Membrane proteins and protein-protein interactions in marine cyanobacteria

Fallen Kai Yik Teoh

This thesis is presented as partial fulfilment to the requirements for a
Doctorate of Philosophy

Department of Molecular Sciences,

Faculty of Science and Engineering,

Macquarie University, NSW, Australia

May 2019

Table of Contents

List of figures

Chapter 4

List of tables

List of Supplementary info

Declaration

I declare that the work in this thesis was conducted by me under the supervision of Prof. Ian Thomas Paulsen. The assistance and contribution of others have been appropriately acknowledged. The work presented here has not been previously submitted for any other degree.

Fallen Kai Yik Teoh

Acknowledgements

Conference proceedings

Oral presentations

- Teoh, F., Ostrowski, M., Paulsen, I. Comparative membrane proteomics of marine cyanobacteria. AMSA (Australian Marine Science Association), Darwin, July 2017.

Poster presentations

- Teoh, F., Ostrowski, M., Paulsen, I. Adaptation strategies inferred from resource allocation in cyanobacterial membrane proteomes. ASM (The Australian Society for Microbiology), Queensland, July 2018.
- Teoh, F., Ostrowski, M., Paulsen, I. Adaptation strategies inferred from resource allocation in cyanobacterial membrane proteomes. AMSA (Australian Marine Science Association), Sydney, March 2018.
- Teoh, F., Ostrowski, M., Paulsen, I. Comparative membrane proteomics of marine cyanobacteria. ANZSMS26 (Australian and New Zealand Society for Mass Spectrometry Conference), Adelaide, July 2017.

Achievements

Awarded ARC Laureate fellowship PhD scholarship to undertake the Doctor of Philosophy degree.

Contributions

Chapter 2:

Comparative membrane proteomics reveal contrasting adaptation strategies for coastal and oceanic marine *Synechococcus* cyanobacteria. (Manuscript in preparation for submission)

The work was conceptualized by Teoh, Paulsen, Ostrowski. All experimental work including design, troubleshooting, data analysis was conducted solely by Teoh. The manuscript was written by Teoh with contributions from Shah and Paulsen.

Chapter 3:

Protein-protein interactions of marine *Synechococcus* phosphatases and response regulators.

The work was conceptualised by Paulsen, Shah and Teoh. All experimental work and data analysis was conducted by Teoh. Cyanobacterial DNA, the initial starting material was provided from Shah. Primers were designed by Teoh, checked by Shah before sent for synthesis. Size exclusion chromatography was conducted with the assistance from Shah. The thesis writing of Chapter 3 was conducted by Teoh with the advice from Shah and Paulsen.

Chapter 4:

Defining the protein-protein interaction network of protein phosphatases in marine cyanobacteria. (Manuscript in preparation for submission)

The work was conceptualised by Teoh, Paulsen and Shah. All experimental work and data analysis was conducted by Teoh. Cyanobacterial DNA, the initial starting material was provided from Shah. Primers were designed by Teoh, checked by Shah before sent for synthesis. Size exclusion chromatography was conducted with the assistance from Shah. The manuscript was written by Teoh with contributions from Paulsen and Shah.

Abstract

Marine cyanobacteria *Synechococcus* are globally distributed in the world oceans. They play a vital role in primary production and the carbon biogeochemical cycles, as well as significant ecological roles in the ocean ecosystem. The cell membrane is the primary permeability barrier for ions, molecules and/or substrates enter/leave a cell. A comparative membrane proteomic study was conducted on 4 different marine *Synechococcus* namely CC9311 (Clade I), CC9605 (Clade II), WH8102 (Clade III) and CC9902 (Clade IV) representing the 4 most abundant *Synechococcus* clades in the world oceans to investigate adaptation strategies of marine cyanobacteria in nutrient acquisition. Comparative membrane proteomic analyses revealed distinct adaptation strategies among the 4 marine *Synechococcus* in nutrient acquisition. One protein annotated as protein phosphatase 2C was detected to be highly expressed in the coastal mesotrophic strains *Synechococcus* sp. CC9311 and CC9902. This suggests a pivotal, but as yet undefined, functional role for this protein phosphatase in cyanobacteria.

The high expression level of this putative protein phosphatase in coastal *Synechococcus* sparked an interest in further investigation of the functional roles of protein phosphatases in marine cyanobacteria. Therefore, a total of nine putative protein phosphatases including serine/threonine, tyrosine and histidine phosphatases identified in marine *Synechococcus* sp. CC9311 were chosen for a protein-protein interaction study to investigate the proteins that interact with phosphatases and elucidate their functional roles. Pull-down assay was conducted on phosphatases and response regulators, the potential interacting proteins with phosphatases. For the 9 protein phosphatases and 17 response regulators, their genes were PCR amplified, cloned and transformed to test for protein expression and solubility. Two of the phosphatase proteins namely Sync_1857 and Sync_2828, and 9 response regulators were successfully purified to conduct *in vitro* pull-down assays. The pull-down assay coupled with mass spectrometry analysis shows that 7 out of the 9 response regulators interact with both of the phosphatases.

Both purified phosphatase proteins were used to conduct a protein-protein interaction study against the cytoplasmic lysate of *Synechococcus* sp. CC9311 using an affinity purification mass spectrometry approach. Mass spectrometry analyses show that the protein phosphatase Sync_1857 interacts with a periplasmic urea ABC transporter,

suggesting phosphorylation modification may be crucial for the activation of the transport of urea in marine cyanobacteria. Additionally, a broad range of proteins were found to interact with Sync_1857 including proteins involved in amino acid metabolism and translation suggesting it may play a broad role in cell signalling regulation. Meanwhile, Sync_2828 was found to interact with a range of stress-related proteins, suggesting a potential role in stress response regulation in cyanobacterial cells.

Abbreviations

ABC	ATP-binding cassette
ACN	Acetonitrile
ADP	Adenosine diphosphate
AGC	Automatic gain control
AP-MS	Affinity purification mass spectrometry
APC	Allophycocyanin
AqpZ	Aquaporin Z
Asp	Aspartate
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid
C	Carbon
c-region	Carboxyl-terminal region
ca.	Circa
CBP	Calmodulin binding peptide
Cd	Cadmium
CFB	Cytophaga-Flavobacteria-Bacteroides
CHAPS	3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate
Chl a_2	Divinyl chlorophyll a
Chl b_2	Divinyl chlorophyll b
CM	Cytoplasmic membrane
Co	Cobalt
CO_2	Carbon dioxide
CTD	C-terminal domain
Cu	Copper
Cyt b_6f	Cytochrome b_6f
DIP	Dissolved inorganic phosphate
DMF	Dimethyl formamide
DNA	Deoxyribonucleic acid
DOC	Dissolved organic carbon
DOP	Dissolved organic phosphorus
DsPTP	Dual-specificity protein tyrosine phosphatase
DTT	Dithiothreitol

ESI	Electrospray ionization
ETC	Electron transport chain
FA	Formic acid
Fe	Iron
g	Gravity
GSH	Glutathione
GST	Glutathione S-transferase
GTP	Guanosine-5'-triphosphate
HCD	High energy collisional dissociation
His	Histidine
His-GFP	Histidine-green fluorescent protein
HK	Histidine kinase
HL	High light
HNLC	High nutrient-low chlorophyll
HP	Histidine phosphatase
HPLC	High performance liquid chromatography
IAA	Iodoacetamide
ICAT	Isotope-coded affinity tag
IMPs	Integral membrane proteins
IP	Immunoprecipitation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
iTRAQ	Isobaric peptide tags for relative and absolute quantification
ITS	Internal transcribed spacer
LB	Luria-Bertani
LEC	Linear electron flow
LIC	Ligation-independent cloning
LL	Low light
LMM	Low molecular mass
LMW PTP	Low-molecular-weight protein tyrosine phosphatase
MALDI	Matrix-assisted laser desorption ionization
MFS	Major facilitator superfamily
Mg	Magnesium
MgATP	Magnesium adenosine triphosphate

MIP	Major intrinsic protein
Mn	Manganese
MS/MS	Tandem mass spectrometry
m/z	Mass-to-charge
N	Nitrogen
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
NBDs	Nucleotide-binding domains
NDH-I	Nicotinamide adenine dinucleotide phosphate dehydrogenase
NH_4^+	Ammonium
Ni	Nickel
Ni-NTA	Nickel nitrilotriacetic acid
NO_3^-	Nitrate
NO_2^-	Nitrite
NSAF	Normalized spectral abundance factor
OM	Outer membrane
OPP	Oxidative pentose phosphate
OPS	Oxygenic-photosynthesis-specific
ORF	Open reading frame
P	Phosphorus
PBS	Phycobilisome
PC	Phycocyanin
PCB	Phycocyanobilin
PCR	Polymerase chain reaction
PE	Phycoerythrin
PEB	Phycoerythrobilin
PEP	Phosphoenolpyruvate
PHP	Polymerase and histidinol family of phosphoesterase
pI	Isoelectric point
POC	Particulate organic carbon
POTRA	Polypeptide Translocation-Associated
PPIs	Protein-protein interactions
PPM	Metal-dependent protein phosphatase
PPP	Phosphoprotein phosphatase
PRC	Photosynthetic reduction of carbon

PSP	Protein serine/threonine phosphatase
psi	Pounds per square inch
PSI	Photosystem I
PSII	Photosystem II
PTM	Post-translational modifications
PTP	Protein tyrosine phosphatase
PTS	Phosphotranferase
PUB	Phycourobilin
RC	Reaction centre
REC	Receiver
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RR	Response regulator
RRI	Response regulator class I
RRII	Response regulator class II
RRIII	Response regulator class III
RRIV	Response regulator class IV
rRNA	Ribosomal ribonucleic acid
S	Sulfur
S-layer	Surface layer
SDH	Succinate dehydrogenase
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sec	Secretory
SEC	Size exclusion chromatography
Ser	Serine
SILAC	Stable isotope labelling by/with amino acids
SLH	Surface layer homology
SOB	Super optimal broth
SOC	Super optimal broth with catabolite repression
SOW	Synthetic Ocean Water
SpC	Spectral count
SRP	Signal recognition particle
Tat	Twin arginine translocation

TBDTs	TonB-dependent transporters
TCEP	Tris(2-carboxyethyl) phosphine
TCRS	Two-component regulatory system
TE	Terminal energy
Thr	Threonine
TM	Thylakoid membrane
TMDs	Transmembrane domains
TMT	Tandem mass tag
tRNA	Transfer ribonucleic acid
Tyr	Tyrosine
X-gal	5-bromo-4-chloro-3-indolyl-β-D-1-thiogalactopyranoside
YFP	Yellow fluorescent protein
Zn	Zinc

Chapter 1

Introduction

1.1 Cyanobacteria

Cyanobacteria are a large phylum of prokaryotes that existed at least 2.5 billion years ago (Kasting, 2001, Kasting and Siefert, 2002, Noffke, 2008, Blank and Sanchez-Baracaldo, 2010). They are the first organisms to carry out oxygenic photosynthesis which resulted in the transition of the biosphere from anoxic to oxic (Des Marais, 2000, Falkowski, 2006) and played a vital role in the evolution of the life on Earth (Kasting and Siefert, 2002; Demoulin et al. 2019).

Cyanobacteria, first described in 1829 by Gomont (Collins, 1891), constitute a single but large phylogenetically diverse group within the domain Eubacteria (Castenholz and Waterbury, 1989). The genome size of cyanobacteria varies, ranging from 1.6Mb to 11.6 Mb (Shih et al. 2013). It is estimated that around 8000 species of cyanobacteria exist in the world, with more than half of the species remaining to be described (Guiry, 2012).

Cyanobacteria utilize water soluble phycobiliproteins and chlorophyll *a* as well as photosystem I (PSI) and II (PSII) to perform oxygenic photosynthesis (Anderson and Toole, 1998). They have a membranous compartment called the thylakoid containing the PSI and PSII reaction centres where photosynthesis occurs. The reaction centres capture solar energy to reduce electron acceptors, ultimately producing the reduced nicotinamide adenine dinucleotide phosphate (NADPH). NADPH can be used as reducing agent in various biosynthetic processes, including CO_2 fixation via the Calvin-Benson cycle.

1.1.1 Significance of cyanobacteria

Cyanobacteria have been found in diverse habitats on Earth including environments that are hot and cold, acidic and alkaline, marine, freshwater, saline, dry, terrestrial etc. (Stainer and Cohen-Bazire, 1977, Whitton, 2012). They are one of the most morphologically distinct groups among prokaryotes (Huang et al., 2002, Scanlan et al., 2009, Wegener et al., 2010, Takahashi et al., 2011) and marine cyanobacteria are major contributors to ocean biogeochemical cycling, playing a pivotal role in primary production to support resident planktonic food webs (Armengol et al. 2019). Moreover, they are the progenitors of chloroplasts in higher plants (Kasting, 2001, Kasting and Siefert, 2002) and important in the evolutionary study of living organisms on the Earth. In addition, cyanobacteria are economically important and have shown their potential applications in diverse fields such as pharmaceutical (Koehn et al., 1992, Jaki et al., 2000, Romay et al., 2003), food (Gantar

and Svircev, 2008), biomedical (Rastogi et al., 2015), biomaterial (Sudesh et al., 2002), bioremediation (Abed and Köster, 2005), biofuel (Lindblad, 1999, Antal and Lindblad, 2005), and biofertilizer (Kaushik and Venkataraman, 1979) industries.

1.2 Marine picocyanobacteria

Marine picocyanobacteria are numerically the most abundant primary producers on Earth, consisting of two distinct genera, *Prochlorococcus* and *Synechococcus* (Scanlan et al., 2009). The genomes of marine picocyanobacteria are comprised of a single circular chromosome without plasmids. Genome size of *Prochlorococcus* ranges from 1.64 to 2.7Mb meanwhile *Synechococcus* ranges from 2.2 to ~2.86Mb (Scanlan et al., 2009). The small size of picocyanobacterial genomes is believed to have provided a strong selective advantage for marine picocyanobacteria to colonize oligotrophic ocean waters as it allows more economical energy consumption for nutrient uptake (Strehl et al., 1999). Besides that, small size of picocyanobacteria cells (high surface area to volume ratio) allows them to maximize nutrient uptake and to extract nutrients at the minuscule bulk concentrations, characteristic of the ocean environment (Perez-Sepulveda et al., 2018).

Although *Prochlorococcus* strains are a sister clade to *Synechococcus*, these two cyanobacteria exhibit distinct ecology, physiology and distribution (Partensky et al., 1999, Partensky and Garczarek, 2010). For example, they utilise distinct photosynthetic apparatus to harvest light (Li and Wood, 1988, Olson et al., 1990, Campbell et al., 1994). *Prochlorococcus* employs divinyl chlorophylls *a* and *b* (chl a_2 and chl b_2) light-harvesting antenna whereas *Synechococcus* employs the phycobilisome (PBS) as its major light harvesting antenna (Waterbury et al., 1986, Scanlan et al., 2009). P*rochlorococcus* and *Synechococcus* differ in their ecology and biogeographical distribution patterns. Although, these organisms co-occur in oceanic waters, *Prochlorococcus* typically dominates tropical and subtropical oligotrophic waters between the latitudes of 45^0N and 45^0S (Campbell et al., 1994, Partensky et al., 1999), while, *Synechococcus* has a widespread distribution from the equator to the polar regions, and from coastal waters to open oceans (Six et al., 2007, Pittera et al., 2014, Sohm et al. 2016, Pittera et al., 2017). *Synechococcus* is often found in the upper mixed layer in the ocean, while *Prochlorococcus* shows a strong depth partitioning. The high light (HL)-adapted ecotype of *Prochlorococcus* mostly dominates

the upper part of the euphotic zone while the low light (LL)-adapted ecotype occupies the bottom of the euphotic layer (Moore et al., 1998).

1.3 *Synechococcus* **spp.**

Marine cyanobacteria from the genus *Synechococcus* are important contributors to global primary productivity and biogeochemical cycling of carbon (Glover et al., 1985, Li, 1994, Agawin et al., 2000). Marine *Synechococcus* spp. contribute 5 to 30% of primary production in the surface waters of world's oceans (Waterbury et al., 1979, Chen et al., 2004). *Synechococcus* spp. are small (ca. 0.6-2.1μm) coccoid or rod shape cells (Ernst et al., 2003) that divide by transverse binary fission in a single plane (Waterbury and Rippka, 1989). They are often found either as solitary cells or in small clusters or pairs (Stockner et al., 2000).

Previous studies have found that marine *Synechococcus* spp. exhibited two modes of cell cycle regulation (Armbrust et al., 1989, Binder and Chisholm, 1990, Binder and Chisholm, 1995, Liu et al., 1999). The first type is consistent with the slow-growth paradigm of the prokaryotic cell cycle, known as bimodal cell cycle. The second type is asynchronous division which involve asynchronous initiation of chromosome replication, the presence of multiple chromosome copies at low growth rates and variability in chromosome copy number among cells in the population.

Synechococcus contains light harvesting phycobilisomes made up of phycobiliproteins such as phycocyanin (PC), allophycocyanin (APC) and phycoerythrin (PE). *Synechococcus* can be divided into three major pigment types based on their phycobilisomes composition: type I contains only phycocyanobilin (PCB), type II comprises of both PCB and phycoerythrobilin (PEB), while type III contains PCB, PEB and phycourobilin (PUB) (Six et al., 2007). *Synechococcus* consists of strains rich in the pigment phycoerythrin (PE), rendering its representatives a variety of range, reddish, pink and purple colours, and strains rich in phycocyanin (PC), giving the organisms blue-green colour in appearance. It can be easily differentiated from other small picophytoplankton by epifluorescence microscopy or flow cytometry as it fluoresces orange when excited by blue light (Waterbury et al., 1979, Murphy and Haugen, 1985, Olson et al., 1988).

1.3.1. Distribution of *Synechococcus* spp. in the marine environment

Synechococcus spp. are widespread in the global ocean including estuarine, coastal and open ocean environments (Dufresne et al., 2008, Scanlan et al., 2009, Sohm et al., 2016, Kent et al., 2019). Previous studies have shown that *Synechococcus* populations are higher in upwelling and deep mixing regimes as well as over continental shelves (Olson et al., 1990, Lindell and Post, 1995, Moore et al., 1995, DuRand et al., 2001, Zwirglmaier et al., 2007), which are nutrient-rich environments compared to the oligotrophic ocean gyres, predominated by *Prochlorococcus* (Partensky et al., 1999, Johnson et al., 2006).

