major subunits of PSI were all present at a near 1:1 ratio similar to that observed in the literature (Jordan et al., 2001) with the exception of PsaD. PsaD, alongside PsaC and PsaE, is involved in the formation of the ferredoxin (2 Fe) /flavodoxin (0 Fe) docking site. The ratios of PsaD:PSI is closer to 2 in samples PS+, PS- and NF+ and closer to 3.5 in sample NF-. The F_a and F_b Fe_4S_4 containing protein PsaC decreases in abundance relative to PSI in response to iron deficiency (Figure 41 and Table A5). The remaining Fe_4S_4 cluster, F_x, is liganded between PsaA and PsaB. PsaA displays a smaller iron stress decrease whilst PsaB shows no significant change in concentration (Table A5).

F_0F_1 TYPE ATP SYNTHASE - Our results did not show the expected stoichiometry for chloroplastic F_0F_1 type ATP-synthase of $\beta_3\alpha_3\gamma\delta abc_{10-14}$. I observed a distinct lack of membrane bound subunits a, b and c whilst non-membrane proteins α and β are observed to be deplete relative to the membrane-extrinsic proteins γ and δ.

NITROGENASE - The ratio of the heterodimeric proteins NifK:NifD of 2:2 is well conserved across all treatments. However the relative abundance of NifH to both NifD and NifK shows a distinct iron stress response whereby iron replete samples, PS+ and NF+, are present at $\approx$3:1 whilst iron stressed cultures show ratios of $\approx$4:1 to $\approx$8:1 for PS- and NF- respectively. This change is driven predominantly by a decrease in the abundance of iron-rich NifK and NifD relative to NifH. Given the theoretically optimal configuration of NifH:NifDK of 5:1 (Kustka, 2003a), our observations of an increased NifH:NifDK ratio under iron stress may be an iron compensatory response where iron stress forces a more iron-efficient nitrogenase configuration.

RUBISCO - As with the nitrogenase stoichiometry, the ratio of the RuBisCO subunits in iron replete conditions is near to the expected stoichiometry of 8 RbcS : 8 RbcL (Figure 41). Iron deplete samples however show a distinct departure from this idealised stoichiometry and instead are observed at ≈10:5 stoichiometry of RbcS:RbcL. Neither subunit contains any iron so the reason for this change in stoichiometry is unclear.

COMPLEX RATIOS Given the majority of subunits are observed at consistent expected ratios I can be confident in my protein recovery and analysis methods. Further to this, for each respective multi-subunit complex I can derive an average complex concentration. Normalising these concentrations to the concentration of PSII, I see the ratios of PSI:PSII and IsiA:PSII closely match those observed in culture by Richier et al. (2012) but are less similar to ratios observed in environmental *Trichodesmium*. As observed in Richier et al. (2012) both our PSI:PSII and IsiA:PSII ratios show an iron-stress responses. In addition I show iron-stress related changes in nitrogenase and Cytochrome b_6f (Figure 43). Of note are the dramatic changes in IsiA:PSI ratios where our two high iron treatments show an IsiA:PSI ratios of ≈0-0.15 whilst our low iron treatments show ≈4.0-4.6. The observations of a reduced PSI concentration alongside a concomitant increase in IsiA suggest an iron stress induced association of IsiA with PSI. This allows cells to decrease the absolute abundance of PSI whilst increasing each PSI complex's light harvesting ability (Yeremenko et al., 2004; Chauhan et al., 2011). Resulting in a decreased PSI-associated Fe-burden (Richier et al., 2011) whilst maintaining cellular function.

5.5.4 *Differentially Abundant Proteins*

Statistical analysis by two-way ANOVA with a Bonferroni correction identified a total of 210 differentially abundant proteins across the 4 treatments (PS+, PS-, NF+, NF-) (Table 17 and Figure 44). The most significant driver of change in the proteome was observed between sampling time points, however samples collected during the photosynthetic period had more unique proteins observed (Figure 39). Between PS+ and NF+ a total of 137 proteins showed statistically different abundances whilst 109 different proteins were identified between PS- and NF- (Table 17) Comparisons between the iron replete and deplete proteomes identified fewer differentially abundant proteins with PS- and PS+, and NF- and NF+ having 50 and 111 differentially abundant proteins respectively. Of the total 210 differentially abundant proteins 86 were identified to only have a temporal response and 38 appear to be affected exclusively by iron availability (Figure 44).

Hierarchical clustering of all differentially abundant proteins reiterates the observations that the greatest change in the proteome is be-

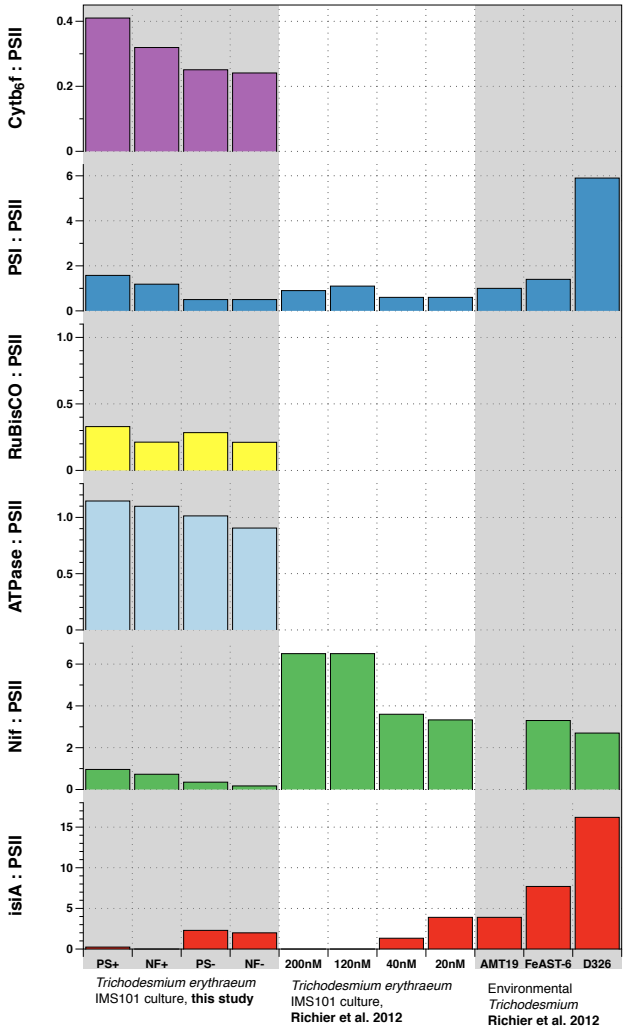

Figure 43: Stoichiometric ratios of select multi-subunit protein complexes relative to the mean concentration of PSII. Data are shown for *Trichodesmium erythraeum* IMS101 for the 4 treatments from this study in addition to culture data and environmental data from Richier et al. (2012).

Treatment Comparison	Differentially observed proteins
PS- vs. PS+	50
NF- vs. NF+	111
PS+ vs. NF+	137
PS- vs. NF-	109
Cumulative	210

Table 17: Summary of differentially abundant proteins as determined by two-way ANOVA followed by a Bonferroni multiple-comparison correction.

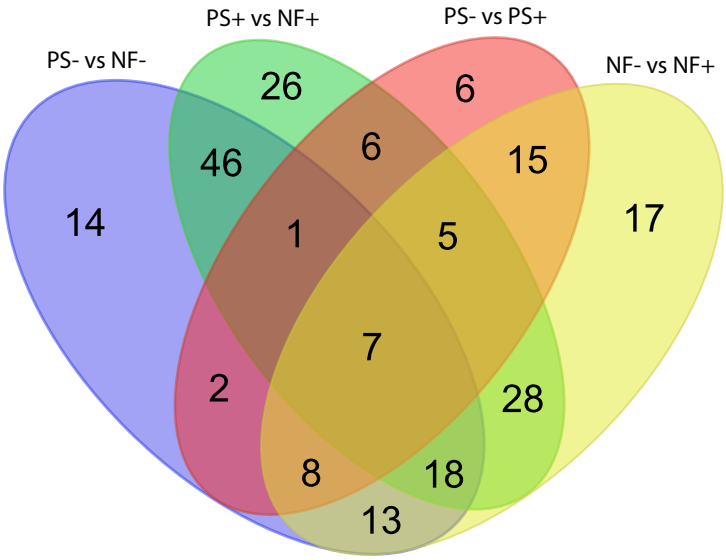

Figure 44: Venn diagram showing the distribution of the cumulative proteins observed at significantly different abundances between the four treatments (PS-, PS+, NF- and NF+).

tween the putatively photosynthesising (PS-, PS+) and nitrogen fixing (NF-, NF+) time points (Figure 45) as opposed to between iron replete (PS+, NF+) or deplete (PS-, NF-) treatments. The observed proteomic response is clustered into three broad groups, proteins more abundant in iron deplete conditions (PS-, NF-), proteins more abundant in iron replete conditions (PS+, NF+), and proteins more abundant during the period of maximal nitrogen fixation (NF-, NF+) (Figure 45).

5.5.4.1 *Iron stress response*

ACQUISITION - An acquisitional response to iron sufficiency is observed in the increased concentrations of the iron storage protein, Ferritin (Tery_4282). Ferritin was observed at 99.0 fmol.ug^{-1} and 124.9 fmol.ug^{-1} in samples PS+ and NF+ respectively. Conversely, ferritin was not observed in either iron deplete samples. This increase in cellular iron storage proteins suggests a luxury uptake response to elevated iron concentrations. Heme oxygenase (Tery_0335), a protein responsible for catalysing the degradation of heme and the subsequent release of iron, is observed at elevated concentrations between samples NF-/NF+ and a presence/absence change observed between PS-/PS+. These observations may signify a degree of intracellular iron re-purposing or highlight an unexplored iron uptake mechanism for marine cyanobacteria whereby hemeoproteins could be directly acquired from the environment.

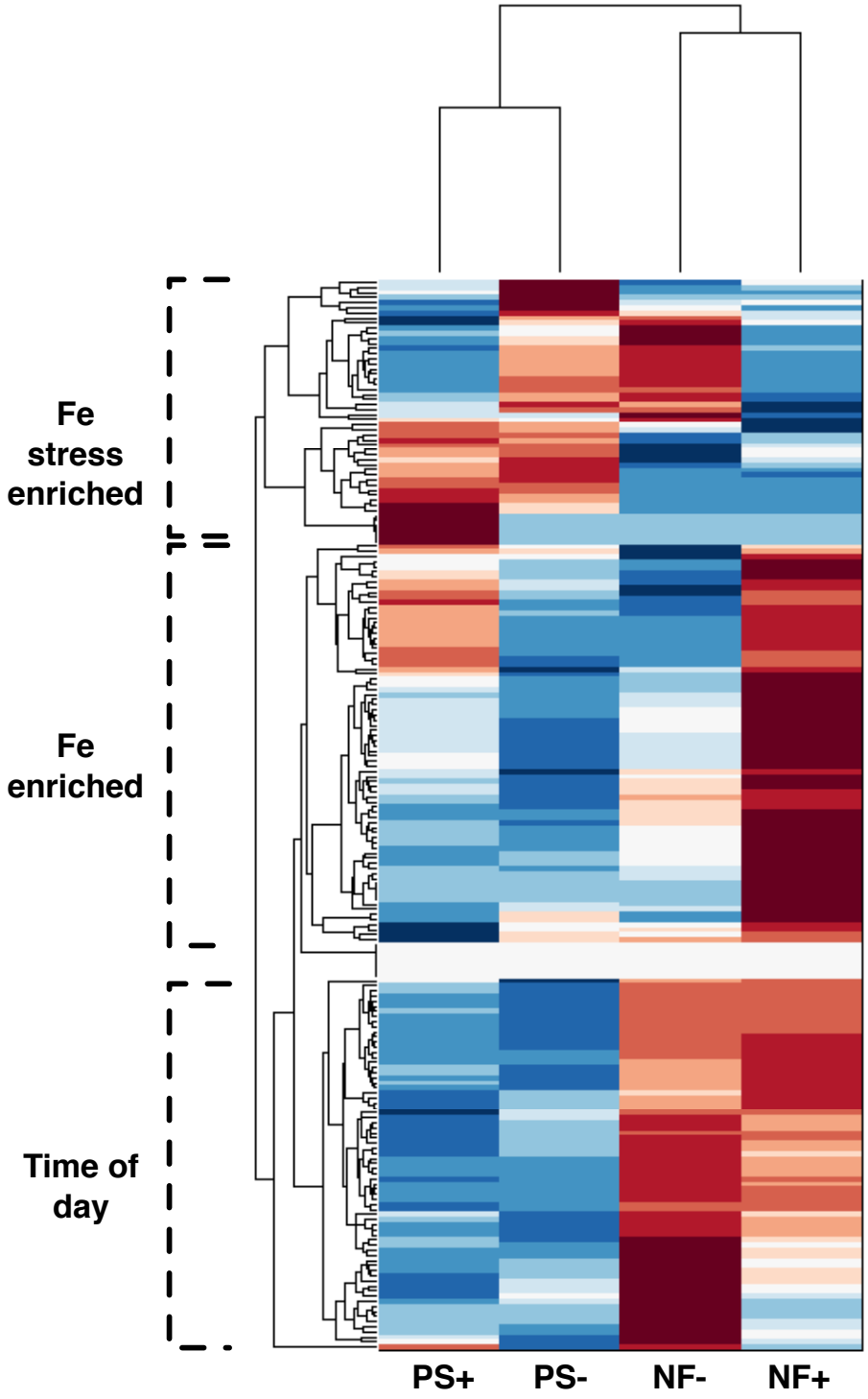

Figure 45: Hierarchical clustering of all 210 differentially abundant proteins. Between treatment clustering demonstrates a stronger influence from time-of-day changes than from iron replete/deplete treatments. Within treatment clustering highlights three distinct groups of proteins. Red and blue colours indicate and increase or decrease in relative protein abundance respectively whilst white indicates an absence of data for a given protein in a given treatment.

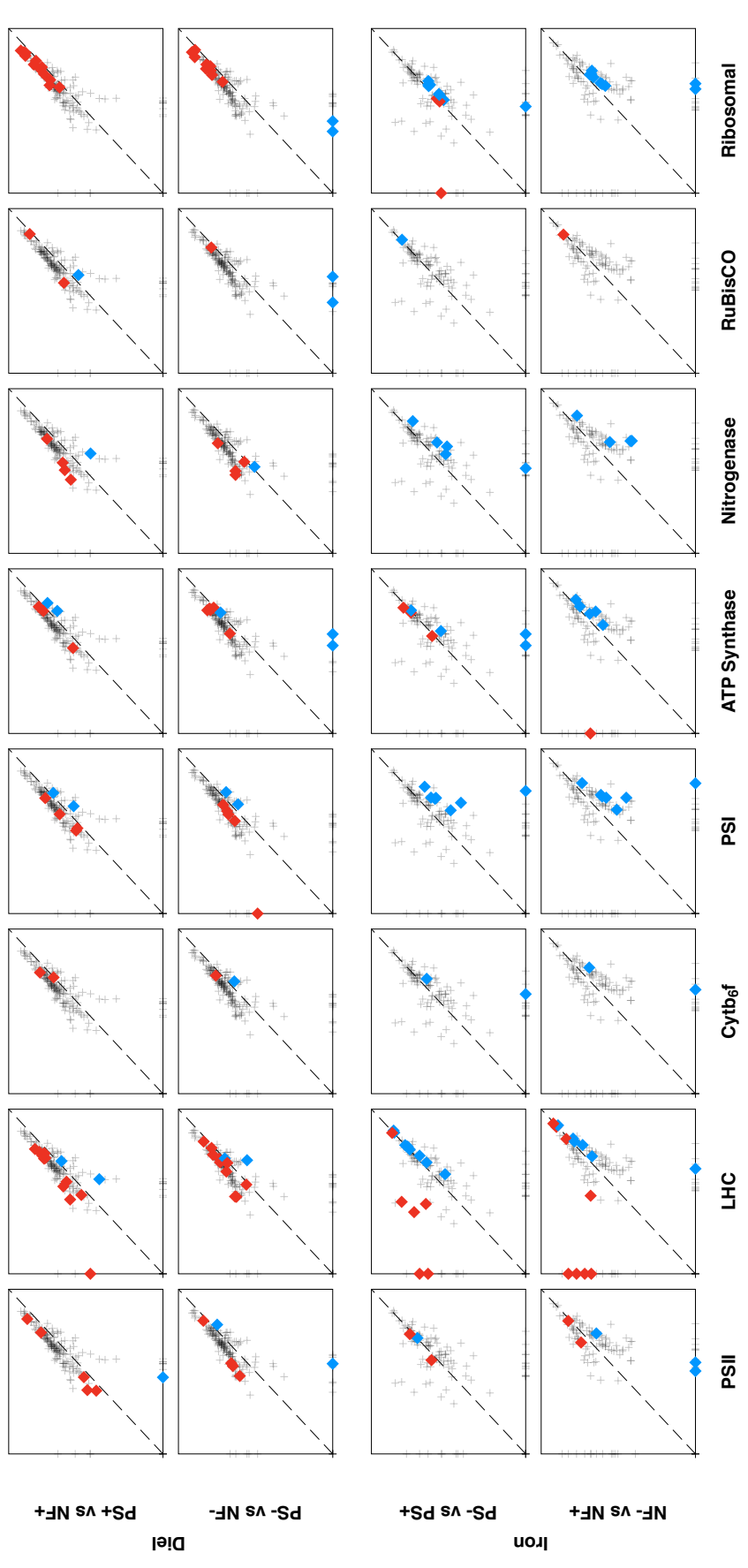

Figure 46: Scatter plots of proteins showing significantly different concentrations between our four comparisons. First and second rows show diel variability (PS- vs NF- and PS+ vs NF+) whilst third and fourth rows show iron-stress variability (PS- vs PS+ and NF- vs NF+). All significantly differing proteins are shown as grey crosses whilst protein groups are highlighted on particular plots, here red diamonds indicate a statistically significant positive fold change whilst blue indicates a negative fold change.

A number of transporter proteins were observed with higher abundances in the iron deplete conditions, these include an Fe^{3+} inner membrane transport protein FutA (Tery_3377) showing a 2.38 and 2.15 fold change in PS-/PS+ and NF-/NF+ respectively. This elevation in Fe^{3+} uptake protein suggests samples PS- and NF- were severely iron deplete and were responding to this iron deficiency by increasing their efforts to acquire Fe. Interestingly it suggests the FutA protein is more abundant during Fe stress conditions rather than being more abundant when Fe is plentiful.

A similar increase in abundance of a Phosphonate ABC transporter protein PhnD (Tery_4993) and an L-amino acid transporter protein Aapj (Tery_1998) are observed in iron deplete samples. The reason for these increases are unclear. With regards to the phosphonate uptake protein, all cultures were cultivated in an excess of DIP so as to prevent any P stress induced proteomic changes. Although unlikely to have occurred, had the cultures been grown to a point of Fe and P co-limitation it would likely be the iron replete cultures, where growth rates were higher, would have reached this point first. As such, if anything I would expect the relative abundance of this phosphonate uptake protein to be the inverse of our observations.

The majority of proteins observed at elevated concentrations during iron deplete conditions were uncharacterised (Table A5). Of these proteins, Tery_3823, Tery_3825 and Tery_3826 are of particular interest given their sequential position in the *Trichodesmium erythraeum* IMS101 genome and large increase in abundance during iron deplete conditions (Table A5 and Figure 40).

Tery_3823 is a 162 aa long protein containing within it the lipocalin 5 conserved domain. Lipocalin's are a fairly diverse group of proteins involved in the transport of hydrophobic molecules such as steroids, bilins, retinoids and lipids (Flower et al., 2000). Siderocalin's are a sub-group of the lipocalin family specifically involved in siderophore scavenging (Goetz et al., 2002). Goetz et al. (2002) observed the scavenging of the siderophore enterobactin by siderocalin proteins. Enterobactin is the strongest known Fe^{3+} siderophore requiring the enzyme ferrienterobactin esterase to remove the iron. No homolog of ferrienterobactin esterase is present in the *Trichodesmium erythraeum* IMS101 genome. Tery_3824 contains a TrkA conserved domain predicted to be involved in inorganic ion transport, Tery_3825 contains a lipase/esterase conserved domain whilst Tery_3826 has a conserved flavin-binding monooxygenase known to be involved in siderophore production (Frederick et al., 2011).

Although *Trichodesmium* are not thought to produce siderophores they are known to acquire iron from externally derived siderophores (Roe et al., 2012; Achilles and Church, 2003) and culture isolates of *Trichodesmium erythraeum* IMS101 are generally not axenic. Given that the mechanism of siderophore uptake in *Trichodesmium* is currently

unknown, I suggest this group of strongly iron-regulated proteins may be involved in *Trichoesmium* siderophore acquisition and further investigation into their function is warranted.

COMPENSATION - A number of the major iron containing proteins involved in photosynthesis (Table 2) demonstrated a substantial iron stress induced change in abundance (Table A5). These include the Fe_4S_4 cluster containing proteins of PSI, PsaA (4 Fe, Tery_4669) and PsaC (8 Fe, Tery_0454)) both of which were less abundant in PS- and NF- samples (Figure 46 and Table A5). Cytochrome b_6 (3 Fe, Tery_1137) and apocytochrome f (1 Fe, Tery_1798) of the Cytochrome b_6f complex were both observed at lower abundances in iron deplete conditions, however the rieske proteins (1 Fe, Tery_1799) did not show a differential iron response. Protein Tery_1765 shows a significant reduction in concentration during iron deplete conditions, this protein shows strong homology to the Cytochrome b_6f associated protein PetP which currently does not have a known function. Within photosytem II PsbD (1 Fe, Tery_1230), PsbV (1 Fe, Tery_2687) and PsbE (1 Fe, Tery_3504) did not show any change in abundance despite having iron binding co-factors. Cytochrome c_6 was present in samples PS+ and NF+ but whilst its Fe free, Cu containing, functional replacement plastocyanin was observed in all samples however showed no significant iron stress response. The well studied iron stress induced protein, IsiA (Tery_1667), showed a $\approx$13 fold increase between samples PS+ and PS-. IsiA was observed in sample NF- at similar concentrations to PS- however was not observed in sample NF+ so no fold change was calculable.

The reduction of the iron-rich PSI complex, independent of a change in PSII concentration, is in agreement with the findings of Richier et al. (2012) where they observed a shift in PSI:PSII stoichiometry under iron limitation (Figure 47). A reduction in the number of complete PSI complexes is supported by our observed increase in IsiA concentration (Figure 47). Here IsiA is thought to support the formation of a PSI tetramer, increasing the efficiency of each PSI unit by $\approx$3 fold (Yeremenko et al., 2004; Chauhan et al., 2011; Bibby et al., 2001)(Figure 47). This observed iron compensatory strategy aims to maintain cellular function but at a reduced cellular iron requirement. The observed reduction in Cytochrome b_6f complex alongside this PSI reduction is expected given Cytochrome b_6f's role in electron transport between the two photosystems and it's high Fe requirement.

The nitrogenase metalloproteins, NifH, NifD and NifK, were all more abundant in both iron replete treatment with the more iron rich proteins, NifD and NifK showing larger decrease in iron deplete treatments when compared with NifH (Table A5)(Figure 47). Alongside these functional nitrogenase proteins were significant decreases

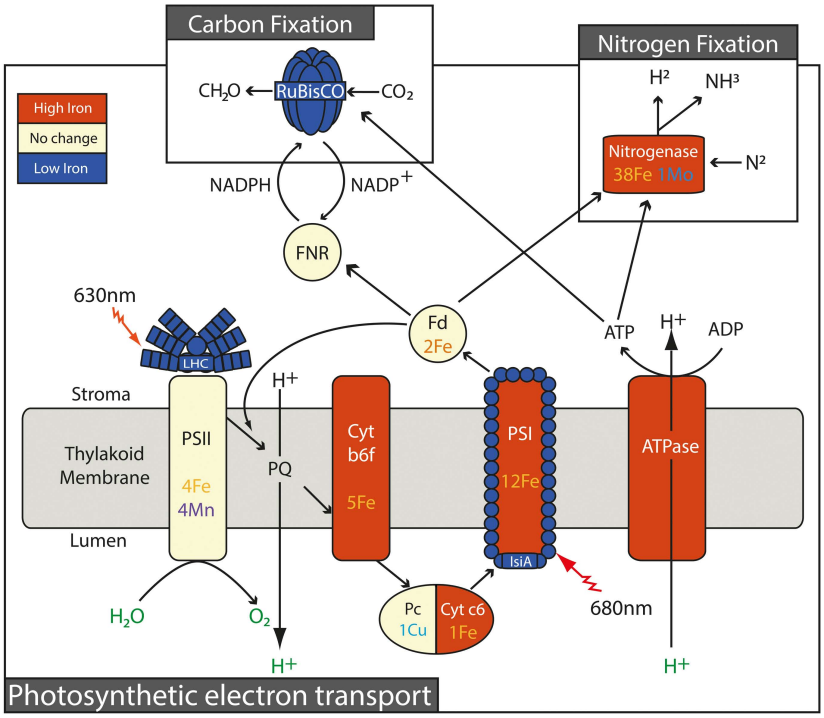

Figure 47: Simplified schematic model illustrating iron-stress induced changes to the flow of electrons through the principle components in photosynthesis, nitrogen fixation, respiration and carbon fixation in cyanobacteria. Complexes coloured in red are observed to be more abundant under iron-replete conditions, complexes coloured blue are more abundant under iron-deplete conditions and complexes coloured yellow show no iron-induced change in abundance. The metal composition of each complex is also shown. Abbreviations include - PSII - photosystem II, LHC - Phycobilisome light harvesting complex, Cytb_6f - cytochrome b_6f complex, PSI - photosystem I, ATPase - ATP synthase, Pc - plastocyanin, Cyt c6 - cytochrome c6, PQ - plastoquinone, Fd - ferredoxin, FNR - ferredoxin:NADP$^+$ reductase, IsiA - iron stress induced protein, ATP - adenosine triphosphate, ADP - adenosine diphosphate, NADP$^+$/H - Nicotinamide adenine dinucleotide phosphate, RuBisCO - Ribulose-1,5-bisphosphate carboxylase/oxygenase.

in the abundance of NifW and an absence of NifX in iron deplete conditions. A reduction in the iron-rich nitrogenase proteins NifD, NifK and NifH is a well documented iron retrenchment response observed at either the transcriptional level (Shi et al., 2007), the protein level (Richier et al., 2012) or in observed activity (Berman-Frank et al., 2007). The function of the nitrogenase stabilising protein NifW currently remains unclear. Whilst I observe a concomitant decrease in abundance of both the functional nitrogenase proteins (NifH, D and K) alongside NifW the ratio of NifW to either NifD or NifK varies with iron treatment. During the active nitrogen fixation period, the

iron deplete (NF-) shows a ratio of NifW:NifD/K of ≈2 whilst the iron replete sample (NF+) has a ratio of ≈1. This suggests a previously unexplored iron compensatory strategy whereby for NifW could potentially act to maintain nitrogenase functionality under iron deplete conditions.

The increased abundance of nitrogenase proteins observed in iron replete samples was mirrored by a similar increase in the abundance of two thioredoxin proteins, Tery_3311 with strong homology to *Synechocystis* TrxA (Uniprot ID - P52231) and Tery_0945 showing strong homology to *Nostoc* sp. PCC 7120 TrxB (Uniprot ID - P20857). Thioredoxin's regulate key cellular functions such as photosynthesis, are thought to aid in combating oxidative stress (Yoshida et al., 2005; Meyer et al., 2008) but also may potentially play a role in cell-to-cell communication (Meng et al., 2010).

Nickel superoxide dismutase (Tery_0891) along with a protein believed to be a peroxiredoxin (Tery_0162, strong homology to K9QPS9 peroxiredoxin of *Nostoc* sp. ATCC 29411), both of which are thought to be involved in anti-oxidant defence, show an inverse relationship with the abundance of nitrogenase proteins. An inverse relationship between nitrogenase abundance and the anti-oxidant defence proteins required for protecting this oxygen sensitive enzyme complex hints at either an iron induced increase in oxidative stress similar to that observed in *Anabaena* sp. (Latifi et al., 2005) or an increased cellular effort to protect the nitrogenase proteins from oxidative stress.

RETRENCHMENT - In both iron deplete samples (PS- and NF-) a significant number of ribosomal, proteolysis and protein folding proteins all show lower concentrations when compared to iron replete samples (PS+ and NF+) (Table A5). These reduced protein abundances match the decrease in growth-rates observed in the cell count data (Figure 38) and suggest an iron deficiency induced limitation of growth. Similarly a systematic decrease in the abundance of all ATP synthase subunit proteins was observed between NF- and NF+ however this was not mirrored between PS- and PS+ (Table A5). The general decrease in cellular functions inferred from decreases in components of photosynthesis, nitrogenase, ribosomal proteins and ATP synthase all suggest a general iron retrenchment response whereby *Trichodesmium* have exhausted their acquisitional and compensatory adaptations to iron deficiency and so can be considered to have been iron *limited* at the point of sampling. As a result of this iron limitation our low iron cultures had a markedly lower growth rate (0.36±0.03 d^{-1}) compared with that of the high iron cultures (0.43±0.02 d^{-1}) (Student's t-test, n=3, $P<0.05$).

5.5.4.2 *Diel cycle response*

Between the earlier (PS- and PS+) and later (NF- and NF+) time points I observe a systematic increase in the majority of complexes involved in photosynthesis. The oxygen evolving activities of PSII are presumed to be elevated given the increased concentrations of 10 out of 13 observed PSII proteins (Table A4). Cytochrome b_6 of the Cytochrome b_6f complex was more abundant in samples NF- and NF+ relative to PS- and PS+ respectively. Of the observed PSI proteins, only PsaD (Tery_3791) and PsaL (Tery_1204) showed a significant increase in abundance during the nitrogen fixation period with the remaining PSI proteins showing no significant change in abundance. The inferred increase in PSII activity, potentially leading to increased oxygen evolution, is in contrast to previous studies whereby nitrogen fixation is only thought possible if PSII activity results in no net production of oxygen (Berman-Frank et al., 2001b; Küpper and Ferimazova, 2004). However, I am perhaps incorrect to presume an increased abundance of PSII proteins signifies increased PSII activity. Nitrogen fixation is an extremely energetically costly cellular process requiring 16 ATP molecules per mole of fixed N_2 (Equation 1), this energetic cost must somehow be met either by increased ATP synthase activity driven by PSII proton generation, energy release from glycolysis or from cyclic PSI flow. Küpper and Ferimazova (2004) observed a uniform distribution of PSII fluorescence across diazocyte and non-diazocyte cells noting a distinction between active PSII and inactive PSII. These putative diazocyte cells regulate their PSII activity by rapidly coupling or uncoupling their phycobiliproteins during the diazotrophic period. Additionally the 'cellular state' characteristic of the nitrogen fixation period was observed to have a particularly large PSII-associated antenna. Our observations, showing elevated concentrations of phycobilisome proteins, phycocyanin proteins and the phycobilisome linker polypeptide (Table A4), support this idea. Given the cellular homogeneity of our analysed protein samples there is no way to differentiate between protein from diazocyte and non-diazocyte regions of the trichome. Complementing the hypothesised uncoupling of phycobiliproteins in diazocyte cells, where PSII proteins are thought to be conserved but inactive, perhaps our observed increase in PSII abundance is from non-diazocyte cells only. Here I suggest an upregulation of PSII activity in non-diazocyte cell regions helps compensate for the inactivation of diazocyte PSII with this inactivation potentially supported by the previously discussed alternate PsbA-OEC-PsbO complex. The nitrogen fixation localised in the diazocyte cells could then be fuelled by energy generated in non-diazocyte cell regions via intercellular energy transfer. Given that the sample preparation method homogenises protein from both diazocyte and non-diazocyte cells, elucidating localisation of these PSII protein increases is unfortunately beyond the scope of this study.