Genus *Synechococcus* is polyphyletic and genetically highly diverse (.1.1). Based on 16S rRNA sequences, marine *Synechococcus* members form a well-defined clade termed cluster 5, which is divided into 3 sub-clusters: 5.1, 5.2 and 5.3 (Scanlan et al., 2009, Huang et al., 2012). Sub-cluster 5.1 is the dominant *Synechococcus* group with phycoerythrin (PE) enriched strains, comprising more than 20 taxonomically distinct clades based on a range of conserved markers, such as the 16S-23S internal transcribed spacer region (ITS), *rpoC*, *cpeAB*, *narB* and *ntcA* genes (Palenik, 1994, Toledo and Palenik, 1997, Ferris and Palenik, 1998, Fuller et al., 2003, Ahlgren and Rocap, 2006, Dufresne et al., 2008, Scanlan et al., 2009, Huang et al., 2012,).
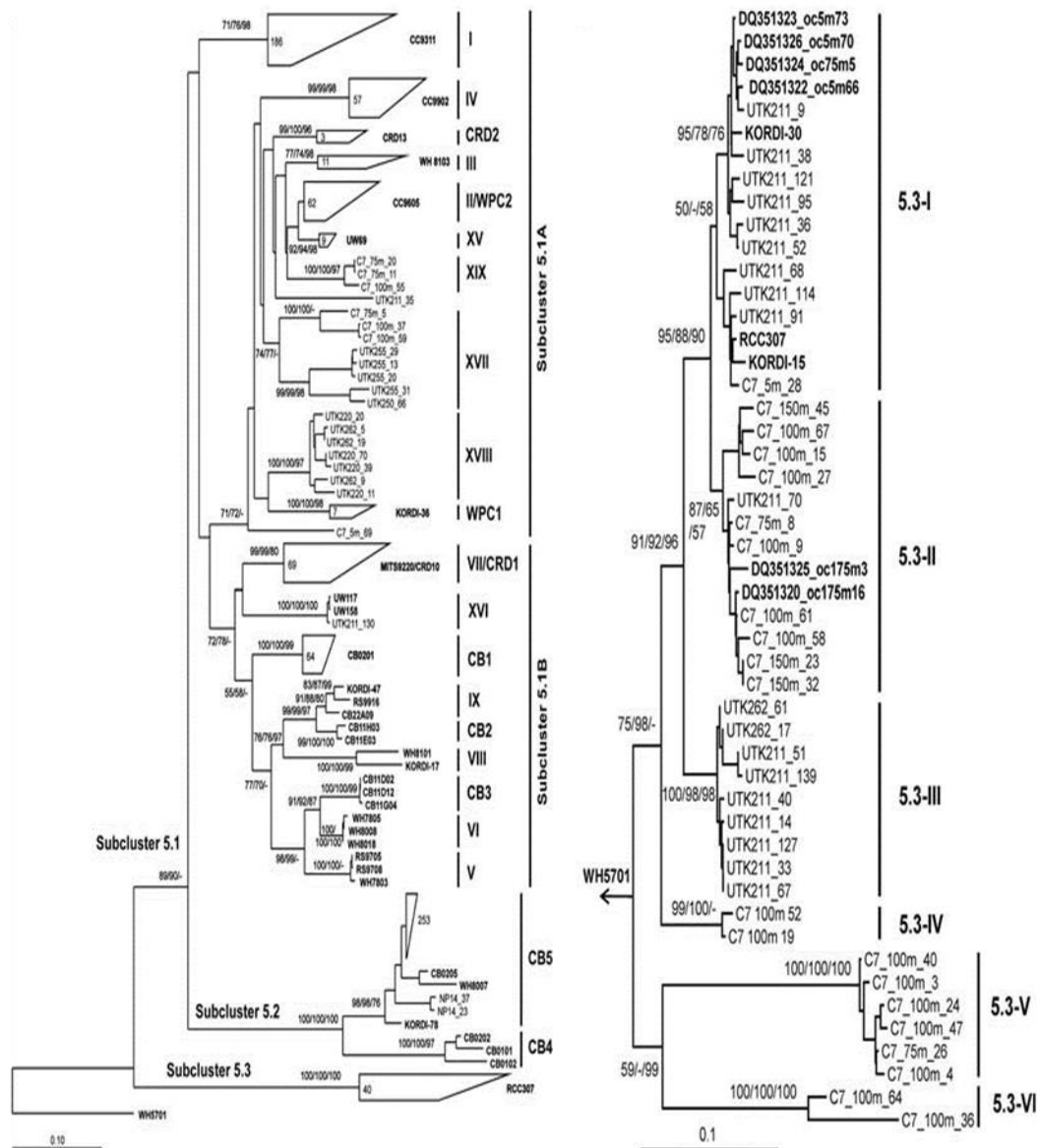
Figure 1.1 Phylogenetic tree of marine cyanobacteria *Synechococcus* spp.

Phylogenetic tree based on 16S-23S rRNA ITS sequences (866bp, without tRNAs), showing the relationships among *Synechococcus* genotypes. (reproduced from Huang et al. 2012, copyright license number: 4577940220592).

Clades I and IV of sub-cluster 5.1 predominate in coastal and higher-latitude regions (Figure 1.2) (Ferris and Palenik, 1998, Toledo and Palenik, 2003, Zwirglmaier et al., 2007, Zwirglmaier et al., 2008). Clade III members are abundant in tropical and subtropical oceanic gyres, often co-occurring with clade II, which is the most broadly distributed *Synechococcus* species (Ferris and Palenik, 1998, Toledo and Palenik, 2003, Ahlgren and Rocap, 2006, Fuller et al., 2006, Penno et al., 2006, Zwirglmaier et al., 2008, Scanlan et

6

al., 2009, Tai and Palenik, 2009). Clades V, VI, VII, VIII (Dufresne et al., 2008) and IX are generally detected in low abundance (Zwirglmaier et al., 2008, Post et al., 2011). Meanwhile, clades XI-XIV have been detected in region such as the Gulf of Aqaba (Penno et al., 2006, Post et al., 2011), clades XV-XVI in Sargasso Sea (Ahlgren and Rocap, 2006). Clade CRD1 and CRD2 are highly abundant in the Costa Rica dome, equatorial Pacific and subtropical North Atlantic, which are upwelling waters with low iron availability (Huang et al., 2012, Ahlgren et al., 2014). Clades WPC1 and WPC2 have been found distributed in East China and East Sea (Choi and Noh, 2009), and clades CB1-CB3 in the summer season of Chesapeake Bay (Chen et al., 2006, Cai et al., 2010).
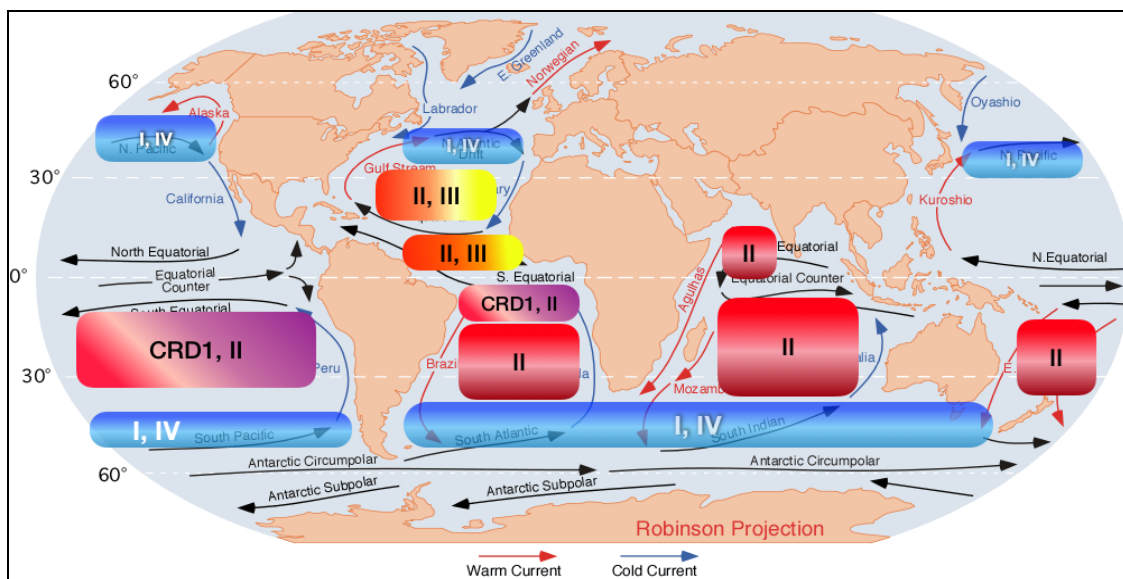
Figure 1.2 Simplified model of the global distribution of dominant *Synechococcus* spp. clades in different oceanic provinces.
Clades I, II, III and IV are found to be the most abundant clades of *Synechococcus* spp. sampling around the world oceans. (Ostrowski et al. unpublished).

Little is known about the biogeography of sub-cluster 5.2 and 5.3 compare to 5.1. Sub-cluster 5.2 typically comprise phycocyanin-enriched (PC-type), euryhaline strains, which have been observed in coastal and estuarine environment, such as the Chesapeake Bay (Chen et al., 2006, Cai et al., 2010) and Hong Kong (Xia et al., 2015). Sub-cluster 5.3 members are phycoerythrin-enriched (PE-type) strains, found in open ocean such as East China, Mediterranean, Sargasso and South China Seas (Ahlgren and Rocap, 2006, Dufresne et al., 2008, Choi and Noh, 2009, Huang et al., 2012).

1.3.2 Factors influencing the distribution of *Synechococcus* spp.

Synechococcus spp. distribution is influenced by abiotic parameters including light, temperature and nutrients, and biotic factors such as viral lysis and grazing. However, these factors can also intercept, for example, a well-mixed water body is generally colder and richer in nutrients than a stratified water body and may also have different grazing communal members. It is therefore difficult to differentiate important factors influencing the distribution and colonization of *Synechococcus* spp. in marine ecological niches.

In addition to spatial differentiation, the composition of *Synechococcus* communities has been shown to vary seasonally (Flombaum et al., 2013, Sohm et al., 2016, Varkey et al., 2018). The Gulf of Aqaba is dominated by clade II *Synechococcus* during spring bloom and subsequently shifts to clade III and XII in summer. Clade III abundance is low in winter while clade XII continues to dominate in winter (Post et al., 2011). A similar seasonal shift is observed in the Sydney Harbour estuary where the predominant *Synechococcus* lineages change from clades I and IV in late austral winter to clade II in austral summer (Varkey et al., 2018). To gain insights into how marine *Synechococcus* has become a widespread colonizer of the oceans, it is critical to understand the factors that define niche selection in different lineages.

1.3.2.1 Abiotic

Abiotic factors that influence the distribution of *Synechococcus* spp. include macronutrients, micronutrients or trace metals, temperature, light and salinity. Stratification can also be one of the factors that influence their distribution, however, it is not discussed in this Chapter as one distinct section, considering this factor ultimately influence the nutrient concentrations which will be discussed in sections 1.3.2.1a and 1.3.2.1b.

1.3.2.1a Macronutrients (Nitrogen, Phosphorus, Sulfur)

The succession of marine cyanobacteria in the oceanic waters is highly correlated to nutrient availability (Post et al., 2002; Sohm et al. 2016). Carbon (C), nitrogen (N), phosphorus (P) and sulfur (S) are essential macronutrients needed for sustaining

proliferation of cyanobacteria. Carbon is readily available to cyanobacteria from dissolved carbon dioxide in the oceanic water and is easily diffusible into their cells.

The ocean N and P cycles are complex as this is influenced by several factors including vertical water mixing, cyclonic or anticyclonic eddies (Twining et al., 2010) and biological processes that may change the availability of N and P resources. Although the average C:N:P ratio of oceanic particulate matter is close to the Redfield ratio (molar C:N:P of 106:16:1) (Redfield, 1958), the ratio varies noticeably in selected oceanic regions, at depths, and through seasons (Copin-Montegut and Copin-Montegut, 1983, Hebel and Karl, 2001). The seasonal succession of different *Synechococcus* strains had been associated with local and temporal variation in elemental resources (Post et al., 2011), suggesting that *Synechococcus* strains have distinct requirements and uptake strategies for nutrients.

Vast areas of the open ocean are nutrient poor (oligotrophic), containing low levels of N, P and iron which can be limiting for carbon fixation and growth (Beardall et al., 2001). The bio-availability of nitrogen is a key environmental factor governing photosynthetic activity in diverse open ocean and coastal marine ecosystems. Light harvesting antenna – PBS, is an N-rich protein complex representing a major cellular investment which constitutes up to 50% of the soluble proteins in *Synechococcus* cells (Grossman et al., 1993, Apt et al., 1995). Therefore, the distribution of marine *Synechococcus* spp. is also linked to N availability (Olson et al., 1990, Partensky et al., 1999). *Synechococcus* population density is high in the transition zone of equatorial Pacific and high nutrient-low chlorophyll (HNLC) region in which concentrations of nitrate and nitrite is high (Partensky et al., 1999). Marine *Synechococcus* spp. lack the ability to fix molecular dinitrogen like other marine cyanobacteria such as *Trichodesmium* and *Crocosphaera* (Scanlan et al., 2009). However, they are reported to utilize a broad range of N resources such as nitrate (NO_3^-), nitrite (NO_2^-), ammonium (NH_4^+), urea and amino acids, employing systems associated with active uptake of these compounds (Gilbert et al., 1986, Paerl, 1991, Lindell and Post, 1995, Collier et al., 1999).

Nitrate is the most abundant dissolved form of inorganic nitrogen in global oceans (Paerl et al., 2008). Nitrate concentrations are high in deep ocean and coastal areas, while it is low in surface waters of open ocean (Paerl et al., 2008). Nitrate can be assimilated by most marine *Synechococcus* spp. (Moore et al., 2002, Fuller et al., 2003, Scanlan et al., 2009) via nitrate reductase encoded by the *narB* gene (Rubio et al., 1996). However, some

Synechococcus strains such as RS9917 and PROS-7-1 lack the *narB* gene and nitrate transporters (Dufresne et al., 2008).

Eddy and advection events make important contributions to the proliferation of *Synechococcus* in the open ocean upper surface (Falkowski et al., 1991, McGillicuddy et al., 2003, Palter et al., 2005), as these events cause fluctuations in nitrate concentrations which stimulate phytoplankton growth (Glover et al., 2007). *Synechococcus* spp. blooms have been observed in the Sargasso Sea annually due to nutrient enhancement from deep mixing (DuRand et al., 2001) and small nitrate perturbations into stratified surface waters (Glover et al., 1988). They showed higher nitrate uptake and enhanced photosynthesis in response to small nitrate perturbations. In N deprivation conditions, *Synechococcus* can degrade the PBS proteins (Wyman et al., 1985, Grossman et al., 1993), which act as internal N storage.

Open ocean surface water is depleted of low dissolved inorganic phosphate (DIP) and hence may limit phytoplankton growth (Wu et al., 2000, Karl and Bjorkman, 2001, Lomas et al., 2004, Thingstad et al., 2005). During summer and fall, surface water of the Atlantic Ocean is particularly depleted of dissolved inorganic nitrogen and phosphorus which are present in just nanomolar levels (Steinberg et al., 2001). P is an important building block of DNA, RNA, lipids and intracellular energy (ATP) in cells with diverse roles in cellular metabolic pathways. During phosphorus stress, *Synechococcus* showed increase production of alkaline phosphatases for phosphate utilization (Tetu et al., 2009). Oceanic *Synechococcus* spp. have high affinity uptake systems for both nitrate and phosphate (Moutin et al., 2002). Their nutrient uptake kinetics showed that they can outcompete heterotrophic bacteria and algae to obtain both N and P resources (Moutin et al., 2002), thus making them well adapted to live in the oligotrophic open ocean.

Sulfur is an essential component of proteins and sulfolipids in cyanobacteria. There is currently limited research on the sulfur requirements of marine picocyanobacteria. However, a study on euryhaline strains of *Synechococcus* sp. PCC7002 showed that the expression of outer-membrane porins SomA/SomB was increased under conditions of nitrogen and sulfur starvation (Sauer et al., 2001). High expression of porins has been suggested to enhance cells' ability to scavenge nutrients from their surroundings (Sauer et al., 2001, Schwarz and Forchhammer, 2005).

1.3.2.1b Micronutrients (Iron and trace metals)

Trace metals play a critical role in the metabolism of cyanobacteria, thus their availability influences the *Synechococcus* spp. community structure, physiology and the cycling of carbon and nutrient elements in the oceanic water (Saito et al., 2004, Palenik et al., 2006, Stuart et al., 2009, Twining et al., 2010, Moore et al. 2013). In contrast to heterotrophs, cyanobacteria have an additional requirement for iron as the photosynthetic apparatus comprises several iron-dependent electron transport components. During photosynthesis, a linear electron transport chain requires approximately 24 atoms of iron (Michel and Pistorius, 2004). Under iron stress condition, the overall photosynthetic capacity of cyanobacteria decreases, with a greater decline of electron transport capacity in PSI than PSII and Cyt b_6f, owing to its higher demand for iron (three 4Fe-4S centers (12 Fe atoms in total)) (Schrader et al., 2011).

The availability of iron is limited in the open ocean and some coastal/continental shelf waters (Behrenfeld and Milligan, 2013). The prevalence of several *Synechococcus* strains such as clades II and III in iron-depleted oligotrophic open water suggests the use of adaptation mechanisms to cope with low iron availability (Sohm et al. 2016). Strain WH8102 (clade III) maintains lower levels of PSI to reduce iron requirement (Bailey et al., 2008). Furthermore, this strain does not contain *isiA* gene, which functions as an antenna for PSI under iron deplete conditions (Rivers et al. 2009). Therefore, instead of increasing PSI activity through the development of larger antenna complexes to overcome the low levels of PSI relative to PSII, WH8102 tends to bypass electron flow to PSI via an oxidase-dependent electron valve (Bailey and Grossman, 2008). This helps maintain the photochemical activity of PSII at high irradiance levels when the passage of electrons through PSI become limiting. Meanwhile, *Synechococcus* strains in HNLC (high nitrate, low chlorophyll) regions such as clade CRD1 have high PSII:PSI ratio and increased expression level of IsiA protein in excess of PSI (Schrader et al., 2011).

Zinc (Zn) is needed as a cofactor for phosphatase enzymes such as alkaline phosphatase, a metalloenzyme which cleaves phosphor-monoester bonds enabling microorganisms to acquire phosphorus from organic P compounds under low availability of inorganic orthophosphate (Shaked et al., 2006). In the Pacific and Atlantic oceans, where both phosphate and Zn concentrations are low (Bruland, 1980, Wu et al., 2000, Lohan et al., 2002), the synthesis of alkaline phosphatase enzymes might be limited, thus restricting orthophosphate regeneration through organic compounds or enzymatic hydrolysis and

influences the proliferation of phytoplankton (Shaked et al., 2006). However, the open ocean strain *Synechococcus* sp. WH8102 has the capability to growth in Zn-depleted media (Barnett et al., 2014), perhaps owing to its high efficiency of zinc uptake via the zinc ABC transporter periplasmic binding protein, ZnuA.

Marine cyanobacteria are sensitive to copper (Cu) (Mann et al., 2002, Stuart et al., 2009) and cadmium (Cd) toxicity (Payne and Price, 1999). Copper is a vital component in respiratory proteins and oxidases e.g. plastocyanin and cytochrome c_6 in cyanobacteria (Giner-Lamia et al., 2012). However, free Cu^{2+} concentrations found in natural environments can have negative effects on phytoplankton including the reduction of the photosynthetic rates (Baron et al., 1995), cell division (Brand et al., 1986), interference of utilisation of other essential trace metals such as manganese (Sunda, 1989, Sunda and Hunstman, 1998) and disruption of enzyme function by producing hydroxyl radicals and binding to −SH groups (Stauber and Florence, 1985, Brown et al., 1994) etc.

Copper mostly occurs as organic complexes in seawater (Coale and Bruland, 1988, Kozelka and Bruland, 1998, Tang et al., 2001), however, while free Cu^{2+} concentrations are commonly below 1pM in both open ocean and unpolluted coastal waters, however, higher concentrations have been found in some coastal regions (Moffett et al., 1997, Kozelka and Bruland, 1998, Tang et al., 2001). *Synechococcus* spp. encounter the influx of free cupric ions in cells by producing strong organic copper binding ligands to avoid cell damage (Moffett, 1995, Moffett et al., 1997, Mann et al., 2002).

The case studies in the Costa Rica Dome showed that *Synechococcus* can be co-limited by cobalt (Co) and Fe (Saito et al., 2005, Ahlgren et al., 2014). Co is believed to be the potential substitute for Zn-based metalloenzymes (Cullen et al., 1999, Saito et al., 2004). However, the cobalt-zinc substitution of carbonic anhydrase has been found in eukaryotic phytoplankton but not in cyanobacteria (Sunda and Hunstman, 1995, Saito et al., 2002). As opposed to eukaryotic phytoplankton species, it is suspected that the requirement for cobalt in *Synechococcus* is more highly connected to vitamin B_{12} biosynthesis (Saito et al., 2002, Saito et al., 2003). The ability of marine *Synechococcus* spp. to utilize organic cobalt ligand complexes (Ahlgren et al., 2014) additionally provides them competitive advantage to eukaryotic phytoplankton such as diatoms in the Co-limited ocean.

Nickel (Ni) quotas vary between *Synechococcus* strains (Dupont et al., 2008). Overall, Ni requirement is higher in the *Synechococcus* spp. inhabiting the surface water gyres than

those at depths due to the increased expression of Ni superoxide dismutase to overcome the oxidative stress from reactive oxygen species (ROS) produced by light-dependent photochemical reactions (Twining et al., 2010).

Manganese (Mn) is a vital transition metal for all living organisms (Frausto da Silva and Williams, 2001, Hansch and Mendel, 2009), playing an important role as a redox-active cofactor in cyanobacterial cells and actively involved in PSII light-dependent reactions (Brandenburg et al., 2017). In PSII, Mn is incorporated into a Mn_4O_5Ca cluster, which mediates the splitting of water into oxygen, protons, and electrons. Thus, Mn limitation can influence primary productivity, (Salomon and Keren, 2011) while elevated concentrations of Mn is harmful to cyanobacteria (Brandenburg et al., 2017). Mn concentration fluctuates in the ocean and varies with depth and seasons (Bruland et al., 1994, Morel, 2008, Sunda, 2012). Mn transporter, Mnx, has been identified to play an important role in Mn homeostasis of freshwater *Synechocystis* sp. PCC6803 strain (Brandenburg et al., 2017). However, the mechanism involved in Mn acquisition and intracellular homeostasis in marine *Synechococcus* has not been defined yet.

1.3.2.1c Temperature

Synechococcus spp. colonize a wide latitudinal range from tropical to polar regions, with major ecotypes displaying different temperature niche preferences (Zwirglmaier et al., 2008, Pittera et al., 2014, Sohm et al. 2016, Varkey et al., 2016). Since temperature is a major factor that controls growth and photosynthetic rates (Huner et al., 1998), these temperature ecotypes require specific physiological adaptations that enable regulation of photochemistry under different temperature conditions.