An increase in ATP synthase subunits proteins is observed during the later time period with 6 of the 8 ATP-synthase subunit proteins showing elevated concentrations in samples NF- and NF+ . This potential increase in ATP synthesis is thought to driven by the observed abundance of PSII and meet the high energy requirements of nitrogen fixation. Together these observations suggest PSII activity is responsible for supplying nitrogenase with its required energy although it is unclear whether this is from uniform upregulation of PSII or preferential upregulation in non-diazocyte cells.

A moderate increase in the abundance of the NifH (Tery_4136) protein is observed between the earlier (PS) and later (NF) time periods in both iron replete and iron deplete conditions. The α-subunit of dinitrogen reductase, NifD, on the other hand shows a 3 fold decrease between PS- and NF- suggesting a possible diel cycling of this protein during iron deficient growth (Table A4 and Figure 46). The β-subunit of this heterodimeric enzyme system, NifK, shows a similar decrease in absolute abundance however the change is not statistically significant. Under iron replete conditions no such diel variability is observed in either NifD or NifK abundance. This iron-associated change in NifD and NifK abundance is in direct contrast to that seen by Saito et al. (2011) in *Crocosphaera watsonii*. Here all three functional nitrogenase proteins were observed at elevated concentrations during the diazotrophic period. The diazotrophic strategy employed by *Trichodesmium* is markedly different than that of *C. watsonii*. *C. watsonii* alternate between these two antagonistic processes over a 24 hour cycle whilst *Trichodesmium* fine-tune the two processes over a much shorter time period (Figure 5). This reduced time period could limit *Trichodesmium*'s protein recycling ability as suggested by nitrogenase protein abundances observed under iron replete conditions where a small increase is observed for NifH and NifD, and NifK shows no significant change. The sharp decrease in NifD and NifK abundance under iron deplete conditions may be a response to the seemingly severe iron stress experienced by the cultures (see Section 5.5.4.1). Under such iron stress conditions *Trichodesmium* are known to preferentially sacrifice their diazotrophic ability in order to preserve their photosynthetic function (Küpper and Ferimazova, 2004; Berman-Frank et al., 2007; Richier et al., 2012; Shi et al., 2007).

The psuedocyclic electron transport around PSII described above, whilst not resulting in the net evolution of oxygen, does liberate superoxide molecules and as such will require an increased abundance of anti-oxidant defence proteins (Kustka, 2003a; Kana, 1993). NiSOD, a putative peroxiredoxin (Tery_0162) and the two thioredoxins TrxA (Tery_3311) and TrxB (Tery_0945) are all observed at elevated concentrations during the later time period (NF) suggesting an increased effort to mitigate oxidative stress.

Numerous proteins involved in cellular respiration are observed at decreased concentrations during the later time period (NF- or NF+). These include proteins involved in glycolysis (Tery_3376, Tery_5066, Tery_2793, Tery_4099, Tery_1687, Tery_1380) and the reductive pentose phosphate pathway (Tery_3654, Tery_4410, Tery_4099, Tery_3376) both of which are involved in energy production/release.

The presumed increase in intracellular fixed N abundance, associated with the presumed active period of nitrogen fixation, would require a degree of cellular N allocation. *Trichodesmium* are thought to store fixed N in cyanophycin granules (Finzi-Hart et al., 2009), the biosynthesis of which is catalysed by cyanophycin synthetase (Ziegler et al., 1998). I observe cyanophycin synthetase (Tery_1965) at low concentrations ($\approx$5-7 fmol.ug^{-1}) and to be marginally, but non-significantly, more abundant during the earlier time period (PS) suggesting a limited role in fixed N sequestration. An alternative storage strategy for intracellular N could be direct assimilation into protein complexes. It is thought that $\approx$85% of a microbial organisms organic N is protein associated (Laws, 1991; Geider and Roche, 2002). The potential sequestration of fixed N into complete proteins could explain some of the observed increases in individual protein concentrations described above. During the later period, where the rate of nitrogen fixation is thought to be highest, I see an increase in the abundance of ribosomal proteins (both small and large subunits) involved in protein biosynthesis and chaperonin proteins involved in protein folding (Table A4 and Figure 46).

Whilst it must be noted that protein abundance is not a measure of enzyme activity, increases in protein abundance such as those detailed above must serve some biological purpose. As such, I propose a revised understanding of the nitrogen fixation strategy conducted by *Trichodesmium* whereby concentrations of both nitrogenase proteins but also photosystem II proteins increase over the diazotrophic period as a means of energy supply and N sequestration (Figure 48). Our conceptual understanding is complimentary to the diel cycling observed by Berman-Frank et al. (2001b) and Küpper and Ferimazova (2004) where N_{Fix} and C_{Fix} occur at different periods of the photoperiod and is achieved by the inactivation of PSII via phycobilisome decoupling however I am unable to localise either process to diazocyte or non-diazocyte cells.

5.6 CONCLUSION

By conducting whole cell proteomic analysis of a defined iron-stress culture experiment we've been able to detail an elaborate *Trichodesmium erythraeum* IMS101 iron-stress response. Ultimately aimed at maintaining cellular function these proteomic changes fall into

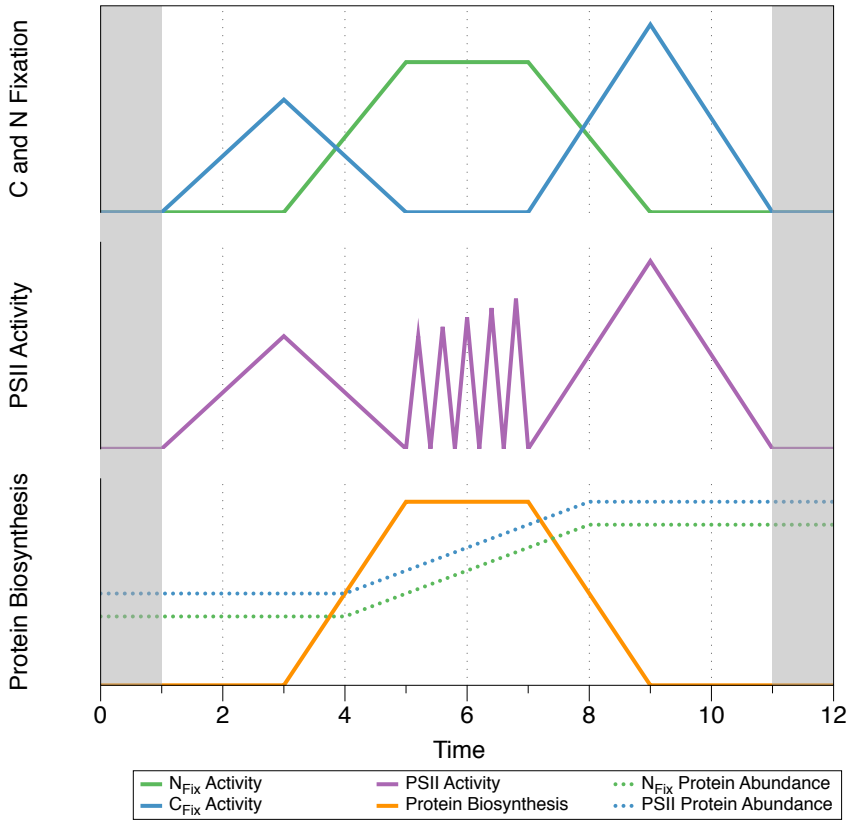

Figure 48: Conceptual model of a revised nitrogen fixation / photosynthesis temporal strategy. C_{Fix} (blue) and N_{Fix} (green) alternate as described by Berman-Frank et al. (2001b), during the periods of C_{Fix} the PSII have fully coupled phycobilisome proteins complexes (purple), during periods of nitrogen fixation the phycobilisome coupling alternates as described in Küpper and Ferimazova (2004) resulting in either active or inactive PSII. The nitrogen fixation period is characterised by a period of increased protein biosynthesis (orange) acting as a means of intracellular N sequestration and resulting in an increased abundance of photosynthetic proteins (blue, dashed) alongside an increase between early and late C_{Fix} rates (blue).

three distinct mechanisms, acquisition, compensation and retrenchment (Straus, 2004). Attempts to increase iron acquisition in iron-deplete conditions are mirrored by luxury uptake and elevated intracellular iron storage during iron replete conditions. The photosynthetic apparatus is observed to undergo a substantial proteomic shift potentially maintaining photosynthetic activity in the face of nutrient scarcity. Meanwhile currently under-explored strategies for potentially maintaining diazotrophic activity under iron-scarcity are identified in an increased anti-oxidant defence and the elevated abundance of NifW. Described as a nitrogenase stabilising protein, the

exact function of NifW remains unknown however the observed anti-oxidant defence mechanisms are similar to those observed between *Trichodesmium* growing diazotrophically or on fixed nitrogen (Sandh et al., 2011). Despite these numerous attempts to maintain cellular function our physiological data suggest severe iron limitation and a number of proteomic retrenchment strategies are observed to support this conclusion.

The counter-intuitive observations of an increased abundance of the oxygen-evolving PSII complex during the diazotrophic period, along with the systematic increase in ribosomal protein abundances has prompted a re-think of *Trichodesmium* diazotrophic strategy. I offer a small revision and substantial extension to the diazotrophic strategy proposed by Berman-Frank et al. (2001b). Our observations support an inactivation of PSII suggested by Küpper and Ferimazova (2004) alongside potential localisation along a trichome rather than the decrease in protein abundance observed in *Crocosphaera watsonii* by Saito et al. (2011). Furthermore our data suggests a substantial increase in protein biosynthesis during this period, given the assumed increase in nitrogen fixation during this time period alongside no change in cyanophycin synthetase abundance I theorise this to be a form of intracellular storage for newly fixed nitrogen.

6

SYNTHESIS

6.1 INTRODUCTION

Throughout this study I have investigated the interactive ecophysiology of *Trichodesmium* both in the environment and in culture. Chapter 3 focussed on *Trichodesmium*'s biogeographical environmental adaptability where I defined distinct niches within the Atlantic ocean meanwhile Chapter 4 and Chapter 5 described the elemental and proteomic flexibility by which *Trichodesmium* are able to adapt to this niche.

This final chapter, whilst presenting minimal new experimental or observational data, will instead attempt to draw together the previous three chapters. In doing so I expand our focus beyond P and Fe limitation. Presenting nutrient uptake rates for N, P and Fe for the first time, I highlight an observed discrepancy in *Trichodesmium*'s growth characteristics observed in both our data and that found in the literature whereby P derived growth rates appear to be vastly over estimated when compared with corresponding Fe or N derived rates. Additionally, utilizing a bioinformatical approach I aim to characterise *Trichodesmium*'s metallo-proteins, generate a predicted protein-derived *Trichodesmium* metallome, and compare this to our observed intracellular metallome. In doing so I show strong similarity between *Trichodesmium*'s metallo-proteomic and metallomic composition and go on to further explore *Trichodesmium*'s dependence on Fe but also on numerous other potentially limiting trace metals.

Finally I synthesise my findings into the description of discreet *Trichodesmium* phenotypes observed in the (sub)-tropical Atlantic, summarising the key findings of this thesis and addressing areas of potential future study.

6.2 THE PROBLEM WITH P UPTAKE

Compared to other phytoplankton, *Trichodesmium* are widely thought to be slow growing with natural populations thought to have a growth rate of $\approx$0.1 d^{-1} (LaRoche and Breitbarth, 2005) whilst culture populations have been observed to grow at $\approx$0.3-0.4 d^{-1} (Hutchins et al., 2007) (Chapter 5). In steady state, where the growth of an organism occurs at a constant rate, the intracellular quota of any given nutrient must be balanced by uptake of that same nutrient. Throughout D361 uptake rates were assessed for N, P and Fe using $^{15}N_2$, ^{33}P and ^{55}Fe isotopes respectively (see Chapter 2 for further methods)

whilst only N_{Fix} rates were assessed on AMT21. From these nutrient uptake rates and their corresponding intracellular concentrations presented in Chapter 4 I was able to derive nutrient specific growth rates for these environmental populations (Figure 49 and Figure 50) N_{Fix} derived growth rates for AMT21 (0.016-0.064 day^{-1}) and D361 (0.0003-0.056 day^{-1}) were lower than the growth rates observed in culture (0.3-0.4 day^{-1}) (Figure 49 and Figure 50). Iron specific growth rates observed during D361 (0.0003-0.032 day^{-1}) were again lower than those observed in cultures however they were similar to the N derived growth rates of both AMT21 and D361. Growth rates derived from the ^{33}P uptake rate and intracellular P concentration were observed to be at least two orders of magnitude higher than those calculated for either N or Fe and ranged from 0.87 - 29.0 day^{-1} (Figure 49 and Figure 50).

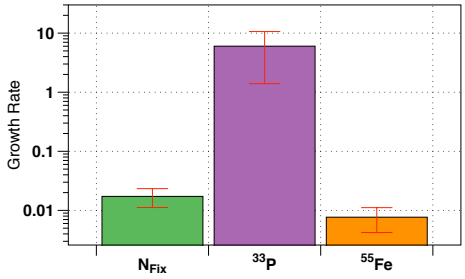

Figure 49: Logarithmic plot showing the mean N_{Fix}, ^{33}P and ^{55}Fe derived *Trichodesmium* growth rates observed during D361. Error bars represent standard error.

Should these rates to be maintained for any length of time the intracellular composition of *Trichodesmium* would deviate significantly and become substantially enriched in P. I observe no such evidence of this in our assessment of the intracellular composition of environmental populations of *Trichodesmium* (Chapter 4). The duration of these P uptake assays were relatively short whereby colonies were incubated with the addition of 0.01 nM ^{33}P (representing 0.0079-0.045% of the ambient DIP pool) for 3 hours. It is therefore conceivable that an elevated uptake rate could be maintainable over the duration of the experiment, however it is unlikely that it could be maintained over longer periods. Additionally with the ^{33}P amendments representing such a small fraction of the ambient DIP pool, it is unclear why P uptake would be elevated during the experimental time-course and a P-enrichment induced over-plus or luxury uptake response seems unlikely.

In order to further investigate this anomaly a comparison of literature intracellular concentrations and uptake rates for the three elements in question was performed, including data from Sañudo Wilhelmy et al. (2001); Post et al. (2002); Lugomela et al. (2002); Mulhol-

land et al. (2002); Carpenter et al. (2004); Capone et al. (2005); Krauk et al. (2006); White et al. (2006); Tovar-Sanchez and Sañudo Wilhemy (2006); Mulholland and Bernhardt (2006); Hynes (2009); Orchard et al. (2010a); Orchard (2010); Sohm et al. (2011a); Rubin et al. (2011); Roe et al. (2012); Nuester et al. (2012) (Table A8). Data were aggregated for intracellular N, P and Fe as well as uptake rates from the $^{15}N_2$ method, the acetylene reduction method, ^{33}P radio-tracer method, Alkaline Phosphatase activity method (for DOP) and the ^{55}Fe method (Figure 50).

Considering this extensive data set, our data are comparable with literature values (Figure 50) for all parameters. What becomes evident in our data collected during D361 and corroborated in the literature, is that P uptake appears to be happening vastly in excess of the rate which would be expected on the basis of observed growth rates and *Trichodesmium*'s known intracellular N:P (Figure 50). Indeed P uptake, on a molar basis, is shown to be occurring at approximately the same rate as N_{Fix} (N uptake). Clearly these observations represent an anomalous P uptake rate requiring further explanation.

Quorum sensing, a cell density signalling system; shown to trigger elevated APase activity in *Trichodesmium* consortia (Van Mooy et al., 2012) may explain our observation of elevated uptake. In performing both APase activity and ^{33}P uptake bioassays, a discrete number of colonies are picked then resuspended in a small volume of media (usually filtered seawater) after which the assays are spiked with trace-concentrations of DI^{33}P or MUF-P (detailed methods can be found in Chapter 2). During D361 this resulted in a *Trichodesmium* abundance of 400 colonies.L^{-1} in each bioassay. In contrast the highest environmental *Trichodesmium* abundance observed during D361 was 5 colonies.L^{-1}. At minimum this would account for an 80-fold increase in relative consortia density and potentially trigger a quorum sensing response along with any other density-related signalling pathway systems, which we hypothesise could induce an enhanced P uptake.

Whilst a quorum sensing response seems plausible as to why these ^{33}P uptake rates are elevated, it provides no indication as to the fate of this P. Whilst the analytical method employed measures the uptake rate of ^{33}P it is also reliant on the fact that the *Trichodesmium* colonies retain this P for at minimum the length of the incubation, this being considered it is unlikely that colonies are rapidly taking-up and recycling/exuding the ^{33}P. Synthesis of polyphosphate granules have been demonstrated as a P storage mechanism in *Trichodesmium erythraeum* (Janson et al., 1995; Orchard, 2010) however qualitative TEM observation for samples collected during D361 showed no evidence for polyP accumulation in either positively or negatively buoyant colonies nor any evident of polyP bodies in *Trichodesmium erythraeum* IMS101 culture (Sargent, 2013). Unfortunately no TEM observations

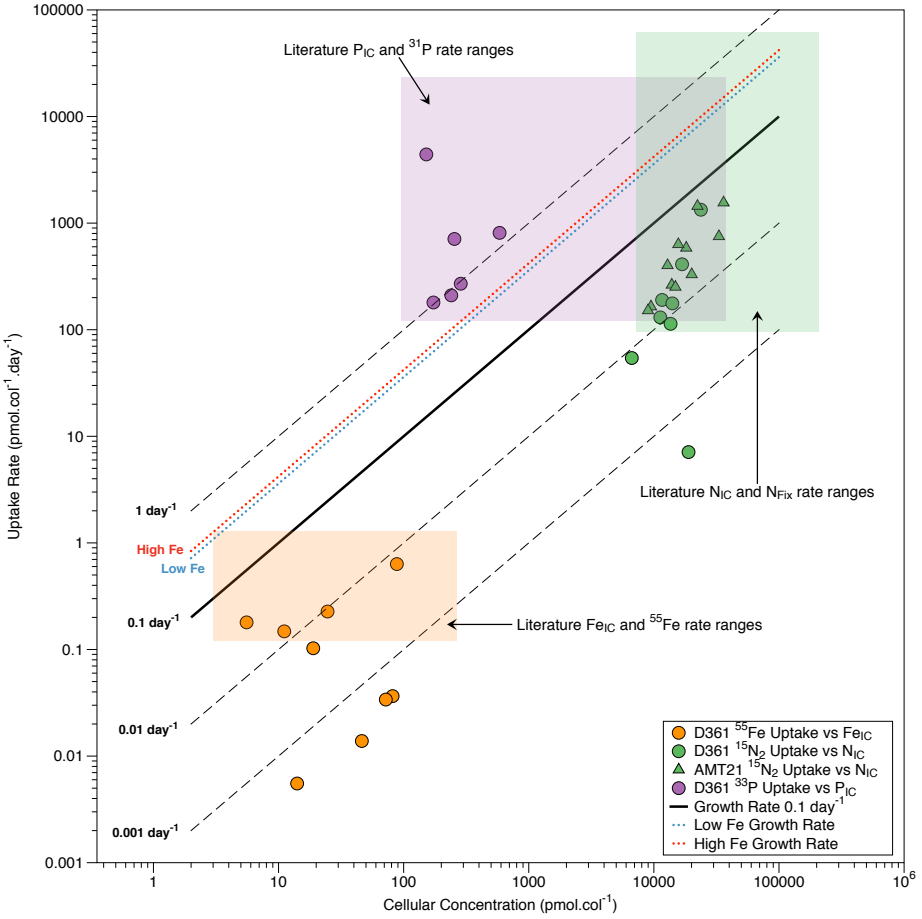

Figure 50: Scatter plot showing uptake rate versus intracellular concentration for N (green), P (purple) and Fe (orange) during D361 (circles) and AMT21 (triangles). Specific growth rates of 0.001-1 day^{-1} are displayed as hatched or solid black lines whilst growth rates observed in *Trichodesmium erythraeum* IMS101 cultures grown at high or low iron (see Chapter 5) are shown with red and blue dotted lines respectively. Literature ranges for uptake rate and intracellular concentration are depicted as shaded regions for N (green), P (purple) and Fe (orange). Literature sources are summarised in Table A8 and include Sañudo Wilhelmy et al. (2001); Post et al. (2002); Lugomela et al. (2002); Mulholland et al. (2002); Carpenter et al. (2004); Capone et al. (2005); Krauk et al. (2006); White et al. (2006); Tovar-Sanchez and Sañudo Wilhemy (2006); Mulholland and Bernhardt (2006); Hynes (2009); Orchard et al. (2010a); Orchard (2010); Sohm et al. (2011a); Rubin et al. (2011); Roe et al. (2012); Nuester et al. (2012).

were made for AMT21 collected data so I am unable to comment on the presence or absence of PolyP accumulation.

A major P containing constituent of microbial cells is that of its genomic material, here it has been demonstrated that upwards of 50% of total cyanobacterial P is genomically bound (Bertilsson et al., 2003). Sargent (2013) has shown *Trichodesmium erythraeum* IMS101 to be polyploid. Documented in numerous other cyanobacteria (Griese et al., 2011), polyploidy is the condition by which an organism contains more than one genome copy per cell. *Synechocystis* PCC 6803 is highly polyploid with ≈58 genome copies per cell during its stationary phase and upwards of ≈218 genome copies per cell in the exponential phase (Griese et al., 2011). This growth strategy, involving a variable degree of polyploidy, insures chromosome presence in daughter cells when cells differentiate but burdens the cell with an increased cellular P requirement (Sargent, 2013). The observed polyploidic nature of *Trichodesmium* is inconsistent with our current understanding of oligotrophic microbes, where conservation of resources is thought to be of paramount importance (Swan and Tupper, 2013). Polyploidy in *Trichodesmium* may potentially act as an intracellular P storage mechanism where a rapid increase in genomic material could serve as a repository for freshly acquired P. Whilst being a novel concept, it is not at odds with observations of polyploidy in *Synechocystis* where the degree of ploidy is known to increase during the exponential growth phase (Griese et al., 2011), however no such similar trend was observed in *Trichodesmium* (Sargent, 2013). Moreover during the 3 hour duration of the ^{33}P experiment the observed P derived growth rates suggest a turnover of between 11-360% of total intracellular P and so a rapid change in the degree of polyploidy cannot explain our observations. It is clear that a reassessment and further investigation of the P uptake and intracellular cycling in *Trichodesmium* is required with current literature P uptake rates appearing to be significantly overestimated.

6.3 LINKING ELEMENTAL AND PROTEOMIC COMPOSITION

Using the proteomic data presented in Chapter 5 I took a sub-sample consisting of the top 425 most abundant proteins observed in each treatment (PS-, PS+, NF+ and NF-) which accounted for between 90-98% of the total protein concentration. The amino-acid sequence for each of these proteins was submitted to the Protein Homology/AnalogY Recognition Engine (PHYRE-2) (Kelley and Sternberg, 2009). PHYRE-2 relies on the observation that the number of protein-folds in nature appears to be limited and that many different remotely homologous protein sequences adopt remarkably similar structures (Baker and Sali, 2001). This allows the structure of an unknown protein to be predicted from its amino-acid sequence using this so called

template-based homology modelling or fold-recognition. This homology based strategy is reliant upon the pre-existing catalogue of protein folds to search against, the PHYRE-2 server uses the library of known protein structures supplied by the Structural Classification of Protein (SCOP) (Murzin and Brenner, 1995) and the Protein Data Bank (PDB) (Berman et al., 2000). Upon completion the PHYRE-2 server will match each submitted amino-acid sequence to a PDB entry which most closely matches the unknown sequence. Using the information contained in these PDB entries I was able to predict the type and number of associated ligands and from these extract the stoichiometric metal signature for that protein. This stoichiometric metal signature, when multiplied by our observed protein concentrations supplied us with a prediction of the concentration of each metal present in that protein. Considering the relative infancy of predicting conserved protein-folds involved in metal-binding, there exists the possibility our method may underestimate the predicted protein-metal concentration of certain less well studied metals. Proteins beyond these top 425 were neglected due to their low abundance and diminishing effect on the predicted-metallome as a whole. This extrapolated metal stoichiometry was then compared with the mean elemental stoichiometry discussed in Section 4.3.5 of Chapter 4 alongside a more detailed comparison against two locations from AMT 21 representative of a putative iron replete phenotype (TN004) and an iron deplete phenotype (TN022). Data are presented as relative stoichiometries as I have no robust way of normalising between the predicted culture metallo-proteome and the observed environmental metallome. The significance of these stoichiometric changes for each metal was assessed by first converting the calculated per-protein metal concentration into a percentage of the total protein-derived metal concentrations for each metal of interest. Changes in these percentages were then assessed using a two-way ANOVA followed by a Bonferroni correction to account for multiple comparisons (further details can be found in Chapter 5). Multiple comparison adjusted P-values are shown in Table A7 (located in Appendix B) where a P-value of <0.05 was deemed significant.

Of the top 425 proteins surveyed across each of our 4 samples 101-121 proteins were found to contain metal-binding co-factors, many of which were protein with a well known metal composition such as the iron containing protein involved in nitrogenase and photosynthesis, the correct prediction of which provides confidence in our method. These metal containing proteins account for between 24-28% of the total proteins surveyed, which is in good agreement with previous predictions, whereby one quarter to one third of all proteins are thought to be metalloproteins (Dupont et al., 2008; Waldron and Robinson, 2009) (Figure 51).

Within our identified metallo-proteins 18 different metal ions were predicted to be bound, with the top 5 most abundant being Mg, Fe, Ca, Zn and Na (Table 18). Of the metals to which I have intracellular concentrations, metalloproteins containing Fe, Zn, Mn, Ni, Cu, Mo, As, Co and Cd are identified whilst no proteins were predicted to contain V. Comparison between direct intracellular stoichiometry and protein-derived metal stoichiometry show marked similarities with regards to Fe, Ni, Cu, Zn, Cd and Co (Figure 52). Protein derived estimates of Mo and As closely resemble intracellular stoichiometries observed during D361 however the elevated observed Mo, As and V during AMT21 are not mirrored in the metallo-proteomic estimates furthering my mistaken uptake hypothesis outlined in Chapter 4. Conversely, the relative concentration of Mn in the protein-derived estimates appears to overestimate compared with the direct environmental intracellular stoichiometries (Figure 52).

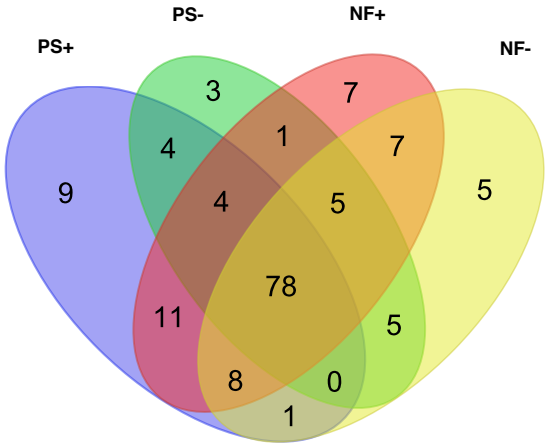

Figure 51: Venn diagram depicting all metallo-proteins observed in the top 425 most abundant protein from each of our 4 treatments (PS+, PS-, NF+ and NF-).

6.3.1 *Iron, Manganese and Copper*

Trichodesmium's nutritional requirements for, and functional uses of iron have been discussed in detail in the preceding chapters of this thesis. Iron is of pivotal importance in both nitrogen fixation and photosynthesis alongside numerous additional metals such as manganese and copper, both of which are found in trace amounts in the marine environment (Bruland, 1983; Byrne et al., 1988).

The protein-derived iron concentration shows a dramatic decrease when grown under iron stress (Figure 53), of the 23 proteins assessed, 13 were observed at significantly reduced concentration in the low iron treatments, details of which can be found in Chapter 5. Signif-

Metal	NF+	PS+	NF-	PS-	Number of PDB structures
Mg	42	38	40	40	8808
Fe	22	21	17	15	5985
Ca	19	18	19	19	7443
Zn	15	13	12	10	9130
Na	14	15	11	10	4558
K	8	8	7	9	1590
Mn	7	9	8	7	2115
Ni	5	5	5	4	838
Cu	3	3	3	2	1111
Hg	3	3	2	1	504
La	1	1	1	1	8
Mo	1	1	1	1	172
As	1	2	2	1	NA
Co	1	1	1	1	862
Yt	1	1	1	1	NA
Pt	1	1	0	0	146
Cd	1	1	0	0	715
V	0	0	0	0	114
Al	0	0	0	0	150

Table 18: Table listing the number of metal binding proteins ordered from highest count to lowest. Alongside these are the number of Protein Data Bank (PDB) protein structure entries that contain that metal (obtained from http://metalweb.cerm.unifi.it/statistics/summary/ on the 8th April 2014).