Mackey et al. (2013) showed that temperature adaptation in tropical strain *Synechococcus* sp. WH8102 involved state transitions and change in the abundance of photosynthetic proteins (Mackey et al., 2013). When grown at higher temperature, WH8102 elevated major photosynthetic proteins involved in the electron transport chain (ETC), such as subunits of PSI, PSII and cytochrome b_6f, and their light-harvesting PBS pigment proteins (Mackey et al., 2013). This increase in the number of ETC proteins suggests the need to boost overall photosynthetic electron flux to meet metabolic demands for NADPH during rapid growth. At optimal temperature, higher rates of C fixation and N assimilation generate a stronger need for reductant, hence *Synechococcus* spp. has the highest demand

for NADPH to maintain cellular metabolic needs. In contrast, the demand for NADPH decrease under low temperature, linear ETC declines and the PBS associates with photosystem I.

Another adaption, regulation of membrane fluidity is also a strategy of *Synechococcus* spp. for acclimating to variations in temperature (Varkey et al., 2016, Pittera et al., 2018). In contrast to freshwater cyanobacteria, marine *Synechococcus* sp. WH7803 membranes are almost devoid of fatty acid C18 chains, comprising mainly C14 and C16 chains (Merritt et al., 1991) with no more than two unsaturation per chain (Pittera et al., 2018). In cold environments, shortening of *Synechococcus* sp. WH7803 acyl chains induces membrane thinning, along with cold-induced acyl desaturation, enhancing membrane fluidity of cells (Pittera et al., 2018). These activities are believed to help maintaining efficient photosynthetic activity and enhance competitiveness of marine *Synechococcus* spp. to proliferate in the ocean during seasonal change. In addition, it is suspected that xanthophylls play a role in *Synechococcus* adaptation to temperature change in its living environment. In cyanobacterium *Cylindrospermopsis raciborskii,* xanthophylls are shown to increase when growing at low temperature to compensate for elevated lipid desaturation and alter membrane dynamics (Varkonyi et al., 2002). However, it is unclear whether this applies to marine *Synechococcus* strains growing at low temperatures.

Previous studies showed that temperate strains have better tolerance to temperature fluctuations, for example clade I strains predominantly seen in cold, mesotrophic waters can also grow in tropical waters (Ahlgren et al., 2014) at temperatures up to 30^0C (Pittera et al., 2014). In contrast, clade III strains dominant in tropical waters has a narrow temperature spectrum (Pittera et al., 2014), cannot withstand colder living environments (Pittera et al., 2014).

1.3.2.1d Light

The versatility of marine *Synechococcus* spp. has been related to its ability to grow over a wide range of habitats where the light intensities and spectral quality can vary dramatically over relatively short time-scales (Kana and Gilbert, 1987, Waterbury et al., 1986, Olson et al., 1990, Li, 1994, Partensky et al., 1999, Scanlan et al., 2009, Li and Wood, 1988, Crosbie et al., 2003). Light intensity differs with depth of the water column, with distance from the coast and also with biotic factors. In surface waters the light intensity can reach

up to 2,000 μmol quanta m^{-2}s^{-1} (Mackey et al., 2013). To acclimate to the change of light intensities and spectrum, *Synechococcus* spp. alter the composition or abundance of PBS, that is low light intensities stimulate the synthesis of PBS and extension of the length of the rods of PBS. Moreover, *Synechococcus* spp. can increase their cellular phycoerythrin levels to enhance light absorption capacity. Furthermore, PBS associates with PSII, elevating the input of light from PBS antenna into PSII reaction center for photosynthetic activity (Kana and Gilbert, 1987).

Synechococcus spp. of sub-cluster 5.1 and 5.3 contain phycoerythrin (PE) as a major photosynthetic pigment, making them more adapted to blue-green light and enabling their predominance in open ocean and coastal waters (Ong and Glazer, 1991, Morel et al., 1993, Moore et al., 1995, Six et al., 2007, Dufresne et al., 2008). In contrast, phycocyanin (PC)-rich strains of sub-cluster 5.2 are more abundant in highly turbid estuarine waters in which red and orange light prevail (Stomp et al., 2007).

Chromophore phycoerythobilin (PEB) is found in all PE-rich *Synechococcus* (Wood et al. 1999), but phycourobilin (PUB) is not found in all marine cyanobacteria (Yona et al. 2014). The ratio of PUB to PEB varies between strains: that is, open ocean strains have high PUB:PEB ratio (Six et al., 2004) while, coastal strains have low PUB:PEB ratio (Palenik, 2001) or no PUB (Haverkamp et al., 2008). PEB pigments enable *Synechococcus* to absorb green wavelength light (~500-550nm) (Ong and Glazer, 1991, Morel et al., 1993, Moore et al., 1995), hence PEB is higher in strains that are prevalent in coastal waters where green light predominates. In the open ocean, surface waters (0-30m) are illuminated primarily with blue and green wavelengths, however, they seem to have high preference for blue light by maintaining high levels of PUB (Campbell and Iturriaga, 1988, Olson et al., 1990, Wood et al., 1998). *Synechococcus* relies on light to survive, hence, changing the spectral quality of its' environment represents a selective factor expected to foster evolution of phenotypes adapted to different spectral environments (Wood, 1985, Stomp et al., 2004).

1.3.2.1e Salinity

Salinity has been suggested to affect the growth, photosynthetic capacity and nitrogenase activity of planktonic cyanobacteria (Kim et al., 2018). Most sub-cluster 5.1 PE-type *Synechococcus* strains are sensitive to salinity changes with their population density

positively correlating with salinity (Herdman et al., 2001, Xia et al., 2015). An exception to this is *Synechococcus* MW02 of clade IX, isolated from Hong Kong waters (Xia et al., 2015), which has the ability to survive in low-salinity waters. The ability to tolerate fluctuations in salinity is primarily seen in *Synechococcus* subcluster 5.2, which are prevalent in euryhaline environments (Fuller et al., 2003, Xia et al., 2015).

1.3.2.2 Biotic

The distribution and population density of *Synechococcus* spp. is not solely manipulated by abiotic factors, but also biotic factors including grazing by protists, viral infection and symbiosis with other organisms in the ocean waters.

1.3.2.2a Grazing

As an important phytoplankton underpinning marine food webs, *Synechococcus* is grazed upon by heterotrophic protists (Quevedo and Anadon, 2001, Calbet et al., 2008, Apple et al., 2011). There is some controversy about the primary grazers of *Synechococcus* spp: that is, some studies have found that small, heterotophic nanoflagellates are the primary consumers of *Synechococcus* (Caron et al., 1999, Christaki et al., 2001, Sato et al., 2007), while others have suggested that dinoflagellates, small (<20-μm) aloricate ciliates, or larger ciliates and appendicularians are the major contributors to the grazing mortality rates of *Synechococcus* (Dolan and Simek, 1999, Quevedo and Anadon, 2001).

The relationships between grazers and picocyanobacteria can be quite complicated as grazers diminish picophytoplankton biomass, relaxing nutrient competition. Simultaneously, they are also involved in nutrient regeneration, increasing availability of nutrients in the open ocean, hence promoting the growth rates of phytoplankton. The impact of grazing is highly dependent on growth rates of *Synechococcus* members (Xia et al., 2015). During summer, the population density of *Synechococcus* remains high since growth rate exceeds grazing mortality, while in winter, the grazing significantly impacts *Synechococcus* members due to slow growth rates being unable to keep pace with the grazing rates.

1.3.2.2b Virus infection

Besides the threat of grazers, marine cyanobacteria are susceptible to infection by viruses. Despite growing research on viruses of microbes, virus-*Synechococcus* interactions remain poorly understood (Weitz and Wilhelm, 2012). *Synechococcus* may be infected by cyanophage via lysogeny, wherein the virus integrates its genomes into the host chromosomes and remains as a prophage within the host. Lysogeny protects a virus from environmental factors such as UV light or proteolytic digestion that may cause damage to the viral capsid or nucleic acid. When cyanophage is induced, the host cell lyses and thus affecting *Synechococcus* populations (Suttle, 2000, Mann, 2003). Viral lysis can cause up to 14% of mortality of *Synechococcus* cells in the open ocean on a daily basis (Suttle and Chan, 1994). In coastal waters, cyanophage concentrations co-vary with *Synechococcus* abundance on temporal and spatial scales (Suttle and Chan, 1993, Waterbury and Valois, 1993, Marston and Sallee, 2003). Cyanophage titres are typically high during summer and low during in winter (Jiang and Paul, 1994, Cochran and Paul, 1998, Wang et al., 2011).

Genomes of many cyanophages contain genes related to photosystems (Sullivan et al. 2006, Labrie et al., 2013, Puxty et al., 2014) and carbon metabolism (Millard et al., 2009, Thompson et al., 2011). For example, S-PM2 and S-WHM1 phages contain photosystem related genes such as *psbA* and *psbD* (Mann et al., 2003, Millard et al., 2004), which could interfere with the photosynthetic physiology of marine cyanobacteria. Mann et al. (2003) found that the D1 protein of S-PM2 is similar to the D1 protein of the marine *Synechococcus* sp. WH8102, suggesting that S-PM2 might have acquired the gene horizontally from its *Synechococcus* host (Mann et al., 2003). The expression of virus-encoded D1 and D2 proteins in infected cells can resume the repair cycle of PSII when the host cells are in photoinhibition mode.

Cyanophages modulate cyanobacterial population and microbial communities, and influence the global nutrient cycles. However, the biochemical mechanisms involved remains unknown and the impact of viral-mediated nutrient turnover in the oceans remains to be fully investigated.

1.3.2.2c Association with heterotrophic bacteria

Isolates of marine *Synechococcus* spp. have always been found associated with heterotrophic bacteria. Examples of heterotrophic bacteria that are potentially associated

with *Synechococcus* spp. include: Cytophaga-Flavobacteria-Bacteroides (CFB) group, Alphaproteobacteria (mainly from the *Roseobacter* clade), Gammaproteobacteria (mainly from the Alteromonadales and Pseudomonadales) and Actinobacteria. All these heterotrophic bacteria have their own lifestyles and different living association with *Synechococcus* spp. For example, Flavobacteria, an important class of Bacteroidetes tend to form aggregate or attached to the *Synechococcus* cells, whereas genus *Aquiluna* of Actinobacteria prefer free-living lifestyle, and *Roseobacters* show different association patterns depending on their growth phase (Zheng et al., 2018).

These heterotrophic bacteria can utilise the dissolved organic carbon (DOC) and particulate organic carbon (POC) released from *Synechococcus* spp. as their organic carbon sources (Buchan et al., 2014, Arandia-Gorostidi et al., 2017). In return, *Synechococcus* spp. obtain some essential micronutrients, such as amino acids, trace metals and probably vitamins from heterotrophic bacteria (Hayashi et al., 2011, Christie-Oleza et al., 2015b). Besides that, heterotrophic bacteria have been found to help to reduce the harmful ROS generated in *Prochlorococcus* during photosynthesis by generation of hydrogen peroxide (Morris et al., 2008, Morris et al., 2011). However, it is unknown whether this is the case for *Synechococcus* as well. The metabolic exchange of nutrients among *Synechococcus* and heterotrophic bacteria and how this may shape the distribution and population of *Synechococcus* is not well understood.

1.4 Membrane proteins of cyanobacteria

Cyanobacteria are Gram-negative bacteria with a membrane organization consisting of an outer membrane (OM), a periplasmic space with a peptidoglycan layer and a cytoplasmic membrane (CM). Cyanobacteria also share many features of Gram-positive bacteria: they have similar peptidoglycan composition as Gram-positive bacteria and are considerably thicker than most Gram-negative bacteria (Figure 1.3) (Jurgens and Weckesser, 1985, Hoiczyk and Hansel, 2000). In addition, cyanobacteria have an internal thylakoid membrane (TM) which provides added complexity to their biology, compared with other Gram-negative bacteria and hence presenting interesting questions in prokaryotic cell membrane biology (Gantt, 1994).
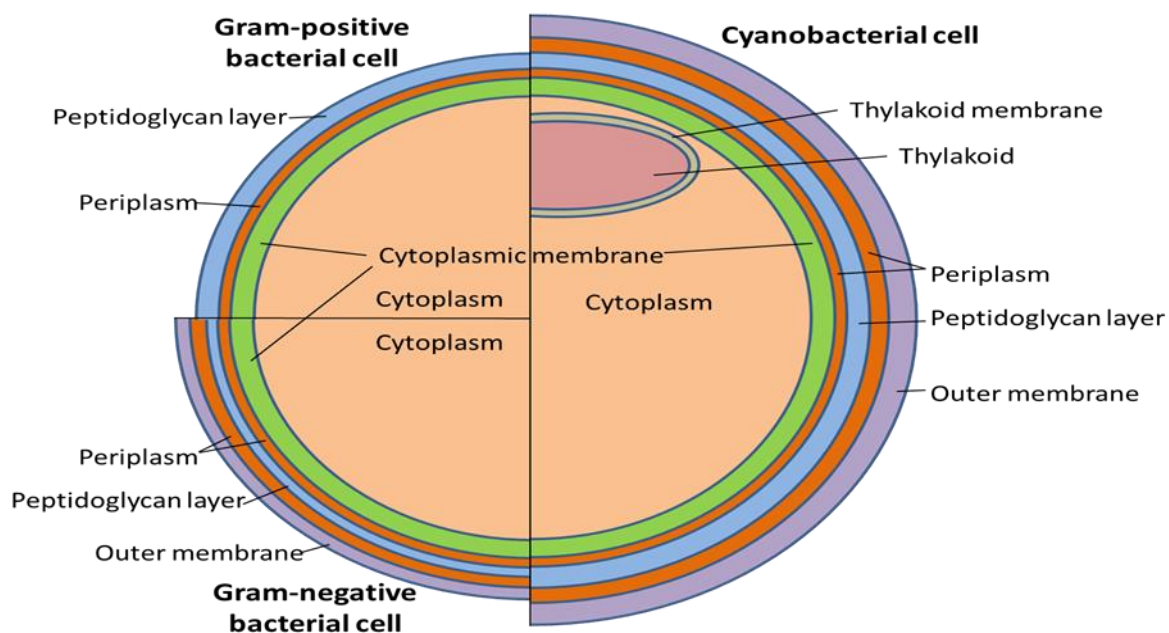
Figure 1.3 Comparison of the cell envelope of Gram-positive bacteria, Gram-negative bacteria and cyanobacterial cell.

Thus far, the membrane proteomes of marine *Synechococcus* have not been studied. However, there are several published studies of the membrane proteomes of freshwater cyanobacteria such as *Anabaena* sp. PCC7120 (Sen et al., 2015, Shvarev and Maldener, 2018), and *Synechocystis* sp. PCC6803 (Huang et al., 2002, Srivastava et al., 2005, Pisareva et al., 2007, Goncalves et al., 2018). Much of our knowledge regarding membrane context of cyanobacteria is derived from freshwater cyanobacteria. Thus, there is a crucial need to explore membrane proteomes of marine cyanobacteria due to their very dynamic ocean environments.

1.4.1 Outer Membrane

Traditional biochemical and genetic approaches have yielded a wealth of knowledge relating to the function of outer membrane proteins in Gram-negative bacteria, however, few studies have been undertaken for cyanobacteria as there are limited studies conducted on outer membrane proteins in cyanobacteria. Thus far, the information available on the structure-function of cyanobacterial integral outer membrane proteins is primarily derived

from predictions based on gene sequences and protein domains similarity with other proteins of known structure and function (Nakao et al., 2010).

Outer membrane (OM) is the exterior layer of cyanobacteria consisting of lipids, polysaccharides and proteins. It serves as an efficient permeability barrier that protects Gram-negative bacteria from various harmful compounds, such as antibiotics, detergents and disinfectants (Nikaido and Vaara, 1985), while allowing the nonspecific passage of small hydrophilic nutrients through porin channels (Paulsen et al., 1998). In Gram-negative bacteria, including cyanobacteria, porins are highly expressed proteins found in the OM (Schulz, 1993, Koebnik et al., 2000). In the *Synechocystis* genome, six putative porin genes have been identified, meanwhile in marine *Synechococcus* the numbers vary from strain to strain, for example strain WH8102 isolated from the oligotrophic open ocean water has four genes coding for porin proteins while the coastal strain CC9311 has six genes encoding porin proteins. Similar to freshwater *Synechocystis*, the molecular mass of a typical porin of marine *Synechococcus* is about 50-70kDa, which is larger than most common Gram-negative bacterial porins, which is between 30 and 40kDa (Hoiczyk and Hansel, 2000). The porins have a conserved surface layer homology (SLH) domain tightly associated to the peptidoglycan layer, and stabilizing the cell wall (Olabarria et al., 1996). They have low hydrophobicity, folding into β-barrel conformations through an interaction with lipopolysaccharide in outer membrane (Sen and Nikaido, 1991).

The Omp85 family proteins are essential components for the assembly of Gram-negative bacteria OM (Genevrois et al., 2003, Voulhoux et al., 2003). It is a large membrane protein family, highly conserved in prokaryotes and eukaryotes. In bacteria, Omp85 proteins play a key role in outer membrane biogenesis, assisting outer membrane protein insertion and folding by an unknown mechanism (Gentle et al., 2004). The Omp85 family proteins are composed of a C-terminal β-barrel domain and an N-terminal domain with POTRA (Polypeptide Translocation-Associated) motifs (Dastvan et al., 2016). The cyanobacterial homolog of Toc75 is a member of Omp85 family and is essential for cyanobacteria (Reumann et al., 1999). However, its exact role in cyanobacteria remains unclear (Huang et al., 2004).

Pilus proteins have also been detected in the outer membrane proteomic study of *Synechocystis* (Huang et al., 2004), showing their capacity for sporadic motility of twitching for locomotion. In contrast, most marine *Synechococcus* are not motile, most motile strains identified in *Synechococcus* genera belong to clade III (Rocap et al., 2002).

Unlike freshwater cyanobacteria which glide on surfaces (Hoiczyk and Baumeister, 1995, Bhaya et al., 1999), clade III *Synechococcus* performs unique swimming motility dependent on the SwmA and SwmB proteins (McCarren and Brahamsha, 2009). SwmA is a 130-kDa glycoprotein that forms a paracrystalline surface layer (S-layer) (McCarren et al., 2005). However, whether S-layer plays a role in motility remains unclear. SwmB contains multiple repeated domains, constituting more than 1% of the genome size and is currently one of the longest bacterial ORFs been found (Palenik et al., 2003).

Besides that, TolC, a component of efflux pump, can also be found to reside in the membrane of Gram-negative bacteria including cyanobacteria. TolC is part of the complex architecture of an efflux pump which made up of three parts: (i) the outer membrane TolC protein and (ii) membrane fusion protein(s) connecting TolC to (iii) specific cytoplasmic membrane efflux pumps (Sharff et al., 2001). This complex machinery allows TolC to export small molecules up to large proteins directly from the cell interior across the two membranes to the extracellular environment without involvement of periplasm. TolC is also involved in the secretion of the S-layer protein in freshwater cyanobacteria *Synechocystis* (Oliveira et al., 2016).

Lipoproteins are an important structural component of outer and cytoplasmic membrane. They play a variety of roles in bacterial physiology such as envelope stability, cell division, signal transduction, transport and protein folding etc. (Zuckert, 2014). Lipoproteins are synthesized in the cytoplasm as prepro-lipoprotein precursors with N-terminal domain containing a signal peptide (von Heijne, 1989) and C-terminal domain containing a four-amino-acid motif known as "lipobox" (Sankaran and Wu, 1993). The majority of prepro-lipoprotein precursors are translocated via general secretion (Sec) translocon into the CM with the assist of YidC chaperone and become mature lipoproteins after post-translational modifications (PTMs) (Nakayama et al., 2012). In *E. coli,* outer membrane-directed lipoproteins are sorted via the Lol transport system to the OM (Zuckert, 2014). However, *lolABCDE* sequence homologues are absent in cyanobacteria such as freshwater *Synechocystis*, *Anabaena* (Kaneko et al., 2001) and marine *Synechococcus*.

1.4.2 Cytoplasmic Membrane

The cyanobacterial cytoplasmic membrane (CM) is equipped with a variety of proteins involved in a broad range of functions which can be divided into three large categories: (i) proteins controlling uptake and export of solutes and proteins, (ii) proteins anchored to the membrane for structural purposes, (iii) enzymes involved in metabolism or signal transduction (MacCoss and Yates, 2001).

Many transport systems in the plasma membrane are coupled to the input of energy. There are several classes of the transport system in CM: (i) facilitated diffusion systems, (ii) secondary transporters which typically couple the uphill transport of nutrients with the downhill influx of H^+ or Na^+, (iii) active transport systems that use periplasmic binding proteins, that are energized by adenosine triphosphate (ATP) hydrolysis and (iv) group translocation systems, for example, phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS) which phosphorylates sugar by using PEP concomitant with their translocation. However, PTS systems are absent in marine *Synechococcus*, hence will not be discussed in here.

1.4.2.1 Facilitated diffusion systems

In *Synechocystis* sp. PCC6803, aquaporin Z (AqpZ) transporter has been identified (Akai et al., 2011, Akai et al., 2012). It consists of major intrinsic protein (MIP) domain and plays a role to maintain osmotic balance by transporting water molecules by passive diffusion across the cytoplasmic membrane. The *aqpZ* gene is absent in almost all marine *Synechococcus*, however, it is still present in some *Synechococcus* sub-cluster 5.2 strains such as WH5701 and PCC7001.

1.4.2.2 Secondary transporters

Many bacterial transporters couple the uphill transport of nutrients with the downhill influx of H^+, Na^+, or phosphate. There are more than 30 subtype of secondary transporters class found in marine *Synechococcus* (Elbourne et al., 2017). Among these, the major facilitator superfamily (MFS) is the largest class of secondary transporters and it plays a role for transportation of a broad spectrum of substrates with diverse physiochemical properties by utilizing the energy stored in electrochemical gradient across the membrane

(Pao et al., 1998). MFS can possess either 12, 14 or 24 transmembrane α-helix fold, and transport a huge variety of ligands including monosaccharides, drugs, enzyme cofactors, peptides, oligosaccharides, iron chelates, nucleotides, and inorganic cations and anions by uniport, symport or antiport mechanisms (Lee et al., 2016). An NRT2-type transporter encoded by *nrtP* (also known as *napA*) has been found to be a bispecific transporter for nitrate and nitrite in euryhaline strain *Synechococcus* sp. PCC7002 (Sakamoto et al. 1999), however, substrate selectivity for marine Synechococcus remains to be justified.

1.4.2.3 ATP-dependent transporters

ATP-binding cassette (ABC) transporters comprise a large and diverse family of membrane-spanning proteins that couples ATP hydrolysis to transport various substrates, ranging from ions to proteins, across membranes (Higgins, 1992, Holland and Blight, 1999, Locher et al., 2002, Schmitt and Tampe, 2002, Hopfner and Tainer, 2003). They are the most common transporters in cyanobacteria (Scanlan et al., 2009).

The core prokaryote ABC transporters consist of the ABC-binding domains (which reside in the cytoplasmic face of CM, also known as nucleotide-binding domains (NBDs)) and the transmembrane domains (TMDs). All bacteria exhibit two major groups of ABC transporters. The first group of ABC transporters is importer which includes a substrate binding protein acting as receptor for the substrate in the periplasm. Bacterial ABC importers import diverse substrates including peptides, amino acids, mono- and oligosaccharides, ions, metals and vitamins, for example, UrtABCDE for urea uptake, PstSABC for phosphate uptake, FutABC for ferric ion uptake (Valladares et al., 2002, Hudek et al., 2016, Polyviou et al., 2018). The second group comprise of ABC exporter that exports substrates such as polysaccharides, lipids, peptides and proteins, including toxins, drugs and virulence factors, for example HlyB ABC protein exporter (Fath and Kolter, 1993). Meanwhile, not all proteins that contain ABC domain functions as transporters, for example, Uvr family which involved in nucleotide excision repair and drug resistance (Davidson et al., 2008).

The basic catalytic cycle of ABC transporters starting from the ground state comprises of a series of steps. These include the binding of substrate-binding proteins (for importers) or the direct binding of a substrate to the TMDs (for exporters), binding of two MgATP molecules to the ATPases, dimerization of the ATPases, switching of the TMDs from

closed state to the open state or from open state to the close state (depending on transporter type), ATP hydrolysis, phosphate, ADP and transport substrate release concomitant with ATPase dissociation to reset the transporter to ground state for the next cycle. Highly conserved domains of Walker A (also known as P-loop) and Walker B motifs in ATPase play a major role in dimerization of ATPases in the transportation system (Dawson and Locher, 2006, Davidson and Maloney, 2007, Jones et al., 2009, Rees et al., 2009, Wilkens, 2015). In addition to ABC transporters, there are other classes of ATP-driven transporters including P-type ATPases, frequently involved in transport of metal cations (Phung et al., 1994), and F- or V-type ATPAses involved in proton and/or sodium ion flux (Frasch, 1994, Pogoryelov et al., 2007, Dibrova et al., 2010).

1.4.2.4 Protein secretion

The secretory (Sec) and twin arginine translocation (Tat) pathways are the bacterial secretion systems most commonly used to translocate proteins across the cytoplasmic membrane (Natale et al., 2008). The Sec and Tat pathways are highly conserved mechanisms of protein secretion, and have been identified in all domains of life, from prokaryotes to eukaryotes (Robinson and Bolhuis, 2004, Papanikou et al., 2007). The Sec system has been implicated in both the secretion of unfolded proteins across the CM and the insertion of membrane proteins into the CM. Meanwhile, the Tat system primarily involved in the secretion of folded and/or cofactor containing proteins (Natale et al., 2008). Generally, proteins that will be secreted into the periplasm or extracellular by the Sec pathway contain SecB-specific signal sequences, while majority of the proteins translocase into the CM encompass a signal recognition particle (SRP)-specific signal sequence (Papanikou et al., 2007). Two components of the Sec machinery, SecY and SecA are found in both the cytoplasmic and thylakoid membrane of freshwater *Synechococcus* PCC7942 (Nakai et al. 1993), indicating the importance of Sec pathway in translocation of proteins to the cyanobacterial membrane. However, it remains unknown whether the mechanisms involved for protein translocation in CM are similar to those in the TM.

As not all proteins can be secreted in their unfolded state, some proteins such as redox factors require post-translational modifications (PTMs) after synthesis (Green and Mecsas, 2016). Therefore, these proteins must be folded and modified in the cytoplasm prior to secretion in their 3-dimensional state. The Tat secretion pathway allows transportation of

folded proteins by forming a membrane-spanning channel to translocate these folded proteins into periplasm or out of the cell. The Tat pathway of protein secretion contains 2 or 3 subunits: TatA and/or TatB and TatC. Proteins that are secreted via the Tat pathway contain a twin arginine motif at the N-terminal signal peptide which is recognized by TatB and TatC (Muller, 2005). There are two type of Tat transcolase system: TatAC and TatABC. Most of the Gram-positive bacteria comprise TatAC system which lack TatB, except *Actinomycetes* (Schaerlaekens et al., 2001) while, the majority of Gram-negative bacteria possess TatABC translocase system (Barnett et al., 2011). A high number of marine *Synechococcus* spp. have just a single TatA/B homologue, however some strains have two separate TatA/B homologues, hence it is unclear whether marine *Synechococcus* contains TatABC or TatAC-type translocase system for proteins translocation (Barnett et al., 2011).

There are many components of signal transduction cascades present in the cytoplasmic membrane. These are often based on phosphorylation cascades involving two-component regulatory systems (TCRS): a membrane-bound sensor histidine kinase (HK) and a cytoplasmic response regulators (RR) protein (Stock et al., 2000). Upon reception of extracellular stimuli, the HK catalyzes ATP-dependent auto-phosphorylation of its conserved histidine residue and then transfer phosphoryl groups to a conserved aspartate residue within the receiver domain of the RR protein. For example, histidine kinase Hik16, Hik33, Hik 34 and Hik41 of freshwater *Synechocytis* sp. 6803 act as sensors in the perception of salt stress (Marin et al., 2003). Overall, TCRS are not well-understood in cyanobacteria especially in marine settings; further exploration is explicitly needed to understand the cell signalling regulation in cyanobacteria.

1.4.3 Periplasm

There are relatively fewer studies conducted on cyanobacterial periplasmic proteins. Therefore, the composition and dynamics of the cyanobacterial periplasm remain largely uncharacterized. The periplasm is a multipurpose compartment 'sandwiched' between the outer and cytoplasmic membrane whose distinct reducing compartment allows more efficient and diverse mechanisms of protein oxidation, folding and quality control. Besides that, it comprises vital environmental sensing modules and facilitates communication among the extracellular environment, the cell envelope and the cytoplasm (Beveridge,

1999) by allowing complex nano-machines to span the cell envelope. For example, TonB-dependent iron uptake system mediated by TonB in plasma membrane extends into the periplasm to bind with iron substrates which transport through the TonB-dependent transporters (TBDTs) localized in outer membrane (Mirus et al., 2009). TonB is commonly seen in freshwater cyanobacteria but absent in most of the marine cyanobacteria (Tang et al., 2012).

Generally, periplasmic proteins contain a typical signal peptide composed of three segments: a positive charged N-terminus followed by a central hydrophobic region and a more polar C-terminus containing the cleavage site (c-region) (von Heijne, 1985). A previous study of periplasmic proteomics of *Synechocystis* sp. 6803 demonstrated that periplasmic proteins are dominated by proteins that contain Sec signal (Fulda et al., 2001). The periplasmic space of Gram-negative bacteria is enriched with hydrolytic enzymes, such as phosphatases, glycosylases, nucleotidases, and proteases (Han et al., 2014), as well as ABC periplasmic binding proteins (Fulda et al., 2001, Han et al., 2014). In addition, carbonic anhydrase, a zinc metalloenzyme involved in inorganic carbon uptake by converting CO_2 into bicarbonate (Price et al., 1992) has been found in the periplasm of freshwater *Synechocystis* (Fulda et al., 2001). It plays a pivotal role in regulating pH homeostasis, respiratory gas exchange, photosynthesis and ion transport in cyanobacteria (Tashian, 1989, Coleman, 1991, Badger and Price, 1994).

Bacterial cells have developed a sophisticated system of molecular chaperones and proteases for protein folding and quality control (Wickner et al., 1999, Yamamoto, 2001, Maurizi, 2002). Proteases are enzymes that catalyze the cleavage of peptide bonds and are required for processing of the trafficked proteins in the cell envelope. For example, Deg-family proteases are periplasm-associated serine endopeptidases associated with cell envelope stress response (Clausen et al., 2011). The Deg-proteases are well-conserved in marine *Synechococcus* (Stuart et al., 2014). A predicted member of the Deg-protease, SYNW2176 protein in oligotrophic marine *Synechococcus* sp. WH8102 have been shown to be highly expressed under copper stress, suggesting its involvement in protein quality control during cell envelope stress (Stuart et al., 2014). Proteases cooperate with chaperones which assist in folding, ensuring conformational integrity and prevent aggregation under stress conditions. For example, peptidyl-prolyl cis/trans isomerase found in periplasm of *Synechocystis* sp. PCC6803 (Fulda et al., 2001) assists proper folding of secreted proteins (Hayano et al., 1991).

1.4.4 Thylakoid membrane

The thylakoid membrane (TM) is the most widely studied membrane system of cyanobacterium and has been a prominent research focus for more than four decades (Williams, 1988, Wada et al., 1993; Keren et al., 2005a,b, Srivastava et al., 2005). TM is a more homogenous system in contrast to CM which has diverse cellular roles: TM comprises mainly of photosynthetic machinery and the main site of respiratory electron transport, dedicating to the energetics of electron transport (Lea-Smith et al., 2013). The co-existence of photosynthesis and respiration in the same membrane in cyanobacterium allows flexible and versatile electron transport pathways, but this could be a challenge for cyanobacterium to monitor the modes of electron transport per the physiological requirement of the cell. The primary reactions of photosynthesis are mediated by photosystem II (PSII), photosystem I (PSI) complexes and cytochrome b_6f complex (Cyt b_6f) as well as an ATPase localized in TM (schematic of light-dependent reactions of photosynthesis is shown in Figure 1.4). However, some of these components are also found in CM (Zak et al., 2001).
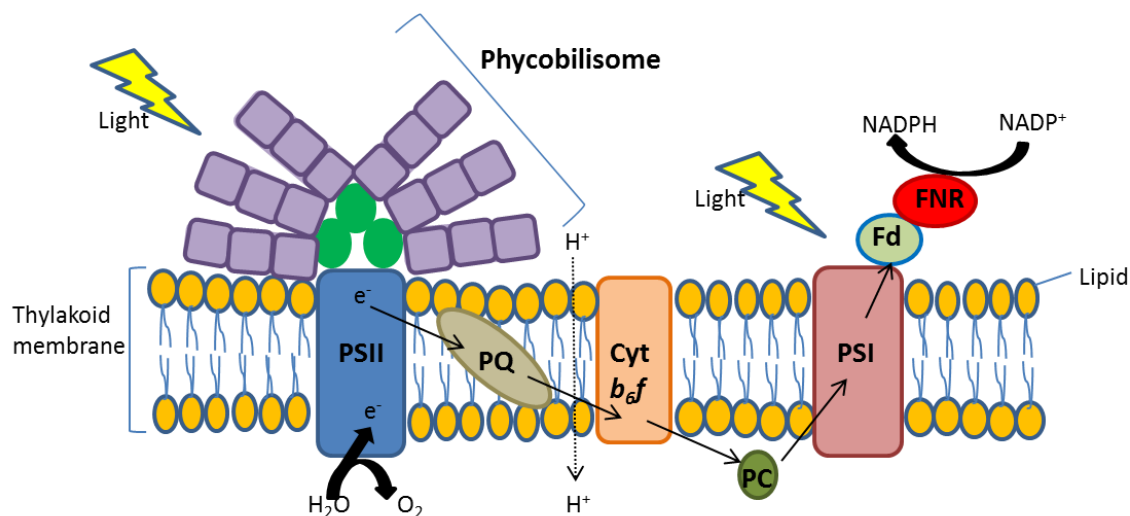
Figure 1.4 Light-dependent reactions of cyanobacterial photosynthesis at the thylakoid membrane.

Photosynthetic light harvesting and electron transfer are mediated by three complexes embedded in the thylakoid membrane: PSII, Cyt *b6f* and PSI. Phycobilisome absorbs light and transfer the excitation energy to the reaction centres in PSII. Electrons extracted from H_2O in PSII are transferred through a linear electron flow (LEF) to $NADP^+$. PQ transfers electrons from the PSII reaction centre to the Cyt b_6f and carries protons across the thylakoid membrane. Cyt b_6f removes the electrons from the reduced PQ and transfer the

electrons to PSI through PC. Electron transfer from PSI to NADP$^+$ requires Fd and FNR. NADPH generated is used in cellular biochemical pathway of cyanobacteria such as Calvin cycle. Abbreviations: PSII, photosystem II; PQ, plastoquinone, Cyt b_6f; PC, plastocyanin; PSI, photosystem I; Fd, ferredoxin; FNR, ferredoxin NADP$^+$ oxidoreductase.

Both PSI and PSII consist of a core complex of proteins and pigments energetically linked to phycobilisome (PBS) (Sidler 1994). The PSII complex catalyses the oxidation of water during photosynthesis in cyanobacterium by splitting water into reducing equivalents and molecular oxygen using photon energy. The dimeric PSII complex isolated from the thermophilic cyanobacterium *Thermosynechococcus elongatus* shows that it consists of 17 integral and 3 peripheral membrane proteins, 35 Chlorophyll a, 2 pheophytin molecules, 12 carotenoids, 2 haem molecules, one non-haem iron, 2 calcium ions, 1-2 chloride ions, 3 plastoquinones, >25 lipid molecules and a CaMn$_4$ metal cluster (Guskov et al., 2009, Umena et al., 2011, Suga et al., 2015).

The core complex of PSII, D1 and D2 reaction centre (RC) subunits consist of five transmembrane α-helices, binding with chlorophyll, pheophytin and plastoquinone co-factors involved in transmembrane light-induced separation of charge (Rappaport and Diner, 2008). Cyanobacteria have multiple *psbA* variants which code for the D1 protein (Sugiura et al., 2010, Kiss et al., 2012, Vinyard et al., 2013). Expression of these genes varies depend on the environmental change (Sugiura et al., 2010, Kiss et al., 2012). For example, up-regulation of *psbA$_3$* gene in *Thermosynechococcus elongatus* that occurs under high light conditions is indicative of a photo-protection mechanism (Kos et al., 2008, Loll et al., 2008, Sander et al., 2010). The D1 and D2 reaction centre (RC) are adjacent to CP43 and CP47 protein, each containing six transmembrane α-helices, binding to chlorophyll *a* (16 molecules in CP47 and 13 molecules in CP43) and β-carotene. Surrounding these sub-units, are the 13 low molecular mass (LMM) (<10kDa) subunits of PSII (Guskov et al., 2009), and PsbO, PsbU and PsbV, which are involved in stability of the CaMn$_4$ cluster (Roose et al., 2007).

PSII is particularly prone to photo-oxidative damage as the water-splitting reaction catalysed by this complex inexorably leads to the generation of ROS that could damage the complex. Thus, a PSII repair cycle operates to substitute damaged protein subunits, D1 and, to a lesser extent, D2 by de novo protein synthesis (Nixon et al., 2005). However,

when the rate of photo-inactivation and damage to D1 is beyond the capacity for repair, photo-inhibition would occur. This subsequently reduces the efficiency of PSII photochemistry (Aro et al., 1993).

Unlike PSII, PSI is rather stable and composed of 11-12 subunits (Xu et al., 2001). PsaA and PsaB are the two main large subunits of PSI, which function as a scaffold for many small subunits (PsaC, D, E, F, I, J, K, L, M and PsaX) (Xu et al., 2001). Cyanobacterial PSI mediates the light-induced electron transfer from plastocyanin on the luminal side to ferredoxin on the cytoplasmic side of thylakoids. In addition, PSI cycles electrons back into the intersystem electron transport chain at the level of PQ or the cytochrome b_6f complex. The cyclic PSI generates a ΔpH across the thylakoid electron flow drives the formation of a trans-thylakoid ΔpH, which can be subsequently used for ATP generation without the production of NADPH (Bailey and Grossman, 2008). It is important to maintain a high ATP:NADPH ratio to balance cellular growth (Munekage et al., 2004). The ratios of PSII and PSI vary among cyanobacteria strains (Campbell et al., 1998, Bailey et al., 2008) and this has been suspected to be highly relevant to their living environment. In addition, PSI plays a vital role in pigment biosynthesis, nitrate, nitrite, sulfite metabolism and CO_2 fixation (Lea-Smith et al., 2016).

Carotenoids (carotenes and xanthophylls) are ubiquitous constituents of cyanobacteria which serve as modulators of membrane micro-viscosity and antioxidants. They protect cyanobacteria PSII and PSI from over excitation through PBS and harmful reactive oxygen species (ROS) produced from water splitting of photosynthesis that damage photosystem of cells (Packer et al., 1981). A carotenoid molecule has a rod-like structure, which often terminates with polar groups (Gruszecki, 1999). The molecular architecture of these pigments is responsible for their localization and orientation in the TM (Loll et al., 2005), forming a bridge between various photosynthetic apparatus proteins (Umena et al., 2011). Free carotenes and xanthophylls occupy the hydrophobic region of membrane bilayers and recent study found that carotenoid is synthesized in cytoplasmic membranes and hence precursors of β-carotene are found to be more abundant in CM than TM (Zakar et al., 2016). β-carotene has been shown to be important for the assembly of PSII and PSI trimer, whereas xanthophylls stabilize those (Zakar et al., 2016). Besides that, it is suspected that carotenoids are also required for the assembly or maintenance of the complete PBS structure (Toth et al., 2015).

PBS is found to be highly mobile over the TM surface (Mullineaux et al., 1997). It is the major light harvesting antenna system of most cyanobacteria including *Synechococcus* spp. PBS complexes are about 3-7 MDa in size and constitute of a central core and lateral rods (Adir, 2005) and able to associate or dissociate with either PSI or PSII (Mackey et al., 2013). The core is made up of trimeric allophycocyanin (APC) and terminal energy acceptors (TE) while the rods are made up of hexamers of the α- and β subunits - containing phycobiliproteins such as phycocyanin (PC) and phycoerythrin (PE) stabilized within the PBS complex by linker polypeptides. The PBS rods absorb excitation energy from sunlight and funnel that energy into the APC of the PBS core.

A major fraction of respiratory electron transport occurs in the TM in cyanobacterium, where respiratory electron transport chains share several components such as plastoquinone, cytochrome b_6f and plastocyanin/cytochrome c (Lea-Smith et al., 2013) with the photosynthetic electron transport chain. Different cyanobacteria species have different allocation and morphology of respiratory enzymes. For example, *Synechococcus* sp. PCC6301 harbours high levels of cytochrome c oxidase in the CM, whereas, it is found to be distributed in the CM and TM in *Synechocystis* sp. PCC6803 (Peschek et al., 1989, Peschek et al., 1994).