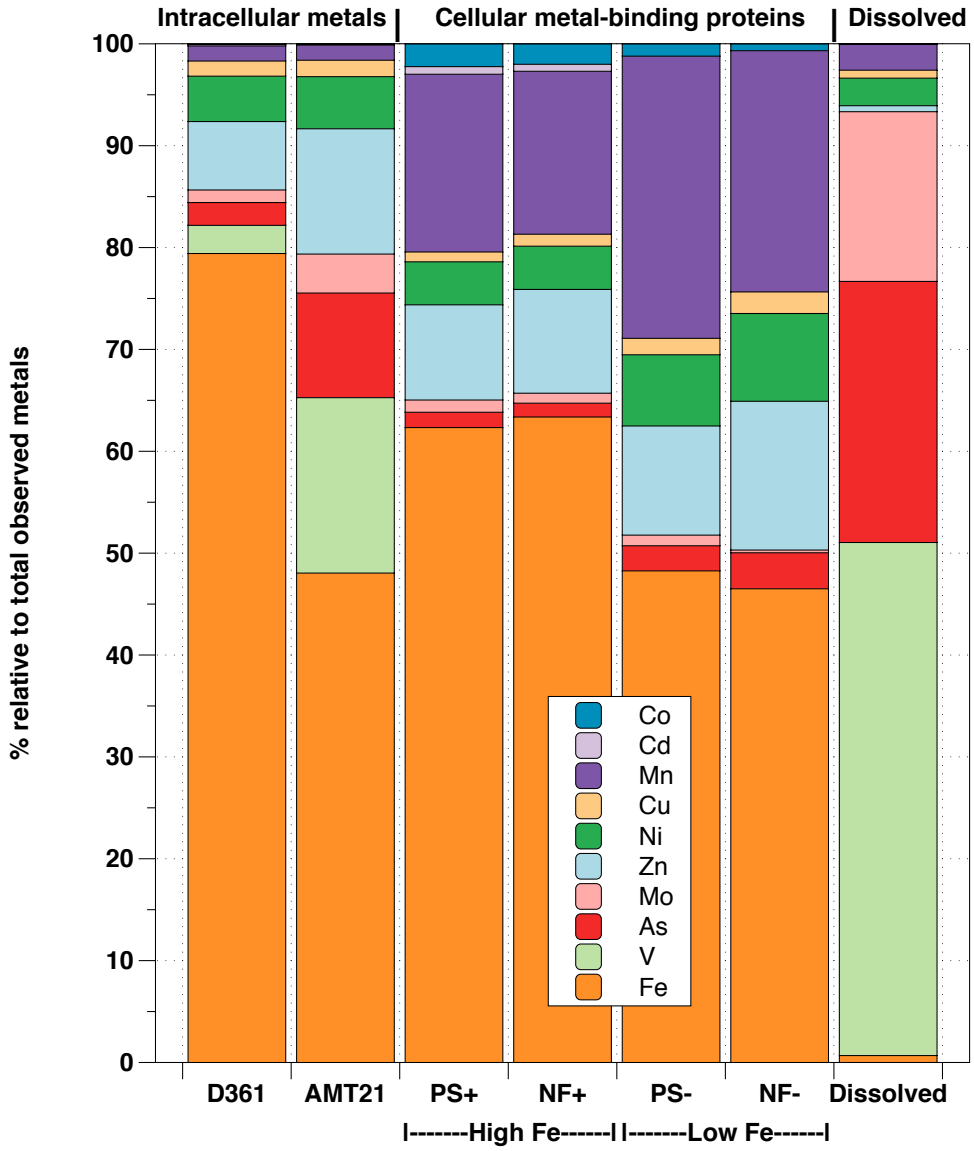

Figure 52: Mean intracellular metal composition for D361 and AMT21 compared to protein-derived metal stoichiometries from PS+, NF+, PS- and NF-. Data displayed for Fe, V, As, Mo, Zn, Ni, Cu, Mn, Cd and Co. By comparison the mean dissolved stoichiometry of each metal as observed in the Atlantic (except As and V which are taken from global literature values (Bruland, 1983; Byrne et al., 1988)) is displayed, sources as follows Fe, Zn, Ni, Cu, Mn, Cd and Co - D361 and AMT21 average (this study), Mo - Surface water equatorial values from GEOTRACES Intermediate Data Product section GA02.

icant stoichiometric changes are seen for the three primary nitrogenase proteins (NifH, NifD and NifK) alongside PsaC, the iron-sulfur center containing protein of PSI.

In Chapter 4 I used our observed nitrogen fixation rates to predict a minimum of ≈ 12-37% of the observed intracellular Fe pool was attributable to the nitrogenase proteins. By comparison, our protein derived iron concentration range from ≈40-58%, this is closer to the values predicted by Richier et al. (2012) or Kustka (2003a) of 38-90%. If I again consider the potential underestimation attributed to the $^{15}N_2$ nitrogen fixation method, discussed in Section 4.3.5 (Mohr et al., 2010; Groß kopf and LaRoche, 2012), then there is potentially very good agreement between protein-derived and activity-derived estimates of nitrogenase-Fe (Figure 53).

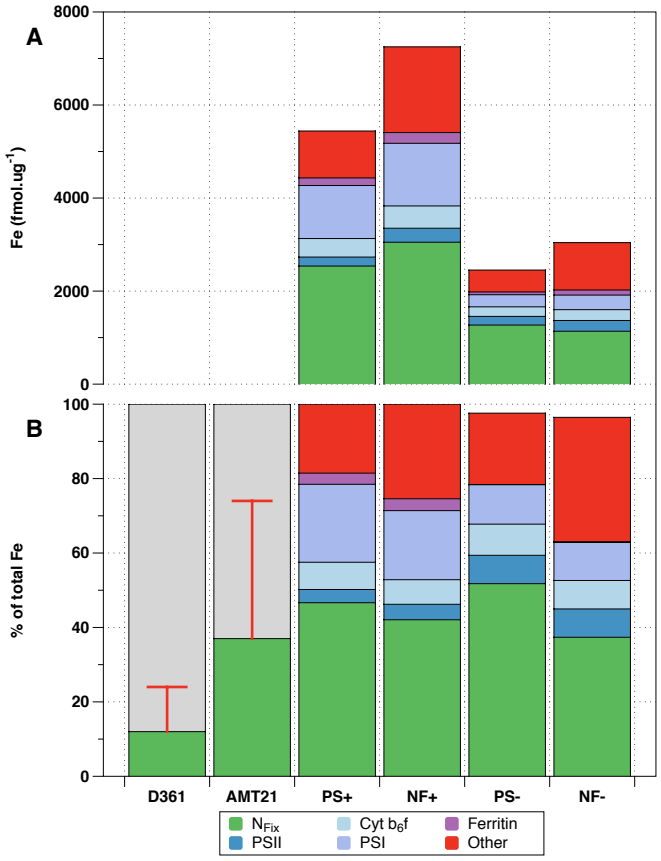

Figure 53: A) Predicted protein derived Fe concentration expressed as fmol Fe.ug^{-1} total protein B) Comparison between metallo-protein derived intracellular Fe partitioning for culture samples PS+, PS-, NF+, NF- and iron partitioning calculated from intracellular concentrations and nitrogen fixation rates for D361 and AMT21 (see Section 4.3.5). Positive error bars for D361 and AMT21's N_{Fix} values correct for the potential two-fold underestimation discussed in Section 4.3.5.

Manganese-dependent enzymes include, amongst others, oxidoreductases, transferases, hydrolases, lyases, isomerases and ligases (Roth et al., 2013). Within our *Trichodesmium* proteomes I identify 9 Mn containing proteins with the oxygen evolving complex of PsbA1/2 and PsbA3 (see Section 5.5.3) accounting for 51-60% of the total intracellular Mn pool (Figure 54). A significant increase in this percentage is observed between samples PS- and NF- for PsbA3 only (Section 5.5.4.2 and Table A7). The stoichiometric contribution glutamine synthetase makes to the total Mn pool is significantly elevated during the earlier time period (Section 5.5.4.2 and Table A7).

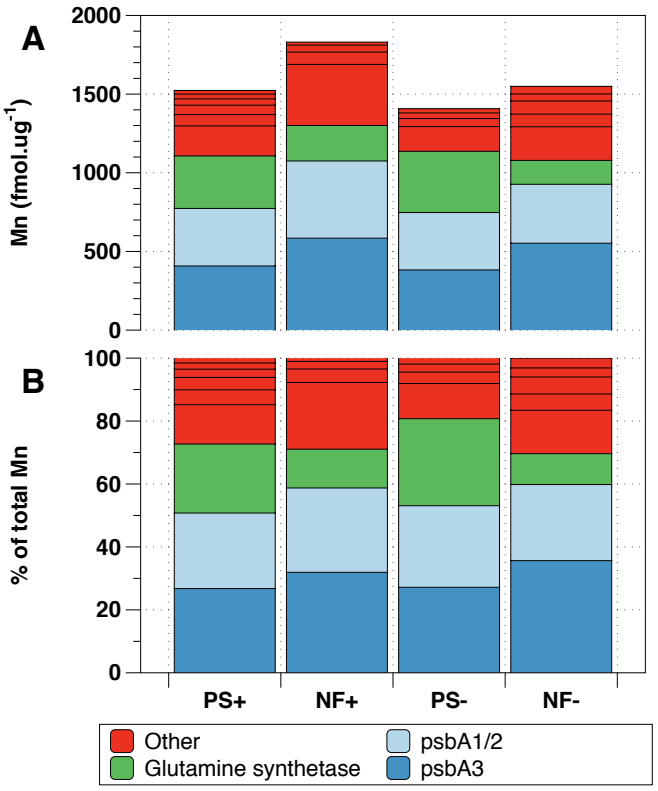

Figure 54: A) Predicted protein derived Mn concentration expressed as fmol Mn.ug^{-1} total protein B) Stoichiometric comparison of Mn-containing protein identified in the top 425 most abundant proteins observed in each treatment (PS+, PS-, NF+ and NF-). PSII isofrom proteins, PsbA1/2 and PsbA3, are shown in light blue and blue respectively, glutamine synthetase is shown in green whilst the remaining 6 Mn containing proteins are shown in red.

Three Cu proteins were identified containing 1 Cu ion each, plastocyanin (Tery_2563) discussed in Chapter 4 and Chapter 5, an NmrA-like protein (Tery_3563) involved in transcriptional regulation and the thioredoxin protein (Tery_0945) putatively identified as TrxB in Chapter 5. This TrxB varies across the diel cycle and so causes a change in total protein-related Cu content (Figure 55 and Table A4 and Ta-

ble A7). TrxB also shows significant stoichiometric change between NF- and NF+ alongside a similar change in plastocyanin's contribution to total Cu and suggests a potential nitrogen fixation associated change in cellular Cu requirement.

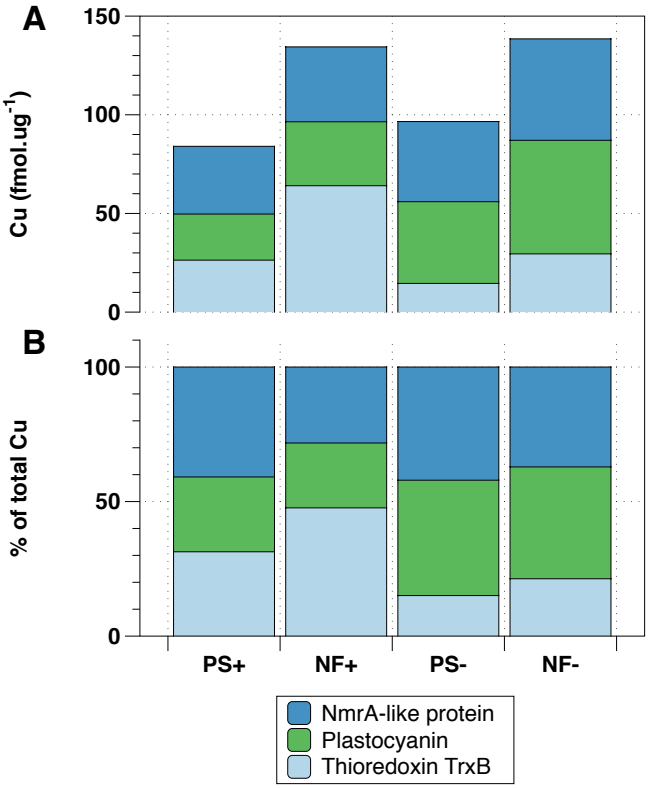

Figure 55: A) Predicted protein derived Cu concentration expressed as fmol Cu.ug^{-1} total protein B) Stoichiometric comparison of Mn-containing protein identified in the top 425 most abundant proteins observed in each treatment (PS+, PS-, NF+ and NF-). An NmrA like protein is shown in dark blue (Tery_3563), plastocyanin is show in green (Tery_2563) and the putative thioredoxin TrxB is shown in light blue (Tery_0945).

6.3.2 *Vanadium, Arsenic and Molybdenum*

Despite our observations of substantially elevated vanadium, arsenic and molybdenum in some samples collected during AMT21 I see few proteins that could be responsible for such enrichments. Conversely, relative to our other metals of interest, dissolved V, As and Mo are highly enriched in the marine environment (Figure 52) (Bruland, 1983; Byrne et al., 1988).

Of the proteins assessed, none were identified as being vanadium binding. Nuester et al. (2012) observed a similar intracellular enrichment in vanadium where they suggested this enrichment could be

due to vanadium nitrogenase for diazotrophy or vanadium haloperioxidases as a potential oxidative-defence mechanism. The lack of identified vanadium binding proteins in our most abundant 425 proteins suggests one of two things. Either, the observed intracellular vanadium is not protein associated, as was suggested in Section 4.3.3 of Chapter 4 where inadvertent V enrichment is considered as a symptom of P stress. An alternative explanation could be that the library of known vanadium binding protein folds is incomplete, despite their being 114 V-containing PDB entries (Table 18), making prediction of a protein's vanadium content unreliable at this present time.

Two uncharacterised proteins (Tery_3509 and Tery_3508) are predicted to bind an arsenic containing cacodylate ion ($C_2H_6AsO_2$). Tery_3509's predicted structure matches PDB entry 3FCU describing an integrin-like protein whilst Tery_3508 matches entry 2JFS describing a hydrolase-like protein.

Finally, the only protein found to contain molybdenum in any treatments was NifD, the alpha subunit of the nitrogenase MoFe protein, which contains the Fe_7S_8Mo cofactor. To some extent this validates our assumption in Section 4.3.6.1 where I assumed NifD to be the sole Mo-containing protein of significant abundance and was therefore only able to account for a maximum of $\approx$33% of the observed intracellular molybdenum. A lack of other Mo-containing proteins affords us no further ability to account for this remaining intracellular Mo.

These predicted V-, As- and Mo-containing proteins (or lack thereof) are unable to adequately explain our observed V, As and Mo enrichments discussed in Chapter 4. Unfortunately, within the scope of this study we're unable to fully explain these enrichments directly and so can only hypothesise the, P-stress induced, mistaken-uptake of vanadate, arsenate and molybdate described in (Section 4.3.5).

6.3.3 *Nickel and Zinc*

A total of 5 Ni-dependent proteins were observed across our 4 treatments the most abundant of which were Tery_2842, a nitrogen regulatory protein P-II containing 1 Ni ion per protein closely followed by Tery_0891, a NiSOD protein again containing 1 Ni ion per protein as discussed in Chapter 4 and Chapter 5. I see a varying predicted protein-derived Ni concentration (Figure 56) suggesting a potential diel variability in intracellular Ni content. This diel variability is not readily attributable to any one protein given the only statistically significant stoichiometric change is that of the nitrogen regulatory protein P-II when compared between PS- and NF- (Table A7).

A large number of zinc-containing proteins were observed across all 4 treatments (Table 18). Despite discussion in Chapter 4 alkaline phosphatase, the Zn-containing protein responsible for P-ester acqui-

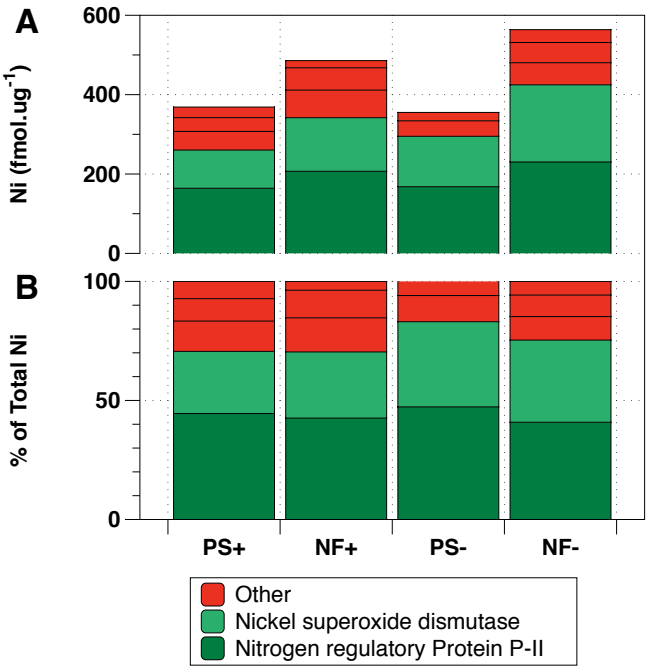

Figure 56: A) Predicted protein derived Ni concentration expressed as fmol Ni.ug^{-1} total protein B) Stoichiometric comparison of Ni-containing protein identified in the top 425 most abundant proteins observed in each treatment (PS+, PS-, NF+ and NF-). A Ni-containing nitrogen regulatory protein P-II (Tery_2842) is shown in dark green, nickel superoxide dismutase (Tery_0891) is shown in light green whilst the 3 remaining Ni-containing proteins are shown in red.

sition in *Trichodesmium* was not present in the 425 proteins assessed. Given all treatments are grown in P-replete media this is perhaps unsurprising. The total protein-derived zinc concentration is variable across all 4 treatments (Figure 57) due in part to the a protein annotated as 'Ferritin and Dps' (Tery_4282, requiring 2 Zn ions, in addition to 1 Fe, per protein) and having strong homology to DpsA in *Synechococcus*. DpsA is a DNA binding hemeoprotein which displays strong variability with iron concentration (Figure 53). The 3'-5' exonuclease protein (Tery_1621) meanwhile showed a significant time-of-day difference (Table A7). Accounting for between 14-30% of the total protein-derived Zn pool is Tery_1794, an uncharacterised protein of unknown function identified as containing 4 Zn ions. The PHYRE-2 predicted protein structure most closely matches entry 3LS1 describing the cyanobacterial photosynthetic protein PsbQ, similarly a BLAST search for this protein identifies strong homology to numerous other cyanobacterial PsbQ proteins. PsbQ is a subunit protein of photosystem II and is thought to enhance PSII's oxygen evolving ability (Balsera et al., 2003).

6.3.4 *Cadmium and Cobalt*

Cadmium and cobalt are present at vastly lower intracellular concentrations than any other metal assessed (Chapter 4). This observation is matched in regards to Cd or Co containing proteins, both metals are predicted to be bound in only one protein each. Subunit 4 of the cytochrome b_6f complex is the sole Cd containing protein (Yamashita et al., 2007) which shows a presence/absence response between high/low iron treatments. The only Co containing protein identified in the top 425 most abundant proteins was a Class II Fructose-bisphosphate aldolase (Tery_4099) which is abundant in the high iron treatments, within low iron treatments this Co containing enzyme was replaced with the metal free, functional replacement Class I Fructose-bisphosphate aldolase(Tery_1687) suggesting a possible link between iron stress and intracellular cobalt concentration.

6.4 TRICHODESMIUM PHENOTYPES IN THE CENTRAL ATLANTIC

Having established distinct biogeochemical provinces of the (sub)-tropical Atlantic within which *Trichodesmium* are found (Chapter 3) then observed and linked their cellular metal concentration to their metallo-protein complement (Chapter 4 and Chapter 5) I can begin to differentiate multiple *Trichodesmium* phenotypes in this region.

Here I focus on two locations sampled during D361, a 'High iron' location (Stn. 14 - 3.33 °N , 26.81 °W) and a 'Low iron' location (Stn. 13 - 1.15 °N, 26.04 °W) thought to reside either side of the interface

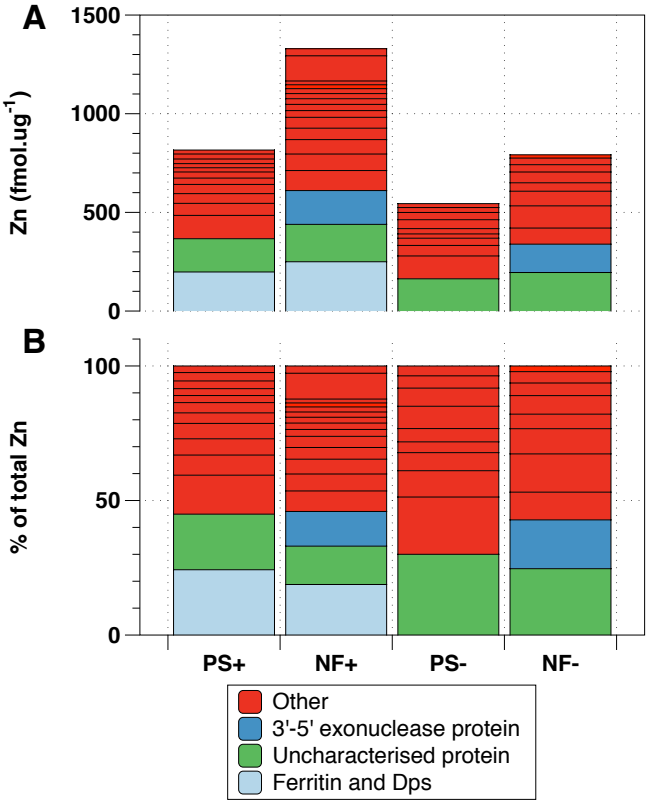

Figure 57: A) Predicted protein derived Zn concentration expressed as fmol Zn.ug^{-1} total protein B) Stoichiometric comparison of Zn-containing protein identified in the top 425 most abundant proteins observed in each treatment (PS+, PS-, NF+ and NF-). A 3'-5'exonuclease protein (Tery_1621) is shown in dark blue, Ferritin and Dps (Tery_4282) is shown in light blue, a Zn-containing uncharacterised protein (Tery_1794) is shown in green and the remaining Zn-containing proteins are shown in red.

between the HPHFe (ITCZ influenced) region and HPLFe (SASG) region (Figure 58).

Chosen to best match the biogeochemical parameters established in the *Trichodesmium* iron-stress culture experiment (Chapter 5) these stations show a large change in DFe (Low iron - 0.25 nM DFe, High iron - 0.65 nM DFe) with minimal change to DIP (Low iron - 41.9 nM DIP, High iron - 14.0 nM DIP) and DIN (Low iron - 12.5 nM DIN, High iron - 3.8 nM DIN). In addition *Trichodesmium* abundance is similar between the two locations (1.6 and 1.4 colonies.L^{-1} respectively).

The physiological characteristics of *Trichodesmium* colonies at these locations can now be considered in terms of a series of biological feedbacks responding to an environmental perturbation. The observed change in DFe concentration, influenced by the ITCZ's location (Schlosser et al., 2014) appears to drive a large increase (from 0.013 to 0.64 pmol.col^{-1}d^{-1}) in the observed ^{55}Fe uptake rate (Figure 58). This elevated rate of Fe acquisition has in turn potentially allowed the colonies to become enriched in Fe and so alter their Fe:P ratio from 101 to 270 mmol Fe.mol^{-1} P and would allow a change in *Trichodesmium*'s proteomic composition as detailed in Chapter 5. One such change is the up-regulation of the nitrogenase proteins involved in nitrogen fixation and the subsequent increase in observed nitrogen fixation rate (Figure 58).

These variations, driven by a change in DFe concentration form the basis for two distinct *Trichodesmium* phenotypes. Observed at similar abundances yet with different elemental and proteomic compositions these two phenotypes have different Fe uptake and N_{Fix} rates and so will have differing effects on the biogeochemistry of the Atlantic.

6.5 SUMMARY OF KEY IDEAS

In Chapter 3 I describe the realised niche within which *Trichodesmium* are most abundant. Put simply this is the boundary region between *Trichodesmium* induced Fe and P oligotrophy within the larger regions of N oligotrophy that is the central Atlantic. I suggest *Trichodesmium* are constrained to these boundary regions as a result of their Fe and P requirement, an idea complimented by the observed ITCZ-associated seasonal *Trichodesmium* migration described in Schlosser et al. (2014). Secure within this N deficient niche I suggest that a primary driver of *Trichodesmium*'s ecophysiology and hence biogeography therefore becomes the acquisition of their required Fe and P. This driver is potentially shared by other diazotrophs as suggested by our observations of non-*Trichodesmium* nitrogen fixation in the western SASG where similar biogeochemical conditions were observed. Chapter 4 and Chapter 5 highlighted some of the strategies and trade-off's *Trichodesmium* employ for achieving these goal whilst the sharp Fe and P curtains observed in the Atlantic act as both the niche boundaries

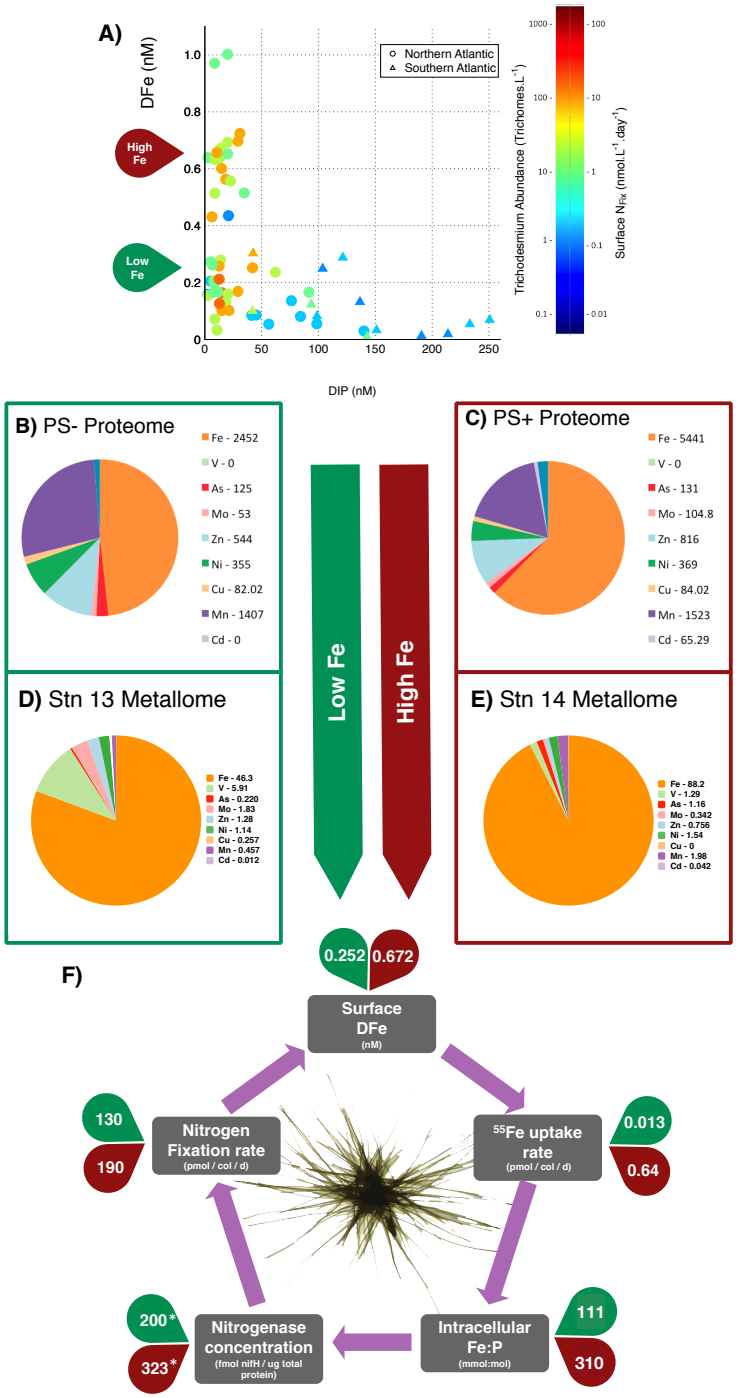

Figure 58: Case study of *Trichodesmium* phenotypes in the (sub)-tropical Atlantic. A) Modified resource-ratio plot from Chapter 3 showing surface nitrogen fixation rate (nmol.L^{-1}.day^{-1}) plotted against DIP and DFe concentrations (nM). Data from the SASG are indicated with triangles. B) and C) show proteomic-derived cellular metal composition for culture treatments PS- (B) and PS+ (C). Plots D) and E) show environmental intracellular metallomic composition from a low iron (D, Stn 13) and high iron (E, Stn 14). F) depicts *Trichodesmium* as a system of biological and environmental feedbacks upon which I have overlaid data discussed throughout this study. Culture data is denoted with an asterisks.

but also as testament to the proliferation of the diazotrophs, and *Trichodesmium* specifically (Chapter 3) (Schlosser et al., 2014; Capone, 2014).

In recent years our understanding of the role DOP plays in *Trichodesmium*'s growth and distribution has dramatically increased (Dyhrman et al., 2006; Dyhrman and Benitez-Nelson, 2009). However widespread analysis of this substantial P source is still not routinely performed. Whilst the importance this P source plays in the (sub)-tropical Atlantic is slowly being recognised (Mather et al., 2008; Lomas et al., 2010) we are still a long way from fully understanding its effect on the diazotrophic community. Furthermore the current body of research looking at P uptake and APase activity is in need of review. To our knowledge all currently published P uptake rates, when compared to N_{Fix} rates show a significant divergence from the N:P ratio's observed in *Trichodesmium* and likely significantly overestimates *Trichodesmium*'s P derived growth rates.

Furthering the work of Nuester et al. (2012), Sañudo Wilhelmy et al. (2001) and Tovar-Sanchez and Sañudo Wilhemy (2006) I re-iterate the dominance of iron in *Trichodesmium*'s metallome. I show *Trichodesmium* to have a unique elemental composition, in-keeping with its cyanobacterial lineage and its known cellular processes. Whilst demonstrating this diazotrophs resilience to environmental variability I also highlight potentially costly mis-steps in the form of potentially mistaken As, V and Mo acquisition. Elemental enrichments of Mo, V and As in P-deficient waters, without any observed proteomic demand, appear to serve little purpose and may indeed be toxic (Banci and Bertini, 2013).

6.6 FINAL THOUGHTS

Planktonic bacteria dominate surface ocean biomass and have major roles in global carbon and nutrient cycling (Glöckner et al., 2012). The cyanobacteria *Prochlorococcus* and *Synechococcus* along with the alphaproteobacteria *Pelagibacter* (belonging to the SAR11 clade) are perhaps the best studied given their cosmopolitan distribution and our ability to grow them in culture. These organisms are highly abundant in the (sub)-tropical oceans where they are thought to have adapted to resource-scarcity by having small, highly streamlined genomes (Swan and Tupper, 2013) (Figure 59). Similarly the unicellular diazotroph UCYN-A appears to have undergone significant genome and metabolic streamlining (Tripp et al., 2010) as a means of surviving in these oligotrophic conditions. At only 1.44 megabases UCYN-A has a very small genome characteristic of its symbiotic nature (Figure 59), in addition to this UCYN-A lacks a functional PSII complex (Zehr et al., 2008) and is unable to synthesise numerous essential amino-acids and purines (Tripp et al., 2010). The Black Queen hypothesis proposed by

Morris et al. (2012) stems from this symbiotic strategy of reductive evolution. In oligotrophic conditions gene loss is proposed to confer a selective advantage by allowing an organism to conserve its limited resources, provided the gene's function is dispensable to the individual yet maintained by the community. An example Morris et al. (2012) draw upon is the absence of *katG* genes in *Prochlorococcus*, *KatG* encodes for catalase-peroxidase, a large Fe-containing protein thought to be the primary defence against external hydrogen peroxide. They suggest that by dispensing of this gene *Prochlorococcus* are able to maintain a selective advantage in Fe-limited conditions provided the detoxification of hydrogen peroxide by other *KatG* containing organisms prevents the build-up of hydrogen peroxide (Morris et al., 2012). Here certain marine microbes are hypothesised to be 'helpers' where their so called 'leaky functions' (cellular processes that alter the extracellular environment) can be leveraged by organisms lacking such functions, termed 'beneficiaries' (Morris et al., 2012), a concept not dis-similar to the 'sloppy' feeding in zooplankton (Hygum, 1997).

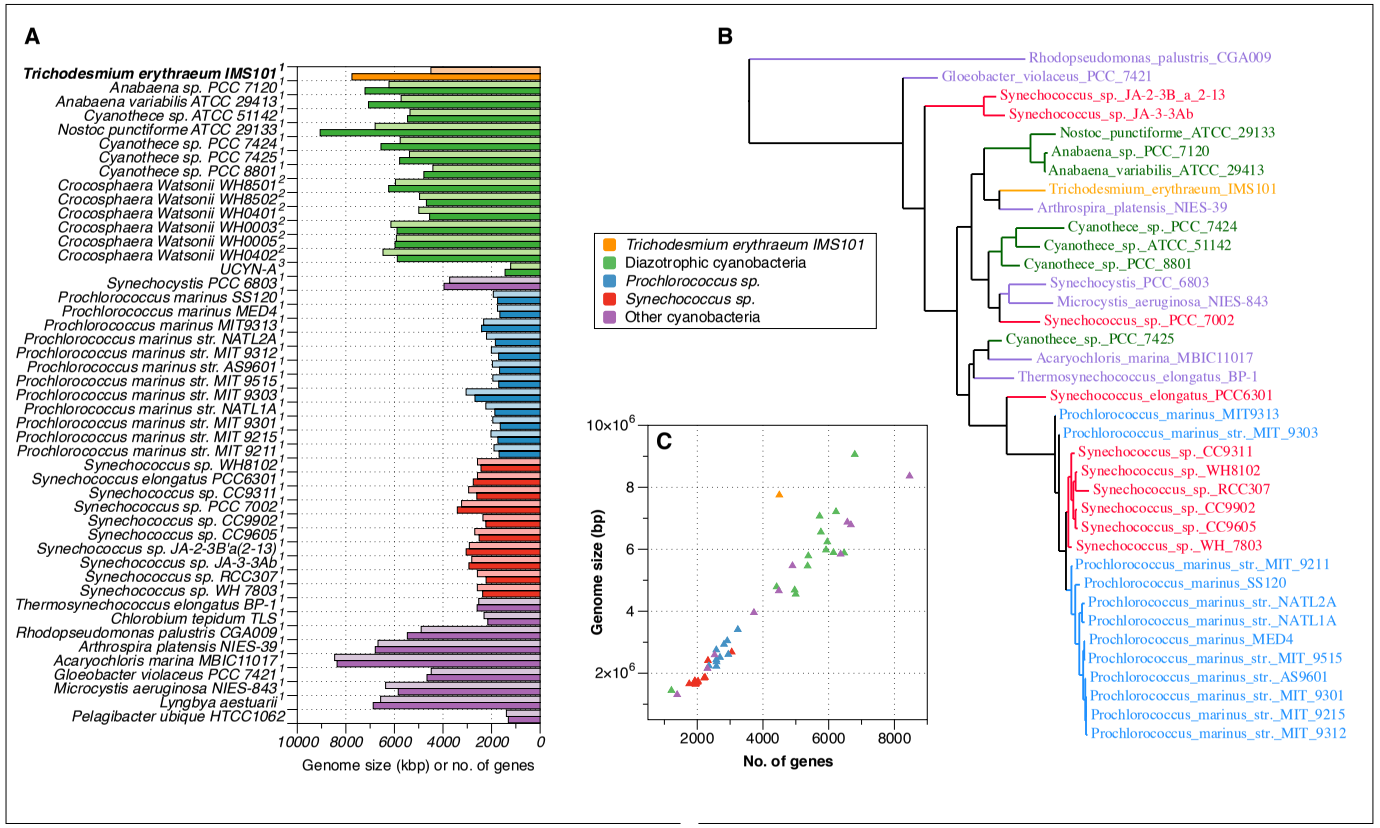

Figure 59: A) Genome size comparison for select cyanobacterial genomes. Colour grouping is as per legend with darker bars show genome size in kbp whilst lighter bars show the number of protein-encoding genes present in each genome. B) 16S ribosomal protein phylogenetic tree showing the evolutionary history of cyanobacteria. C) Scatter plot comparing genome size (bp) against number of genes. Data sources, as indicated on plot A, are as follows - [1]Kazusa Cyanobase database (http://genome.microbedb.jp/cyanobase/), [2]Bench et al. (2013) and [3]Tripp et al. (2010).

In stark contrast to these genetically streamlined organisms are the genomes of the free-living diazotrophic cyanobacteria which range from 4.55 - 9.05 Mbp with that of *Trichodesmium erythraeum* IMS101 being 7.21 Mbp (Figure 59). Of these bacterial genomes, a broadly conserved relationship is observed between genome size and the number of protein encoding genes (Figure 59) with the percentage of coding DNA in UCYN-A of 81% and ≈97% for both *Pelagibacter* and *Prochlorococcus* (Tripp et al., 2010). *Trichodesmium* deviates substantially from this with only ≈60% of the genome thought to be protein coding DNA. With the exception of the endosymbiotic UCYN-A all marine diazotrophs have substantially larger genomes than their non-diazotrophic cousins with *Trichodesmium*'s genome appearing particularly bloated. A larger microbial genome is thought to allow for a larger more complex microbial cell. Indeed *Trichodesmium* cells (225 - 6750 μm^3) are significantly larger than *Prochlorococcus* cells (0.8 μm^3) but also substantially larger than the unicellular diazotroph *Crocosphaera*'s cells (7.1 μm^3) (Bergman et al., 2012; Janson et al., 1995). The very nature of diazotrophy requires a more complex cell than a non-diazotrophic counterpart. Whilst the nitrogenase genes and the proteins they encode provide a diazotroph with access to N_2 there are numerous other genes required for maintaining a cellular environment conducive to the process of N_{Fix}. In *Trichodesmium*, genes involved in protection from oxidative stress, diazocyte differentiation and maintenance and presumably a substantial amount of transcriptional and translational regulation in order to fine-tune *Trichodesmium*'s unique spatio-temporal N_{Fix} strategy (Berman-Frank et al., 2001b) are all additional requirements. Some of this regulation may be unique to *Trichodesmium* due to their N_{Fix} strategy. As is being discovered in other organisms (Shabalina and Spiridonov, 2004), these regulatory sequences could be located in the non-coding fraction of *Trichodesmium*'s genome, and may go some way to explaining *Trichodesmium*'s substantial amount of non-coding DNA.

Through their ability to fix nitrogen, diazotrophs have often been thought of as specialists in the context of global microbial oceanography. Whilst this diazotrophic ability does allow them to survive in N deficient environments where others organisms cannot, I suggest that's where their specialism ends. Once secure in N-limited waters *Trichodesmium*, by virtue of their larger genome, become generalists accessing their required Fe and P through any number of nutrient acquisition, compensation or retrenchment processes outlined throughout this text. I therefore suggest the reason for *Trichodesmium* larger genome size may be considered to be two fold, firstly the relative security diazotrophs experience in N-limited waters does not necessitate the same degree of genome streamlining observed in species such as *Prochlorococcus* or *Synechococcus* (DNA accounts for ≈0.5-3% of dry weight and contains significant amounts of N and P (Geider

and Roche, 2002)) and secondly maintaining a larger genome brings with it a potentially increased ability to acquire nutrients or adapt to nutrient deficiency. Chapter 4 illustrates the degree of intracellular plasticity *Trichodesmium* are able to achieve within this Fe-P niche. The changes to *Trichodesmium*'s proteome by which they are able to achieve intracellular Fe plasticity are discussed in Chapter 5 where I see an iron-stress induced down-regulation of nitrogenase proteins and the IsiA supported protection of PSI. Furthermore the genome of *Trichodesmium* is thought to be expanding (Larsson et al., 2011), here the genome is open to gene gain. A growing genome containing large amounts of non-coding nucleotides is indicative of a strong adaptive potential (Larsson et al., 2011). Throughout this study we've demonstrated *Trichodesmium* strong adaptive potential, both in regards to their biogeography, elemental composition and proteomic plasticity.

At a community level, I propose *Trichodesmium* assume the role of 'helpers' where a do-it-all genome is not only beneficial to themselves but also to their colonial 'beneficiaries' and potentially the wider community. A living-arrangement such as this is undoubtedly bidirectional. *Trichodesmium*'s apparent inability to synthesise siderophores coupled with their observed ability to utilize third-party siderophores (Roe et al., 2012; Achilles and Church, 2003) and our potential identification of a siderophore uptake pathway (Section 5.5.4.1) all suggests a similar strategy albeit in reverse with *Trichodesmium* being the beneficiaries of their epibionts.

At inception the goal of this project was to directly compare *Trichodesmium*'s intracellular composition to a limited set of specific Fe and P stress 'biomarker' proteins in the environment. In some regards I fall short of this goal having to make inferences from culture-based proteomic data to interpret our observed intracellular metal concentrations measured on field collected samples. However this initial goal, utilising AQUA proteomics, would have provided data for only 13 proteins and insight on Fe and P requirements alone. By utilising proteomic technologies not readily available at the onset of this project I have instead delivered a much more detailed understanding of *Trichodesmium*'s response to iron-stress, as well as insights into the requirements for 18 other metals. In doing so I identify key differences in *Trichodesmium*'s environmental metallome and its cultured metallo-proteome, highlighting the importance of Ni, Mn, Cu, Zn and others in maintaining *Trichodesmium*'s cellular function, all of which are found at potentially limiting concentrations in the marine environment (Bruland, 1983; Byrne et al., 1988). Although unable to directly comment on *Trichodesmium*'s potential response to elevated temperature or pCO_2, I instead show the dominant control nutrient availability has on the biogeography of this globally important diazotroph. In doing so I highlight the plasticity *Trichodesmium* display in their elemental and proteomic composition in response to envi-

ronmental change. These observed adaptations over relatively small spatial (hundreds of miles) and temporal (days) scales hint at the potential for *Trichodesmium* to utilise a large and hence relatively non-reduced genome to facility adaptation to impending global change and its predicted potential influences on the biogeochemistry of the oceans.

A

APPENDIX 1 - INTRACELLULAR METALS TABLES

A.1 MEASURED BLANK VALUES

The values obtained for instrument, reagent and filter blanks during D361 and AMT21 intracellular Trichodesmium analysis are summarised in Table A1, A2 & A3. Multiple blanks were analysed for each and so the value is given as a mean and standard deviation of the mean, with the number of blanks in each case shown in the column titles.

Analyte (resolution)	Instrument blank (±σ, n = 17)	Reagent Blank (±σ, n = 15)	Filter Blank (±σ, n = 17)	Sample Mean (±σ, n = 51)
Cd111(LR)	0 ± 0	0.001 ± 0.001	0.042 ± 0.129	0.045 ± 0.058
Mo98(LR)	0.005 ± 0.001	0.043 ± 0.016	0.083 ± 0.049	0.393 ± 0.259
P31(MR)	0.498 ± 0.096	3.936 ± 2.086	7.358 ± 5.645	58.025 ± 24.102
V51(MR)	0.003 ± 0.001	0.007 ± 0.003	0.02 ± 0.01	0.4 ± 0.393
Mn55(MR)	0.013 ± 0.01	0.023 ± 0.022	0.104 ± 0.103	0.304 ± 0.298
Fe56(MR)	0.289 ± 0.105	0.887 ± 1.503	4.139 ± 2.167	16.517 ± 12.577
Co59(MR)	0.001 ± 0	0.002 ± 0.002	0.007 ± 0.007	0.013 ± 0.008
Ni60(MR)	0.011 ± 0.012	0.022 ± 0.009	0.667 ± 0.435	1.199 ± 0.332
Cu65(MR)	0.024 ± 0.007	0.105 ± 0.085	0.203 ± 0.148	0.428 ± 0.414
Zn66(MR)	0.851 ± 0.749	1.28 ± 0.972	1.474 ± 0.782	2.6 ± 1.79
As75(HR)	0.012 ± 0.009	0.158 ± 0.158	0.217 ± 0.123	4.739 ± 3.701

Table A1: Summary of instrument, reagent and filter blanks observed during D361 ICP-MS analysis. Mean ppb values are displayed alongside the standard deviation of the mean. Number of replicates per column are shown in the column headings.

Analyte (resolution)	Instrument blank (±σ, n = 17)	Reagent Blank (±σ, n = 12)	Filter Blank (±σ, n = 20)	Sample Mean (±σ, n = 36)
Cd111(LR)	0.002 ± 0.002	0.002 ± 0.002	0.004 ± 0.011	0.113 ± 0.075
Mo98(LR)	0.024 ± 0.007	0.129 ± 0.053	0.099 ± 0.065	4.979 ± 3.806
P31(MR)	1.746 ± 0.286	5.656 ± 1.138	5.772 ± 2.327	310.542 ± 216.659
V51(MR)	0.005 ± 0.002	0.013 ± 0.004	0.014 ± 0.012	14.264 ± 18.536
Mn55(MR)	0.035 ± 0.046	0.059 ± 0.042	0.102 ± 0.073	1.647 ± 1.761
Fe56(MR)	0.507 ± 0.204	1.201 ± 0.818	2.614 ± 1.95	48.237 ± 43.22
Co59(MR)	0.001 ± 0.002	0.003 ± 0.001	0.003 ± 0.002	0.052 ± 0.04
Ni60(MR)	0.249 ± 0.299	0.302 ± 0.23	0.239 ± 0.138	4.305 ± 2.036
Cu65(MR)	0.039 ± 0.089	0.079 ± 0.051	0.32 ± 0.315	1.463 ± 0.835
Zn66(MR)	0.875 ± 0.48	1.7 ± 1.392	2.45 ± 2.031	10.037 ± 6.818
As75(HR)	0.002 ± 0.001	0.008 ± 0.003	0.013 ± 0.007	24.186 ± 43.114

Table A2: Summary of instrument, reagent and filter blanks observed during AMT21 ICP-MS analysis part 1. Mean ppb values are displayed alongside the standard deviation of the mean. Number of replicates per column are shown in the column headings.

Analyte (resolution)	Instrument blank (±σ, n = 17)	Reagent Blank (±σ, n = 12)	Filter Blank (±σ, n = 18)	Sample Mean (±σ, n = 37)
Cd111(LR)	0.001 ± 0.001	0.001 ± 0	0.019 ± 0.057	0.073 ± 0.04
Mo98(LR)	0.025 ± 0.002	0.111 ± 0.031	0.156 ± 0.149	3.552 ± 3.213
P31(MR)	0.664 ± 0.116	2.256 ± 0.373	5.589 ± 6	259.012 ± 118.192
V51(MR)	0.005 ± 0.001	0.012 ± 0.003	0.066 ± 0.132	10.941 ± 10.075
Mn55(MR)	0.015 ± 0.031	0.024 ± 0.012	0.177 ± 0.205	1.097 ± 0.617
Fe56(MR)	0.442 ± 0.244	1.058 ± 0.283	6.252 ± 6.911	30.293 ± 32.814
Co59(MR)	0.001 ± 0	0.002 ± 0	0.01 ± 0.01	0.032 ± 0.023
Ni60(MR)	0.036 ± 0.01	0.066 ± 0.021	0.416 ± 0.429	3.111 ± 2.69
Cu65(MR)	0.074 ± 0.091	0.154 ± 0.2	0.742 ± 0.995	2.459 ± 5.073
Zn66(MR)	1.058 ± 1.074	3.934 ± 2.101	7.852 ± 8.857	12.356 ± 17.522

Table A3: Summary of instrument, reagent and filter blanks observed during AMT21 ICP-MS analysis part 2. Mean ppb values are displayed alongside the standard deviation of the mean. Number of replicates per column are shown in the column headings.

B

APPENDIX 2 - PROTEOMIC TABLES

B.1 SIGNIFICANTLY DIFFERENT PROTEINS

Table A4 - Proteins observed at significantly different concentrations between either PS- and NF- or PS+ and NF+. Fold-change values are calculated with values >1 displayed in red and values <1 displayed in blue. Significant fold changes are denoted with an asterisk (*).
Table A5 - Proteins observed to be differentially abundant between iron deplete (PS- or NF-) and replete condition (PS+ or NF+). Fold-change values are displayed with values >1 in red and values <1 in blue. Significant fold changes are denoted with an asterisk (*).

B.2 IDENTIFIED METALLO-PROTEINS

Table A6 - Proteins identified as containing metal-biding cofactors. Identification was achieved using the PHYRE-2 search engine described in the text. PHYRE-2 confidence scores (%) and sequence coverage (%) are shown alongside identified PDB structure and finally per-protein metal composition.

Table A7 - Statistical significance of per-metal stoichiometric changes as asssed by two-way ANOVA with a Bonferroni correction. Table shows multiple comparison corrected P-values where values in red indicate statistically significant stoichiometric changes ($P < 0.05$) whilst values in blue indicate non-significant change ($P > 0.05$).

B.3 INTRACELLULAR AND UPTAKE LITERATURE DATA

Table A8 - Summary of literature data for intracellular concentrations and uptake rate for N, P and Fe observed in environmental *Trichodesmium* colonies. Units for intracellular N, P or Fe are all $pmol.col^{-1}$ whilst for N, P or Fe uptake units are in $pmol.col^{-1}.day^{-1}$.

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	f.c. (NF-/PS-)	f.c. (NF+/PS+)
Q117P0	Tery_0882	Arginine biosynthetic process	Argininosuccinate synthase	54.07	n.d.	46.92	31.93	n.d.	n.d.*
Q11AB4	Tery_0043	Arginine biosynthetic process	Arginine biosynthesis bifunctional protein ArgJ	36.05	56.19	66.02	99.91	1.55*	1.51
Q112Z7	Tery_2198	ATP catabolic process	ATP synthase gamma chain	204.01	259.70	188.48	249.19	1.27*	1.32*
Q112Z5	Tery_2200	ATP catabolic process	ATP synthase subunit delta	197.39	248.50	177.33	238.35	1.25*	1.34
Q112Z6	Tery_2199	ATP catabolic process	ATP synthase subunit alpha	238.52	175.88	195.86	206.54	0.73*	1.05*
Q10YA8	Tery_3707	ATP catabolic process	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha	n.d.	n.d.	40.04	n.d.	n.d.	n.d.*
Q112Z1	Tery_2204	ATP catabolic process	ATP synthase subunit a	n.d.	n.d.	64.76	n.d.	n.d.*	n.d.*
Q10Z38	Tery_3385	ATP catabolic process	ATP synthase subunit beta	170.80	113.57	159.38	151.70	0.66	0.95*
Q112Z3	Tery_2202	ATP catabolic process	ATP synthase subunit b′	169.69	211.88	175.67	274.23	1.24*	1.56
Q112Z4	Tery_2201	ATP catabolic process	ATP synthase subunit b	66.00	62.90	59.64	95.05	0.95*	1.59
Q10Z37	Tery_3386	ATP catabolic process	ATP synthase epsilon chain	44.51	87.16	73.59	166.31	1.95*	2.25*
Q112L9	Tery_2334	ATP catabolic process	Gas vesicle protein GvpN	n.d.	110.28	n.d.	n.d.	n.d.	n.d.*
Q10YF7	Tery_3654	Carbohydrate metabolic process	Phosphoribulokinase	84.22	63.74	105.57	65.29	0.75	0.61*
Q10VU5	Tery_4661	Carbon metabolism	S-adenosylmethionine synthase	115.14	56.10	145.95	122.96	0.48*	0.84
Q110V6	Tery_2787	Cellular iron ion homeostasis	Ferritin and Dps	59.90	109.33	64.13	106.18	1.82*	1.65*
Q115D5	Tery_1610	Chromosome condensation	Bacterial nucleoid protein Hbs	31.44	42.77	38.14	77.64	1.36	2.03*
Q113J6	Tery_2088	Cobalamin biosynthetic process	Cob(I)yrinic acid a c-diamide adenosyltransferase	n.d.	n.d.	44.13	n.d.	n.d.	n.d.*
Q116S1	Tery_1141	CTP Biosynthetic process	Nucleoside diphosphate kinase	164.61	238.41	176.08	199.23	1.44*	1.13
O52659	Tery_1666	Electron transport	Flavodoxin	100.47	147.22	52.94	51.36	1.46*	0.97
Q112L5	Tery_2339	Gas vesicle organisation	Gas vesicle synthesis GvpLGvpF	80.66	84.08	45.69	91.43	1.04	2*
Q117I5	Tery_0948	Gluconeogenesis	D-fructose 1_6-bisphosphatase class 2/sedoheptulose 1_7-bisphosphatase	188.24	185.86	177.45	210.20	0.98	1.18*
Q10XZ7	Tery_3834	Glutamine biosynthetic process	Glutamine synthetase	129.78	50.63	111.29	75.17	0.39*	0.67*
Q10Z46	Tery_3376	Glycolysis	Phosphoglycerate kinase	165.14	97.08	164.47	97.67	0.58*	0.59
Q10UV4	Tery_5066	Glycolysis	Pyruvate dehydrogenase (Lipoamide)	70.32	36.50	58.90	31.38	0.51*	0.53
Q110V4	Tery_2793	Glycolysis	Enolase	51.99	13.35	39.20	14.17	0.25*	0.36
Q10XB7	Tery_4099	Glycolysis	Fructose-bisphosphate aldolase	61.56	44.07	195.42	229.83	0.71*	1.17*
Q114W6	Tery_1687	Glycolysis	Fructose-bisphosphate aldolase	343.56	389.64	19.49	n.d.	1.13	n.d.*
Q115Y7	Tery_1380	Glycolysis	Triosephosphate isomerase	44.33	84.27	57.49	n.d.	1.9*	n.d.
Q118Z2	Tery_0476	GTP catabolic process	Elongation factor Tu	157.12	110.14	163.75	113.31	0.7*	0.69*
Q119L5	Tery_0335	Heme oxidation	Heme oxygenase (Decyclizing)	n.d.	26.16	27.16	75.43	n.d.	2.77*
Q117Z0	Tery_0756	Inositol phosphate dephosphorylation	Inositol-1(Or 4)-monophosphatase	26.89	29.37	11.83	40.43	1.09	3.41*

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	f.c. (NF-/PS-)	f.c. (NF+/PS+)
Q117I0	Tery_0954	Intracellular signal transduction	Two component transcriptional regulator_ winged helix family	54.30	100.15	59.52	88.18	1.84*	1.48*
Q10WP7	Tery_4335	Isoleucine biosynthetic pathway	Ketol-acid reductoisomerase	148.23	186.97	130.52	122.43	1.26*	0.93
Q112B9	Tery_2440	Metabolic process	Glutathione S-transferase-like	29.83	54.73	29.68	59.83	1.83	2.01*
Q10XY3	Tery_3849	Metabolic process	Carbonate dehydratase	115.88	81.57	118.18	101.43	0.7*	0.85
Q110D9	Tery_2977	Metabolic process	3-dehydroquinate synthase	43.71	n.d.	44.16	14.48	n.d.*	0.32*
Q10WV1	Tery_4276	Metabolic process	Helicase-like	n.d.	n.d.	n.d.	n.d.	n.d.*	n.d.
Q10WH8	Tery_4408	Metabolic process	Ribulose 1 5-bisphosphate carboxylase small subunit	254.18	368.85	270.77	335.78	1.45*	1.24
Q110W7	Tery_2774	Metabolic process	Peptidase M16-like	n.d.	n.d.	n.d.	n.d.	n.d.*	n.d.*
Q10VN9	Tery_4723	NA	TPR repeat	104.99	128.81	91.46	119.34	1.22	1.3*
Q10XJ6	Tery_4005	NA	RNA-binding region RNP-1	66.47	87.97	50.78	86.90	1.32	1.71*
Q116J4	Tery_1229	NA	UPF0367 protein Tery_1229	64.74	78.51	68.04	97.69	1.21*	1.43*
Q110Z7	Tery_2743	NA	UPF0234 protein Tery_2743	43.39	60.25	38.34	73.08	1.38*	1.9*
Q113T4	Tery_1989	NA	Uncharacterized protein	39.50	55.34	42.38	81.37	1.4*	1.92
Q110V7	Tery_2786	NA	Uncharacterized protein	45.52	45.66	28.10	59.83	1*	2.12*
Q111D7	Tery_2698	NA	Uncharacterized protein	29.81	39.98	26.87	59.62	1.34*	2.21
Q110I9	Tery_2918	NA	Uncharacterized protein	58.28	49.31	n.d.	76.51	0.84*	n.d.*
Q112A3	Tery_2459	NA	Clostridial hydrophobic	26.85	n.d.	42.27	n.d.	n.d.	n.d.*
Q11A97	Tery_0068	NA	Uncharacterized protein	44.30	n.d.	37.74	n.d.	n.d.*	n.d.
Q110C7	Tery_2992	NA	Translation initiation factor IF-1	n.d.	n.d.	n.d.	42.49	n.d.*	n.d.*
Q118X5	Tery_0493	NA	Uncharacterized protein	108.41	169.56	105.75	159.19	1.56*	1.5*
Q112M7	Tery_2325	NA	Gas vesicle structural protein	90.32	132.80	71.58	105.37	1.47*	1.47*
Q10WQ2	Tery_4329	NA	Microcompartments protein	84.73	124.02	82.63	118.75	1.46	1.43*
Q10VA5	Tery_4894	NA	Uncharacterized protein	130.59	176.69	139.35	203.78	1.35	1.46*
Q119I2	Tery_0373	NA	Uncharacterized protein	83.92	129.70	79.21	128.13	1.54	1.61*
Q10UZ3	Tery_5019	NA	ThiJ/PfpI	76.51	132.29	88.02	129.38	1.72*	1.46*
Q114K7	Tery_1810	NA	Uncharacterized protein	89.59	145.54	85.44	130.86	1.62	1.53*
Q116Q3	Tery_1160	NA	Elongation factor Ts	121.55	179.60	101.46	171.67	1.47	1.69*
Q10YP5	Tery_3542	NA	Uncharacterized protein	22.00	61.94	31.67	77.81	2.81	2.45*
Q10WC2	Tery_4465	NA	Carbohydrate-selective porin OprB	65.92	25.59	37.73	33.17	0.38	0.87*
Q118N0	Tery_0600	NA	Uncharacterized protein	44.77	89.03	26.91	77.61	1.98	2.88*
Q115G7	Tery_1577	NA	NADH-ubiquinone/plastoquinone oxidoreductase_ chain 6	32.84	81.11	46.49	52.52	2.46	1.12*

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	f.c. (NF-/PS-)	f.c. (NF+/PS+)
Q10WC1	Tery_4466	NA	Carbohydrate-selective porin OprB	53.45	8.08	25.42	8.80	0.15	0.34*
Q10VJ8	Tery_4773	NA	Uncharacterized protein	37.54	78.00	41.85	65.88	2.07*	1.57
Q10VI5	Tery_4788	NA	Peptidoglycan-binding domain 1	n.d.	50.85	34.74	56.28	n.d.*	1.62*
Q10Y07	Tery_3823	NA	Uncharacterized protein	83.90	124.63	38.66	46.11	1.48*	1.19*
Q10Z45	Tery_3377	NA	Extracellular solute-binding protein family 1	243.11	285.62	102.08	132.83	1.17*	1.3*
Q10Z50	Tery_3372	NA	Beta-Ig-H3/fasciclin	70.57	115.84	38.72	63.98	1.64*	1.65*
Q113S7	Tery_1998	NA	Extracellular solute-binding protein family 3	165.42	141.56	111.67	61.55	0.85*	0.55
Q10Y15	Tery_3815	NA	Uncharacterized protein	109.22	149.82	42.59	107.70	1.37*	2.52
Q10UX6	Tery_5038	NA	1-Cys peroxiredoxin	149.21	231.44	152.15	185.16	1.55*	1.21
Q114E5	Tery_1873	NA	SPFH domain Band 7 family protein	161.18	246.23	141.93	187.32	1.52*	1.31*
Q10W05	Tery_4593	NA	Uncharacterized protein	131.46	134.18	111.14	81.96	1.02	0.73*
Q10UZ2	Tery_5020	NA	Uncharacterized protein	166.26	253.00	147.79	213.49	1.52*	1.44*
Q114K8	Tery_1809	NA	Uncharacterized protein	154.82	206.59	124.60	161.81	1.33*	1.29
Q10XY0	Tery_3852	NA	Microcompartments protein	187.73	308.14	196.16	227.07	1.64*	1.15
Q10YS6	Tery_3508	NA	Uncharacterized protein	125.32	204.69	107.48	155.47	1.63	1.44*
Q118W9	Tery_0500	NA	Conserved TM helix	142.67	220.89	114.09	125.48	1.54	1.09*
Q10ZM9	Tery_3167	NA	Serine–glyoxylate transaminase	113.17	94.17	117.38	46.33	0.83*	0.39
Q10XY1	Tery_3851	NA	Microcompartments protein	128.93	193.08	126.41	125.85	1.49*	0.99
Q10XZ9	Tery_3832	NA	Uncharacterized protein	39.05	86.54	42.16	47.82	2.21*	1.13*
Q9KJL5	Tery_4141	NA	ORF2	47.58	74.97	72.86	116.93	1.57*	1.6*
Q112X5	Tery_2220	NA	Uncharacterized protein	n.d.	25.80	25.59	79.26	n.d.*	3.09*
Q117T0	Tery_0830	NA	Uncharacterized protein	22.57	63.82	n.d.	n.d.	2.82	n.d.*
Q10UY6	Tery_5026	NA	Uncharacterized protein	27.30	38.65	25.28	74.49	1.41*	2.94*
O34106	Tery_4136	Nitrogen fixation	Nitrogenase iron protein	156.89	200.94	256.70	323.61	1.28	1.26*
O84929	Tery_4137	Nitrogen fixation	Nitrogenase protein alpha chain	52.98	17.29	104.82	111.98	0.32*	1.06
Q9KJL4	Tery_4142	Nitrogen fixation	Nitrogenase-stabilizing/protective protein NifW	35.91	46.14	64.07	106.40	1.28*	1.66*
Q111K6	Tery_2623	One-carbon metabolic process	Adenosylhomocysteinase	158.00	135.79	163.28	118.22	0.85	0.72*
Q118G0	Tery_0677	Oxidation-reduction process	Glutaredoxin	n.d.	n.d.	n.d.	28.88	n.d.*	n.d.*
Q10YF4	Tery_3658	Oxidation-reduction process	Ferredoxin–NADP reductase	90.54	54.57	97.93	69.43	0.6	0.7*
Q111H3	Tery_2657	Oxidation-reduction process	Nitroreductase	64.81	102.00	73.78	113.02	1.57*	1.53*
Q117N2	Tery_0891	Oxidation-reduction process	Putative nickel-containing superoxide dismutase (NISOD)	127.06	194.51	96.28	135.01	1.53	1.4*