Cyanobacteria also possess several alternative electron donor complexes and oxidases. Among these, NADPH dehydrogenase (NDH-1) complexes and succinate dehydrogenase (SDH) play a pivotal role as the major input of electrons into the respiratory electron transport chain (Cooley and Vermaas, 2001, Ogawa and Mi, 2007). Moreover, NDH-1 complexes are also responsible for carbon dioxide (CO_2) uptake and consist of multiple types of proteins including hydrophilic and hydrophobic domain subunits, as well as oxygenic-photosynthesis-specific (OPS) domain subunits (Aryal et al., 2014) reside in the TM.

There is a controversy regarding the exact morphology of thylakoid membranes. It remains elusive whether cyanobacterial thylakoid membranes are physically contiguous with the cytoplasmic membrane (Rajalahti et al., 2007, Pisareva et al., 2011). It is therefore unclear whether the presence of some proteins in both membranes is due to the physical contact of both membranes, or the biogenesis of these proteins is different, or if they have different functional roles in different membrane compartment. It is believed that a number of hypothetical proteins that are conserved in cyanobacteria and also in chloroplasts (Kaneko et al. 1996, Trautmann et al. 2012) contribute to the photosynthetic activity (Battchikova et

al. 2015). However, whether those proteins are localized in TM or just work in close proximity with TM proteins remain to be investigated.

1.5 Proteomics as tools for cyanobacterial studies

Study of the entire set of proteins in a cell or in an organism, termed as proteomics (Wilkins et al., 1996) is a powerful and sensitive approach in biological studies owing to the advance development of mass spectrometry technology (Yates 2011). In general terms, a mass spectrometer measures mass-to-charge ratios (*m/z*) and the abundances of gas-phase ions. It can be divided into three main components: (i) a source that converts molecules into gas-phase ions, (ii) an analyser that separates ions according to their *m/z* ratio (MacCoss and Yates, 2001) and (iii) a detector that records the number of ions at each *m/z* (Aebersold and Mann, 2003). Common ionization sources used in proteomics research are ESI (electrospray) and MALDI (matrix-assisted laser desorption ionization). Proteomics analysis postulates an insight into global protein expressions from identification to quantitation, from localization to function, from individual to network systems.

The application of proteomics has made a deep impact on the cyanobacterial research: many cyanobacteria strains have been extensively characterized such as *Synechocystis* 6803 (Huang et al., 2002, Battchikova et al., 2010, Wegener et al., 2010), *Anabaena* (*Nostoc*) sp. PCC7120 (Stensjo et al., 2007, Ow et al., 2009, Shvarev and Maldener, 2018), *Cyanothece* sp. ATCC 51142 (Aryal et al., 2012). Meanwhile, the proteomics study on marine picocyanobacteria have started to emerge (Fuszard et al., 2012, Christie-Oleza et al., 2015b, Christie-Oleza et al., 2015a, Pandhal et al., 2007) since the completion of the genome sequence of several marine picocyanobacteria *Synechococcus* and *Prochlorococcus*.

1.5.1 Shotgun proteomics

Shotgun proteomics is a term derived from DNA shotgun sequencing (McDonald and Yates, 2002) and represents a methodology in which proteins are digested with proteases to yield a mixture of peptides. Peptides are then separated, ionised, fragmented and

detected using a mass spectrometer (Eng et al., 1994, MacCoss and Yates, 2001, Han et al., 2008).

In most approaches of shotgun proteomics and tandem mass spectrometry, the protein sample is proteolysed with trypsin enzyme and the resulting peptide mixtures are resolved onto nano-HPLC (high performance liquid chromatograph) columns coupled to a tandem mass spectrometer. Most tandem mass spectrometer measure *m/z* in two steps: the first step is called survey scan or MS1 scan, which the *m/z* of the intact ionised peptide (referred as parent peptide) is monitored. Following separation of peptide(s) of interest (most often based on their abundance in the survey scan), the peptides are fragmented and the *m/z* ratios of the resulting fragments are measured by the mass spectrometer. This second spectrum is known as MS/MS or MS2 spectrum. Fragmentation of the peptide ion along the peptide bond results in diagnostic fragments that can be used, with the assistance of software tools using database searching algorithms (Eng et al., 1994, Tabb et al., 2005, Tabb et al., 2007), to deduce the amino acid sequence of the parent peptide. These identified peptides are then employed, again with bioinformatics tools, to reconstitute the protein composition of a sample.

The advance development of tandem MS not only allow proteins to be analysed qualitatively but also quantitatively. Relative quantitation of protein abundances can be carried out using labelling or label-free peptides approaches. Currently available labelling quantitative approaches include metabolic labelling such as SILAC (stable isotope labelling by/with amino acids) (Ong et al., 2002) and chemical labelling such as ICAT (isotope-coded affinity tag) (Sethuraman et al., 2004), iTRAQ (isobaric peptide tags for relative and absolute quantification) (Wiese et al., 2007) and TMT (tandem mass tag) (Rauniyar et al., 2012). Meanwhile, label-free quantitation can be based on precursor signal intensity or on spectral counting (Bondarenko et al., 2002, Chelius and Bondarenko, 2002, Griffin et al., 2003, Wang et al., 2003, Porteus et al., 2011).

The advancement of quantitative proteomics allows for comparative differential protein expression analysis – a process that assesses the repertoire of proteins expressed (estimation based on peptides quantitation) under different conditions or in a specific fraction of the microbial cellular components. Cyanobacteria are exposed to drastic changes in environment and this has driven the interest of biologists to investigate the proteome changes because of environmental perturbations. For example, change of carbon dioxide concentration (Battchikova et al., 2010, Wegener et al., 2010), diurnal (Aryal et

al., 2012, Guerreiro et al., 2016), temperature (Varkey et al., 2016), light intensities (Pandhal et al., 2007, Xiong et al., 2015, Mackey et al., 2017), nutrient resources (Wegener et al., 2010, Cox and Saito, 2013, Mackey et al., 2015) etc. Besides that, shotgun proteomics has been applied for biotechnological applications such as the study of biofuel production in cyanobacteria (Qiao et al., 2012, Tian et al., 2013).

Proteins can be modified by multiple posttranslational modifications (PTMs), creating a PTM code that controls the function of proteins in space and time. Owing to the development of mass spectrometry, we can analyse the PTMs of proteins in cells. PTMs are covalent processing events that modify the physico-chemical properties of a protein, eliciting several possible consequences such as change of enzyme activity, oligomerization state, protein-protein interactions (PPIs), sub-cellular localization or half-life. There are currently more than 200 kinds of PTMs including phosphorylation (Sanders et al., 1989, Mann, 1994), methylation (Lauber et al., 2009), acetylation (Lauber et al., 2009), N-glycosylation (Abu-Qarn et al., 2008), O-glycosylation (Abu-Qarn et al., 2008) etc.

Cyanobacteria use an extensive network of PTMs to transmit signals and to coordinate cellular functions such as phosphorylation (Yang et al., 2013, Angeleri et al., 2016), acetylation (Mo et al., 2015), S-glutahionylation (Chardonnet et al., 2015) and malonylation (Ma et al., 2017). Our current understanding of the functional role of PTMs in cellular process, protein structural modification, PPIs, in marine *Synechococcus* are limited owing to the lacks of PTM studies in marine *Synechococcus*. Therefore, extensive characterization of PTMs in marine cyanobacteria is crucially required to enhance our understanding of cell physiology in marine *Synechococcus*.

Due to the hydrophobic nature of membrane proteins in cells, the membrane proteomic study is significantly challenging. Membrane proteome studies are considerably under-represented in conventional bottom-up proteomic analyses, which generally favour soluble, abundant and easy-to-digest proteins and peptides (Griffin and Schnitzer, 2011, Zhou et al., 2013). Further, integral membrane proteins (IMPs) extracted from lipid bilayers tend to precipitate and form aggregates, making the isolation of membrane proteins difficult further complicating the analysis (Bertone and Snyder, 2005). To enhance the membrane protein solubility, membrane proteins extracted from cells are treated with relatively high concentrations of detergents such as SDS (sodium dodecyl sulfate), CHAPS (3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate), Triton-X100 or co-solvent such as urea; however, high concentration of SDS and urea are not

compatible with trypsin digestion or interfere with MS analysis (Chen et al., 2007). To reduce the sample complexity, membrane extracts can be separated by gel-free approach such as size-exclusion chromatography and/or isoelectric focusing (Kasting and Siefert 2002) or gel-based approach (one or two-dimensional SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis)). Additional purification methods such as acetone precipitation are required but it could lead to sample loss particularly for low molecular mass proteins.

1.5.2 Affinity-purification mass spectrometry (AP-MS)

The current trends of proteomics in cyanobacteria focus mainly on shotgun proteomics, studying protein expression in cyanobacteria under various growing conditions (Guerreiro et al., 2016, Mackey et al., 2017). While shotgun proteomics analysis is useful to understand the change of cellular proteomes in defined time and conditions, it does not provide information on the mechanisms of cell metabolism and the interaction of proteins in cells. Many biological processes in intracellular cells involve PPIs. Therefore, the comprehensive study of PPIs is vital to enhance our understanding of the role of the proteins in cells.

With modifications to purification approaches, MS can be made fully compatible with affinity purification, empowering characterization of PPIs in a near-physiological context (Chepelev et al., 2008). The basic principle of affinity purification involves immobilization of a ligand onto a solid support (most often agarose or magnetic beads) and using this coupled ligand to capture target protein(s) (Roque and Lowe, 2008). Once purified, proteins can be processed for direct analysis by MS. There are many types of ligands available to be used in affinity purification, including DNA and RNA molecules (most often oligonucleotides) (Yang and Chow, 2011), chemicals (Bantscheff et al., 2007, Raida, 2011), peptides (Camperi et al., 2014), proteins (Müller and Hemphill, 2011, Slobodin and Gerst, 2011, Tsai et al., 2011, Wei et al., 2011) or lipids (Bieberich, 2011).

One of the commonly used affinity purification method referred as immunopurification or immunoprecipitation (IP) involves the use of antibodies that can be targeted to the protein of interest itself ("bait" protein), or to standardize fusion moiety referred to as "epitope tag". An advantage of IP is the possibility of in-solution incubation, in which the targeted protein is purified in its native condition from cell or tissue lysates and that multiple

isoforms may be interrogated simultaneously, providing that the antibody can react with all its isoforms. However, it is a challenge to find a suitable antibody that bind to the endogenous protein. Even though there are several antibodies commercially available, they represent a very small percentage of the proteome, thereby limiting the number of endogenous proteins that can be immunoprecipitated. In addition, the cost of IP is high in contrast to other approaches.

Affinity-purification using epitope tags have been used extensively in the PPI (Ashburner et al. 2000) studies of living organisms. Epitope tagging comprises of fusing an open reading frame (ORF) of choice to DNA sequences encoding a peptide or protein tag that can be purified efficiently on a support material. The epitope tag can be fused to either the N- or C-terminus of a protein (or even incorporated into the middle of the protein sequence) (Dekker et al., 2008), and the construct is transformed into appropriate host organisms. There are various types of available epitope tags (or tag combinations) used in PPI studies. For example, green fluorescent protein (Tombolini et al., 1997, Basset et al., 2000), yellow fluorescent protein (YFP) (Van der Henst et al., 2010), histidine (His) (Battesti and Bouveret, 2008), glutathione S-transferase (GST) (Tulk et al., 2000, Thurston et al., 2009), calmodulin binding peptide (CBP) (Battesti and Bouveret, 2008) and histidine-green fluorescent protein (His-GFP) (Stanisławska-Sachadyn et al., 2006). One major advantage of epitope tagging is that multiple proteins can be tagged with the same epitope tag and purified in an identical manner. Therefore, background contaminants should be consistent across all purifications, allowing the use of streamlined control experiments to generate reproducible and comparable results. Some epitopes such as fluorescence tags offer additional benefits by facilitating localization studies in parallel to AP-MS studies (Kittler et al., 2005, Hubner et al., 2010, Galan et al., 2011,). This could help to determine whether the fusion protein is functional, so that the interactions detected are likely biologically relevant.

Protein-protein interaction studies are technically challenging and it is a difficult task in the proteomics field. We have seen application of AP-MS focusing mainly in model organisms such as yeast, human, *Drosophila*, *Arabidopsis*, and *Escherichia coli* (Blagoev et al., 2003, Gully et al., 2003, Rubio et al., 2005, Veraksa et al., 2005, Krogan et al., 2006). The applications of AP-MS to study PPIs is still in its infancy stage in cyanobacteria, however, it has a big potential to be used as a tool to explore the protein signalling and two-component system regulatory response of cyanobacteria.

1.6 Scope of this thesis

The cell membrane is the primary barrier that ions, nutrients, substrates need to cross to enter into the cell. Therefore, it is an important mediator between intracellular and extracellular environments of the cell. Generally, the cell membrane and its protein components function as a transport gateway for ions, nutrients, substrates into and out of the cells, maintain cell integrity, facilitate signalling between cells as well as being involved in enzyme catalysis. The diverse roles of the cell membrane are intriguing and hence have triggered our interest to getting insight into membrane proteins in marine *Synechococcus* which live in the dynamic environment of the ocean.

The functional operation of biological process and systems is dependent on the inter-relationships between proteins. Understanding these interactions is essential for understanding the dynamic properties of biological processes such as metabolic pathways, signalling cascades, DNA transcription and replication, DNA translation and others. Therefore, my thesis intends to deepen our understanding of the findings from membrane proteomics study by analysing the protein-protein interactions of marine cyanobacteria and to elucidate their functional roles in the biological systems.

The broad aims of this thesis are to:

  i.   explore and reveal the membrane proteomes of marine *Synechococcus*;

  ii.   identify the protein-protein interactions (PPIs) of marine *Synechococcus* and elucidate their functional roles in cells;

  iii.   generate a protein-protein interactions network and understand the biological systems of marine *Synechococcus*.

Chapter 2 investigates the membrane proteomes of marine *Synechococcus*. Chapter 3 investigates the protein-protein interactions of marine *Synechococcus* phosphatases and response regulators. Chapter 4 investigates the protein-protein interactions among phosphatases and cytoplasmic lysates of marine *Synechococcus*.

References

ABED, R. M. M. & KÖSTER, J. 2005. The direct role of aerobic heterotrophic bacteria associated with cyanobacteria in the degradation of oil compounds. *Int Biodeter Biodegr,* 55**,** 29-37.

ABU-QARN, M., EICHLER, J. & SHARON, N. 2008. Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. *Curr Opin Struct Biol,* 18**,** 544-550.

ADIR, N. 2005. Elucidation of the molecular structures of components of the phycobilisome- reconstructing a giant. *Photosynth Res,* 85**,** 15-32.

AEBERSOLD, R. & MANN, M. 2003. Mass spectrometry-based proteomics. *Nature,* 422**,** 198-207.

AGAWIN, N. S. R., DUARTE, C. M. & AGUSTI, S. 2000. Nutrient and temperature control of the contribution of picoplankton to phytoplankton biomass and production. *Limnol Oceanogr,* 45**,** 591-600.

AHLGREN, N. A., NOBLE, A., PATTON, A. P., ROACHE-JOHNSON, K., JACKSON, L., ROBINSON, D., MCKAY, C., MOORE, L. R., SAITO, M. A. & ROCAP, G. 2014. The unique trace metal and mixed layer conditions of the Costa Rica upwelling dome support a distinct and dense community of *Synechococcus. Limnol Oceanogr,* 59**,** 2166-2184.

AHLGREN, N. A. & ROCAP, G. 2006. Culture isolation and culture-independent clone libraries reveal new marine *Synechococcus* ecotypes with distinctive light and N physiologies. *Appl Environ Microbiol,* 72**,** 7193-204.

AKAI, M., ONAI, K., KUSANO, M., SATO, M., REDESTIG, H., TOYOOKA, K., MORISHITA, M., MIYAKE, H., HAZAMA, A., CHECCHETTO, V., SZABO, I., MATSUOKA, K., SAITO, K., YASUI, M., ISHIURA, M. & UOZUMI, N. 2011. Plasma membrane aquaporin AqpZ protein is essential for glucose metabolism during photomixotrophic growth of *Synechocystis* sp. PCC 6803. *J Biol Chem,* 286**,** 25224-35.

AKAI, M., ONAI, K., MORISHITA, M., MINO, H., SHIJUKU, T., MARUYAMA, H., ARAI, F., ITOH, S., HAZAMA, A., CHECCHETTO, V., SZABO, I., YUKUTAKE, Y., SUEMATSU, M., YASUI, M., ISHIURA, M. & UOZUMI, N. 2012. Aquaporin AqpZ is involved in cell volume regulation and sensitivity to osmotic stress in *Synechocystis* sp. strain PCC 6803. *J Bacteriol,* 194**,** 6828-36.

ANDERSON, L. K. & TOOLE, C. M. 1998. A model for early events in the assembly pathway of cyanobacterial phycobilisomes. *Mol Microbiol,* 30**,** 467-474.

ANGELERI, M., MUTH-PAWLAK, D., ARO, E. M. & BATTCHIKOVA, N. 2016. Study of O-Phosphorylation Sites in Proteins Involved in Photosynthesis-Related Processes in *Synechocystis* sp. Strain PCC 6803: Application of the SRM Approach. *J Proteome Res,* 15**,** 4638-4652.

ANTAL, T. K. & LINDBLAD, P. 2005. Production of H2 by sulphur-deprived cells of the unicellular cyanobacteria *Gloeocapsa alpicola* and *Synechocystis* sp. PCC 6803 during dark incubation with methane or at various extracellular pH. *J Appl Microbiol,* 98**,** 114-20.

APPLE, J. K., STROM, S. L., PALENIK, B. & BRAHAMSHA, B. 2011. Variability in protist grazing and growth on different marine *Synechococcus* isolates. *Appl Environ Microbiol,* 77**,** 3074-84.

APT, K. E., COLLIER, J. L. & GROSSMAN, A. R. 1995. Evolution of the Phycobiliproteins. *J. Mol. Biol.,* 248**,** 79-96.

ARANDIA-GOROSTIDI, N., WEBER, P. K., ALONSO-SAEZ, L., MORAN, X. A. & MAYALI, X. 2017. Elevated temperature increases carbon and nitrogen fluxes between phytoplankton and heterotrophic bacteria through physical attachment. *ISME J,* 11**,** 641-650.

ARMBRUST, E. V., BOWEN, J. D., OLSON, R. J. & CHISHOLM, S. W. 1989. Effect of Light on the Cell Cycle of a Marine *Synechococcus* Strain. *Appl Environ Microbiol,* 55**,** 425-432.

ARMENGOL, L., CALBET, A., FRANCHY, G., RODRIGUEZ-SANTOS, A. & HERNANDEZ-LEON, S. 2019. Planktonic food web structure and trophic transfer efficiency along a productivity gradient in the tropical and subtropical Atlantic Ocean, *Sci Rep*, 9, 2044.

ARO, E.-M., VIRGIN, I. & ANDERSSON, B. 1993. Photoinhibition of Photosystem II. Inactivation, protein damage and turnover. *Biochim Biophys Acta*1143**,** 113-134.

ARYAL, U. K., STOCKEL, J., WELSH, E. A., GRITSENKO, M. A., NICORA, C. D., KOPPENAAL, D. W., SMITH, R. D., PAKRASI, H. B. & JACOBS, J. M. 2012. Dynamic proteome analysis of *Cyanothece* sp. ATCC 51142 under constant light. *J Proteome Res,* 11**,** 609-19.

ARYAL, U. K., XIONG, Y., MCBRIDE, Z., KIHARA, D., XIE, J., HALL, M. C. & SZYMANSKI, D. B. 2014. A Proteomic Strategy for Global Analysis of Plant Protein Complexes. *Plant Cell,* 26**,** 3867-3882.

ASHBURNER, M., BALL, C. A., BLAKE, J. A., BOTSTEIN, D., BUTLER, H., CHERRY, J. M., DAVIS, A. P., DOLINSKI, K., DWIGHT, S. S., EPPIG, J. T., HARRIS, M. A., HILL, D. P., ISSEL-TARVER, L., KASARSKIS, A., LEWIS, S., MATESE, J. C., RICHARDSON, J. E., RINGWALD, M., RUBIN, G. M. & SHERLOCK, G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet,* 25**,** 25-9.

BADGER, M. R. & PRICE, G. D. 1994. The role of carbonic anhydrase in photosynthesis. *Annu Rev Plant Physiol,* 45**,** 369-92.

BAILEY, S. & GROSSMAN, A. 2008. Photoprotection in Cyanobacteria: Regulation of Light Harvesting†. *Photochem Photobiol,* 84**,** 1410-1420.