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	f.c. (NF-/PS-)	f.c. (NF+/PS+)
Q10ZA9	Tery_3311	Oxidation-reduction process	Thioredoxin	32.61	48.90	48.75	80.13	1.49*	1.64*
Q11A18	Tery_0162	Oxidation-reduction process	Redoxin	287.60	434.48	266.75	327.58	1.51	1.22*
Q117I6	Tery_0945	Oxidation-reduction process	Thioredoxin	14.55	29.51	26.33	64.05	2.02*	2.43*
Q10VG2	Tery_4815	Oxidative Phosphorylation	NAD(P)H-quinone oxidoreductase subunit H	24.87	n.d.	44.03	n.d.	n.d.	n.d.*
Q10YT4	Tery_3500	Oxidative Phosphorylation	NAD(P)H-quinone oxidoreductase subunit K	25.02	34.51	44.96	94.29	1.37	2.09*
Q10XH6	Tery_4030	Oxidoreductase	Glyceraldehyde 3-phosphate dehydrogenase (NADP+)	162.74	235.72	226.53	176.58	1.44*	0.77*
Q111M3	Tery_2606	Pentose-phosphate shunt	Ribulose-phosphate 3-epimerase	14.19	19.84	39.95	87.24	1.39	2.18*
Q10WH6	Tery_4410	Photorespiration	Ribulose bisphosphate carboxylase large chain	194.06	141.97	177.17	166.08	0.73*	0.93
Q116P5	Tery_1168	Photosynthesis	Protein thf1	55.25	121.74	58.42	91.58	2.2*	1.56*
Q116S5	Tery_1137	Photosynthesis - Cyt b_6 f	Cytochrome b6	84.20	116.33	124.04	198.53	1.38	1.6*
Q10XZ6	Tery_3835	Photosynthesis - LHC	Allophycocyanin_ beta subunit	166.54	197.05	144.94	214.94	1.18*	1.48*
Q117G1	Tery_0983	Photosynthesis - LHC	Phycobilisome protein	397.06	495.93	411.67	508.52	1.24*	1.23*
Q117G0	Tery_0984	Photosynthesis - LHC	Phycobilisome protein	332.65	491.27	322.54	461.00	1.47	1.42*
Q10YG0	Tery_3650	Photosynthesis - LHC	Phycobilisome protein	256.30	358.60	273.07	377.48	1.39*	1.38*
Q10XB2	Tery_4104	Photosynthesis - LHC	Phycobilisome linker polypeptide	162.96	213.96	179.39	220.42	1.31*	1.22*
Q10UW7	Tery_5049	Photosynthesis - LHC	Phycocyanin_ beta subunit	317.00	381.82	302.75	354.00	1.2	1.16*
Q10UW8	Tery_5048	Photosynthesis - LHC	Phycocyanin_ alpha subunit	317.72	473.19	300.72	481.30	1.48*	1.6
Q117F0	Tery_0996	Photosynthesis - LHC	Phycobilisome protein	369.71	505.58	406.36	475.30	1.36*	1.16
Q111Y2	Tery_2484	Photosynthesis - LHC	Photosystem antenna protein-like	147.76	205.17	13.30	n.d.	1.38*	n.d.*
Q114Y6	Tery_1667	Photosynthesis - LHC	Photosystem antenna protein-like	258.20	295.68	20.48	n.d.	1.14*	n.d.*
Q10YG1	Tery_3649	Photosynthesis - LHC	Allophycocyanin beta subunit	360.39	469.88	373.38	507.28	1.3*	1.35*
Q10YG2	Tery_3647	Photosynthesis - LHC	Phycobilisome 7.8 kDa linker polypeptide allophycocyanin-associated core	115.65	156.16	141.98	222.34	1.35	1.56*
Q117F9	Tery_0985	Photosynthesis - LHC	Phycobilisome linker polypeptide	84.65	102.65	106.70	138.40	1.21*	1.29*
Q10XT2	Tery_3909	Photosynthesis - LHC	Phycobilisome linker polypeptide	179.45	239.37	184.41	290.73	1.33*	1.57*
Q10XB1	Tery_4105	Photosynthesis - LHC	Phycobilisome linker polypeptide	220.11	324.15	220.72	286.09	1.47*	1.29*
Q117E8	Tery_0998	Photosynthesis - LHC	Phycobilisome protein	395.27	580.36	369.71	546.28	1.46*	1.47*
Q117E7	Tery_0999	Photosynthesis - LHC	Phycobilisome linker polypeptide	196.95	225.74	207.49	258.58	1.14*	1.24*
Q10Y37	Tery_3791	Photosynthesis - PSI	Photosystem I protein PsaD	92.25	160.12	204.76	238.72	1.73	1.16*
Q116L4	Tery_1204	Photosynthesis - PSI	Photosystem I reaction center subunit XI	n.d.	n.d.	171.63	236.45	n.d.*	1.37*
Q11A00	Tery_0182/3	Photosynthesis - PSII	Photosystem Q(B) protein 1	91.18	93.63	91.43	122.66	1.02*	1.34
Q116J3	Tery_1230	Photosynthesis - PSII	Photosystem II D2 protein	151.49	181.43	139.52	193.95	1.19	1.39*

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	f.c. (NF-/PS-)	f.c. (NF+/PS+)
Q10YT0	Tery_3504	Photosynthesis - PSII	Cytochrome b559 subunit alpha	82.63	168.99	98.21	161.81	2.04*	1.64
Q10VK7	Tery_4763	Photosynthesis - PSII	Photosystem Q(B) protein 2	95.59	138.08	101.92	146.08	1.44	1.43*
Q11A16	Tery_0164	Photosynthesis - PSII	Photosystem II oxygen evolving complex protein PsbP	73.62	109.10	68.39	87.98	1.48	1.28*
Q110N3	Tery_2868	Photosynthesis - PSII	Photosystem II reaction center protein H	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.*
Q118V6	Tery_0513	Photosynthesis - PSII	Photosystem II 44 kDa subunit reaction center protein	127.39	167.36	129.44	107.05	1.31*	0.82*
Q10XS5	Tery_3917	Photosynthesis - PSII	Photosystem II manganese-stabilizing protein PsbO	179.52	297.30	152.43	265.31	1.65	1.74*
Q114S4	Tery_1740	Photosynthesis - PSII	Photosystem II 11 kD protein	67.86	84.39	51.15	156.78	1.24*	3.06*
Q10YS8	Tery_3506	Photosynthesis - PSII	Photosystem II reaction center protein L	n.d.	n.d.	n.d.	46.46	n.d.*	n.d.*
Q10ZL6	Tery_3183	Primary metabolic process	Sigma 54 modulation protein / SSU ribosomal protein S30P	71.87	125.51	93.52	90.11	1.74*	0.96
Q10WQ5	Tery_4326	Protein folding	10 kDa chaperonin	116.95	177.34	154.06	245.24	1.51*	1.59*
Q10WQ4	Tery_4327	Protein folding	60 kDa chaperonin 2	134.96	129.12	145.30	175.42	0.95	1.2*
Q10UZ9	Tery_5013	Protein phosphorylation	Putative anti-sigma regulatory factor_ serine/threonine protein kinase	34.54	n.d.	n.d.	n.d.	n.d.*	n.d.
Q10XW6	Tery_3869	Proteolysis	WD-40 repeat	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.*
Q110Q8	Tery_2842	Regulation of catlytic activity	Nitrogen regulatory protein P-II	167.98	230.26	164.09	206.86	1.37*	1.26*
Q10UZ7	Tery_5015	Regulation of transcription	Stage II sporulation protein	75.14	120.62	85.64	99.88	1.6*	1.16
Q119G2	Tery_0394	Response to oxidative stress	Peptide methionine sulfoxide reductase MsrA	31.52	65.29	49.49	87.82	2.07*	1.77*
Q10UV8	Tery_5062	Response to stress	Stress protein	96.22	115.22	83.20	126.00	1.19*	1.51*
Q10UW4	Tery_5053	Response to stress	Stress protein	19.23	27.74	21.11	50.91	1.44	2.41*
Q10W85	Tery_4503	Response to stress	UspA	99.60	136.95	113.75	158.97	1.37	1.39*
Q110B9	Tery_3000	Ribosomal protein - Large	50S ribosomal protein L5	91.69	120.10	108.60	142.50	1.3	1.31*
Q110A7	Tery_3012	Ribosomal protein - Large	50S ribosomal protein L3	142.92	176.48	151.87	185.22	1.23*	1.21*
Q110D4	Tery_2985	Ribosomal protein - Large	50S ribosomal protein L13	68.11	100.79	77.60	129.78	1.47*	1.67*
Q110C1	Tery_2998	Ribosomal protein - Large	50S ribosomal protein L6	100.47	148.30	121.75	167.62	1.47*	1.37*
Q119S7	Tery_0264	Ribosomal protein - Large	50S ribosomal protein L1	83.61	130.82	93.16	146.52	1.56	1.57*
Q119S6	Tery_0265	Ribosomal protein - Large	50S ribosomal protein L11	161.55	231.67	157.21	220.21	1.43	1.4*
Q110B2	Tery_3007	Ribosomal protein - Large	50S ribosomal protein L22	111.76	156.21	117.43	185.99	1.39*	1.58*
Q115S4	Tery_1460	Ribosomal protein - Large	50S ribosomal protein L9	98.70	152.82	108.60	154.99	1.54	1.42*
Q110C2	Tery_2997	Ribosomal protein - Large	50S ribosomal protein L18	44.70	79.80	70.07	99.04	1.78*	1.41*
Q110C4	Tery_2995	Ribosomal protein - Large	50S ribosomal protein L15	78.19	106.54	110.23	148.14	1.36*	1.34*
Q119S8	Tery_0263	Ribosomal protein - Large	50S ribosomal protein L10	72.32	103.36	94.30	170.11	1.42	1.8*
Q110B5	Tery_3004	Ribosomal protein - Large	50S ribosomal protein L29	53.56	56.55	52.88	90.14	1.05	1.7*

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	f.c. (NF-/PS-)	f.c. (NF+/PS+)
Q110A8	Tery_3011	Ribosomal protein - Large	50S ribosomal protein L4	77.03	96.54	89.13	129.13	1.25	1.44*
Q110D2	Tery_2987	Ribosomal protein - Large	50S ribosomal protein L17	78.08	113.25	95.31	144.26	1.45*	1.51*
Q10ZG7	Tery_3242	Ribosomal protein - Large	50S ribosomal protein L27	39.16	65.04	50.00	97.10	1.66	1.94*
Q114D6	Tery_1882	Ribosomal protein - Large	50S ribosomal protein L20	47.24	n.d.	46.80	78.88	n.d.	1.68*
Q110B8	Tery_3001	Ribosomal protein - Large	50S ribosomal protein L24	47.84	n.d.	61.80	n.d.	n.d.	n.d.*
Q10XV0	Tery_3887	Ribosomal protein - Small	30S ribosomal protein S6	88.89	145.04	99.76	143.29	1.63	1.43*
Q110C3	Tery_2996	Ribosomal protein - Small	30S ribosomal protein S5	112.54	173.06	116.13	184.07	1.53*	1.58
Q110B3	Tery_3006	Ribosomal protein - Small	30S ribosomal protein S3	78.30	116.40	86.19	131.81	1.48	1.52*
Q111T6	Tery_2536	Ribosomal protein - Small	30S ribosomal protein S16	53.47	115.90	69.07	116.34	2.16*	1.68*
Q118Z1	Tery_0477	Ribosomal protein - Small	30S ribosomal protein S10	92.48	144.10	100.72	141.83	1.55*	1.4*
Q116Q2	Tery_1161	Ribosomal protein - Small	30S ribosomal protein S2	121.52	185.30	116.05	165.17	1.52*	1.42*
Q110D5	Tery_2984	Ribosomal protein - Small	30S ribosomal protein S9	93.87	147.44	100.45	146.60	1.57*	1.45*
Q119M5	Tery_0324	Ribosomal protein - Small	30S ribosomal protein S4	15.54	56.31	30.48	41.84	3.62*	1.37*
Q110B6	Tery_3003	Ribosomal protein - Small	30S ribosomal protein S17	48.66	70.35	64.15	101.51	1.44	1.58*
Q118Z5	Tery_0473	Ribosomal protein - Small	30S ribosomal protein S12	44.05	n.d.	n.d.	98.97	n.d.*	n.d.*
Q11A13	Tery_0169	Ribosomal protein - Small	30S ribosomal protein S15	n.d.	n.d.	n.d.	n.d.	n.d.*	n.d.*
Q10Y49	Tery_3778	Translational termination	Ribosome-recycling factor	50.80	127.39	62.50	134.87	0	0

Table A4: Proteins observed at significantly different concentrations between either PS- and NF- or PS+ and NF+. Fold-change values are caculated with values >1 displayed in red and values <1 displayed in blue. Significant fold changes are denoted with an asterisk (*).

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	f.c. (PS-/PS+)	f.c. (NF-/NF+)
Q11AB4	Tery_0043	Arginine biosynthetic process	Arginine biosynthesis bifunctional protein ArgJ	36.04	56.19	66.01	99.91	0.54	0.56*
Q10YA8	Tery_3707	ATP catabolic process	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha	n.d.	n.d.	40.04	n.d.	n.d.*	n.d.
Q112Z1	Tery_2204	ATP catabolic process	ATP synthase subunit a	n.d.	n.d.	64.75	n.d.	n.d.	n.d.*
Q10Z38	Tery_3385	ATP catabolic process	ATP synthase subunit beta	170.79	113.56	159.38	151.70	1.07	0.74*
Q112Z3	Tery_2202	ATP catabolic process	ATP synthase subunit b′	169.69	211.87	175.66	274.23	0.96*	0.77
Q112Z4	Tery_2201	ATP catabolic process	ATP synthase subunit b	65.99	62.89	59.63	95.04	1.1	0.66*
Q10Z37	Tery_3386	ATP catabolic process	ATP synthase epsilon chain	44.50	87.15	73.58	166.30	0.6	0.52*
Q112L9	Tery_2334	ATP catabolic process	Gas vesicle protein GvpN	n.d.	110.28	n.d.	n.d.	n.d.	n.d.*
Q112Z6	Tery_2199	ATP catabolic process	ATP synthase subunit alpha	238.52	175.88	195.86	206.53	1.21*	0.85
Q10VU5	Tery_4661	Carbon metabolism	S-adenosylmethionine synthase	115.14	56.09	145.95	122.95	0.78*	0.45*
Q5GMY2	Tery_4282	Cellular iron ion homeostasis	Ferritin and Dps	n.d.	n.d.	98.99	124.89	n.d.*	n.d.*
Q115D5	Tery_1610	Chromosome condensation	Bacterial nucleoid protein Hbs	31.44	42.76	38.14	77.64	0.82	0.55*
Q113J6	Tery_2088	Cobalamin biosynthetic process	Cob(I)yrinic acid a c-diamide adenosyltransferase	n.d.	n.d.	44.13	n.d.	n.d.*	n.d.
Q116S1	Tery_1141	CTP Biosynthetic process	Nucleoside diphosphate kinase	164.60	238.41	176.08	199.23	0.93	1.19*
Q114I4	Tery_1834	D-gluconate metabolic process	6-phosphogluconate dehydrogenase decarboxylating	87.07	55.63	93.34	92.29	0.93	0.6*
Q113K0	Tery_2083	dUMP biosynthetic process	Deoxyuridine 5′-triphosphate nucleotidohydrolase	16.30	n.d.	n.d.	33.05	n.d.	n.d.*
O52659	Tery_1666	Electron transport	Flavodoxin	100.47	147.22	52.93	51.35	1.89*	2.86*
Q112L5	Tery_2339	Gas vesicle organisation	Gas vesicle synthesis GvpLGvpF	80.65	84.07	45.68	91.43	1.76*	0.91
Q10XB7	Tery_4099	Glycolysis	Fructose-bisphosphate aldolase	61.56	44.06	195.42	229.82	0.31	0.19*
Q114W6	Tery_1687	Glycolysis	Fructose-bisphosphate aldolase	343.55	389.63	19.49	n.d.	17.62*	n.d.*
Q115Y7	Tery_1380	Glycolysis	Triosephosphate isomerase	44.32	84.27	57.48	n.d.	0.77*	n.d.*
Q10W80	Tery_4509	GTP catabolic process	Elongation factor G 2	28.31	17.94	50.16	49.99	0.56	0.35*
Q119L5	Tery_0335	Heme oxidation	Heme oxygenase (Decyclizing)	n.d.	26.16	27.15	75.42	n.d.	0.34*
Q114P2	Tery_1772	Integrin-mediated signaling pathway	Na-Ca exchanger/integrin-beta4	n.d.	n.d.	n.d.	n.d.	n.d.*	n.d.
Q10WP7	Tery_4335	Isoleucine biosynthetic pathway	Ketol-acid reductoisomerase	148.23	186.96	130.52	122.43	1.13	1.52*
Q10WV1	Tery_4276	Metabolic process	Helicase-like	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.*
Q10WH8	Tery_4408	Metabolic process	Ribulose 1 5-bisphosphate carboxylase small subunit	254.17	368.84	270.77	335.77	0.93*	1.09
Q110W7	Tery_2774	Metabolic process	Peptidase M16-like	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.*
Q117R7	Tery_0847	Methionine biosynthetic process	5-methyltetrahydropteroyltriglutamate–homocysteine methyltransferase	n.d.	n.d.	26.27	34.89	n.d.	n.d.*
Q10Y07	Tery_3823	NA	Uncharacterized protein	83.90	124.63	38.65	46.10	2.17	2.7*
Q10Z45	Tery_3377	NA	Extracellular solute-binding protein family 1	243.11	285.62	102.08	132.82	2.38	2.15*

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	F.C. (PS-/PS+)	F.C. (NF-/NF+)
Q10Z50	Tery_3372	NA	Beta-Ig-H3/fasciclin	70.56	115.83	38.71	63.97	1.82	1.81*
Q113S7	Tery_1998	NA	Extracellular solute-binding protein family 3	165.41	141.56	111.66	61.54	1.48	2.29*
Q10Y04	Tery_3826	NA	Dimethylaniline monooxygenase (N-oxide forming)	43.98	31.94	10.95	21.54	4.01*	1.48*
Q10Y05	Tery_3825	NA	Uncharacterized protein	72.26	69.73	19.97	22.09	3.61	3.15*
Q10Y15	Tery_3815	NA	Uncharacterized protein	109.21	149.82	42.59	107.69	2.56	1.39*
Q112G8	Tery_2388	NA	PHA accumulation regulator DNA-binding-like	208.07	238.19	n.d.	99.82	n.d.	2.38*
Q10UX6	Tery_5038	NA	1-Cys peroxiredoxin	149.20	231.43	152.15	185.15	0.98*	1.24*
Q114E5	Tery_1873	NA	SPFH domain Band 7 family protein	161.18	246.23	141.93	187.31	1.13	1.31*
Q10W05	Tery_4593	NA	Uncharacterized protein	131.46	134.18	111.14	81.96	1.18*	1.63*
Q10UZ2	Tery_5020	NA	Uncharacterized protein	166.26	252.99	147.78	213.49	1.12	1.18*
Q114K8	Tery_1809	NA	Uncharacterized protein	154.82	206.59	124.60	161.81	1.24*	1.27*
Q10XY0	Tery_3852	NA	Microcompartments protein	187.72	308.13	196.15	227.07	0.95*	1.35*
Q10YS6	Tery_3508	NA	Uncharacterized protein	125.31	204.69	107.48	155.46	1.16	1.31*
Q118W9	Tery_0500	NA	Conserved TM helix	142.66	220.89	114.08	125.47	1.25*	1.76*
Q10ZM9	Tery_3167	NA	Serine–glyoxylate transaminase	113.17	94.17	117.38	46.33	0.96*	2.03*
Q10XY1	Tery_3851	NA	Microcompartments protein	128.93	193.07	126.41	125.84	1.01*	1.53*
Q10XZ9	Tery_3832	NA	Uncharacterized protein	39.04	86.53	42.15	47.81	0.92*	1.8
Q9KJL5	Tery_4141	NA	ORF2	47.57	74.96	72.85	116.93	0.65	0.64*
Q112X5	Tery_2220	NA	Uncharacterized protein	n.d.	25.80	25.58	79.25	n.d.	0.32*
Q10X55	Tery_4163	NA	Uncharacterized protein	24.91	36.84	8.20	n.d.	3.03	n.d.*
Q117T0	Tery_0830	NA	Uncharacterized protein	22.56	63.81	n.d.	n.d.	n.d.	n.d.*
Q10UY6	Tery_5026	NA	Uncharacterized protein	27.30	38.65	25.27	74.48	1.08	0.51*
Q114P9	Tery_1765	NA	Uncharacterized protein	10.37	n.d.	61.50	52.70	0.16	n.d.*
Q117A6	Tery_1046	NA	Uncharacterized protein	20.94	32.12	n.d.	n.d.	n.d.	n.d.*
Q10XX0	Tery_3865	NA	Uncharacterized protein	n.d.	n.d.	26.32	38.17	n.d.	n.d.*
Q114J2	Tery_1825	NA	Uncharacterized protein	n.d.	40.33	n.d.	n.d.	n.d.	n.d.*
O34106	Tery_4136	Nitrogen fixation	Nitrogenase iron protein	156.88	200.93	256.70	323.61	0.61*	0.62*
O84930	Tery_4138	Nitrogen fixation	MoFe protein of nitrogenase beta subunit	33.86	18.29	88.56	111.54	0.38	0.16*
O84929	Tery_4137	Nitrogen fixation	Nitrogenase protein alpha chain	52.98	17.29	104.82	111.97	0.5*	0.15*
Q9KJL4	Tery_4142	Nitrogen fixation	Nitrogenase-stabilizing/protective protein NifW	35.90	46.14	64.06	106.39	0.56*	0.43*
Q9KJL6	Tery_4140	Nitrogen fixation	NifX	n.d.	n.d.	35.41	n.d.	n.d.*	n.d.*

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	F.C. (PS-/PS+)	F.C. (NF-/NF+)
Q117N2	Tery_0891	Oxidation-reduction process	Putative nickel-containing superoxide dismutase (NISOD)	127.06	194.50	96.28	135.01	1.31	1.44*
Q10ZA9	Tery_3311	Oxidation-reduction process	Thioredoxin	32.61	48.89	48.75	80.13	0.66*	0.61*
Q11A18	Tery_0162	Oxidation-reduction process	Redoxin	287.60	434.48	266.75	327.57	1.07	1.32*
Q117I6	Tery_0945	Oxidation-reduction process	Thioredoxin	14.54	29.50	26.32	64.04	0.55	0.46*
Q115G5	Tery_1579	Oxidative Phosphorylation	NAD(P)H-quinone oxidoreductase subunit 1	n.d.	n.d.	45.56	n.d.	n.d.*	n.d.
Q10YT4	Tery_3500	Oxidative Phosphorylation	NAD(P)H-quinone oxidoreductase subunit K	25.01	34.50	44.96	94.29	0.55	0.36*
Q10XH6	Tery_4030	Oxidoreductase	Glyceraldehyde 3-phosphate dehydrogenase (NADP+)	162.73	235.71	226.53	176.57	0.71*	1.33
Q10Y11	Tery_3819	Oxidoreductase	Beta-ketoacyl synthase	42.21	14.74	n.d.	n.d.	n.d.*	n.d.*
Q111M3	Tery_2606	Pentose-phosphate shunt	Ribulose-phosphate 3-epimerase	14.19	19.83	39.94	87.24	0.35	0.22*
Q116S5	Tery_1137	Photosynthesis - Cyt b_6 f	Cytochrome b6	84.20	116.33	124.04	198.52	0.67*	0.58*
Q116S6	Tery_1136	Photosynthesis - Cyt b_6 f	Cytochrome b6-f complex subunit 4	n.d.	n.d.	65.28	78.54	n.d.*	n.d.*
Q117F0	Tery_0996	Photosynthesis - LHC	Phycobilisome protein	369.70	505.57	406.36	475.30	0.9	1.06*
Q111Y0	Tery_2486	Photosynthesis - LHC	Phycobilisome linker polypeptide	87.37	109.09	18.74	26.46	4.66*	4.12
Q111Y2	Tery_2484	Photosynthesis - LHC	Photosystem antenna protein-like	147.76	205.16	13.30	n.d.	11.1	n.d.*
Q114Y6	Tery_1667	Photosynthesis - LHC	Photosystem antenna protein-like	258.20	295.68	20.47	n.d.	12.6	n.d.*
Q111Y1	Tery_2485	Photosynthesis - LHC	Photosystem antenna protein-like	79.21	106.05	n.d.	n.d.	n.d.*	n.d.*
Q10YG1	Tery_3649	Photosynthesis - LHC	Allophycocyanin beta subunit	360.39	469.87	373.38	507.28	0.96	0.92*
Q10YG2	Tery_3647	Photosynthesis - LHC	Phycobilisome 7.8 kDa linker polypeptide allophycocyanin-associated core	115.65	156.16	141.98	222.33	0.81*	0.7*
Q117F9	Tery_0985	Photosynthesis - LHC	Phycobilisome linker polypeptide	84.64	102.64	106.70	138.40	0.79*	0.74*
Q10XT2	Tery_3909	Photosynthesis - LHC	Phycobilisome linker polypeptide	179.44	239.36	184.40	290.73	0.97*	0.82*
Q10XB1	Tery_4105	Photosynthesis - LHC	Phycobilisome linker polypeptide	220.11	324.15	220.72	286.09	0.99*	1.13*
Q117E8	Tery_0998	Photosynthesis - LHC	Phycobilisome protein	395.26	580.35	369.71	546.27	1.06	1.06*
Q117E7	Tery_0999	Photosynthesis - LHC	Phycobilisome linker polypeptide	196.95	225.74	207.49	258.57	0.94	0.87*
Q114L7	Tery_1798	Photosynthesis - LHC	Apocytochrome f	36.93	n.d.	65.28	82.12	0.56	n.d.*
Q111Y3	Tery_2483	Photosynthesis - LHC	Photosystem antenna protein-like	115.71	142.21	n.d.	n.d.	n.d.	n.d.*
Q10VT8	Tery_4668	Photosynthesis - PSI	Photosystem I P700 chlorophyll a apoprotein A2	57.85	54.39	123.21	129.21	0.46*	0.42*
Q10Y37	Tery_3791	Photosynthesis - PSI	Photosystem I protein PsaD	92.24	160.11	204.75	238.71	0.45*	0.67*
Q115B4	Tery_1634	Photosynthesis - PSI	Photosystem I reaction center protein PsaF subunit III	55.48	65.18	128.33	135.96	0.43*	0.47*
Q10VT7	Tery_4669	Photosynthesis - PSI	Photosystem I P700 chlorophyll a apoprotein A1	28.71	33.86	76.04	78.05	0.37*	0.43*
Q119B1	Tery_0454	Photosynthesis - PSI	Photosystem I iron-sulfur center	18.15	22.18	104.44	129.23	0.17*	0.17*
Q117D3	Tery_1014	Photosynthesis - PSI	Photosystem I reaction center subunit IV	69.43	68.94	129.78	145.37	0.53*	0.47*

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	F.C. (PS-/PS+)	F.C. (NF-/NF+)
Q116L4	Tery_1204	Photosynthesis - PSI	Photosystem I reaction center subunit XI	n.d.	n.d.	171.62	236.44	n.d.*	n.d.*
Q110N3	Tery_2868	Photosynthesis - PSII	Photosystem II reaction center protein H	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.*
Q118V6	Tery_0513	Photosynthesis - PSII	Photosystem II 44 kDa subunit reaction center protein	127.38	167.35	129.44	107.05	0.98	1.56*
Q10XS5	Tery_3917	Photosynthesis - PSII	Photosystem II manganese-stabilizing protein PsbO	179.51	297.29	152.42	265.31	1.17	1.12*
Q114S4	Tery_1740	Photosynthesis - PSII	Photosystem II 11 kD protein	67.86	84.38	51.14	156.77	1.32*	0.53
Q110N4	Tery_2867	Photosynthesis - PSII	Protein PsbN	n.d.	n.d.	n.d.	32.38	n.d.	n.d.*
Q10YS8	Tery_3506	Photosynthesis - PSII	Photosystem II reaction center protein L	n.d.	n.d.	n.d.	46.46	n.d.	n.d.*
Q10ZL6	Tery_3183	Primary metabolic process	Sigma 54 modulation protein / SSU ribosomal protein S30P	71.86	125.51	93.52	90.10	0.76	1.39*
Q10WQ5	Tery_4326	Protein folding	10 kDa chaperonin	116.95	177.34	154.06	245.24	0.75	0.72*
Q10WQ4	Tery_4327	Protein folding	60 kDa chaperonin 2	134.96	129.12	145.30	175.42	0.92*	0.73*
Q119S1	Tery_0270	Protein folding	60 kDa chaperonin 1	124.29	101.49	146.25	164.71	0.84	0.61*
Q10VW7	Tery_4635	Protein import	Protein translocase subunit SecA	4.95	n.d.	12.90	36.44	0.38	n.d.*
Q112J2	Tery_2362	Proteolysis	Peptidase S8 and S53 subtilisin kexin sedolisin	195.24	173.65	163.34	177.56	1.19*	0.97
Q10XW6	Tery_3869	Proteolysis	WD-40 repeat	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.*
Q10Y67	Tery_3755	Proteolysis	ATP-dependent zinc metalloprotease FtsH	26.88	18.19	35.65	50.03	0.75*	0.36
Q116P6	Tery_1167	Regulation of transcription	Stage II sporulation protein	11.47	n.d.	22.13	43.42	0.51	n.d.*
Q110C4	Tery_2995	Ribosomal protein - Large	50S ribosomal protein L15	78.19	106.54	110.22	148.13	0.7	0.71*
Q10X90	Tery_4126	Ribosomal protein - Large	50S ribosomal protein L28	n.d.	n.d.	37.85	n.d.	n.d.	n.d.*
Q119S8	Tery_0263	Ribosomal protein - Large	50S ribosomal protein L10	72.31	103.35	94.29	170.11	0.76	0.6*
Q110B5	Tery_3004	Ribosomal protein - Large	50S ribosomal protein L29	53.55	56.55	52.87	90.14	1.01*	0.62*
Q110A8	Tery_3011	Ribosomal protein - Large	50S ribosomal protein L4	77.03	96.53	89.12	129.12	0.86	0.74*
Q110D2	Tery_2987	Ribosomal protein - Large	50S ribosomal protein L17	78.08	113.24	95.30	144.26	0.81	0.78*
Q10ZG7	Tery_3242	Ribosomal protein - Large	50S ribosomal protein L27	39.16	65.04	49.99	97.10	0.78	0.66*
Q114D6	Tery_1882	Ribosomal protein - Large	50S ribosomal protein L20	47.23	n.d.	46.79	78.88	1	n.d.*
Q110B8	Tery_3001	Ribosomal protein - Large	50S ribosomal protein L24	47.83	n.d.	61.80	n.d.	0.77	n.d.*
Q10Y90	Tery_3728	Ribosomal protein - Large	50S ribosomal protein L33	n.d.	n.d.	n.d.	n.d.	n.d.*	n.d.
Q110B6	Tery_3003	Ribosomal protein - Small	30S ribosomal protein S17	48.65	70.35	64.15	101.50	0.75	0.69*
Q118Z5	Tery_0473	Ribosomal protein - Small	30S ribosomal protein S12	44.04	n.d.	n.d.	98.96	n.d.	n.d.*
Q11A13	Tery_0169	Ribosomal protein - Small	30S ribosomal protein S15	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.*
Q10VK9	Tery_4761	Transmembrane transport	MscS Mechanosensitive ion channel	64.52	97.43	61.63	66.13	1.04	1.47*
Q10V17	Tery_4993	Transmembrane transport	Phosphonate ABC transporter periplasmic phosphonate-binding protein	35.25	42.98	14.68	n.d.	2.4	n.d.*

Uniprot ID	Cyanobase ID	Biological Process	Protein Description	PS-	NF-	PS+	NF+	F.C. (PS-/PS+)	F.C. (NF-/NF+)

Table A5: Proteins observed to be differentially abundant between iron deplete (PS- or NF-) and replete condition (PS+ or NF+). Fold-change values are displayed with values >1 in red and values <1 in blue. Significant fold changes are denoted with an asterisk (*).