BAILEY, S., MELIS, A., MACKEY, K. R., CARDOL, P., FINAZZI, G., VAN DIJKEN, G., BERG, G. M., ARRIGO, K., SHRAGER, J. & GROSSMAN, A. 2008. Alternative photosynthetic electron flow to oxygen in marine *Synechococcus. Biochim Biophys Acta,* 1777**,** 269-76.

BANTSCHEFF, M., EBERHARD, D., ABRAHAM, Y., BASTUCK, S., BOESCHE, M., HOBSON, S., MATHIESON, T., PERRIN, J., RAIDA, M., RAU, C., READER, V., SWEETMAN, G., BAUER, A., BOUWMEESTER, T., HOPF, C., KRUSE, U., NEUBAUER, G., RAMSDEN, N., RICK, J., KUSTER, B. & DREWES, G. 2007. Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nat Biotechnol,* 25**,** 1035-1044.

BARNETT, J. P., ROBINSON, C., SCANLAN, D. J. & BLINDAUER, C. A. 2011. The Tat protein export pathway and its role in cyanobacterial metalloprotein biosynthesis. *FEMS Microbiol Lett,* 325**,** 1-9.

BARNETT, J. P., SCANLAN, D. J. & BLINDAUER, C. A. 2014. Identification of major zinc-binding proteins from a marine cyanobacterium: insight into metal uptake in oligotrophic environments. *Metallomics,* 6**,** 1254-68.

BARON, M., ARELLANO, J. B. & GORGE, J. L. 1995. Copper and photosystem II- A controversial relationship. *Physiol Plant* 94**,** 174-180.

BASSET, A., KHUSH, R. S., BRAUN, A., GARDAN, L., BOCCARD, F., HOFFMANN, J. A. & LEMAITRE, B. 2000. The phytopathogenic bacteria *Erwinia carotovora* infects *Drosophila* and activates an immune response. *PNAS,* 97**,** 3376-3381.

BATTCHIKOVA, N., ANGELERI, M. & ARO, E.-M. 2015. Proteomic approaches in research of cyanobacterial photosynthesis. *Photosynth Res,* 126**,** 47-70.

BATTCHIKOVA, N., VAINONEN, J. P., VORONTSOVA, N., KERANEN, M., CARMEL, D. & ARO, E.-M. 2010. Dynamic changes in the proteome of *Synechocystis* 6803 in response to CO2 limitation revealed by quantitative proteomics. *J Proteome Res,* 9**,** 5896-5912.

BATTESTI, A. & BOUVERET, E. 2008. Improvement of bacterial two-hybrid vectors for detection of fusion proteins and transfer to pBAD-tandem affinity purification, calmodulin binding peptide, or 6-histidine tag vectors. *Proteomics,* 8**,** 4768-4771.

BEARDALL, J., YOUNG, E. & ROBERTS, S. 2001. Approaches for determining phytoplankton nutrient limitation. *Aquat Sci,* 63**,** 44-69.

BEHRENFELD, M. J. & MILLIGAN, A. J. 2013. Photophysiological expressions of iron stress in phytoplankton. *Ann Rev Mar Sci,* 5**,** 217-46.

BERTONE, P. & SNYDER, M. 2005. Prospects and Challenges in Proteomics. *Plant Physiol,* 138**,** 560-562.

BEVERIDGE, T. J. 1999. Structures of Gram-Negative Cell Walls and Their Derived Membrane Vesicles. *J Bacteriol ,* 181**,** 4725-4733.

BHAYA, D., WATANABE, N., OGAWA, T. & GROSSMAN, A. 1999. The role of an alternative sigma factor in motility and pilus formation in the cyanobacterium *Synechocystis* sp. strain PCC6803. *PNAS,* 96**,** 3188-3193.

BIEBERICH, E. 2011. Lipid Vesicle-mediated Affinity Chromatography using Magnetic Activated Cell Sorting (LIMACS): a Novel Method to Analyze Protein-lipid Interaction. *J Vis Exp,* 26**,** pii:265.

BINDER, B. J. & CHISHOLM, S. W. 1990. Relationship between DNA Cycle and Growth Rate in *Synechococcus* sp. Strain PCC 6301. *J Bacteriol ,* 172**,** 2313-2319.

BINDER, B. J. & CHISHOLM, S. W. 1995. Cell Cycle Regulation in Marine *Synechococcus* sp. Strains. *Appl Environ Microbiol,* 61**,** 708-717.

BLAGOEV, B., KRATCHMAROVA, I., ONG, S. E., NIELSEN, M., FOSTER, L. J. & MANN, M. 2003. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat Biotechnol,* 21**,** 315-8.

BLANK, C. E. & SANCHEZ-BARACALDO, P. 2010. Timing of morphological and ecological innovations in the cyanobacteria--a key to understanding the rise in atmospheric oxygen. *Geobiology,* 8**,** 1-23.

BONDARENKO, P. V., CHELIUS, D. & SHALER, T. A. 2002. Identification and relative quantitation of protein mixtures by enzymatic digestion followed by capillary reversed-phase liquid chromatography-tandem mass spectrometry. *Anal. Chem.,* 74**,** 4741-4749.

BRAND, L. E., SUNDA, W. G. & GUILLARD, R. R. L. 1986. Reduction of marine phytoplankton reproduction rates by copper and cadmium. *J. Exp. Mar. Biol. Ecol.,* 96**,** 225-250.

BRANDENBURG, F., SCHOFFMAN, H., KURZ, S., KRAMER, U., KEREN, N., WEBER, A. P. & EISENHUT, M. 2017. The *Synechocystis* Manganese Exporter Mnx Is Essential for Manganese Homeostasis in Cyanobacteria. *Plant Physiol,* 173**,** 1798-1810.

BROWN, N. L., LEE, B. T. O. & SILVER, S. 1994. Bacterial transport and resistance to copper. *In:* SIGEL, H. & SIGEL, A. (eds.) *Metal ions in biological systems* Marcel Dekker.

BRULAND, K. 1980. Oceanographic distributions of cadmium, zinc, nickel, and copper in the North Pacific. *Earth Planet Sci Lett,* 47**,** 176-198.

BRULAND, K. W., ORIANS, K. J. & COWEN, J. P. 1994. Reactive trace metals in the stratified central North Pacific. *Geochim. Cosmochim. Acta,* 58**,** 3171-3182.

BUCHAN, A., LECLEIR, G. R., GULVIK, C. A. & GONZÁLEZ, J. M. 2014. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat Rev Microbiol,* 12**,** 686-698.

CAI, H., WANG, K., HUANG, S., JIAO, N. & CHEN, F. 2010. Distinct patterns of picocyanobacterial communities in winter and summer in the Chesapeake Bay. *Appl Environ Microbiol,* 76**,** 2955-60.

CALBET, A., TREPAT, I., ALMEDA, R., SALÓ, V., SAIZ, E., MOVILLA, J. I., ALCARAZ, M., YEBRA, L. & SIMÓ, R. 2008. Impact of micro- and nanograzers on phytoplankton assessed by standard and size-fractionated dilution grazing experiments. *Aquat Microb Ecol,* 50**,** 145-156.

CAMPBELL, D., HURRY, V., CLARKE, A. K., GUSTAFSSON, P. & OQUIST, G. 1998. Chlorophyll Fluorescence Analysis of Cyanobacterial Photosynthesis and Acclimation. *Microbiol Mol Biol Rev,* 62**,** 667-683.

CAMPBELL, L. & ITURRIAGA, R. 1988. Identification of *Synechococcus* spp. in the Sargasso Sea by immunofluorescence and fluorescence excitation spectroscopy performed on individual cells. *Limnol. Oceanogr.,* 33**,** 1196-1201.

CAMPBELL, L., NOLLA, H. A. & VAULOT, D. 1994. The importance of *Prochlorococcus* to community structure in the central North Pacific ocean. *Limnol. Oceanogr.,* 39**,** 954-961.

CAMPERI, S. A., MARTINEZ-CERON, M. C., GIUDICESSI, S. L., MARANI, M. M., ALBERICIO, F. & CASCONE, O. 2014. Peptide affinity chromatography based on combinatorial strategies for protein purification. . *In:* LABROU, N. E. (ed.) *Protein Downstream Processing: Design, Development and Application of High and Low-Resolution Methods.* New York: Humana Press.

CARON, D. A., PEELE, E. R., LIM, E. L. & DENNETT, M. R. 1999. Picoplankton and nanoplankton and their trophic coupling in surface waters of the Sargasso Sea south of Bermuda. *Limnol. Oceanogr.,* 44**,** 259-272.

CASTENHOLZ, R. W. & WATERBURY, J. B. 1989. Preface. Bergey's manual of systematic bacteriology. *Williams and Wilkins, Baltimore,* 3**,** 1710-1727.

CHARDONNET, S., SAKR, S., CASSIER-CHAUVAT, C., LE MARECHAL, P., CHAUVAT, F., LEMAIRE, S. D. & DECOTTIGNIES, P. 2015. First proteomic study of S-glutathionylation in cyanobacteria. *J Proteome Res,* 14**,** 59-71.

CHELIUS, D. & BONDARENKO, P. V. 2002. Quantitative Profiling of Proteins in Complex Mixtures Using Liquid Chromatography and Mass Spectrometry. *J Proteome Res,* 1**,** 317-323.

CHEN, E. I., COCIORVA, D., NORRIS, J. L. & YATES, J. R., 3RD 2007. Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. *J Proteome Res,* 6**,** 2529-38.

CHEN, F., WANG, K., KAN, J., BACHOON, D. S., LU, J., LAU, S. & CAMPBELL, L. 2004. Phylogenetic diversity of *Synechococcus* in the Chesapeake Bay revealed by Ribulose-1,5- bisphosphate carboxylase-oxygenase (RuBisCO) large subunit gene (rbcL) sequences. *Aquat Microb Ecol,* 36**,** 153-164.

CHEN, F., WANG, K., KAN, J., SUZUKI, M. T. & WOMMACK, K. E. 2006. Diverse and unique picocyanobacteria in Chesapeake Bay, revealed by 16S-23S rRNA internal transcribed spacer sequences. *Appl Environ Microbiol,* 72**,** 2239-43.

CHEPELEV, N., CHEPELEV, L., ALAMGIR, M. D. & GOLSHANI, A. 2008. Large-Scale Protein-Protein Interaction Detection Approaches: Past, Present and Future. *Biotechnol Biotech Eq,* 22**,** 513-529.

CHOI, D. H. & NOH, J. H. 2009. Phylogenetic diversity of *Synechococcus* strains isolated from the East China Sea and the East Sea. *FEMS Microbiol Ecol,* 69**,** 439-48.

CHRISTAKI, U., GIANNAKOUROU, A., WAMBEKE, F. V. & GREGORI, G. 2001. Nanoflagellate predation on auto- and heterotrophic picoplankton in the oligotrophic Mediterranean Sea. *J Plankton Res,* 23**,** 1297-1310.

CHRISTIE-OLEZA, J. A., ARMENGAUD, J., GUERIN, P. & SCANLAN, D. J. 2015a. Functional distinctness in the exoproteomes of marine *Synechococcus*. *Environ Microbiol,* 17**,** 3781-3794.

CHRISTIE-OLEZA, J. A., SCANLAN, D. J. & ARMENGAUD, J. 2015b. "You produce while I clean up", a strategy revealed by exoproteomics during *Synechococcus-Roseobacter* interactions. *Proteomics,* 15**,** 3454-62.

CLAUSEN, T., KAISER, M., HUBER, R. & EHRMANN, M. 2011. HTRA proteases: regulated proteolysis in protein quality control. . *Nat Rev Mol Cell Biol,* 12**,** 152-62.

COALE, K. H. & BRULAND, K. W. 1988. Copper complexation in the Northeast Pacific. *Limnol. Oceanogr.,* 33**,** 1084-1101.

COCHRAN, P. K. & PAUL, J. H. 1998. Seasonal abundance of lysogenic bacteria in a subtropical estuary. *Appl Environ Microbiol,* 64**,** 2308-2312.

COLEMAN, J. R. 1991. The molecular and biochemical analyses of CO_2-concentrating mechanisms in cyanobacteria and microalgae. *Plant Cell Environ,* 14**,** 861-867.

COLLIER, J. L., BRAHAMSHA, B. & PALENIK, B. 1999. The marine cyanobacterium *Synechococcus* sp. WH7805 requires urease (urea amiohydrolase, EC 3.5.1.5) to utilize urea as a nitrogen source- molecular-genetic and biochemical analysis of the enzyme. *Microbiol Mol Biol Rev,* 145**,** 447-459.

COLLINS, F. S. 1891. Notes on New England marine algae V. *Bull Torr Bot Club* 18**,** 335-341.

COOLEY, J. W. & VERMAAS, W. F. 2001. Succinate dehydrogenase and other respiratory pathways in thylakoid membranes of *Synechocystis* sp. strain PCC 6803: capacity comparisons and physiological function. *J Bacteriol,* 183**,** 4251-8.

COPIN-MONTEGUT, C. & COPIN-MONTEGUT, G. 1983. Stoichiometry of carbon, nitrogen, and phosphorus in marine particulate matter. *Deep Sea Res,* 30**,** 31-46.

COX, A. D. & SAITO, M. A. 2013. Proteomic responses of oceanic *Synechococcus* WH8102 to phosphate and zinc scarcity and cadmium additions. *Front Microbiol,* 4**,** 387.

CROSBIE, N. D., POCKL, M. & WEISSE, T. 2003. Dispersal and Phylogenetic Diversity of Nonmarine Picocyanobacteria, Inferred from 16S rRNA Gene and cpcBA-Intergenic Spacer Sequence Analyses. *Appl Environ Microbiol,* 69**,** 5716-5721.

CULLEN, J. T., LANE, T. W., MOREL, F. M. M. & SHERRELL, R. M. 1999. Modulation of cadmium uptake in phytoplankton by seawater CO_2 concentration. *Nature,* 402**,** 165-167.

DASTVAN, R., BROUWER, E. M., SCHUETZ, D., MIRUS, O., SCHLEIFF, E. & PRISNER, T. F. 2016. Relative Orientation of POTRA Domains from Cyanobacterial Omp85 Studied by Pulsed EPR Spectroscopy. *Biophys J,* 110**,** 2195-206.

DAVIDSON, A. L., DASSA, E., ORELLE, C. & CHEN, J. 2008. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol Mol Biol Rev,* 72**,** 317-64.

DAVIDSON, A. L. & MALONEY, P. C. 2007. ABC transporters: how small machines do a big job. *Trends Microbiol,* 15**,** 448-55.

DAWSON, R. J. & LOCHER, K. P. 2006. Structure of a bacterial multidrug ABC transporter. *Nature,* 443**,** 180-5.

DEKKER, C., STIRLING, P. C., MCCORMACK, E. A., FILMORE, H., PAUL, A., BROST, R. L., COSTANZO, M., BOONE, C., LEROUX, M. R. & WILLISON, K. R. 2008. The interaction network of the chaperonin CCT. *EMBO J,* 27**,** 1827-1839.

DEMOULIN, C., LARA, Y.J., CORNET, L., FRANCOIS, C., BAURAIN, D., WILMOTTE, A. & JAVAUX, E. 2019. Cyanobacteria evolution: Insight from the fossil record. Free Radic Biol Med, 140, 206-223.

DES MARAIS, D. J. 2000. When Did Photosynthesis Emerge on Earth? *Science,* 289**,** 1703-1705.

DIBROVA, D. V., GALPERIN, M. Y. & MULKIDJANIAN, A. Y. 2010. Characterization of the N-ATPase, a distinct, laterally transferred Na^+-translocating form of the bacterial F-type membrane ATPase. *Bioinformatics,* 26**,** 1473-6.

DOLAN, J. R. & SIMEK, K. 1999. Diel periodicity in *Synechococcus* populations and grazing by heterotrophic nanoflagellates- Analysis of food vacuole contents. *Limnol Oceanogr,* 44**,** 1565-1570.

DUFRESNE, A., OSTROWSKI, M., SCANLAN, D. J., GARCZAREK, L., MAZARD, S., PALENIK, B. P., PAULSEN, I. T., DE MARSAC, N. T., WINCKER, P., DOSSAT, C., FERRIERA, S., JOHNSON, J., POST, A. F., HESS, W. R. & PARTENSKY, F. 2008. Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol,* 9**,** R90.

DUPONT, C. L., BARBEAU, K. & PALENIK, B. 2008. Ni Uptake and Limitation in Marine *Synechococcus* Strains. *Appl Environ Microbiol,* 74**,** 23-31.

DURAND, M. D., OLSON, R. J. & CHISHOLM, S. W. 2001. Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. *Deep-Sea Res Pt II,* 48**,** 1983-2003.

ELBOURNE, LIAM D. H., TETU, S. G., HASSAN, K. A. & PAULSEN, I. T. 2017. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Research,* 45**,** D320-D324.

ENG, J. K., MCCORMACK, A. L. & YATES, J. R. 1994. An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *American Society for Microbiology,* 5**,** 976-989.

ERNST, A., BECKER, S., WOLLENZIEN, U. I. A. & POSTIUS, C. 2003. Ecosystem-dependent adaptive radiations of picocyanobacteria inferred from 16S rRNA and ITS-1 sequence analysis. *Microbiology,* 149**,** 217-228.

FALKOWSKI, P. G. 2006. Tracing Oxygen's Imprint on Earth's Metabolic Evolution. *Science,* 311**,** 1724-1725.

FALKOWSKI, P. G., ZIEMANN, D., KOLBER, Z. & BIENFANG, P. K. 1991. Role of eddy pumping in enhancing primary production in the ocean. *Nature,* 352**,** 55-58.

FATH, M. J. & KOLTER, R. 1993. ABC Transporters-Bacterial Exporters. *Microbiological Reviews,* 57**,** 995-1017.

FERRIS, M. J. & PALENIK, B. 1998. Niche adaptation in ocean cyanobacteria. *Nature,* 396**,** 226-228.

FLOMBAUM, P., GALLEGOS, J. L., GORDILLO, R. A., RINCON, J., ZABALA, L. L., JIAO, N., KARL, D. M., LI, W. K., LOMAS, M. W., VENEZIANO, D., VERA, C. S., VRUGT, J. A. & MARTINY, A. C. 2013. Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus. Proc Natl Acad Sci U S A,* 110**,** 9824-9.

FRASCH, W. D. 1994. The F-type ATPase in Cyanobacteria: Pivotal Point in the Evolution of a Universal Enzyme. *In:* BRYANT, D. A. (ed.) *The Molecular Biology of Cyanobacteria.*: Luwer Academic

FRAUSTO DA SILVA, J. J. R. & WILLIAMS, R. J. P. 2001. *The Biological Chemistry of the Elements* Oxford, Oxford University Press.

FULDA, S., HUANG, F., NILSON, F., HAGEMANN, M. & NORLING, B. 2001. Proteomics of *Synechocystis* sp. strain PCC 6803 Identification of periplasmic proteins in cells grown at low and high salt concentrations. *Eur. J. Biochem.,* 267**,** 5900-5907.

FULLER, N. J., MARIE, D., PARTENSKY, F., VAULOT, D., POST, A. F. & SCANLAN, D. J. 2003. Clade-Specific 16S Ribosomal DNA Oligonucleotides Reveal the Predominance of a Single Marine *Synechococcus* Clade throughout a Stratified Water Column in the Red Sea. *Appl Environ Microbiol,* 69**,** 2430-2443.

FULLER, N. J., TARRAN, G. A., YALLOP, M., ORCUTT, K. M. & SCANLAN, D. J. 2006. Molecular analysis of picocyanobacterial community structure along an Arabian Sea transect reveals distinct spatial separation of lineages. *Limnol. Oceanogr.,* 51**,** 2515-2526.

FUSZARD, M. A., WRIGHT, P. C. & BIGGS, C. A. 2012. Comparative quantitative proteomics of *Prochlorococcus* ecotypes to a decrease in environmental phosphate concentrations. *Aquatic Biosystems,* 8**,** 1-8.

GALAN, J. A., PARIS, L. L., ZHANG, H.-J., ADLER, J., GEAHLEN, R. L. & TAO, W. A. 2011. Proteomic Studies of Syk-Interacting Proteins Using a Novel Amine-Specific Isotope Tag and GFP Nanotrap. *Journal of The American Society for Mass Spectrometry,* 22**,** 319-328.

GANTAR, M. & SVIRCEV, Z. 2008. Microalgae and Cyanobacteria: Food for Thought. *J Phycol,* 44**,** 260-8.

GANTT, E. 1994. *The molecular biology of cyanobacteria.*, Kluwer, Dordrecht.

GENEVROIS, S., STEEGHS, L., ROHOLL, P., LETESSON, J.-J. & VAN DER LAY, P. 2003. The Omp85 protein of *Neisseria meningitidis* is required for lipid export to the outer membrane. *EMBO J,* 22**,** 1780-1789.