Uniprot ID	Protein Description	PS+	NF+	PS-	NF-	Confidence (%)	Coverage (%)	PDB Entry	Metals
O34106	Nitrogenase iron protein	256.70	323.61	156.89	200.94	100	75	2afh	4 Fe,
O84929	Nitrogenase protein alpha chain	104.82	111.98	52.98	17.29	100	69	1qh8	1 Mg, 7 Fe, 1 Mo,
O84930	MoFe protein of nitrogenase beta subunit	88.56	111.54	33.87	18.30	100	55	1m1n	2 Ca, 8 Fe,
Q10UU3	Transketolase_ central region	70.68	74.39	68.93	80.13	100	41	1ni4	1 K,
Q10UV4	Pyruvate dehydrogenase (Lipoamide)	58.90	31.38	70.32	36.50	100	24	1qs0	1 Mg,
Q10UV8	Stress protein	83.20	126.00	96.22	115.22	100	53	3ibz	2 Ca,
Q10UW4	Stress protein	21.11	50.91	19.23	27.74	100	44	3ibz	2 Ca,
Q10UZ9	Putative anti-sigma regulatory factor_ serine/threonine protein kinase	n.d.	n.d.	34.54	n.d.	99.9	19	1th8	1 Mg,
Q10V72	Two component transcriptional regulator_ winged helix family	26.86	29.21	22.23	39.08	100	40	1ys7	1 Mg,
Q10VI4	Uncharacterized protein	48.04	58.42	42.97	49.67	43.7	70	3cx5	2 Fe,
Q10VI5	Peptidoglycan-binding domain 1	34.74	56.28	#VALUE!	50.85	99.8	44	3bkh	1 Ni,
Q10VK7	Photosystem Q(B) protein 2	101.92	146.08	95.59	138.08	100	90	2axt	4 Mg, 4 Mn, 1 Fe,
Q10VM5	Cysteine synthase	22.40	24.33	23.97	24.42	100	40	4l3v	1 Fe,
Q10VT7	Photosystem I P700 chlorophyll a apoprotein A1	76.04	78.06	28.72	33.87	100	88	1jbo	46 Mg, 1 Ca, 4 Fe,
Q10VT8	Photosystem I P700 chlorophyll a apoprotein A2	123.22	129.21	57.86	54.40	100	89	1jbo	41 Mg,
Q10VU0	Photosystem antenna protein-like	129.33	119.50	141.22	113.35	100	82	2axt	32 Mg,
Q10VU5	S-adenosylmethionine synthase	145.95	122.96	115.14	56.10	100	60	1rg9	2 Mg, 1 K,
Q10VU6	Uncharacterized protein	26.72	18.08	12.22	n.d.	100	48	3c2q	1 Ni,
Q10VX2	NAD-dependent epimerase/dehydratase	29.57	21.51	29.75	30.17	100	24	1z45	2 Na,
Q10VY5	GTP cyclohydrolase 1	32.03	55.04	36.54	54.60	100	54	1wpl	1 Zn,
Q10VZ2	Delta-aminolevulinic acid dehydratase	39.16	32.30	36.15	54.85	100	63	2c1h	1 Mg,
Q10W27	Uncharacterized protein	n.d.	n.d.	21.11	32.45	100	17	3cu2	1 Ni, 1 Ca,
Q10W62	Rhodanese-like	36.72	28.38	38.70	33.35	100	24	3ipp	1 Na,
Q10W80	Elongation factor G 2	50.17	50.00	28.32	17.94	100	47	2xex	1 K,
Q10W97	Two component transcriptional regulator_ LuxR family	16.68	25.55	19.90	21.76	100	27	1zn2	1 Mg,
Q10WM6	FeS assembly ATPase SufC	15.89	30.18	21.59	34.13	100	52	2d2f	2 Mg,
Q10WP7	Ketol-acid reductoisomerase	130.52	122.43	148.23	186.97	100	54	4kqx	6 Mg,
Q10WQ4	60 kDa chaperonin 2	145.30	175.42	134.96	129.12	100	58	1kp8	1 Mg, 1 K,
Q10WQ5	10 kDa chaperonin	154.06	245.24	116.95	177.34	100	49	1p3h	1 Ca,
Q10X42	Uncharacterized protein	31.22	37.22	20.21	31.03	100	77	3jsr	1 K,
Q10X81	Nitrogen fixation protein NifU	34.43	39.75	15.26	32.64	100	45	2z7e	2 Fe,

Uniprot ID	Protein Description	PS+	NF+	PS-	NF-	Confidence (%)	Coverage (%)	PDB Entry	Metals
Q10XA2	Mo-dependent nitrogenase-like	13.81	11.95	11.98	17.98	99.9	19	2ou3	2 Zn,
Q10XB7	Fructose-bisphosphate aldolase	195.42	229.83	61.56	44.07	100	48	1rvg	1 Na, 1 Co, 1 Yt,
Q10XD3	Phage shock protein A_ PspA	14.96	23.02	35.42	22.98	97.2	10	3ghg	1 Ca,
Q10XD4	Phage shock protein A_ PspA	30.21	54.98	25.65	38.47	96	9	3ghg	1 Ca,
Q10XI2	Glucose-6-phosphate isomerase	55.36	49.13	68.81	44.99	100	26	3hjb	4 Ca,
Q10XQ2	Periplasmic binding protein	19.13	23.64	20.22	26.50	100	19	2r79	1 Fe,
Q10XX0	Uncharacterized protein	26.33	38.18	n.d.	n.d.	67.8	23	1jbo	1 Mg,
Q10XY3	Carbonate dehydratase	118.18	101.43	115.88	81.57	100	73	3kwc	1 Zn,
Q10XZ7	Glutamine synthetase	111.29	75.17	129.78	50.63	100	77	3ngo	3 Mn,
Q10YA8	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha / biotin carboxylase	40.04	n.d.	n.d.	n.d.	100	46	3bg5	1 Mn,
Q10YF1	FAD-dependent pyridine nucleotide-disulphide oxidoreductase	10.37	21.18	17.36	n.d.	100	29	4g6g	6 Mg,
Q10YF3	Uncharacterized protein	60.00	78.09	51.00	83.60	85.5	26	2w69	1 Mn,
Q10YI1	4-hydroxy-tetrahydrodipicolinate reductase	33.86	41.23	37.02	58.91	100	28	3ijp	1 Na,
Q10YM6	Exodeoxyribonuclease III	13.24	13.99	17.24	20.58	100	44	2jc4	3 Mg,
Q10YM7	NmrA-like	34.33	37.94	40.62	51.41	100	19	2zcu	1 Cu,
Q10YP5	Uncharacterized protein	31.67	77.81	22.00	61.94	57.3	18	2c42	1 Mg, 1 Ca, 12 Fe,
Q10YS5	Uncharacterized protein	23.94	#VALUE!	16.44	26.71	100	17	2jfs	3 Mn, 1 As,
Q10YS6	Uncharacterized protein	107.48	155.47	125.32	204.69	100	20	3fcu	1 Mg, 2 Ca, 1 As,
Q10YT2	Rubredoxin	17.29	28.77	n.d.	n.d.	100	46	1dx8	1 Zn,
Q10YT4	NAD(P)H-quinone oxidoreductase subunit K	44.96	94.29	25.02	34.51	100	49	2fug	2 Fe,
Q10YT5	NAD(P)H-quinone oxidoreductase subunit J	22.77	34.40	10.62	12.92	100	34	2fug	4 Fe,
Q10Z40	Site-determining protein	12.84	24.86	15.57	23.50	100	45	3q9l	1 Mg,
Q10Z46	Phosphoglycerate kinase	164.47	97.67	165.14	97.08	100	61	1php	1 Mg,
Q10Z47	UspA	30.11	44.49	35.80	44.57	99.9	24	1mjh	1 Mn,
Q10ZB2	IMP dehydrogenase subunit	21.94	n.d.	14.69	n.d.	100	28	1pvn	3 K,
Q10ZB9	Protein serine/threonine phosphatases	9.84	15.85	10.45	16.13	100	34	1txo	3 Mn,
Q10ZC9	LSU ribosomal protein L25P	70.15	88.31	66.57	87.53	100	24	1dfu	4 Mg,
Q10ZD0	Adenylosuccinate synthetase	30.95	n.d.	27.00	n.d.	100	49	1qf5	1 Mg,
Q10ZE4	YUP8H12.25 ((Arabidopsis thaliana))-type protein. Metallo peptidase. MEROPS family M50B	21.44	25.92	21.84	33.44	99.8	39	3b4r	1 Zn,
Q10ZR8	Uncharacterized protein	20.10	n.d.	7.31	n.d.	100	16	2og5	1 Na,
Q10ZT2	Uncharacterized protein	16.57	19.78	13.08	16.71	42.1	25	1m36	1 Zn,

Uniprot ID	Protein Description	PS+	NF+	PS-	NF-	Confidence (%)	Coverage (%)	PDB Entry	Metals
Q110A6	NAD(P)H-quinone oxidoreductase subunit N	35.23	30.70	n.d.	n.d.	71.1	26	1xpj	1 Hg,
Q110C6	Adenylate kinase	28.83	38.60	24.16	34.72	99.9	42	2cdn	1 Mg,
Q110C8	50S ribosomal protein L36	n.d.	n.d.	24.85	n.d.	99.9	84	1dfe	1 Zn,
Q110E0	Caffeoyl-CoA O-methyltransferase	46.92	69.53	39.05	55.72	100	36	3tr6	1 Ni,
Q110K9	Uncharacterized protein	18.87	34.46	17.05	17.32	72.2	12	4ksa	1 Mg,
Q110N4	Protein PsbN	n.d.	32.38	n.d.	n.d.	27.1	32	3kzi	1 Fe,
Q110Q8	Nitrogen regulatory protein P-II	164.09	206.86	167.98	230.26	100	88	1qy7	1 Ni,
Q110S7	HpcH/HpaI aldolase	12.74	23.53	14.75	32.65	100	26	2vwt	1 Mg,
Q110V4	Enolase	39.20	14.17	51.99	13.35	100	58	1iyx	1 Mg,
Q110V6	Ferritin and Dps	64.13	106.18	59.90	109.33	100	17	1vlg	1 Fe,
Q110Z7	UPF0234 protein Tery_2743	38.34	73.08	43.39	60.25	100	42	1ino	1 Na, 2 Hg,
Q111H3	Nitroreductase	73.78	113.02	64.81	102.00	100	23	1zch	1 Ca,
Q111R2	Plastocyanin	23.36	32.41	41.40	57.55	99.9	32	7pcy	1 Cu,
Q111Y1	Photosystem antenna protein-like	n.d.	n.d.	79.22	106.05	100	62	2axt	13 Mg,
Q111Y2	Photosystem antenna protein-like	13.3	n.d.	147.76	205.17	100	64	2axt	26 Mg,
Q111Y3	Photosystem antenna protein-like	n.d.	n.d.	115.71	142.21	100	61	2axt	26 Mg,
Q112C2	ATPase AAA-2	22.46	25.76	17.2	12.25	100	52	1qvr	9 Pt,
Q112I6	von Willebrand factor_ type A	n.d.	10.39	n.d.	18.90	99.9	19	4fx5	1 Na,
Q112M1	Lipoate-protein ligase	25.00	49.21	26.45	30.20	38.6	21	1h32	1 Fe,
Q112M3	Gas vesicle protein GvpN	65.81	69.46	71.13	73.55	99.9	17	4akg	1 Mg,
Q112Z6	ATP synthase subunit alpha	195.86	206.54	238.52	175.88	100	60	4b2q	2 Mg,
Q113J6	Cob(I)yrinic acid a_c-diamide adenosyltransferase	44.13	n.d.	n.d.	n.d.	100	27	1g64	1 Mg,
Q113K0	Deoxyuridine 5′-triphosphate nucleotidohydrolase	n.d.	33.05	16.3	n.d.	100	45	3ehw	2 Mg,
Q113M1	FAD-dependent pyridine nucleotide-disulphide oxidoreductase	22.27	n.d.	13.51	n.d.	100	29	4g6g	6 Mg,
Q113P0	Aminomethyltransferase	25.84	21.78	26.25	n.d.	100	29	1pj6	1 Na,
Q114G8	Acetolactate synthase_ small subunit	45.15	60.34	42.94	65.48	100	45	2pc6	1 Mg,
Q114L6	#VALUE!	42.50	n.d.	16.94	n.d.	100	74	2e76	2 Fe,
Q114L7	Apocytochrome f	65.29	82.13	36.93	n.d.	100	61	1q90	1 Fe,
Q114M1	Uncharacterized protein	84.27	94.71	81.68	97.63	100	45	3ls1	2 Zn,
Q114N9	Methyltransferase type 12	19.63	18.50	18.69	n.d.	99.9	15	4krg	1 Na,
Q114P9	Uncharacterized protein	61.51	52.71	10.37	n.d.	81.3	24	3jtz	1 Na,

Uniprot ID	Protein Description	PS+	NF+	PS-	NF-	Confidence (%)	Coverage (%)	PDB Entry	Metals
Q115B4	Photosystem I reaction center protein PsaF_ subunit III	128.34	135.97	55.48	65.19	100	64	1jbo	1 Mg,
Q115B6	Nitrilase/cyanide hydratase and apolipoprotein N-acyltransferase	21.08	35.98	11.33	23.53	100	38	1ems	1 Na, 2 Hg,
Q115B7	Sodium/hydrogen exchanger	24.70	36.57	26.64	37.06	100	37	4bwz	2 Zn,
Q115C4	3′-5′ exonuclease	19.34	34.27	14.31	28.73	100	25	1yt3	5 Zn,
Q115D5	Bacterial nucleoid protein Hbs	38.14	77.64	31.44	42.77	99.9	33	2iif	5 Mn,
Q115G6	NAD(P)H-quinone oxidoreductase subunit I	34.61	36.72	n.d.	n.d.	100	30	4hea	8 Fe,
Q115L6	Uncharacterized protein	31.05	34.08	19.99	37.09	94.8	20	3c37	1 Zn,
Q115M5	ATP-dependent Clp protease proteolytic subunit	43.55	54.01	42.73	67.72	100	55	3q7h	2 Ca,
Q116B5	dTDP-4-dehydrorhamnose reductase	20.05	26.31	21.92	33.40	100	41	1n2s	1 Mg,
Q116I0	ATP-dependent Clp protease proteolytic subunit	84.38	94.64	67.70	96.86	100	54	3q7h	2 Ca,
Q116I1	ATP-dependent Clp protease proteolytic subunit	57.31	60.08	50.45	72.69	100	44	3q7h	2 Ca,
Q116J3	Photosystem II D2 protein	139.52	193.95	151.49	181.43	100	88	2axt	4 Mg,
Q116J4	UPF0367 protein Tery_1229	68.04	97.69	64.74	78.51	53.3	16	1ybh	1 Mg,
Q116L2	Probable 2-phosphosulfolactate phosphatase	29.23	24.60	18.94	26.95	100	30	1vro	1 Mg,
Q116L4	Photosystem I reaction center subunit XI	171.63	236.45	n.d.	n.d.	100	78	1jbo	3 Mg, 1 Ca,
Q116S1	Nucleoside diphosphate kinase	176.08	199.23	164.61	238.41	100	58	1ucn	1 Ca,
Q116S5	Cytochrome b6	124.04	198.53	84.20	116.33	100	34	2qjk	2 Fe,
Q116S6	Cytochrome b6-f complex subunit 4	65.29	78.54	n.d.	n.d.	100	83	2e74	1 Mg, 1 Cd,
Q116T4	Bifunctional purine biosynthesis protein PurH	17.65	n.d.	19.70	n.d.	100	43	4a10	1 K,
Q116V2	UBA/THIF-type NAD/FAD binding fold	25.35	n.d.	36.69	n.d.	100	40	1jw9	1 Zn,
Q117B1	3-isopropylmalate dehydratase_ small subunit	20.90	25.03	11.83	12.54	100	36	2pkp	1 Zn,
Q117E3	NmrA-like	17.6	19.84	n.d.	n.d.	100	15	1z45	2 Na,
Q117H2	GTP cyclohydrolase 1	15.51	31.54	8.89	n.d.	100	58	1wpl	1 Zn,
Q117I0	Two component transcriptional regulator_ winged helix family	59.52	88.18	54.30	100.15	100	42	2oqr	5 La,
Q117I5	D-fructose 1_6-bisphosphatase class 2/sedoheptulose 1_7-bisphosphatase	177.45	210.20	188.24	185.86	100	84	3roj	2 Mg,
Q117I6	Thioredoxin	26.33	64.05	14.55	29.51	100	37	2trx	1 Cu,
Q117N2	Putative nickel-containing superoxide dismutase (NISOD)	96.28	135.01	127.06	194.51	100	43	1t6u	1 Ni,
Q117T0	Uncharacterized protein	n.d.	n.d.	22.57	63.82	94.6	19	3ls1	2 Zn,
Q117W8	Cob(I)yrinic acid a_c-diamide adenosyltransferase	28.48	28.38	27.99	35.48	100	42	1g64	1 Mg,
Q118C1	Transketolase-like	28.08	n.d.	22.41	n.d.	100	22	3mos	1 Ca, 1 Na,
Q118D3	Stress protein	9.29	19.46	10.67	19.38	100	54	3ibz	2 Ca,

Uniprot ID	Protein Description	PS+	NF+	PS-	NF-	Confidence (%)	Coverage (%)	PDB Entry	Metals
Q118F5	Fructose-1_6-bisphosphatase class 1	34.27	20.82	33.51	40.37	100	41	1nuw	4 Mg,
Q118P5	ATP-dependent Clp protease proteolytic subunit	46.63	58.89	46.51	52.29	100	58	3q7h	2 Ca,
Q118V6	Photosystem II 44 kDa subunit reaction center protein	129.44	107.05	127.39	167.36	100	84	2axt	26 Mg,
Q118X7	NAD-dependent epimerase/dehydratase	43.72	34.93	51.71	54.78	100	22	1z45	2 Na,
Q118X8	GDP-mannose 4_6-dehydratase	25.71	18.71	22.75	21.13	100	54	3ibz	2 Ca,
Q118Z5	30S ribosomal protein S12	n.d.	98.97	44.05	n.d.	100	77	1i94	3 Mg,
Q119B1	Photosystem I iron-sulfur center	104.44	129.23	18.15	22.19	100	94	1jbo	8 Fe,
Q119B5	Transketolase	90.96	98.22	95.21	72.34	100	48	3rim	1 Mg,
Q119F1	HAD-superfamily hydrolase_ subfamily IA_ variant 3	24.13	40.25	14.22	35.54	100	17	3mc1	1 Na,
Q119F6	Methionine aminopeptidase	23.31	28.92	16.35	21.39	100	44	4fuk	2 Zn,
Q119L5	Heme oxygenase (Decyclizing)	27.16	75.43	n.d.	26.16	100	50	1wov	1 Fe,
Q119M5	30S ribosomal protein S4	30.48	41.84	15.54	56.31	100	47	2uub	2 Zn,
Q119P9	Peptidase M1_ membrane alanine aminopeptidase	20.04	n.d.	8.47	n.d.	100	20	3b37	1 Zn, 2 Na,
Q119R0	Transaldolase	89.53	72.62	93.54	60.58	100	55	3hjz	6 Na,
Q119S1	60 kDa chaperonin 1	146.26	164.71	124.29	101.50	100	58	1kp8	1 Mg, 1 K,
Q119T2	Asparaginase	14.59	20.54	9.9	19.51	100	34	4gdu	1 Na,
Q119Y0	Short-chain dehydrogenase/reductase SDR	19.2	21.31	26.82	41.15	100	21	3rih	1 K,
Q11A00	Photosystem Q(B) protein 1	91.43	122.66	91.18	93.63	100	90	2axt	4 Mg, 4 Mn, 1 Fe,
Q11A96	Probable cytosol aminopeptidase	23.37	19.01	26.86	11.6	100	36	3h8g	1 Zn, 1 K, 1 Mn,
Q11AA1	Oxidoreductase_ molybdopterin binding	13.97	15.6	16.73	30.09	100	50	1xdy	1 W,
Q5GMY2	Ferritin and Dps	98.99	124.89	n.d.	n.d.	100	67	2vxx	2 Zn, 1 Fe,

Table A6: Proteins identified as containing metal-biding cofactors. Identification was acheived using the PHYRE-2 search engine described in the text. PHYRE-2 confidence scores (%) and sequence coverage (%) are shown alongside identified PDB structure and finally per-protein metal composition.

Metal Ion	Uniprot ID	Protein Description	PDB Entry	PS- vs NF-	PS- vs PS+	NF- vs NF+	PS+ vs NF+
Fe	O34106	Nitrogenase iron protein	2afh	> 0.9999	0.0002	< 0.0001	> 0.9999
Fe	O84929	Nitrogenase protein alpha chain	1qh8	< 0.0001	> 0.9999	< 0.0001	0.0829
Fe	O84930	MoFe protein of nitrogenase beta subunit	1m1n	0.0002	0.2298	< 0.0001	> 0.9999
Fe	Q10VI4	Uncharacterized protein	3cx5	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q10VK7	Photosystem Q(B) protein 2	2axt	> 0.9999	0.8554	> 0.9999	0.2777
Fe	Q10VM5	Cysteine synthase	4l3v	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q10VT7	Photosystem I P700 chlorophyll a apoprotein A1	1jbo	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q10X81	Nitrogen fixation protein NifU	2z7e	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q10XQ2	Periplasmic binding protein	2r79	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q10YP5	Uncharacterized protein	2c42	< 0.0001	0.0722	< 0.0001	< 0.0001
Fe	Q10YT4	NAD(P)H-quinone oxidoreductase subunit K	2fug	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q10YT5	NAD(P)H-quinone oxidoreductase subunit J	2fug	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q110N4	Protein PsbN	3kzi	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q110V6	Ferritin and Dps	1vlg	> 0.9999	> 0.9999	0.6849	> 0.9999
Fe	Q112M1	Lipoate-protein ligase	1h32	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q114L6	Cytochrome b6-f complex iron-sulfur subunit	2e76	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q114L7	Apocytochrome f	1q90	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q115G6	NAD(P)H-quinone oxidoreductase subunit I	4hea	> 0.9999	0.0335	> 0.9999	> 0.9999
Fe	Q116S5	Cytochrome b6	2qjk	> 0.9999	0.6948	0.9104	> 0.9999
Fe	Q119B1	Photosystem I iron-sulfur center	1jbo	> 0.9999	< 0.0001	< 0.0001	> 0.9999
Fe	Q119L5	Heme oxygenase (Decyclizing)	1wov	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Fe	Q11A00	Photosystem Q(B) protein 1	2axt	> 0.9999	0.8017	> 0.9999	> 0.9999
Fe	Q5GMY2	Ferritin and Dps	2vxx	> 0.9999	0.8784	> 0.9999	> 0.9999
Mn	Q10VK7	Photosystem Q(B) protein 2	2axt	0.0441	> 0.9999	0.9647	0.7811
Mn	Q11A00	Photosystem Q(B) protein 1	2axt	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Mn	Q10XZ7	Glutamine synthetase	3ngo	0.0002	> 0.9999	> 0.9999	0.0089
Mn	Q115D5	Bacterial nucleoid protein Hbs	2iif	> 0.9999	> 0.9999	0.2316	0.0515
Mn	Q10YS5	Uncharacterized protein	2jfs	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Mn	Q10YF3	Uncharacterized protein	2w69	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Mn	Q10YA8	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha / biotin carboxylase	3bg5	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Mn	Q10Z47	UspA	1mjh	> 0.9999	> 0.9999	> 0.9999	> 0.9999

Metal Ion	Uniprot ID	Protein Description	PDB Entry	PS- vs NF-	PS- vs PS+	NF- vs NF+	PS+ vs NF+
Mn	Q11A96	Probable cytosol aminopeptidase	3h8g	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Mn	Q10ZB9	Protein serine/threonine phosphatases	1txo	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Cu	Q117I6	Thioredoxin	2trx	> 0.9999	0.04	0.0006	0.0187
Cu	Q111R2	Plastocyanin	7pcy	> 0.9999	0.0578	0.0301	> 0.9999
Cu	Q10YM7	NmrA-like	2zcu	> 0.9999	> 0.9999	0.5379	0.0731
Ni	Q110Q8	Nitrogen regulatory protein P-II	1qy7	0.0443	0.1902	> 0.9999	> 0.9999
Ni	Q117N2	Putative nickel-containing superoxide dismutase (NISOD)	1t6u	0.079	> 0.9999	0.8837	> 0.9999
Ni	Q110E0	Caffeoyl-CoA O-methyltransferase	3tr6	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Ni	Q10VI5	Peptidoglycan-binding domain 1	3bkh	0.8758	> 0.9999	> 0.9999	> 0.9999
Ni	Q10W27	Uncharacterized protein	3cu2	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Ni	Q10VU6	Uncharacterized protein	3c2q	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Zn	Q5GMY2	Ferritin and Dps	2vxx	> 0.9999	< 0.0001	< 0.0001	> 0.9999
Zn	Q114M1	Uncharacterized protein	3ls1	0.1932	0.062	0.908	> 0.9999
Zn	Q115C4	3′-5′ exonuclease	1yt3	0.0006	> 0.9999	0.5568	0.0129
Zn	Q10XY3	Carbonate dehydratase	3kwc	0.0003	0.3075	> 0.9999	0.1388
Zn	Q119M5	30S ribosomal protein S4	2uub	0.0122	> 0.9999	0.0021	> 0.9999
Zn	Q115B7	Sodium/hydrogen exchanger	4bwz	0.9189	0.2791	0.0675	> 0.9999
Zn	Q119F6	Methionine aminopeptidase	4fuk	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Zn	Q10VY5	GTP cyclohydrolase 1	1wpl	0.5444	> 0.9999	0.3442	> 0.9999
Zn	Q115L6	Uncharacterized protein	3c37	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Zn	Q117H2	GTP cyclohydrolase 1	1wpl	> 0.9999	> 0.9999	0.7702	> 0.9999
Zn	Q10YT2	Rubredoxin	1dx8	> 0.9999	> 0.9999	0.6587	> 0.9999
Zn	Q10ZE4	YUP8H12.25 ((Arabidopsis thaliana))-type protein. Metallo peptidase. MEROPS family M50B	3b4r	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Zn	Q117B1	3-isopropylmalate dehydratase_ small subunit	2pkp	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Zn	Q10ZT2	Uncharacterized protein	1m36	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Zn	Q11A96	Probable cytosol aminopeptidase	3h8g	0.827	> 0.9999	> 0.9999	> 0.9999
Zn	Q117T0	Uncharacterized protein	3ls1	0.0545	< 0.0001	< 0.0001	> 0.9999
Zn	Q10XA2	Mo-dependent nitrogenase-like	2ou3	> 0.9999	> 0.9999	> 0.9999	> 0.9999
Zn	Q116V2	UBA/THIF-type NAD/FAD binding fold	1jw9	0.0428	> 0.9999	> 0.9999	> 0.9999
Zn	Q110C8	50S ribosomal protein L36	1dfe	0.7624	> 0.9999	> 0.9999	> 0.9999
Zn	Q115L6	Uncharacterized protein	3c37	> 0.9999	> 0.9999	> 0.9999	> 0.9999

Metal Ion	Uniprot ID	Protein Description	PDB Entry	PS- vs NF-	PS- vs PS+	NF- vs NF+	PS+ vs NF+
Zn	Q119P9	Peptidase M1_ membrane alanine aminopeptidase	3b37	> 0.9999	> 0.9999	> 0.9999	> 0.9999

Table A7: Statistical significance of per-metal stoichiometric changes as asssed by two-way ANOVA with a Bonferroni correction. Table shows multiple comparison corrected P-values where values in red indicate statistically significant stoichiometric changes ($P < 0.05$) whilst values in blue indicate non-significant change ($P > 0.05$).

Parameter	Minimum	Maximum	Reference
Intracellular N	67000	99000	(Nuester et al., 2012)
Intracellular N	27000	207000	(Carpenter et al., 2004)
Intracellular N	20160	186240	(Mulholland et al., 2002)
Intracellular N	10000	60000	(Sañudo Wilhelmy et al., 2001)
Intracellular N	7140	17860	(White et al., 2006)
Intracellular N	28700	121200	(Krauk et al., 2006)
Intracellular P	740	1340	(Nuester et al., 2012)
Intracellular P	360	2550	(Mulholland et al., 2002)
Intracellular P	250	2500	(Sañudo Wilhelmy et al., 2001)
Intracellular P	96	480	(White et al., 2006)
Intracellular P	980	37530	(Tovar-Sanchez and Sañudo Wilhemy, 2006)
Intracellular P	742	3240	(Krauk et al., 2006)
Intracellular Fe	11.7	22.2	(Nuester et al., 2012)
Intracellular Fe	3	13	(Sañudo Wilhelmy et al., 2001)
Intracellular Fe	5.7	267	(Tovar-Sanchez and Sañudo Wilhemy, 2006)
Acetylene Reduction	240	5160	(Sohm et al., 2011a)
Acetylene Reduction	2880	9840	(Capone et al., 2005)
Acetylene Reduction	96	240	(Post et al., 2002)
Acetylene Reduction	9600	31200	(Hynes, 2009)
Acetylene Reduction	8400	25920	(Lugomela et al., 2002)
Acetylene Reduction	10560	61440	(Mulholland et al., 2002)
^{33}P Uptake	120	1176	(Orchard et al., 2010a)
Alkaline Phosphatase	6048	20160	(Orchard et al., 2010a)
Alkaline Phosphatase	5280	23280	(Orchard et al., 2010b)
^{33}P Uptake	240	7200	(Sohm et al., 2011a)
Alkaline Phosphatase	240	13920	(Orchard, 2010)
$^{55}FeCl_3$	0.12	0.16	(Rubin et al., 2011)
$^{55}FeCl_3$	0.226	0.311	(Roe et al., 2012)
^{55}FeCit	0.972	1.281	(Roe et al., 2012)

Table A8: Summary of literature data for intracellular concentrations and uptake rate for N, P and Fe observed in environmental *Trichodesmium* colonies. Units for intracellular N, P or Fe are all pmol.col^{-1} whilst for N, P or Fe uptake units are in pmol.col^{-1}.day^{-1}.

BIBLIOGRAPHY

Achilles, K. and Church, T. (2003). Bioavailability of iron to trichodesmium colonies in the western subtropical atlantic ocean. *Limnology and Oceanography*.

Adams, D. G. (2000). Heterocyst formation in cyanobacteria. *Current Opinion in Microbiology*, 3(6):618–24.

Alldridge, L. and Metodieva, G. (2008). Proteome profiling of breast tumors by gel electrophoresis and nanoscale electrospray ionization mass spectrometry. *Journal of Proteome Research*, pages 1458–1469.

Ammerman, J. W. (1993). Microbial cycling of inorganic and organic phosphorus in the water column. In *Handbook of methods in aquatic microbial ecology*, pages 621–631. CRC Press.

Arrigo, K. R. (2005). Marine microorganisms and global nutrient cycles. *Nature*, 437:349–355.

Baker, a. R. (2003). Atmospheric deposition of nutrients to the Atlantic Ocean. *Geophysical Research Letters*, 30(24):2296.

Baker, a. R., Adams, C., Bell, T. G., Jickells, T. D., and Ganzeveld, L. (2013). Estimation of atmospheric nutrient inputs to the Atlantic Ocean from 50 N to 50 S based on large-scale field sampling: Iron and other dust-associated elements. *Global Biogeochemical Cycles*, 27(3):n/a–n/a.

Baker, D. and Sali, A. (2001). Protein structure prediction and structural genomics. *Science*, 294:93–6.

Balsera, M., Arellano, J. B., Gutiérrez, J. R., Heredia, P., Revuelta, J. L., and De Las Rivas, J. (2003). Structural analysis of the PsbQ protein of photosystem II by Fourier transform infrared and circular dichroic spectroscopy and by bioinformatic methods. *Biochemistry*, 42:1000–1007.

Banci, L. and Bertini, I. (2013). Metallomics and the Cell: Some Definitions and General Comments. In Banci, L., editor, *Metallomics and the Cell SE - 1*, volume 12 of *Metal Ions in Life Sciences*, pages 1–13. Springer Netherlands.

Baniulis, D., Yamashita, E., Whitelegge, J. P., Zatsman, A. I., Hendrich, M. P., Hasan, S. S., Ryan, C. M., and Cramer, W. a. (2009). Structure-Function, Stability, and Chemical Modification of the Cyanobacterial Cytochrome b6f Complex from Nostoc sp. PCC 7120. *The Journal of biological chemistry*, 284(15):9861–9.

Barcelos e Ramos, J., Biswas, H., Schulz, K. G., LaRoche, J., and Riebesell, U. (2007). Effect of rising atmospheric carbon dioxide on the marine nitrogen fixer Trichodesmium. *Global Biogeochemical Cycles*, 21(2):n/a–n/a.

Bench, S. R., Heller, P., Frank, I., Arciniega, M., Shilova, I. N., and Zehr, J. P. (2013). Whole genome comparison of six Crocosphaera watsonii strains with differing phenotypes. *Journal of Phycology*, 49(4):786–801.

Bergman, B., Sandh, G., Lin, S., Larsson, J., and Carpenter, E. J. (2012). Trichodesmium - a widespread marine cyanobacterium with unusual nitrogen fixation properties. *FEMS microbiology reviews*, pages 1–17.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000). The Protein Data Bank. *Nucleic acids research*, 28(1):235–42.

Berman-Frank, I. and Bidle, K. (2004). The demise of the marine cyanobacterium, Trichodesmium spp via an autocatalyzed cell death pathway. *Limnology and Oceanography*, 49(4):997–1005.

Berman-Frank, I., Cullen, J. T., Shaked, Y., Sherrell, R. M., and Falkowski, P. G. (2001a). Iron availability, cellular iron quotas, and nitrogen fixation in Trichodesmium. *Limnology and Oceanography*, 46(6):1249–1260.

Berman-Frank, I., Lundgren, P., Chen, Y. B., Küpper, H., Kolber, Z., Bergman, B., and Falkowski, P. (2001b). Segregation of nitrogen fixation and oxygenic photosynthesis in the marine cyanobacterium Trichodesmium. *Science*, 294(5546):1534–7.

Berman-Frank, I., Lundgren, P., and Falkowski, P. (2003). Nitrogen fixation and photosynthetic oxygen evolution in cyanobacteria. *Research in microbiology*, 154(3):157–64.

Berman-Frank, I., Quigg, A., and Finkel, Z. (2007). Nitrogen-fixation strategies and Fe requirements in cyanobacteria. *Limnology and Oceanography*, 52(5):2260–2269.

Bertilsson, S., Berglund, O., Karl, D. M., and Chisholm, S. W. (2003). Elemental composition of marine Prochlorococcus and Synechococcus: Implications for the ecological stoichiometry of the sea. *Limnology and Oceanography*, 48(5):1721–1731.

Bibby, T. S., Nield, J., and Barber, J. (2001). Iron deficiency induces the formation of an antenna ring around trimeric photosystem I in cyanobacteria. *Nature*, 412(6848):743–745.

Biemann, K. (1988). Contributions of mass spectrometry to peptide and protein structure. *Biomedical and Environmental Mass Spectrometry*, 16(1-12):99–111.

Biesiadka, J., Loll, B., and Kern, J. (2004). Crystal structure of cyanobacterial photosystem II at 3.2 A resolution: a closer look at the Mn-cluster. *Physical Chemistry Chemical Physics*, pages 4733–4736.

Blackman, F. (1905). Optima and limiting factors. *Annals of Botany*, XIX(Lxxiv).

Blais, M., Tremblay, J.-E., Jungblut, A. D., Gagnon, J., Martin, J., Thaler, M., and Lovejoy, C. (2012). Nitrogen fixation and identification of potential diazotrophs in the Canadian Arctic. *Global Biogeochemical Cycles*, 26(3):n/a–n/a.

Bolier, G., Koningh, M. C. J., Schmale, J. C., and Donze, M. (1992). Differential luxury phosphate response of planktonic algae to phosphorus removal. *Hydrobiologia*, 243-244:113–118.

Bothe, H., Schmitz, O., Yates, M. G., and Newton, W. E. (2010). Nitrogen fixation and hydrogen metabolism in cyanobacteria. *Microbiology and molecular biology reviews : MMBR*, 74(4):529–51.

Bowie, A., Whitworth, D., Achterberg, E., Mantoura, R., and Worsfold, P. (2002). Biogeochemistry of Fe and other trace elements (Al, Co, Ni) in the upper Atlantic Ocean. *Deep Sea Research Part I Oceanographic Research Papers*, 49(4):605–636.

Bowman, B. (1983). Vanadate Uptake in Neurospora crassa Occurs via Phosphate Transport System II. *Journal of Bacteriology*.

Boyd, E. S. and Peters, J. W. (2013). New insights into the evolutionary history of biological nitrogen fixation. *Frontiers in microbiology*, 4(August):201.

Boyd, P. W., Jickells, T., Law, C. S., Blain, S., Boyle, E. A., Buesseler, K. O., Coale, K. H., Cullen, J. J., De Baar, H. J. W., Follows, M., Harvey, M., Lancelot, C., Levasseur, M., Owens, N. P. J., Pollard, R., Rivkin, R. B., Sarmiento, J., Schoemann, V., Smetacek, V., Takeda, S., Tsuda, A., Turner, S., and Watson, A. J. (2007). Mesoscale iron enrichment experiments 1993-2005: synthesis and future directions. *Science*, 315:612–7.

Breitbarth, E., Oschlies, a., and LaRoche, J. (2007). Physiological constraints on the global distribution of Trichodesmium – effect of temperature on diazotrophy. *Biogeosciences*, 4(1):53–61.

Brewer, P. and Riley, J. (1965). The automatic determination of nitrate in sea water. *Deep Sea Research and Oceanographic Abstracts*, 12(6):765–772.

Bricker, T. M., Roose, J. L., Fagerlund, R. D., Frankel, L. K., and Eaton-Rye, J. J. (2012). The extrinsic proteins of Photosystem II. *Biochimica et Biophysica Acta*, 1817(1):121–42.

Bruland, K. W. (1983). Trace elements in Sea-water. In *Chemical Oceanography*, volume 8, pages 157–220. Academic Press.

Bunt, J. (1995). Light and photosynthesis in aquatic ecosystems.

Burris, R. H. (1991). Nitrogenases. *Journal of Biological Chemistry*, 266:9339–9342.

Byrne, R. H., Kump, L. R., and Cantrell, K. J. (1988). The influence of temperature and pH on trace metal speciation in seawater. *Marine Chemistry*, 25:163–181.

Cantley, L. and Josephson, L. (1977). Vanadate is a potent (Na,K)-ATPase inhibitor found in ATP derived from muscle. *Journal of Biological Chemistry*.

Capone, D. G. (1997). Trichodesmium, a Globally Significant Marine Cyanobacterium. *Science*, 276(5316):1221–1229.

Capone, D. G. (2014). An iron curtain in the Atlantic Ocean forms a biogeochemical divide. *Proceedings of the National Academy of Sciences*, 111(4):1231–2.

Capone, D. G., Burns, J. a., Montoya, J. P., Subramaniam, A., Mahaffey, C., Gunderson, T., Michaels, A. F., and Carpenter, E. J. (2005). Nitrogen fixation by Trichodesmium spp.: An important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. *Global Biogeochemical Cycles*, 19(2):n/a–n/a.

Carpenter, E. J., Capone, D. G., and Rueter, J. G. (1992). Marine pelagic cyanobacteria: Trichodesmium and other diazotrophs.

Carpenter, E. J., Subramaniam, A., and Capone, D. G. (2004). Biomass and primary productivity of the cyanobacterium Trichodesmium spp. in the tropical N Atlantic ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 51(2):173–203.

Chappell, P. D. and Webb, E. a. (2010). A molecular assessment of the iron stress response in the two phylogenetic clades of Trichodesmium. *Environmental microbiology*, 12(1):13–27.

Chauhan, D., Folea, I. M., Jolley, C. C., Kouril, R., Lubner, C. E., Lin, S., Kolber, D., Wolfe-Simon, F., Golbeck, J. H., Boekema, E. J., and Fromme, P. (2011). A novel photosynthetic strategy for adaptation to low-iron aquatic environments. *Biochemistry*, 50(5):686–92.

Chen, Y., Zehr, J., and Mellon, M. (1996). Growth and nitrogen fixation of the diazotrophic filamentous nonheterocystous cyanobacterium Trichodesmium sp IMS 101 in defined media Evidence for a circadian rhythm. *Journal of Phycology*.

Collier, R. W. (1985). Molybdenum in the Northeast Pacific Ocean. *Limnology and Oceanography*, 30(6):1351–1354.

Colón-López, M. S., Sherman, D. M., and Sherman, L. A. (1997). Transcriptional and translational regulation of nitrogenase in light-dark- and continuous-light-grown cultures of the unicellular cyanobacterium Cyanothece sp. strain ATCC 51142. *Journal of Bacteriology*, 179:4319–4327.

Compton, J., Mallinson, D., Glenn, C., and Filippelli, G. (2000). Variations in the global phosphorus cycle. In *Marine Authigenesis: From Global to m Microbial*, pages 21–33. SEPM Society for Sedimentary.

Cottrell, J. and London, U. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*.

Cox, J. and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature biotechnology*, 26(12):1367–72.

Craig, R. and Beavis, R. C. (2004). TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*, 20(9):1466–7.

Croot, P. L. (2004). Short residence time for iron in surface seawater impacted by atmospheric dry deposition from Saharan dust events. *Geophysical Research Letters*, 31(23):L23S08.

de Baar, H. (1994). von Liebig's law of the minimum and plankton ecology (1899–1991). *Progress in Oceanography*, 33(4):347–386.

Deutsch, C., Sarmiento, J. L., Sigman, D. M., Gruber, N., and Dunne, J. P. (2007). Spatial coupling of nitrogen inputs and losses in the ocean. *Nature*, 445:163–167.

Dickson, A., Sabine, C., and Christian, J. (2007). *Guide to best practices for ocean CO2 measurements*. PICES Special Publisher.

Dixon, R. and Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nature Reviews: Microbiology*, 2(8):621–31.

Donat, J. and Bruland, K. (1995). Trace elements in the oceans. *Trace elements in natural waters. CRC*.

Dugdale, R. C. and Goering, J. J. (1967). Uptake of new and regenerated forms of nitrogen in primary productivity.

Dupont, C. L., Neupane, K., Shearer, J., and Palenik, B. (2008). Diversity, function and evolution of genes coding for putative Ni-containing superoxide dismutases. *Environmental microbiology*, 10(7):1831–43.

Dupont, C. L., Yang, S., Palenik, B., and Bourne, P. E. (2006). Modern proteomes contain putative imprints of ancient shifts in trace metal geochemistry. *Proceedings of the National Academy of Sciences*, 103:17822–17827.

Dutkiewicz, S., Ward, B. A., Monteiro, F., and Follows, M. J. (2012). Interconnection of nitrogen fixers and iron in the Pacific Ocean: Theory and numerical simulations. *Global Biogeochemical Cycles*, 26(1):1–16.

Dutkiewicz, S., Ward, B. a., Scott, J. R., and Follows, M. J. (2014). Understanding predicted shifts in diazotroph biogeography using resource competition theory. *Biogeosciences Discussions*, 11(5):7113–7149.

Dyhrman, S. and Benitez-Nelson, C. (2009). A microbial source of phosphonates in oligotrophic marine systems. *Nature Geoscience*, 2(September):696–699.

Dyhrman, S. T., Chappell, P. D., Haley, S. T., Moffett, J. W., Orchard, E. D., Waterbury, J. B., and Webb, E. a. (2006). Phosphonate utilization by the globally important marine diazotroph Trichodesmium. *Nature*, 439(7072):68–71.

Eady, R. (2003). Current status of structure function relationships of vanadium nitrogenase. *Coordination Chemistry Reviews*, 237:23–30.

Eppley, R. W. R. and Peterson, B. B. J. (1979). Particulate organic matter flux and planktonic new production in the deep ocean. *Nature*, 282:677–680.

Fabry, V., Seibel, B., Feely, R., and Orr, J. (2008). Impacts of ocean acidification on marine fauna and ecosystem processes. *Ices Journal of Marine Science*, 65:414–432.

Falkowski, P. (1997). Evolution of the nitrogen cycle and its influence on the biological sequestration of CO 2 in the ocean. *Nature*, 387:272–275.

Falkowski, P. G. (2006). Tracing oxygen's imprint on Earth's metabolic evolution. *Science*, 311:1724–1725.

Falkowski, P. G., Fenchel, T., and Delong, E. F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science*, 320(5879):1034–9.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science*, 246:64–71.

Ferreira, F. and Straus, N. (1994). Iron deprivation in cyanobacteria. *Journal of Applied Phycology*, 199(Ps Ii):199–210.

Field, C. B. (1998). Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science*, 281:237–240.

Finzi-Hart, J. a., Pett-Ridge, J., Weber, P. K., Popa, R., Fallon, S. J., Gunderson, T., Hutcheon, I. D., Nealson, K. H., and Capone, D. G. (2009). Fixation and fate of C and N in the cyanobacterium Trichodesmium using nanometer-scale secondary ion mass spectrometry. *Proceedings of the National Academy of Sciences*, 106(15):6345–50.

Flower, D. R., North, a. C., and Sansom, C. E. (2000). The lipocalin protein family: structural and sequence overview. *Biochimica et Biophysica Acta*, 1482(1-2):9–24.

Foster, R. A., Subramaniam, A., Mahaffey, C., Carpenter, E. J., Capone, D. G., and Zehr, J. P. (2007). Influence of the Amazon River plume on distributions of free-living and symbiotic cyanobacteria in the western tropical north Atlantic Ocean. *Limnology and Oceanography*, 52:517–532.

Fraser, M. (2012). *Hydrogen Supersaturation in the North and South Atlantic - A Possible Indicator of Nitrogen Fixation*. PhD thesis, Dalhousie University.

Frederick, R. E., Mayfield, J. a., and DuBois, J. L. (2011). Regulated O2 activation in flavin-dependent monooxygenases. *Journal of the American Chemical Society*, 133(32):12338–41.

Fredriksson, C. and Bergman, B. (1995). Nitrogenase quantity varies diurnally in a subset of cells within colonies of the non-heterocystous cyanobacteria Trichodesmium spp. *Microbiology*, 141(10).

Fu, F., Yu, E., Garcia, N., Gale, J., Luo, Y., Webb, E., and Hutchins, D. (2014). Differing responses of marine N2 fixers to warming and consequences for future diazotroph community structure. *Aquatic Microbial Ecology*, 72(1):33–46.

Gallon, J. (1981). The oxygen sensitivity of nitrogenase : a problem for biochemists and micro-organisms Physi barrier s. *Trends in Biochemical Sciences*, 6:19–23.

Galloway, J., Dentener, F., and Capone, D. (2004). Nitrogen cycles: past, present, and future. *Biogeochemistry*, pages 153–226.

Geider, R. and Roche, J. L. (2002). Redfield revisited : variability of C : N : P in marine microalgae and its biochemical basis. *European Journal of Phycology*, 37(July 2012):1–17.

Gerber, S. a., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences*, 100(12):6940–5.

Geromanos, S. J., Vissers, J. P. C., Silva, J. C., Dorschel, C. A., Li, G.-Z., Gorenstein, M. V., Bateman, R. H., and Langridge, J. I. (2009). The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS. *Proteomics*, 9:1683–1695.

Glantz, S. A. and Slinker, B. K. (1990). *Primer of applied regression and analysis of variance*. McGraw-Hill, Health Professions Division.

Glöckner, F., Gasol, J., McDonough, N., and Calewaert, J. (2012). Marine Microbial Diversity and its role in Ecosystem Functioning and Environmental Change. Technical Report May, European Science Foundation.

Goetz, D. H., Holmes, M. a., Borregaard, N., Bluhm, M. E., Raymond, K. N., and Strong, R. K. (2002). The neutrophil lipocalin NGAL is a bacteriostatic agent that interferes with siderophore-mediated iron acquisition. *Molecular cell*, 10(5):1033–43.

Golbeck, J. H. (1992). Structure and Function of Photosystem I. *Annual Review of Plant Physiology and Plant Molecular Biology*, 43(1):293–324.

Goldman, S. J., Lammers, P. J., Berman, M. S., and Sanders-Loehr, J. (1983). Siderophore-mediated iron uptake in different strains of Anabaena sp. *Journal of Bacteriology*, 156:1144–1150.

Gorbunov, M. Y. and Falkowski, P. G. (2004). Fluorescence induction and relaxation (FIRe) technique and instrumentation for monitoring photosynthetic processes and primary production in aquatic ecosystems. In *Photosynthesis: Fundamental Aspects to Global Perspectives"-Proc. 13th International Congress of Photosynthesis, Montreal, Aug*, pages 1029–1031.

Gradoville, M., White, A., and Böttjer, D. (2014). Diversity trumps acidification: Lack of evidence for carbon dioxide enhancement of Trichodesmium community nitrogen or carbon fixation at Station ALOHA. *Limnol. . . .*, 59(3):645–659.

Graham, W. and Duce, R. (1979). Atmospheric pathways of the phosphorus cycle. *Geochimica et Cosmochimica Acta*.

Grasshoff, K. (1976). *Methods of Seawater Analysis*. John Wiley and Sons.

Greenwood, C., Metodieva, G., Al-Janabi, K., Lausen, B., Alldridge, L., Leng, L., Bucala, R., Fernandez, N., and Metodiev, M. V. (2012). Stat1 and CD74 overexpression is co-dependent and linked to increased invasion and lymph node metastasis in triple-negative breast cancer. *Journal of proteomics*, 75(10):3031–40.

Griese, M., Lange, C., and Soppa, J. (2011). Ploidy in cyanobacteria. *FEMS microbiology letters*, 323(2):124–31.

Groß kopf, T. and LaRoche, J. (2012). Direct and indirect costs of dinitrogen fixation in Crocosphaera watsonii WH8501 and possible implications for the nitrogen cycle.

Groß kopf, T., Mohr, W., Baustian, T., Schunck, H., Gill, D., Kuypers, M. M. M., Lavik, G., Schmitz, R. a., Wallace, D. W. R., and LaRoche, J. (2012). Doubling of marine dinitrogen-fixation rates based on direct measurements. *Nature*, 488(7411):361–4.

Gruber, N. and Sarmiento, J. L. (1997). Global patterns of marine nitrogen fixation and denitrification. *Global Biogeochemical Cycles*, 11(2):235–266.

Guilhaus, M., Selby, D., and Mlynski, V. (2000). Orthogonal acceleration time-of-flight mass spectrometry. *Mass spectrometry reviews*, 19(2):65–107.

Harvey, D. J. (2003). Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates and glycoconjugates.

Hellwegerl, F. and Farley, K. (2003). Greedy algae reduce arsenate. *Limnology and Oceanography*, 48(6):2275–2288.

Hewson, I., Poretsky, R. S., Dyhrman, S. T., Zielinski, B., White, A. E., Tripp, H. J., Montoya, J. P., and Zehr, J. P. (2009). Microbial community gene expression within colonies of the diazotroph, Trichodesmium, from the Southwest Pacific Ocean. *The ISME journal*, 3(11):1286–300.

Ho, T. (2013). Nickel limitation of nitrogen fixation in Trichodesmium. *Limnology and Oceanography*, 58(1):112–120.

Ho, T., Quigg, A., and Finkel, Z. (2003). The Elemental Composition of some Marine Phytoplankton. *Journal of Phycology*, 1159:1145–1159.

Ho, T.-Y. T., Chu, T. T.-H., and Hu, C.-L. C. (2013). Interrelated influence of light and Ni on Trichodesmium growth. *Frontiers in Microbiology*, 4(May):1–6.

Holland, H. D. (2002). Volcanic gases, black smokers, and the great oxidation event. *Geochimica et Cosmochimica Acta*, 66:3811–3826.

Holland, H. D. (2006). The oxygenation of the atmosphere and oceans. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 361:903–915.

Honjo, S. and Manganini, S. J. (1993). Annual biogenic particle fluxes to the interior of the North Atlantic Ocean; studied at 34 N 21 W and 48 N 21 W. *Deep Sea Research Part II: Topical Studies in Oceanography*, 40:587–607.

Hopkinson, B. and Morel, F. (2009). The role of siderophores in iron acquisition by photosynthetic marine microorganisms. *BioMetals*, 22:659–669.

Hudson, R. and Morel, F. (1989). Distinguishing between extra- and intracellular marine phytoplankton. *Limnology and Oceanography*, 34(6).

Hutchins, D. A., Fu, F.-X., Webb, E. a., Walworth, N., and Tagliabue, A. (2013). Taxon-specific response of marine nitrogen fixers to elevated carbon dioxide concentrations. *Nature Geoscience*, 6(7):1–6.

Hutchins, D. A., Fu, F. X., Zhang, Y., Warner, M. E., Feng, Y., Portune, K., Bernhardt, P. W., and Mulholland, M. R. (2007). CO2 control of Trichodesmium N-2 fixation, photosynthesis, growth rates, and elemental ratios: Implications for past, present, and future ocean biogeochemistry. *Limnology and Oceanography*, 52:1293–1304.

Hygum, B. (1997). Dissolved organic carbon released by zooplankton grazing activity-a high-quality substrate pool for bacteria. *Journal of Plankton Research*, 19(1):97–111.

Hynes, A. M., Webb, E. a., Doney, S. C., and Waterbury, J. B. (2012). Comparison of Cultured Trichodesmium (Cyanophyceae) With Species Characterized From the Field1. *Journal of Phycology*, 48(1):196–210.

Hynes, A. M. A. (2009). *Diversity of the marine cyanobacterium Trichodesmium: characterization of the Woods Hole culture collection and quantification of field populations*. PhD thesis, Massachusetts Institute of Technology.

Jacobson, L. and Halmann, M. (1982). Polyphosphate metabolism in the blue-green alga Microcystis aeru-ginosa. *Journal of Plankton Research*, 4(3):481–488.

Janson, S., Bergman, B., Carpenter, E. J., Giovannoni, S. J., and Vergin, K. (1999). Genetic analysis of natural populations of the marine diazotrophic cyanobacterium Trichodesmium. *FEMS Microbiology Ecology*, 30(1):57–65.

Janson, S., Siddiqui, P. J. A., Walsby, A. E., Romans, K. M., Carpenter, E. J., and Bergman, B. (1995). Cytomorphological characterization of the planktonic diazotrophic cyanobacteria Trichodesmiumspp. from the Indian Ocean and Caribbean and Sargasso Seas. *Journal of Phycology*, 31:463–477.

Jickells, T. (2006). The role of air-sea exchange in the marine nitrogen cycle. *Biogeosciences*, 3(3):271–280.

Jickells, T. D., An, Z. S., Andersen, K. K., Baker, A. R., Bergametti, G., Brooks, N., Cao, J. J., Boyd, P. W., Duce, R. A., Hunter, K. A., Kawahata, H., Kubilay, N., LaRoche, J., Liss, P. S., Mahowald, N., Prospero, J. M., Ridgwell, a. J., Tegen, I., and Torres, R. (2005). Global iron connections between desert dust, ocean biogeochemistry, and climate. *Science*, 308(5718):67–71.

Jordan, P., Fromme, P., Witt, H. T., Klukas, O., Saenger, W., and Krauss, N. (2001). Three-dimensional structure of cyanobacterial photosystem I at 2.5 A resolution. *Nature*, 411(6840):909–17.

Kagaya, S., Maeba, E., Inoue, Y., Kamichatani, W., Kajiwara, T., Yanai, H., Saito, M., and Tohda, K. (2009). A solid phase extraction using a chelate resin immobilizing carboxymethylated pentaethylenehexamine for separation and preconcentration of trace elements in water samples. *Talanta*, 79:146–152.

Kana, T. M. (1993). Rapid oxygen cycling in Trichodesmium thiebautii. *Limnology and Oceanography*, 38:18–24.

Karadjova, I. B., Slaveykova, V. I., and Tsalev, D. L. (2008). The biouptake and toxicity of arsenic species on the green microalga Chlorella salina in seawater. *Aquatic Toxicology*, 87(4):264–71.

Karas, M. and Hillenkamp, F. (1988). Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10 000 Daltons. *Analytical Chemistry*, 60:2299–2301.

Kelley, L. a. and Sternberg, M. J. E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols*, 4(3):363–71.

Keren, N., Aurora, R., and Pakrasi, H. (2004). Critical roles of bacterioferritins in iron storage and proliferation of cyanobacteria. *Plant physiology*, 135(July):1666–1673.

Klunder, M., Laan, P., Middag, R., De Baar, H., and van Ooijen, J. (2011). Dissolved iron in the Southern Ocean (Atlantic sector). *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(25-26):2678–2694.

Kolber, Z., Prasil, O., and Falkowski, P. (1998). Measurements of variable chlorophyll fluorescence using fast repetition rate techniques: defining methodology and experimental protocols. *Biochimica et Biophysica Acta*, 1367(1-3):88–106.

Kornberg, A. (1995). Inorganic polyphosphate: toward making a forgotten polymer unforgettable. *Journal of Bacteriology*, 177:491–496.

Kornberg, A., Rao, N. N., and Ault-Riché, D. (1999). Inorganic polyphosphate: a molecule of many functions. *Annual Review of Biochemistry*, 68:89–125.

Kramer, G., Moerland, P. D., Jeeninga, R. E., Vlietstra, W. J., Ringrose, J. H., Byrman, C., Berkhout, B., and Speijer, D. (2012). Proteomic analysis of HIV-T cell interaction: an update. *Frontiers in microbiology*, 3(July):240.

Kranz, S. and Sültemeyer, D. (2009). Carbon acquisition in Trichodesmium: The effect of pCO2 and diurnal changes. *Limnology and Oceanography*, 54(2):548–559.

Kranz, S. a., Wolf-Gladrow, D., Nehrke, G., Langer, G., and Rost, B. r. (2010). Calcium carbonate precipitation induced by the growth of the marine cyanobacteria Trichodesmium. *Limnology and Oceanography*, 55(6):2563–2569.

Kranzler, C., Rudolf, M., Keren, N., Schleif, E., and Schleiff, E. (2013). Chapter Three - Iron in Cyanobacteria. In Research, F. C. and in Botanical, C. C.-C. B. T. A., editors, *Genomics of Cyanobacteria*, volume Volume 65, pages 57–105. Academic Press.

Krauk, J., Villareal, T., and Sohm, J. (2006). Plasticity of N: P ratios in laboratory and field populations of Trichodesmium spp. *Aquatic Microbial Ecology*, 42(1977):243–253.

Küpper, H. and Ferimazova, N. (2004). Traffic Lights in Trichodesmium . Regulation of Photosynthesis for Nitrogen Fixation Studied by Chlorophyll Fluorescence Kinetic Microscopy. *Plant physiology*, 135(August):2120–2133.

Küpper, H., Setlík, I., Seibert, S., Prásil, O., Setlikova, E., Strittmatter, M., Levitan, O., Lohscheider, J., Adamska, I., and Berman-Frank, I. (2008). Iron limitation in the marine cyanobacterium Trichodesmium reveals new insights into regulation of photosynthesis and nitrogen fixation. *The New phytologist*, 179(3):784–98.

Kustka, A. (2003a). A Revised estimate of the iron use efficiency of nitrogen fixation with Special Reference to the marine cyanobacterium Trichodesmium spp. *Journal of Phycology*, 25:12–25.

Kustka, A. (2003b). Iron requirements for dinitrogen- and ammonium-supported growth in cultures of Trichodesmium (IMS 101): Comparison with nitrogen fixation rates and iron : carbon ratios of field populations. *Limnology and Oceanography*, 48(5):1869–1884.

Lamond, A. I., Uhlen, M., Horning, S., Makarov, A., Robinson, C. V., Serrano, L., Hartl, F. U., Baumeister, W., Werenskiold, A. K., Andersen, J. S., Vorm, O., Linial, M., Aebersold, R., and Mann, M. (2012). Advancing Cell Biology Through Proteomics in Space and Time (PROSPECTS). *Molecular and Cellular Proteomics*, 11:O112.017731–O112.017731.

Lange, V., Picotti, P., Domon, B., and Aebersold, R. (2008). Selected reaction monitoring for quantitative proteomics: a tutorial. *Molecular systems biology*, 4(222):222.

LaRoche, J., Boyd, P., McKay, R., and Geider, R. (1996). Flavodoxin as an in situ marker for iron stress in phytoplankton. *Nature*.

LaRoche, J. and Breitbarth, E. (2005). Importance of the diazotrophs as a source of new nitrogen in the ocean. *Journal of Sea Research*, 53(1-2):67–91.

Larsson, J., Nylander, J. A., and Bergman, B. (2011). Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. *BMC Evolutionary Biology*, 11(1):187.

Latifi, A., Jeanjean, R., and Lemeille, S. (2005). Iron Starvation Leads to Oxidative Stress in Anabaena sp. PCC 7120. *Journal of Bacteriology*, 187(18):6596–6598.

Laws, E. A. (1991). Photosynthetic quotients, new production and net community production in the open ocean. *Deep Sea Research Part A. Oceanographic Research Papers*, 38:143–167.

Leuschner, C. and Bricker, T. M. (1996). Interaction of the 33 kDa extrinsic protein with photosystem II: rebinding of the 33 kDa extrinsic protein to photosystem II membranes which contain four, two, or zero manganese per photosystem II reaction center. *Biochemistry*, 35(14):4551–7.

Levitan, O., Brown, C. M., Sudhaus, S., Campbell, D., LaRoche, J., and Berman-Frank, I. (2010). Regulation of nitrogen metabolism in the marine diazotroph Trichodesmium IMS101 under varying temperatures and atmospheric CO2 concentrations. *Environmental microbiology*, 12(7):1899–912.

Li, G.-Z., Vissers, J. P. C., Silva, J. C., Golick, D., Gorenstein, M. V., and Geromanos, S. J. (2009). Database searching and accounting of mul-

tiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. *Proteomics*, 9:1696–1719.

Lin, S., Henze, S., Lundgren, P., Bergman, B., and Carpenter, E. (1998). Whole-cell immunolocalization of nitrogenase in marine diazotrophic cyanobacteria, trichodesmium spp. *Applied and Environmental Microbiology*, 64(8):3052–8.

Lomas, M. W., Burke, a. L., Lomas, D. a., Bell, D. W., Shen, C., Dyhrman, S. T., and Ammerman, J. W. (2010). Sargasso Sea phosphorus biogeochemistry: an important role for dissolved organic phosphorus (DOP). *Biogeosciences*, 7(2):695–710.

Longhurst, A. R. and Harrison, G. W. (1989). The biological pump: Profiles of plankton production and consumption in the upper ocean. *Progress in Oceanography*, 22:47–123.

Lugomela, C., Lyimo, T., and Bryceson, I. (2002). Trichodesmium in coastal waters of Tanzania: diversity, seasonality, nitrogen and carbon fixation. *Hydrobiologia*, 477(477):1–13.

Lundgren, P. and Janson, S. (2005). Unveiling of novel radiations within Trichodesmium cluster by hetR gene sequence analysis. *Applied and Environmental Microbiology*, 71(1):190–196.