GENTLE, I., GABRIEL, K., BEECH, P., WALLER, R. & LITHGOW, T. 2004. The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *J Cell Biol,* 164**,** 19-24.

GILBERT, P. M., KANA, T. M., OLSON, R. J., KIRCHMAN, D. L. & ALBERTE, R. S. 1986. Clonal comparisons of growth and photosynthetic responses to nitrogen availability in marine *Synechococcus* spp. *Journal of Experimental Marine Biology and Ecology,* 101**,** 199-208.

GINER-LAMIA, J., LOPEZ-MAURY, L., REYES, J. C. & FLORENCIO, F. J. 2012. The CopRS Two-Component System Is Responsible for Resistance to Copper in the Cyanobacterium *Synechocystis* sp. PCC 6803. . *Plant Physiol,* 159**,** 1806-1818.

GLOVER, H. E., GARSIDE, C. & TREES, C. C. 2007. Physiological responses of Sargasso Sea picoplankton to nanomolar nitrate perturbations. *J Plankton Res,* 29**,** 263-274.

GLOVER, H. E., PREZELIN, B. B., CAMPBELL, L., WYMAN, M. & GARSIDE, C. 1988. A nitrate-dependent *Synechococcus* bloom in surface. *Nature,* 331**,** 161-163.

GLOVER, H. E., SMITH, A. E. & SHAPIRO, L. 1985. Diurnal variations in photosynthetic rates- comparisons of ultraphytoplankton with a larger phytoplankton size fraction. *J Plankton Res,* 7**,** 519-535.

GONCALVES, C. F., PACHECO, C. C., TAMAGNINI, P. & OLIVEIRA, P. 2018. Identification of inner membrane translocase components of TolC-mediated secretion in the cyanobacterium *Synechocystis* sp. PCC 6803. *Environ Microbiol,* 20**,** 2354-2369.

GREEN, E. R. & MECSAS, J. 2016. Bacterial Secretion Systems: An Overview. *Microbiol Spectr,* 4.

GRIFFIN, N. M. & SCHNITZER, J. E. 2011. Overcoming key technological challenges in using mass spectrometry for mapping cell surfaces in tissues. *Mol Cell Proteomics,* 10**,** 1-14.

GRIFFIN, T. J., LOCK, C., M., LI, X.-J., PATEL, A., CHERVETSOVA, I., LEE, H., WRIGHT, M. E., RANISH, J. A., CHEN, S. S. & AEBERSOLD, R. 2003. Abundance Ratio-Dependent Proteomic Analysis by Mass Spectrometry. *Anal. Chem.,* 75**,** 867-874.

GROSSMAN, A. R., SCHAEFER, M. R., CHIANG, G. G. & COLLIER, J. L. 1993. The Phycobilisome, a Light-Harvesting Complex Responsive to Environmental Conditions. *Microbiological Reviews,* 57**,** 725-749.

GRUSZECKI, W. I. 1999. Carotenoids in Membranes. . *In:* FRANK, H. A., YOUNG, A. J., BRITTON, G. & COGDELL, R. J. (eds.) *The Photochemistry of Carotenoids.* Dodrrecht, The Netherlands: Kluwer Academic.

GUERREIRO, A. C., PENNING, R., RAAIJMAKERS, L. M., AXMAN, I. M., HECK, A. J. & ALTELAAR, A. F. 2016. Monitoring light/dark association dynamics of multi-protein complexes in cyanobacteria using size exclusion chromatography-based proteomics. *J Proteomics,* 142**,** 33-44.

GUIRY, M. D. 2012. How Many Species of Algae Are There? *J Phycol,* 48**,** 1057-63.

GULLY, D., MOINIER, D., LOISEAU, L. & BOUVERET, E. 2003. New partners of acyl carrier protein detected in *Escherichia coli* by tandem affinity purification. *FEBS Letters,* 548**,** 90-96.

GUSKOV, A., KERN, J., GABDULKHAKOV, A., BROSER, M., ZOUNI, A. & SAENGER, W. 2009. Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride. *Nature Structural & Molecular Biology,* 16**,** 334-342.

HAN, M.-J., KIM, J. Y. & KIM, J. A. 2014. Comparison of the large-scale periplasmic proteomes of the *Escherichia coli* K-12 and B strains. *Journal of Bioscience and Bioengineering,* 117**,** 437-442.

HAN, X., ASLANIAN, A. & YATES, J. R., 3RD 2008. Mass spectrometry for proteomics. *Curr Opin Chem Biol,* 12**,** 483-90.

HANSCH, R. & MENDEL, R. R. 2009. Physiological functions of mineral micronutrients (Cu, Zn, Mn, Fe, Ni, Mo, B, Cl). *Current opinion in Plant Biology,* 12**,** 259-266.

HAVERKAMP, T. H. A., SCHOUTEN, D., DOELEMAN, M., WOLLENZIEN, U., HUISMAN, J. & STAL, L. J. 2008. Colorful microdiversity of *Synechococcus* strains (picocyanobacteria) isolated from the Baltic Sea. *The ISME Journal,* 3**,** 397-408.

HAYANO, T., TAKAHASHI, N., KATO, S., MAKI, N. & SUZUKI, M. 1991. Two distinct forms of peptidylprolyl-cis-trans-isomerase are expressed separately in periplasmic and cytoplasmic compartments of *Escherichia coli* cells. *Biochemistry,* 30**,** 3041-3048.

HAYASHI, S., ITOH, K. & SUYAMA, K. 2011. Growth of the Cyanobacterium *Synechococcus leopoliensis* CCAP1405/1 on Agar Media in the Presence of Heterotrophic Bacteria. *Microbes and Environments,* 26**,** 120-127.

HEBEL, D. V. & KARL, D. M. 2001. Seasonal, interannual and decadal variations in particulate matter concentrations and composition in the subtropical North Pacific Ocean. *Deep Sea Research II,* 48**,** 1669-1695.

HERDMAN, M., CASTENHOLZ, R., WATERBURY, J. & RIPPKA, R. 2001. Form-genus XIII. *Synechococcus. In:* BOONE, D. R. & CASTENHOLZ, R. W. (eds.) *Bergey's manual of systematic bacteriology, 2nd ed, vol 1. The Archaea and deeply branching and phototrophic Bacteria.* New York: Spinger.

HIGGINS, C. F. 1992. ABC transporters- from microorganisms to man. *Annual Review of Cell Biology,* 8**,** 67-113.

HOICZYK, E. & BAUMEISTER, W. 1995. Envelope Structure of Four Gliding Filamentous Cyanobacteria. *J Bacteriol ,* 177**,** 2387-2395.

HOICZYK, E. & HANSEL, A. 2000. Cyanobacterial cell walls- News from an unusual prokaryotic envelope. *J Bacteriol ,* 182**,** 1191-1199.

HOLLAND, I. B. & BLIGHT, M. A. 1999. ABC-ATPases, Adaptable Energy Generators Fuelling Transmembrane Movement of a Variety of Molecules in Organisms from Bacteria to Humans. *J. Mol. Biol.,* 293**,** 381-399.

HOPFNER, K.-P. & TAINER, J. A. 2003. Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures. *Current Opinion in Structural Biology,* 13**,** 249-255.

HUANG, F., HEDMAN, E., FUNK, C., KIESELBACH, T., SCHRODER, W. P. & NORLING, B. 2004. Isolation of outer membrane of *Synechocystis* sp. PCC 6803 and its proteomic characterization. *Mol Cell Proteomics,* 3**,** 586-95.

HUANG, F., PARMRYD, I., NILSSON, F., PERSSON, A. L., PAKRASI, H. B., ANDERSSON, B. & NORLING, B. 2002. Proteomics of *Synechocystis* sp. Strain PCC 6803. *Molecular & Cellular Proteomics,* 1**,** 956-966.

HUANG, S., WILHELM, S. W., HARVEY, H. R., TAYLOR, K., JIAO, N. & CHEN, F. 2012. Novel lineages of *Prochlorococcus* and *Synechococcus* in the global oceans. *ISME J,* 6**,** 285-97.

HUBNER, N. C., BIRD, A. W., COX, J., SPLETTSTOESSER, B., BANDILLA, P., POSER, I., HYMAN, A. & MANN, M. 2010. Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J Cell Biol,* 189**,** 739-54.

HUDEK, L., PREMACHANDRA, D., WEBSTER, W. A. & BRAU, L. 2016. Role of Phosphate Transport System Component PstB1 in Phosphate Internalization by *Nostoc punctiforme. Appl Environ Microbiol,* 82**,** 6344-6356.

HUNER, N. P. A., OQUIST, G. & SARHAN, F. 1998. Energy balance and acclimation to light and cold. *Trends in Plant Science,* 3**,** 224-230.

JAKI, B., HEILMANN, J. & STICHER, O. 2000. New Antibacterial Metabolites from the Cyanobacterium *Nostoc commune* (EAWAG 122b). *Journal of Natural Products,* 63**,** 1283-1285.

JIANG, S. & PAUL, J. H. 1994. Seasonal and diel abundance of viruses and occurrence of lysogeny/bacteriocinogeny in the marine environment. *Marine Ecology Progress Series,* 104**,** 163-172.

JOHNSON, Z. I., ZINSER, E. R., COE, A., MCNULTY, N. P., WOODWARD, E. M. & CHISHOLM, S. W. 2006. Niche Partitioning Among *Prochlorococcus* Ecotypes Along Ocean-Scale Environmental Gradients. *Science,* 311**,** 1737-1740.

JONES, P. M., O'MARA, M. L. & GEORGE, A. M. 2009. ABC transporters: a riddle wrapped in a mystery inside an enigma. *Trends in Biochemical Sciences,* 34**,** 520-531.

JURGENS, U. J. & WECKESSER, J. 1985. The fine structure and chemical composition of the cell wall and sheath layers of cyanobacteria. . *Ann Inst Pasteur/Microbiol,* 136**,** 41-44.

KANA, T. M. & GILBERT, P. M. 1987. Effect of irradiances up to 2000 µE m−2 s−1 on marine *Synechococcus* WH7803—I. Growth, pigmentation, and cell composition. *Deep Sea Research* 34**,** 479-495.

KANEKO, T., NAKAMURA, Y., WOLK, C. P., KURITZ, T., SASAMOTO, S., WATANABE, A., IRIGUCHI, M., ISHIKAWA, A., KAWASHIMA, K., KIMURA, T., KISHIDA, Y., KOHARA, M., MATSUMOTO, M., MATSUNO, A., MURAKI, A., NAKAZAKI, N., SHIMPO, S., SUGIMOTO, M., TAKAZAWA, M., YAMADA, M., YASUDA, M. & TABATA, S. 2001. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Research,* 8**,** 205-213.

KANEKO, T., SATO, S., KOTANI, H., TANAKA, A., ASAMIRU, E., NAKAMURA, Y., MIYAJIMA, N., HIROSAWA, M., SUGIURA, M., SASAMOTO, S., KIMURA, T., HOSOUCHI, T., MATSUNO, A., MURAKI, A., NAKAZAKI, N., NARUO, K., OKUMURA, S., SHIMPO, S., TAKEUCHI, C., WADA, T., WATANABE, A., YAMADA, M., YASUDA, M. & TABATA, S. 1996. Sequence Analysis of the Genome of the Unicellular Cyanobacterium *Synechocystis* sp. Strain PCC6803. II. Sequence Determination of the Entire Genome and Assignment of Potential Protein-coding Regions. *DNA Res.,* 3**,** 185-209.

KARL, D. M. & BJORKMAN, K. M. 2001. Phosphorus cycle in seawater- dissolved and particulate pool inventories and selected phosphorus fluxes. *Methods in microbiology,* 30**,** 239-270.

KASTING, J. F. 2001. The Rise of Atmospheric Oxygen. *Science,* 293**,** 819-820.

KASTING, J. F. & SIEFERT, J. L. 2002. Life and the evolution of Earth's atmosphere. *Science,* 296**,** 1066-1068.

KAUSHIK, B. D. & VENKATARAMAN, G. S. 1979. Effect Of Algal Inoculation On The Yield And Vitamin C Content Of Two Varieties Of Tomato. *Plant and Soil,* 52**,** 135-137.

KENT, A. G., BAER, S. E., MOUGINOT, C., HUANG, J. S., LARKIN, A. A., LOMAS, M. W. & MARTINY, A. C. 2019. Parallel phylogeography of *Prochlorococcus* and *Synechococcus*. *ISME J,* 13**,** 430-441.

KEREN, N., LIBERTON, M. & PAKRASI, H. B. 2005a. Photochemical competence of assembled photosystem II core complex in cyanobacterial plasma membrane. *J Biol Chem,* 280**,** 6548-53.

KEREN, N., OHKAWA, H., WELSH, E. A., LIBERTON, M. & PAKRASI, H. B. 2005b. Psb29, a conserved 22-kD protein, functions in the biogenesis of Photosystem II complexes in *Synechocystis* and *Arabidopsis*. *Plant Cell,* 17**,** 2768-81.

KIM, Y., JEON, J., KWAK, M. S., KIM, G. H., KOH, I. & RHO, M. 2018. Photosynthetic functions of *Synechococcus* in the ocean microbiomes of diverse salinity and seasons. *Plos One,* 13.

KISS, É., KÓS, P. B., CHEN, M. & VASS, I. 2012. A unique regulation of the expression of the *psb*A, *psb*D, and *psb*E genes, encoding the D1, D2 and cytochrome b_{559} subunits of the Photosystem II complex in the chlorophyll d containing cyanobacterium *Acaryochloris marina*. *BBA-Bioenergetics,* 1817**,** 1083-1094.

KITTLER, R., PELLETIER, L., MA, C., POSER, I., FISCHER, S., HYMAN, A. A. & BUCHHOLZ, F. 2005. RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells. *PNAS,* 102**,** 2396-2401.

KOEBNIK, R., LOCHER, K. P. & VAN GELDER, P. 2000. Structure and function of bacterial outer membrane proteins- barrels in a nutshell. *Mol Microbiol,* 37**,** 239-253.

KOEHN, F. E., LONGLEY, R. E. & REED, J. K. 1992. Microcolins A and B, new immunosuppressive peptides from the blue-green alga *Lyngbya majuscula*. *Journal of Natural Products,* 55**,** 613-619.

KOS, P. B., DEAK, Z., CHEREGI, O. & VASS, I. 2008. Differential regulation of *psb*A and *psb*D gene expression, and the role of the different D1 protein copies in the cyanobacterium *Thermosynechococcus elongatus* BP-1. *Biochim Biophys Acta,* 1777**,** 74-83.

KOZELKA, P. B. & BRULAND, K. W. 1998. Chemical speciation of dissolved Cu, Zn, Cd, Pb in Narragansett Bay, Rhode Island. *Marine Chemistry,* 60**,** 267-282.

KROGAN, N. J., CAGNEY, G., YU, H., ZHONG, G., GUO, X., IGNATCHENKO, A., LI, J., PU, S., DATTA, N., TIKUISIS, A. P., PUNNA, T., PEREGRIN-ALVAREZ, J. M., SHALES, M., ZHANG, X., DAVEY, M., ROBINSON, M. D., PACCANARO, A., BRAY, J. E., SHEUNG, A., BEATTIE, B., RICHARDS, D. P., CANADIEN, V., LALEV, A., MENA, F., WONG, P., STAROSTINE, A., CANETE, M. M., VLASBLOM, J., WU, S., ORSI, C., COLLINS, S. R., CHANDRAN, S., HAW, R., RILSTONE, J. J., GANDI, K., THOMPSON, N. J., MUSSO, G., ST ONGE, P., GHANNY, S., LAM, M. H., BUTLAND, G., ALTAF-UL, A. M., KANAYA, S., SHILATIFARD, A., O'SHEA, E., WEISSMAN, J. S., INGLES, C. J., HUGHES, T. R., PARKINSON, J., GERSTEIN, M., WODAK, S. J., EMILI, A. & GREENBLATT, J. F. 2006. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature,* 440**,** 637-43.

LABRIE, S. J., FROIS-MONIZ, K., OSBURNE, M. S., KELLY, L., ROGGENSACK, S. E., SULLIVAN, M. B., GEARIN, G., ZENG, Q., FITZGERALD, M., HENN, M. R. & CHISHOLM, S. W. 2013. Genomes of marine cyanopodoviruses reveal multiple origins of diversity. *Environ Microbiol,* 15**,** 1356-1376.

LAUBER, M. A., RUNNING, W. E. & REILLY, J. P. 2009. *B. subtilis* Ribosomal Proteins: Structural Homology and Post-Translational Modifications. *J. Proteome Res.,* 8**,** 4193-4206.

LEA-SMITH, D. J., BOMBELLI, P., VASUDEVAN, R. & HOWE, C. J. 2016. Photosynthetic, respiratory and extracellular electron transport pathways in cyanobacteria. *Biochim Biophys Acta,* 1857**,** 247-55.

LEA-SMITH, D. J., ROSS, N., ZORI, M., BENDALL, D. S., DENNIS, J. S., SCOTT, S. A., SMITH, A. G. & HOWE, C. J. 2013. Thylakoid terminal oxidases are essential for the cyanobacterium *Synechocystis* sp. PCC 6803 to survive rapidly changing light intensities. *Plant Physiol,* 162**,** 484-95.

LEE, J., SANDS, Z. A. & BIGGIN, P. C. 2016. A Numbering System for MFS Transporter Proteins. *Frontiers in Molecular Biosciences,* 3.

LI, W. K. W. 1994. Primary production of prochlorophytes, cyanobacteria and eucaryotic ultraphytoplankton- Measurements from flow cytometric sorting. *Limnol Oceanogr,* 39**,** 169-175.

LI, W. K. W. & WOOD, A. M. 1988. Vertical distribution of North Atlantic ultraphytoplankton- analysis by flow cytometry and epifluorescence microscopy. *Deep Sea Research,* 35**,** 1615-1638.

LINDBLAD, P. 1999. Cyanobacterial H_2 metabolism: knowledge and potential/strategies for a photobiotechnological production of H_2. *Biotechnologia Aplicada,* 16**,** 141-144.

LINDELL, D. & POST, A. F. 1995. Ultraphytoplankton succession is triggered by deep winter mixing in the Gulf of Aqaba (Eilat), Red Sea. *Limnology and Oceanography,* 40**,** 1130-1141.

LIU, H., BIDIGARE, R. R., LAWS, E., LANDRY, M. R. & CAMPBELL, L. 1999. Cell cycle and physiological characteristics of *Synechococcus* (WH7803) in chemostat culture. *Marine Ecology Progress Series,* 189**,** 17-25.

LOCHER, K. P., LEE, A. T. & REES, D. 2002. The *E. coli* BtuCD Structure- A Framework for ABC Transporter Architecture and Mechanism. *Science,* 296**,** 1091-1098.

LOHAN, M. C., STATHAM, P. J. & CRAWFORD, D. W. 2002. Total dissolved zinc in the upper water column of the subarctic North East Pacific. *Deep Sea Research II,* 49**,** 5793-5808.

LOLL, B., BROSER, M., KÓS, P. B., KERN, J., BIESIADKA, J., VASS, I., SAENGER, W. & ZOUNI, A. 2008. Modeling of variant copies of subunit D1 in the structure of photosystem II from *Thermosynechococcus elongatus*. *Biological Chemistry,* 389**,** 609-617.

LOLL, B., KERN, J., SAENGER, W., ZOUNI, A. & BIESIADKA, J. 2005. Towards complete cofactor arrangement in the 3.0 Å resolution structure of photosystem II. *Nature,* 438**,** 1040-4.

LOMAS, M. W., SWAIN, A., SHELTON, R. & AMMERMAN, J. W. 2004. Taxonomic variability of phosphorus stress in Sargasso Sea phytoplankton. *Limnol. Oceanogr.,* 49**,** 2303-2310.

MA, Y., YANG, M., LIN, X., LIU, X., HUANG, H. & GE, F. 2017. Malonylome Analysis Reveals the Involvement of Lysine Malonylation in Metabolism and Photosynthesis in Cyanobacteria. *J Proteome Res,* 16**,** 2030-2043.

MACCOSS, M. J. & YATES, J. R. I. 2001. Proteomics- analytical tools and techniques. *Current Opinion in Clinical Nutrition and Metabolic Care,* 4**,** 369-375.

MACKEY, K. R., PAYTAN, A., CALDEIRA, K., GROSSMAN, A. R., MORAN, D., MCILVIN, M. & SAITO, M. A. 2013. Effect of temperature on photosynthesis and growth in marine *Synechococcus* spp. *Plant Physiol,* 163**,** 815-29.