Luo, Y. W., Doney, S. C., Anderson, L. A., Benavides, M., Berman-Frank, I., Bode, A., Bonnet, S., Bostrom, K. H., Bottjer, B., Capone, D. G., Carpenter, E. J., Chen, Y. L., Church, M. J., Dore, J. E., Falcon, L. I., Fernandez, A., Foster, R. A., Furuya, K., Gomez, F., Gundersen, K., Hynes, A. M., Karl, D. M., Kitajima, S., Langlois, R. J., LaRoche, J., Letelier, R. M., Maranon, E., McGillicuddy, J. D. J., Moisander, P. H., Moore, C., and Zehr, J. (2012). Database of diazotrophs in global ocean: abundance, biomass and nitrogen fixation rates. *Earth System Science Data*, 4:47–73.

Luo, Y.-W., Lima, I. D., Karl, D. M., and Doney, S. C. (2013). Data-based assessment of environmental controls on global marine nitrogen fixation. *Biogeosciences Discussions*, 10(4):7367–7412.

Lyons, T. W., Reinhard, C. T., and Planavsky, N. J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature*, 506(7488):307–315.

Makarov, A. (2000). Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical Chemistry*, 72:1156–62.

Marsay, C. M. (2012). *Particulate trace metals, carbon and nitrogen in the Mesopelagic*. PhD thesis, University of Southampton.

Mather, R. L., Reynolds, S. E., Wolff, G. a., Williams, R. G., Torres-Valdes, S., Woodward, E. M. S., Landolfi, A., Pan, X., Sanders, R., and Achterberg, E. P. (2008). Phosphorus cycling in the North and South Atlantic Ocean subtropical gyres. *Nature Geoscience*, 1(7):439–443.

Measures, C. I. and Vink, S. (1999). Seasonal variations in the distribution of Fe and Al in the surface waters of the Arabian Sea. *Deep Sea Research Part II: Topical Studies in Oceanography*, 46(8-9):1597–1622.

Meng, L., Wong, J. H., Feldman, L. J., Lemaux, P. G., and Buchanan, B. B. (2010). A membrane-associated thioredoxin required for plant growth moves from cell to cell, suggestive of a role in intercellular communication. *Proceedings of the National Academy of Sciences*, 107(8):3900–5.

Meyer, Y., Siala, W., Bashandy, T., Riondet, C., Vignols, F., and Reichheld, J. P. (2008). Glutaredoxins and thioredoxins in plants. *Biochimica et Biophysica Acta*, 1783:589–600.

Michel, K.-P. and Pistorius, E. K. (2004). Adaptation of the photosynthetic electron transport chain in cyanobacteria to iron deficiency: The function of IdiA and IsiA. *Physiologia plantarum*, 120(1):36–50.

Milligan, A. J., Berman-Frank, I., Gerchman, Y., Dismukes, G. C., and Falkowski, P. G. (2007). Light-dependent oxygen consumption in nitrogen-fixing cyanobacteria plays a key role in nitrogenase protection 1. *Journal of Phycology*, 43(5):845–852.

Mills, M. M., Ridame, C., Davey, M., La Roche, J., and Geider, R. J. (2004). Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. *Nature*, 429(6989):292–4.

Milne, A., Landing, W., Bizimis, M., and Morton, P. (2010). Determination of Mn, Fe, Co, Ni, Cu, Zn, Cd and Pb in seawater using high resolution magnetic sector inductively coupled mass spectrometry (HR-ICP-MS). *Analytica Chimica Acta*, 665:200–207.

Minster, J.-F. (1983). Tracers in the sea. In *Geochimica et Cosmochimica Acta*, volume 47, page 1336.

Mitchell, T. P. and Wallace, J. M. (1992). The annual cycle in equatorial convection and sea surface temperature. *Journal of Climate*, 5(10):1140–1156.

Mohr, W., Grosskopf, T., Wallace, D. W. R., and LaRoche, J. (2010). Methodological underestimation of oceanic nitrogen fixation rates. *PloS one*, 5(9):e12583.

Moisander, P. H., Beinart, R. a., Hewson, I., White, A. E., Johnson, K. S., Carlson, C. a., Montoya, J. P., and Zehr, J. P. (2010). Unicellular cyanobacterial distributions broaden the oceanic N2 fixation domain. *Science*, 327(5972):1512–4.

Monteiro, F., Dutkiewicz, S., and Follows, M. J. (2011). Biogeographical controls on the marine nitrogen fixers. *Global Biogeochemical Cycles*, 25(2).

Montoya, J. P., Holl, C. M., Zehr, J. P., Hansen, A., Villareal, T. A., and Capone, D. G. (2004). High rates of N2 fixation by unicellular diazotrophs in the oligotrophic Pacific Ocean. *Nature*, 430:1027–1032.

Montoya, J. P., Voss, M., Kahler, P., and Capone, D. G. (1996). A Simple, High-Precision, High-Sensitivity Tracer Assay for N(inf2) Fixation. *Applied and Environmental Microbiology*, 62(3):986–93.

Moore, C. M., Mills, M. M., Achterberg, E. P., Geider, R. J., LaRoche, J., Lucas, M. I., McDonagh, E. L., Pan, X., Poulton, A. J., Rijkenberg, M. J. A., Suggett, D. J., Ussher, S. J., and Woodward, E. M. S. (2009a). Large-scale distribution of Atlantic nitrogen fixation controlled by iron availability. *Nature Geoscience*, 2(12):867–871.

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., Jickells, T. D., La Roche, J., Lenton, T. M., Mahowald, N. M., Marañón, E., Marinov, I., Moore, J. K., Nakatsuka, T., Oschlies, A., Saito, M. a., Thingstad, T. F., Tsuda, A., and Ulloa, O. (2013). Processes and patterns of oceanic nutrient limitation. *Nature Geoscience*, 6(March):701–710.

Moore, C. M., Mills, M. M., Langlois, R., Milne, A., Achterberg, E. P., La Roche, J., and Geider, R. J. (2008). Relative influence of nitrogen and phosphorus availability on phytoplankton physiology and productivity in the oligotrophic sub-tropical North Atlantic Ocean. *Limnology and Oceanography*, 53(1):291–305.

Moore, C. M., Mills, M. M., Milne, A., Langlois, R., Achterberg, E. P., Lochte, K., Geider, R. J., and La Roche, J. (2006). Iron limits primary productivity during spring bloom development in the central North Atlantic. *Global Change Biology*, 12(4):626–634.

Moore, R. M., Punshon, S., Mahaffey, C., and Karl, D. (2009b). The relationship between dissolved hydrogen and nitrogen fixation in ocean waters. *Deep Sea Research Part I: Oceanographic Research Papers*, 56(9):1449–1458.

Morel, F., Kustka, A., and Shaked, Y. (2008). The role of unchelated Fe in the iron nutrition of phytoplankton. *Limnology and Oceanography*, 53(1):400–404.

Morris, J., Lenski, R., and Zinser, E. (2012). The Black Queen Hypothesis : Evolution of Dependencies through Adaptive Gene Loss. *MBio*.

Mulholland, M. and Bernhardt, P. (2006). Nitrogen fixation and release of fixed nitrogen by Trichodesmium spp. in the Gulf of Mexico. *Limnology and Oceanography*, 51(4):1762–1776.

Mulholland, M., Floge, S., Carpenter, E., and Capone, D. (2002). Phosphorus dynamics in cultures and natural populations of Trichodesmium spp. *Marine Ecology Progress Series*, 239(3):45–55.

Murray, J. W. (2012). Sequence variation at the oxygen-evolving centre of photosystem II: a new class of 'rogue' cyanobacterial D1 proteins. *Photosynthesis research*, 110(3):177–84.

Murzin, A. and Brenner, S. (1995). SCOP: A Structural Classification of Proteins Database for the Investigation of Sequences and Structures. *Journal of Molecular Biology*, pages 536–540.

Najafpour, M. M., Moghaddam, A. N., Allakhverdiev, S. I., and Govindjee (2012). Biological water oxidation: Lessons from Nature. *Biochimica et Biophysica Acta*, 1817:1110–1121.

Navarro-González, R., McKay, C. P., and Mvondo, D. N. (2001). A possible nitrogen crisis for Archaean life due to reduced nitrogen fixation by lightning. *Nature*, 412:61–64.

Neilson, K. a., Ali, N. a., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., Lee, A., van Sluyter, S. C., and Haynes, P. a. (2011). Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics*, 11(4):535–53.

Nuester, J., Newville, M., and Twining, B. S. (2014). Distributions of iron, phosphorus and sulfur along trichomes of the cyanobacteria Trichodesmium. *Metallomics*, pages –.

Nuester, J., Vogt, S., Newville, M., Kustka, A. B., and Twining, B. S. (2012). The unique biogeochemical signature of the marine diazotroph trichodesmium. *Frontiers in microbiology*, 3:150.

Orchard, E. (2010). *Phosphorus Physiology Of The Marine Cyanobacterium Trichodesmium*. PhD thesis, Massachusetts Institute of Technology.

Orchard, E. D., Ammerman, J. W., Lomas, M. W., and Dyhrman, S. T. (2010a). Dissolved inorganic and organic phosphorus uptake in Trichodesmium and the microbial community: The importance of phosphorus ester in the Sargasso Sea. *Limnology and Oceanography*, 55(3):1390–1399.

Orchard, E. D., Benitez-Nelson, C. R., Pellechia, P. J., Lomas, M. W., and Dyhrman, S. T. (2010b). Polyphosphate in Trichodesmium from the low-phosphorus Sargasso Sea. *Limnology and Oceanography*, 55(5):2161–2169.

Orchard, E. D., Webb, E. a., and Dyhrman, S. T. (2009). Molecular analysis of the phosphorus starvation response in Trichodesmium spp. *Environmental microbiology*, 11(9):2400–11.

Orcutt, K. and Rasmussen, U. (2002). Characterization of Trichodesmium spp. by genetic techniques. *Applied and Environmental Microbiology*, 68(5):2236–2245.

Oremland, R. S. and Stolz, J. F. (2003). The ecology of arsenic. *Science*, 300(5621):939–44.

Palter, J. B., Sarmiento, J. L., Gnanadesikan, a., Simeon, J., and Slater, R. D. (2010). Fueling export production: nutrient return pathways from the deep ocean and their dependence on the Meridional Overturning Circulation. *Biogeosciences*, 7(11):3549–3568.

Pandey, K. D., Shukla, S. P., Shukla, P. N., Giri, D. D., Singh, J. S., Singh, P., and Kashyap, A. K. (2004). Cyanobacteria in Antarctica: ecology, physiology and cold adaptation. *Cellular and Molecular Biology*, 50:575–584.

Patey, M. D., Rijkenberg, M. J., Statham, P. J., Stinchcombe, M. C., Achterberg, E. P., and Mowlem, M. (2008). Determination of nitrate and phosphate in seawater at nanomolar concentrations. *Trends in Analytical Chemistry*, 27(2):169–182.

Peers, G. and Price, N. M. (2006). Copper-containing plastocyanin used for electron transport by an oceanic diatom. *Nature*, 441(7091):341–4.

Philander, S., Gu, D., and Lambert, G. (1996). Why the ITCZ is mostly north of the equator. *Journal of Climate*.

Popelkova, H. and Yocum, C. F. (2011). PsbO, the manganese-stabilizing protein: analysis of the structure-function relations that provide insights into its role in photosystem II. *Journal of photochemistry and photobiology. B, Biology*, 104(1-2):179–90.

Post, A., Dedej, Z., Gottlieb, R., Li, H., Thomas, D., El-Absawi, M., El-Naggar, a., El-Gharabawi, M., and Sommer, U. (2002). Spatial and temporal distribution of Trichodesmium spp. in the stratified Gulf of Aqaba, Red Sea. *Marine Ecology Progress Series*, 239:241–250.

Pratte, B. and Thiel, T. (2006). High-Affinity Vanadate Transport System in the Cyanobacterium Anabaena variabilis ATCC 29413. *Journal of Bacteriology*.

Prufert-Bebout, L., Paerl, H. W., and Lassen, C. (1993). Growth, nitrogen fixation, and spectral attenuation in cultivated trichodesmium species. *Applied and Environmental Microbiology*, 59:1367–1375.

Quigg, A., Irwin, A. J., and Finkel, Z. V. (2011). Evolutionary inheritance of elemental stoichiometry in phytoplankton. *Proceedings. Biological sciences / The Royal Society*, 278(1705):526–34.

Raven, J. (1988). The iron and molybdenum use efficiencies of plant growth with different energy, carbon and nitrogen sources. *New Phytologist*.

Redfield, A. (1934). On the proportions of organic derivatives in sea water and their relation to the composition of plankton. *James Johnstone Memorial Volume*.

Redfield, A. C. (1958). The Biological Control of Chemical Factors in the Environment. *American Scientist*, 46:205–221.

Richier, S., Fiorini, S., Kerros, M. E., von Dassow, P., and Gattuso, J. P. (2011). Response of the calcifying coccolithophore Emiliania huxleyi to low pH/high pCO(2): from physiology to molecular level.

Richier, S., Macey, A. I., Pratt, N. J., Honey, D. J., Moore, C. M., and Bibby, T. S. (2012). Abundances of Iron-Binding Photosynthetic and Nitrogen-Fixing Proteins of Trichodesmium Both in Culture and In Situ from the North Atlantic. *PLoS ONE*, 7(5):e35571.

Rödenbeck, C., Keeling, R. F., Bakker, D. C. E., Metzl, N., Olsen, a., Sabine, C., and Heimann, M. (2013). Global surface-ocean p^{CO₂} and sea–air CO₂ flux variability from an observation-driven ocean mixed-layer scheme. *Ocean Science*, 9(2):193–216.

Roe, K. L., Barbeau, K., Mann, E. L., and Haygood, M. G. (2012). Acquisition of iron by Trichodesmium and associated bacteria in culture. *Environmental microbiology*, 14(7):1681–95.

Rogowska-Wrzesinska, A. (2014). Heteromer score: using internal standards to assess the quality of proteomic data. *Proteomics*, pages 1–15.

Romanova, I. M., Smirnova, T. A., Andreev, A. L., Il'ina, T. S., Didenko, L. V., and Gintsburg, A. L. (2002). The Pho regulons of bacteria. *Mikrobiologiia*, 71:581–595.

Romans, K. M., Carpenter, E. J., and Bergman, B. (1994). Buoyancy Regulation In The Colonial Diazotrophic Cyanobacterium Trichodesmium Tenue: Ultrastructure And Storage Of Carbohydrate, Polyphosphate, And Nitrogen. *Journal of Phycology*, 30(6):935–942.

Roth, J., Ponzoni, S., and Aschner, M. (2013). Manganese homeostasis and transport. *Metal ions in life sciences*, 12:169–201.

Roy, N. K., Ghosh, R. K., and Das, J. (1982). Monomeric alkaline phosphatase of Vibrio cholerae. *Journal of Bacteriology*, 150(3):1033–9.

Rubin, M., Berman-Frank, I., and Shaked, Y. (2011). Dust- and mineral-iron utilization by the marine dinitrogen-fixer Trichodesmium. *Nature Geoscience*, 4(August):529–534.

Rubio, L. M. and Ludden, P. W. (2008). Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Annual Review of Microbiology*, 62:93–111.

Ryan-Keogh, T. J., Macey, A. I., Cockshutt, A. M., Moore, C. M., and Bibby, T. S. (2012). The cyanobacterial chlorophyll-binding-protein IsiA acts to increase the in vivo effective absorption cross-section of PSI under iron limitation. *Journal of Phycology*, 48(1):145–154.

Sañudo Wilhelmy, S. a., Kustka, a. B., Gobler, C. J., Hutchins, D. a., Yang, M., Lwiza, K., Burns, J., Capone, D. G., Raven, J. a., and Carpenter, E. J. (2001). Phosphorus limitation of nitrogen fixation by Trichodesmium in the central Atlantic Ocean. *Nature*, 411(6833):66–9.

Saito, M. a., Bertrand, E. M. E., Dutkiewicz, S., Bulygin, V. V., Moran, D. M., Monteiro, F. M., Follows, M. J., Valois, F. W., and Waterbury, J. B. (2011). Iron conservation by reduction of metalloenzyme inventories in the marine diazotroph Crocosphaera watsonii. *Proceedings of the National Academy of Sciences*, 108(6):2184–9.

Saito, M. a., Goepfert, T. J., and Ritt, J. T. (2008). Some thoughts on the concept of colimitation: Three definitions and the importance of bioavailability. *Limnology and Oceanography*, 53(1):276–290.

Saito, M. A. and Moffett, J. W. (2002). Temporal and spatial variability of cobalt in the Atlantic Ocean. *Geochimica Et Cosmochimica Acta*, 66:1943–1953.

Sandh, G., El-Shehawy, R., Díez, B., and Bergman, B. (2009). Temporal separation of cell division and diazotrophy in the marine diazotrophic cyanobacterium Trichodesmium erythraeum IMS101. *FEMS microbiology letters*, 295(2):281–8.

Sandh, G., Ran, L., Xu, L., Sundqvist, G., Bulone, V., and Bergman, B. (2011). Comparative proteomic profiles of the marine cyanobacterium Trichodesmium erythraeum IMS101 under different nitrogen regimes. *Proteomics*, 11(3):406–19.

Sargent, E. (2013). *Describing the fate of diazotroph-derived new nitrogen*. PhD thesis, University of Southampton.

Sarmiento, J. L. and Gruber, N. (2006). *Ocean Biogeochemical Dynamics*. Princeton University Press.

Schlosser, C., Klar, J., Wake, B. D., Snow, J., Honey, D. J., and Woodward, E. M. S. (2014). Seasonal ITCZ migration dynamically controls the location of the (sub-)tropical Atlantic biogeochemical divide. *Proceedings of the National Academy of Sciences*.

Schneegurt, M. and Tucker, D. (2000). Metabolic rythyms of a diazotrophic C cyanobacterium, Cyanothece sp. strain ATCC 51142, heterotrophically grown In continous dark. *Journal of Phycology*, 117:107–117.

Scigelova, M. and Makarov, A. (2006). Orbitrap mass analyzer–overview and applications in proteomics. *Proteomics*, 6 Suppl 2:16–21.

Seargeant, L. E. and Stinson, R. a. (1979). Inhibition of human alkaline phosphatases by vanadate. *The Biochemical journal*, 181(1):247–50.

Shabalina, S. a. and Spiridonov, N. a. (2004). The mammalian transcriptome and the function of non-coding DNA sequences. *Genome biology*, 5(4):105.

Shelley, R. U., Sedwick, P. N., Bibby, T. S., Cabedo-Sanz, P., Church, T. M., Johnson, R. J., Macey, A. I., Marsay, C. M., Sholkovitz, E. R., Ussher, S. J., Worsfold, P. J., and Lohan, M. C. (2012). Controls on dissolved cobalt in surface waters of the Sargasso Sea: Comparisons with iron and aluminum.

Shi, D. and Kranz, S. (2012). Ocean acidification slows nitrogen fixation and growth in the dominant diazotroph Trichodesmium under low-iron conditions. *Proceedings of the National Academy of Sciences*, 109(45).

Shi, T., Sun, Y., and Falkowski, P. G. (2007). Effects of iron limitation on the expression of metabolic genes in the marine cyanobacterium Trichodesmium erythraeum IMS101. *Environmental microbiology*, 9(12):2945–56.

Silva, J., Denny, R., and Dorschel, C. (2005). Quantitative proteomic analysis by accurate mass retention time pairs. *Analytical Chemistry*, 77(7):2187–200.

Silva, J. C., Denny, R., Dorschel, C., Gorenstein, M. V., Li, G.-Z., Richardson, K., Wall, D., and Geromanos, S. J. (2006a). Simultaneous qualitative and quantitative analysis of the Escherichia coli proteome: a sweet tale. *Molecular and cellular proteomics : MCP*, 5(4):589–607.

Silva, J. C., Gorenstein, M. V., Li, G.-Z., Vissers, J. P. C., and Geromanos, S. J. (2006b). Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Molecular and Cellular Proteomics*, 5(1):144–56.

Six, C., Finkel, Z., and Rodriguez, F. (2008). Contrasting photoacclimation costs in ecotypes of the marine eukaryotic picoplankter Ostreococcus. *Limnology and Oceanography*, 53(1):255–265.

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985). Measurement of protein using bicinchoninic acid. *Analytical Biochemistry*, 150(1):76–85.

Sohm, J. and Capone, D. (2006). Phosphorus dynamics of the tropical and subtropical north Atlantic: Trichodesmium spp. versus bulk plankton. *Marine Ecology Progress Series*, 317:21–28.

Sohm, J. a., Subramaniam, A., Gunderson, T. E., Carpenter, E. J., and Capone, D. G. (2011a). Nitrogen fixation by Trichodesmium spp. and unicellular diazotrophs in the North Pacific Subtropical Gyre. *Journal of Geophysical Research*, 116(G3):G03002.

Sohm, J. a., Webb, E. a., and Capone, D. G. (2011b). Emerging patterns of marine nitrogen fixation. *Nature Reviews: Microbiology*, 9(7):499–508.

Staal, M., Meysman, F. J. R., and Stal, L. J. (2003). Temperature excludes N2-fixing heterocystous cyanobacteria in the tropical oceans. *Nature*, 425:504–507.

Stal, L. J. (2009). Is the distribution of nitrogen-fixing cyanobacteria in the oceans related to temperature? *Environmental microbiology*, 11(7):1632–45.

Stephens, W. (1946). Proceedings of the American Physical Society: Minutes of the Meeting of the New York Section at Hamilton. *Physical Review*, 69(11-12):691.

Straub, M., Sigman, D. M., Ren, H., Martínez-García, A., Meckler, a. N., Hain, M. P., and Haug, G. H. (2013). Changes in North Atlantic nitrogen fixation controlled by ocean circulation. *Nature*, 501(7466):200–3.

Straus, N. (2004). Iron Deprivation: Physiology and Gene Regulation. In Bryant, D., editor, *The Molecular Biology of Cyanobacteria SE - 25*, volume 1 of *Advances in Photosynthesis and Respiration*, pages 731–750. Springer Netherlands.

Sugio, S., Hiraoka, B. Y., and Yamakura, F. (2000). Crystal structure of cambialistic superoxide dismutase from porphyromonas gingivalis. *European journal of biochemistry / FEBS*, 267:3487–3495.

Swan, B. and Tupper, B. (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proceedings of the National Academy of Sciences*.

Tang, D. and Morel, F. M. (2006). Distinguishing between cellular and Fe-oxide-associated trace elements in phytoplankton. *Marine Chemistry*, 98(1):18–30.

Thomas, M. K., Kremer, C. T., Klausmeier, C. a., and Litchman, E. (2012). A global pattern of thermal adaptation in marine phytoplankton. *Science*, 338(6110):1085–8.

Thompson, A. W., Foster, R. a., Krupke, A., Carter, B. J., Musat, N., Vaulot, D., Kuypers, M. M. M., and Zehr, J. P. (2012). Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. *Science*, 337(6101):1546–50.

Tilman, D. (1977). Resource Competition between Plankton Algae : An Experimental and Theoretical Approach. *Ecology*, 58(2):338–348.

Tilman, D. (1980). Resources: a graphical-mechanistic approach to competition and predation. *American Naturalist*, 116(3):362–393.

Tovar-Sanchez, a. and Sañudo Wilhelmy, S. a. (2011). Influence of the Amazon River on dissolved and intra-cellular metal concentrations in Trichodesmium colonies along the western boundary of the subtropical North Atlantic Ocean. *Biogeosciences*, 8(1):217–225.

Tovar-Sanchez, A., Sañudo Wilhelmy, S. a., Garcia-Vargas, M., Weaver, R. S., Popels, L. C., and Hutchins, D. a. (2003). A trace metal clean reagent to remove surface-bound iron from marine phytoplankton. *Marine Chemistry*, 82(1-2):91–99.

Tovar-Sanchez, A. and Sañudo Wilhemy, S. (2006). Effects of dust deposition and river discharges on trace metal composition of Trichodesmium spp. in the tropical and subtropical North Atlantic Ocean. *Limnology and Oceanography*, 51(4):1755–1761.

Tripp, H. J., Bench, S. R., Turk, K. a., Foster, R. a., Desany, B. a., Niazi, F., Affourtit, J. P., and Zehr, J. P. (2010). Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium. *Nature*, 464(7285):90–4.

Twining, B. S. and Baines, S. B. (2013). The trace metal composition of marine phytoplankton. *Annual review of marine science*, 5:191–215.

Twining, B. S., Baines, S. B., Bozard, J. B., Vogt, S., Walker, E. a., and Nelson, D. M. (2011). Metal quotas of plankton in the equatorial Pacific Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 58(3-4):325–341.

Tyrrell, T. and Marañón, E. (2003). Large-scale latitudinal distribution of Trichodesmium spp. in the Atlantic Ocean. *Journal of Plankton Research*, 25(4).

Tzeng, C. and Kornberg, A. (1998). Polyphosphate kinase is highly conserved in many bacterial pathogens. *Molecular microbiology*, 29:381–382.

Van Mooy, B. a. S., Fredricks, H. F., Pedler, B. E., Dyhrman, S. T., Karl, D. M., Koblízek, M., Lomas, M. W., Mincer, T. J., Moore, L. R., Moutin, T., Rappé, M. S., and Webb, E. a. (2009). Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature*, 458(7234):69–72.

Van Mooy, B. A. S., Hmelo, L. R., Sofen, L. E., Campagna, S. R., May, A. L., Dyhrman, S. T., Heithoff, A., Webb, E. a., Momper, L., and Mincer, T. J. (2012). Quorum sensing control of phosphorus acquisition in Trichodesmium consortia. *The ISME journal*, 6(2):422–9.

Van Mooy, B. a. S., Rocap, G., Fredricks, H. F., Evans, C. T., and Devol, A. H. (2006). Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments. *Proceedings of the National Academy of Sciences*, 103(23):8607–12.

Vichitphan, K. (2001). *Azotobacter vinelandii Nitrogenase: Effect of Amino-Acid Substitutions at the αGln-191 Residue of the MoFe Protein on Substrate Reduction and CO Inhibition*. PhD thesis, Virginia Polytechnic Institute and State University.

Waldron, K. J. and Robinson, N. J. (2009). How do bacterial cells ensure that metalloproteins get the correct metal? *Nature Reviews: Microbiology*, 7(1):25–35.

Walve, J., Gelting, J., and Ingri, J. (2014). Trace metals and nutrients in Baltic Sea cyanobacteria: Internal and external fractions and potential use in nitrogen fixation. *Marine Chemistry*, 158:27–38.

Ward, B. a., Dutkiewicz, S., Mark Moore, C., and Follows, M. J. (2013). Iron, phosphorus, and nitrogen supply ratios define the biogeography of nitrogen fixation. *Limnology and Oceanography*, 58(6):2059–2075.

Waters Corporation (2011). An overview of the principles of MSE, the engine that drives MS performance. Technical report, Waters Corporation.

Welschmeyer, N. A. (1994). Fluorometric analysis of chlorophyll a in the presence of chlorophyll b and pheopigments. *Limnology and Oceanography*, 39(8):1985–1992.

Wessel, D. and Flügge, U. I. (1984). A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Analytical Biochemistry*, 138(1):141–143.

White, A. E., Spitz, Y. H., Karl, D. M., and Letelier, R. M. (2006). Flexible elemental stoichiometry in Trichodesmium spp. and its ecological implications. *Limnology and Oceanography*, 51(4):1777–1790.

Williams, R. and Follows, M. (2003). Physical Transport of Nutrients and the Maintenance of Biological Production. *Ocean Biogeochemistry*.

Wilson, S., Kolber, Z., and Tozzi, S. (2012). Nitrogen Fixation, Hydrogen cycling. and electron transport kinetics in Trichodesmium Erythraeum (Cyanobacteria) strain IMS101. *Journal of Phycology*.

Wolfe-Simon, F., Grzebyk, D., Schofield, O., and Falkowski, P. G. (2005). The Role and Evolution of Superoxide Dismutases in Algae. *Journal of Phycology*, 41(3):453–465.

Woodward, E. M. S. and Rees, A. P. (2001). Nutrient distributions in an anticyclonic eddy in the northeast Atlantic ocean, with reference to nanomolar ammonium concentrations. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 48:775–793.

Wu, J.-R., Shien, J.-H., Shieh, H. K., Hu, C.-C., Gong, S.-R., Chen, L.-Y., and Chang, P.-C. (2007). Cloning of the gene and characterization of the enzymatic properties of the monomeric alkaline phosphatase (PhoX) from Pasteurella multocida strain X-73. *FEMS Microbiology Letters*, 267:113–120.

Wurl, O., Zimmer, L., and Cutter, G. a. (2013). Arsenic and phosphorus biogeochemistry in the ocean: Arsenic species as proxies for P-limitation. *Limnology and Oceanography*, 58(2):729–740.

Xia, Z., Broadhurst, R. W., Laue, E. D., Bryant, D. a., Golbeck, J. H., and Bendall, D. S. (1998). Structure and properties in solution of PsaD, an extrinsic polypeptide of photosystem I. *European journal of biochemistry / FEBS*, 255(1):309–16.

Xie, S. and Saito, K. (2001). Formation and Variability of a Northerly ITCZ in a Hybrid Coupled AGCM : Continental Forcing and Oceanic – Atmospheric Feedback. *Journal of climate*, pages 1262–1276.

Xu, Q. and Bricker, T. M. (1992). Structural organization of proteins on the oxidizing side of photosystem II. Two molecules of the 33-kDa manganese-stabilizing proteins per reaction center. *The Journal of biological chemistry*, 267:25816–25821.

Yamashita, E., Zhang, H., and Cramer, W. a. (2007). Structure of the cytochrome b6f complex: quinone analogue inhibitors as ligands of heme cn. *Journal of molecular biology*, 370(1):39–52.

Yentsch, C., Yentsch, C., and Perras, J. (1972). Alkaline Phosphatase Activity in the Tropical Marine Blue-green alga, Oscillatoria erythraea ("Trichodesmium"). *Limnology and Oceanography*, pages 772–774.

Yeremenko, N., Kouril, R., Ihalainen, J. A., D'Haene, S., van Oosterwijk, N., Andrizhiyevskaya, E. G., Keegstra, W., Dekker, H. L., Hagemann, M., Boekema, E. J., Matthijs, H. C. P., and Dekker, J. P. (2004). Supramolecular organization and dual function of the IsiA chlorophyll-binding protein in cyanobacteria. *Biochemistry*, 43:10308–10313.

Yoshida, T., Nakamura, H., Masutani, H., and Yodoi, J. (2005). The involvement of thioredoxin and thioredoxin binding protein-2 on cellular proliferation and aging process. *Annals of the New York Academy of Sciences*, 1055:1–12.

Yruela, I. (2013). Transition metals in plant photosynthesis. *Metallomics*.

Zahalak, M., Pratte, B., Werth, K. J., and Thiel, T. (2004). Molybdate transport and its effect on nitrogen utilization in the cyanobacterium Anabaena variabilis ATCC 29413. *Molecular microbiology*, 51(2):539–49.

Zehr, J., Bench, S., Carter, B., Hewson, I., and Niazi, F. (2008). Globally distributed uncultivated oceanic N2-fixing cyanobacteria lack oxygenic photosystem II. *Science*, 322(November):1110–1112.

Ziegler, K., Diener, A., Herpin, C., Richter, R., Deutzmann, R., and Lockau, W. (1998). Molecular characterization of cyanophycin synthetase, the enzyme catalyzing the biosynthesis of the cyanobacterial reserve material multi-L-arginyl-poly-L-aspartate (cyanophycin). *European journal of biochemistry / FEBS*, 254:154–159.