MACKEY, K. R., POST, A. F., MCILVIN, M. R., CUTTER, G. A., JOHN, S. G. & SAITO, M. A. 2015. Divergent responses of Atlantic coastal and oceanic *Synechococcus* to iron limitation. *Proc Natl Acad Sci U S A,* 112**,** 9944-9.

MACKEY, K. R. M., POST, A. F., MCILVIN, M. R. & SAITO, M. A. 2017. Physiological and proteomic characterization of light adaptations in marine *Synechococcus. Environ Microbiol,* 19**,** 2348-2365.

MANN, E. L., AHLGREN, N., MOFFETT, J. W. & CHISHOLM, S. W. 2002. Copper toxicity and cyanobacteria ecology in the Sargasso Sea. *Limnol. Oceanogr.,* 47**,** 976-988.

MANN, N. H. 1994. Protein phosphorylation in cyanobacteria. *Microbiology,* 140**,** 3207-3215.

MANN, N. H. 2003. Phages of the marine cyanobacterial picophytoplankton. *FEMS Microbiology Reviews,* 27**,** 17-34.

MANN, N. H., COOK, A., MILLARD, A., BAILEY, S. & CLOKIE, M. 2003. Bacterial photosynthesis genes in a virus. *Nature,* 424**,** 741-742.

MARIN, K., SUZUKI, I., YAMAGUCHI, K., RIBBECK, K., YAMAMOTO, H., KANESAKI, Y., HAGEMANN, M. & MURATA, N. 2003. Identification of histidine kinases that act as sensors in the perception of salt stress in *Synechocystis* sp. PCC 6803. *PNAS,* 100**,** 9061-9066.

MARSTON, M. F. & SALLEE, J. L. 2003. Genetic Diversity and Temporal Variation in the Cyanophage Community Infecting Marine *Synechococcus* Species in Rhode Island's Coastal Waters. *Appl Environ Microbiol,* 69**,** 4639-4647.

MAURIZI, M. R. 2002. Love it or cleave it: Tough choices in protein quality control. *Nature Structural Biology,* 9**,** 410-412.

MCCARREN, J. & BRAHAMSHA, B. 2009. Swimming motility mutants of marine *Synechococcus* affected in production and localization of the S-layer protein SwmA. *J Bacteriol,* 191**,** 1111-4.

MCCARREN, J., HEUSER, J., ROTH, R., YAMADA, N., MARTONE, M. & BRAHAMSHA, B. 2005. Inactivation of swmA results in the loss of an outer cell layer in a swimming *Synechococcus* strain. *J Bacteriol,* 187**,** 224-30.

MCDONALD, W. H. & YATES, J. R. 2002. Shotgun proteomics and biomarker discovery. *Disease Markers,* 18**,** 99-105.

MCGILLICUDDY, D. J., ANDERSON, L. A., DONEY, S. C. & MALTRUD, M. E. 2003. Eddy-driven sources and sinks of nutrients in the upper ocean: Results from a 0.1° resolution model of the North Atlantic. *Global Biogeochemical Cycles,* 17**,** n/a-n/a.

MERRITT, M. V., ROSENSTEIN, S. P., LOH, C., CHOU, R. H.-S. & ALLEN, M. M. 1991. A comparison of the major lipid classes and fatty acid composition of marine unicellular cyanobacteria with freshwater species. *Archives of Microbiology,* 155**,** 107-113.

MICHEL, K.-P. & PISTORIUS, E. K. 2004. Adaptation of the photosynthetic electron transport chain in cyanobacteria to iron deficiency: The function of IdiA and IsiA. *Physiologia plantarium,* 120**,** 36-50.

MILLARD, A., CLOKIE, M. R. J., SHUB, D. A. & MANN, N. H. 2004. Genetic organization of the psbAD region in phages infecting marine *Synechococcus* strains.. *PNAS,* 101**,** 11007-11012.

MILLARD, A. D., ZWIRGLMAIER, K., DOWNEY, M. J., MANN, N. H. & SCANLAN, D. J. 2009. Comparative genomics of marine cyanomyoviruses reveals the widespread occurrence of *Synechococcus* host genes localized to a hyperplastic region: implications for mechanisms of cyanophage evolution. *Environ Microbiol,* 11**,** 2370-87.

MIRUS, O., STRAUSS, S., NICOLAISEN, K., VON HAESELER, A. & SCHLEIFF, E. 2009. TonB-dependent transporters and their occurrence in cyanobacteria. *BMC Biol,* 7**,** 68.

MO, R., YANG, M., CHEN, Z., CHENG, Z., YI, X., LI, C., HE, C., XIONG, Q., CHEN, H., WANG, Q. & GE, F. 2015. Acetylome analysis reveals the involvement of lysine acetylation in photosynthesis and carbon metabolism in the model cyanobacterium *Synechocystis* sp. PCC 6803. *J Proteome Res,* 14**,** 1275-86.

MOFFETT, J. W. 1995. Temporal and spatial variability of copper complexation by strong chelators in the Sargasso Sea. *Deep Sea Research I,* 42**,** 1273-1295.

MOFFETT, J. W., BRAND, L. E., CROOT, P. L. & BARBEAU, K. A. 1997. Cu speciation and cyanobacterial distribution in harbors subject to anthropogenic Cu inputs. *Limnol. Oceanogr.,* 42**,** 789-799.

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., LAROCHE, J., LENTON, T.M., MAHOWALD, N.M., MARANON, E., MARINOV, I., MOORE, J.K., NAKATSUKA, T., OSCHLIES, A., SAITO, M.A., THINGSTAD, T.F., TSUDA, A. & ULLOA, O. 2013. Process and patterns of oceanic nutrient limitation. *Nat Geosci*, 6, 701-710.

MOORE, L. R., GOERICKE, R. & CHISHOLM, S. W. 1995. Comparative physiology of *Synechococcus* and *Prochlorococcus*- influence of light and temperature on growth, pigments, fluorescence and absorptive properties. *Marine Ecology Progress Series,* 116**,** 259-275.

MOORE, L. R., POST, A. F., ROCAP, G. & CHISHOLM, S. W. 2002. Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*.pdf. *Limnol. Oceanogr.,* 47**,** 989-996.

MOORE, L. R., ROCAP, G. & CHISHOLM, S. W. 1998. Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature,* 393**,** 464-467.

MOREL, A., AHN, Y. H., PARTENSKY, F., VAULOT, D. & CLAUSTRE, H. 1993. *Prochlorococcus* and *Synechococcus*: a comparative study of their optical properties in relation to their size and pigmentation. . *Journal of Marine Research,* 51**,** 617-649.

MOREL, F. M. 2008. The co-evolution of phytoplankton and trace element cycles in the oceans. *Geobiology,* 6**,** 318-24.

MORRIS, J. J., JOHNSON, Z. I., SZUL, M. J., KELLER, M. & ZINSER, E. R. 2011. Dependence of the cyanobacterium *Prochlorococcus* on hydrogen peroxide scavenging microbes for growth at the ocean's surface. *PLoS One,* 6**,** e16805.

MORRIS, J. J., KIRKEGAARD, R., SZUL, M. J., JOHNSON, Z. I. & ZINSER, E. R. 2008. Facilitation of robust growth of *Prochlorococcus* colonies and dilute liquid cultures by "helper" heterotrophic bacteria. *Appl Environ Microbiol,* 74**,** 4530-4.

MOUTIN, T., THINGSTAD, T. F., VAN WAMBEKE, F., MARIE, D., SLAWYK, G., RAIMBAULT, P. & CLAUSTRE, H. 2002. Does competition for nanomolar phosphate supply explain the predominance of the cyanobacterium *Synechococcus*? *Limnol. Oceanogr.,* 47**,** 1562-1567.

MÜLLER, J. & HEMPHILL, A. 2011. Identification of a host cell target for the thiazolide class of broad-spectrum anti-parasitic drugs. *Experimental Parasitology,* 128**,** 145-150.

MULLER, M. 2005. Twin-arginine-specific protein export in *Escherichia coli. . Research in Microbiology,* 156**,** 131-136.

MULLINEAUX, C. W., TOBIN, M. J. & JONES, G. R. 1997. Mobility of photosynthetic complexes in thylakoid membranes. *Nature,* 390**,** 421-424.

MUNEKAGE, Y., HASHIMOTO, M., MIYAKE, C., TOMIZAWA, K.-I., ENDO, T., TASAKA, M. & SHIKANAI, T. 2004. Cyclic electron flow around photosystem I is essential for photosynthesis. *Nature,* 429**,** 579-582.

MURPHY, L. S. & HAUGEN, E. M. 1985. The distribution and abundance of phototrophic ultraplankton in the North Atlantic. *Limnol. Oceanogr.,* 30**,** 47-58.

NAKAO, M., OKAMOTO, S., KOHARA, M., FUJISHIRO, T., FUJISAWA, T., SATO, S., TABATA, S., KANEKO, T. & NAKAMURA, Y. 2010. CyanoBase: the cyanobacteria genome database update 2010. *Nucleic Acids Res,* 38**,** D379-81.

NAKAYAMA, H., KUROKAWA, K. & LEE, B. L. 2012. Lipoproteins in bacteria: structures and biosynthetic pathways. *FEBS J,* 279**,** 4247-68.

NATALE, P., BRÜSER, T. & DRIESSEN, A. J. M. 2008. Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane—Distinct translocases and mechanisms. *BBA-Biomembranes,* 1778**,** 1735-1756.

NIKAIDO, H. & VAARA, M. 1985. Molecular basis of bacterial outer membrane permeability. *Microbiol Rev.,* 49**,** 1-32.

NIXON, P. J., BARKER, M., BOEHM, M., VRIES, R. D. & KOMENDA, J. 2005. FtsH-mediated repair of the photosystem II complex in response to light stress. *Journal of Experimental Botany,* 56**,** 357-363.

NOFFKE, N. 2008. Turbulent lifestyle: Microbial mats on Earth's sandy beaches—Today and 3 billion years ago. *GSA Today,* 18.

OGAWA, T. & MI, H. 2007. Cyanobacterial NADPH dehydrogenase complexes. *Photosynth Res,* 93**,** 69-77.

OLABARRIA, G., CARRASCOSA, J. L., DE PEDRO, M. A. & BERENGUER, J. 1996. A conserved motif in S-layer proteins is involved in peptidoglycan binding in *Thermus thermophilus*. *J Bacteriol ,* 178**,** 4765-4772.

OLIVEIRA, P., MARTINS, N. M., SANTOS, M., PINTO, F., BUTTEL, Z., COUTO, N. A. S., WRIGHT, P. C. & TAMAGNINI, P. 2016. The versatile TolC-like Slr1270 in the cyanobacterium *Synechocystis* sp. PCC 6803. *Environ Microbiol,* 18**,** 486-502.

OLSON, R. J., CHISHOLM, S., W., ZETTLER, E., R. & ARMBRUST, E. V. 1990. Pigments, size, and distribution of *Synechococcus* in the North Atlantic and Pacific Oceans. *Limnol. Oceanogr.,* 35**,** 45-58.

OLSON, R. J., CHISHOLM, S. W., ZETTLER, E., R. & ARMBRUST, E. V. 1988. Analysis of *Synechococcus* pigment types in the sea using single and dual beam flow cytometry. *Deep Sea Research,* 35**,** 425-440.

ONG, L. J. & GLAZER, A. N. 1991. Phycoerythrins of marine unicellular cyanobacteria. I. Bilin types and locations and energy transfer pathways in *Synechococcus* spp. phycoerythrins. *The Journal of Biological Chemistry,* 266**,** 9515-9527.

ONG, S.-E., BLAGOEV, B., KRATCHMAROVA, I., KRISTENSEN, D. B., STEEN, H., PANDEY, A. & MANN, M. 2002. Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Molecular & Cellular Proteomics,* 1**,** 376-386.

OW, S. Y., NOIREL, J., CARDONA, T., TATON, A., LINDBLAD, P., STENSJO, K. & WRIGHT, P. C. 2009. Quantitative Overview of N2 Fixation in *Nostoc punctiforme* ATCC 29133 through Cellular Enrichments and iTRAQ Shotgun Proteomics. *J Proteome Res,* 8**,** 187-198.

PACKER, J. E., MAHOOD, J. S., MORA-ARELLANO, V. O., SLATER, T. F., WILLSON, R. L. & WOLFENDEN, B. S. 1981. Free radicals and singlet oxygen scavengers: reaction of a peroxy-radical with beta-carotene, diphenyl furan and 1,4-diazobicyclo (2,2,2)-octane. *Biochemical and Biophysical Research Communications,* 98**,** 901-906.

PAERL, H. W. 1991. Ecophysiological and Trophic Implications of Light-Stimulated Amino Acid Utilization in Marine Picoplankton. *American Society for Microbiology,* 57**,** 473-479.

PAERL, R. W., FOSTER, R. A., JENKINS, B. D., MONTOYA, J. P. & ZEHR, J. P. 2008. Phylogenetic diversity of cyanobacterial *narB* genes from various marine habitats. *Environ Microbiol,* 10**,** 3377-3387.

PALENIK, B. 1994. Cyanobacterial community structure as seen from RNA polymerase gene sequence analysis. *Appl Environ Microbiol,* 60**,** 3212-3219.

PALENIK, B. 2001. Chromatic adaptation in marine *Synechococcus* strains. *Appl Environ Microbiol,* 67**,** 991-4.

PALENIK, B., BRAHAMSHA, B., LARIMER, F. W., LAND, M., HAUSER, L., CHAIN, P., LAMERDIN, J., REGALA, W., ALLEN, E. E., MCCARREN, J., PAULSEN, I., DUFRESNE, A., PARTENSKY, F., WEBB, E. A. & WATERBURY, J. B. 2003. The genome of a motile marine *Synechococcus*. *Nature,* 424**,** 1037-1042.

PALENIK, B., REN, Q., DUPONT, C. L., MYERS, G. S., HEIDELBERG, J. F., BADGER, J. H., MADUPU, R., NELSON, W. C., BRINKAC, L. M., DODSON, R. J., DURKIN, A. S., DAUGHERTY, S. C., SULLIVAN, A. S., KHOURI, H., MOHAMOUD, Y., HALPIN, R. & PAULSEN, I. T. 2006. Genome sequence of *Synechococcus* CC9311- Insights into adaptation to a coastal environment. *PNAS,* 103**,** 13555-13559.

PALTER, J. B., LOZIER, M. S. & BARBER, R. T. 2005. The effect of advection on the nutrient reservoir in the North Atlantic subtropical gyre. *Nature,* 437**,** 687-92.

PANDHAL, J., WRIGHT, P. C. & BIGGS, C. A. 2007. A Quantitative Proteomic Analysis of Light Adaptation in a Globally Significant Marine Cyanobacterium *Prochlorococcus marinus* MED4. *J Proteome Res,* 6**,** 996-1005.

PAO, S. S., PAULSEN, I. T. & SAIER, M. H. J. 1998. Major facilitator superfamily. *Microbiol Mol Biol Rev,* 62**,** 1-34.

PAPANIKOU, E., KARAMANOU, S. & ECONOMOU, A. 2007. Bacterial protein secretion through the translocase nanomachine. *Nat Rev Microbiol,* 5**,** 839-851.

PARTENSKY, F., BLANCHOT, J. & VAULOT, D. 1999. Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters- a review. *Bull Inst Oceanogr Monaco-Numero Special,* 19**,** 457-476.

PARTENSKY, F. & GARCZAREK, L. 2010. *Prochlorococcus*: advantages and limits of minimalism. *Ann Rev Mar Sci,* 2**,** 305-31.

PAULSEN, I. T., SLIWINSKI, M. K. & SAIER, M. H. J. 1998. Microbial genome analyses: global comparisons of transport capabilities based on phylogenies, bioenergetics and substrate specificities. *J. Mol. Biol.,* 277**,** 573-592.

PAYNE, C. D. & PRICE, N. M. 1999. Effects of cadmium toxicity on growth and elemental composition of marine phytoplankton. *Journal of Phycology,* 35**,** 292-302.

PENNO, S., LINDELL, D. & POST, A. F. 2006. Diversity of *Synechococcus* and *Prochlorococcus* populations determined from DNA sequences of the N-regulatory gene *ntcA*. *Environ Microbiol,* 8**,** 1200-11.

PEREZ-SEPULVEDA, B., PITT, F., N'GUYEN, A. N., RATIN, M., GARCZAREK, L., MILLARD, A. & SCANLAN, D. J. 2018. Relative stability of ploidy in a marine *Synechococcus* across various growth conditions. *Environ Microbiol Rep,* 10**,** 428-432.

PESCHEK, G. A., OBINGER, C., FROMWALD, S. & BERGMAN, B. 1994. Correlation between immuno-gold labels and activities of the cytochrome-c oxidase (aa_3-type) in membranes of salt stressed cyanobactria. *FEMS Microbiol Ecol,* 124**,** 431-438.

PESCHEK, G. A., WASTYN, M., TRNKA, M., MOLITOR, V., FRY, I. V. & PACKER, L. 1989. Characterization of the cytochrome c oxidase in isolated and purified plasma membranes from the cyanobacterium *Anacystis nidulans*. . *Biochemistry,* 28**,** 3057-3063.

PHUNG, L. T., AJLANI, G. & HASELKORN, R. 1994. P-type ATPase from the cyanobacterium *Synechococcus* 7942 related to the human Menkes and Wilson disease gene products. *PNAS,* 91**,** 9651-9654.

PISAREVA, T., KWON, J., OH, J., KIM, S., GE, C., WIESLANDER, A., CHOI, J. S. & NORLING, B. 2011. Model for membrane organization and protein sorting in the cyanobacterium *Synechocystis* sp. PCC 6803 inferred from proteomics and multivariate sequence analyses. *J Proteome Res,* 10**,** 3617-31.

PISAREVA, T., SHUMSKAYA, M., MADDALO, G., ILAG, L. & NORLING, B. 2007. Proteomics of *Synechocystis* sp. PCC 6803. Identification of novel integral plasma membrane proteins. *FEBS Journal,* 274**,** 791-804.

PITTERA, J., HUMILY, F., THOREL, M., GRULOIS, D., GARCZAREK, L. & SIX, C. 2014. Connecting thermal physiology and latitudinal niche partitioning in marine *Synechococcus*. *ISME J,* 8**,** 1221-36.

PITTERA, J., JOUHET, J., BRETON, S., GARCZAREK, L., PARTENSKY, F., MARECHAL, E., NGUYEN, N. A., DORE, H., RATIN, M., PITT, F. D., SCANLAN, D. J. & SIX, C. 2018. Thermoacclimation and genome adaptation of the membrane lipidome in marine *Synechococcus*. . *Environ Microbiol,* 20**,** 612-631.

PITTERA, J., PARTENSKY, F. & SIX, C. 2017. Adaptive thermostability of light-harvesting complexes in marine picocyanobacteria. *ISME J,* 11**,** 112-124.

POGORYELOV, D., REICHEN, C., KLYSZEJKO, A. L., BRUNISHOLZ, R., MULLER, D. J., DIMROTH, P. & MEIER, T. 2007. The oligomeric state of c rings from cyanobacterial F-ATP synthases varies from 13 to 15. *J Bacteriol,* 189**,** 5895-902.

POLYVIOU, D., MACHELETT, M. M., HITCHCOCK, A., BAYLAY, A. J., MACMILLAN, F., MOORE, C. M., BIBBY, T. S. & TEWS, I. 2018. Structural and functional characterization of IdiA/FutA (Tery_3377), an iron-binding protein from the ocean diazotroph *Trichodesmium erythraeum. Journal of Biological Chemistry,* 293**,** 18099-18109.

PORTEUS, B., KOCHARUNCHITT, C., NILSSON, R. E., ROSS, T. & BOWMAN, J. P. 2011. Utility of gel-free, label-free shotgun proteomics approaches to investigate microorganisms. *Applied Microbiology and Biotechnology,* 90**,** 407-416.

POST, A. F., DEDEJ, Z., GOTTLIEB, R., LI, H., THOMAS, D. N., EL-ABSAWI, M., EL-NAGGAR, A., EL-GHARABAWI, M. & SOMMER, U. 2002. Spatial and temporal distribution of *Trichodesmium* spp. in the stratified waters of the Gulf of Aqaba. *Marine Ecology Progress Series,* 239**,** 241-250.

POST, A. F., PENNO, S., ZANDBANK, K., PAYTAN, A., HUSE, S. M. & WELCH, D. M. 2011. Long term seasonal dynamics of synechococcus population structure in the gulf of aqaba, northern red sea. *Front Microbiol,* 2**,** 131.

PRICE, G. D., COLEMAN, J., R. & BADGER, M. R. 1992. Association of carbonic anhydrase activity with carboxysomes isolated from the cyanobacterium *Synechococcus* PCC7942. *Plant Physiol.,* 100**,** 784-793.

PUXTY, R. J., MILLARD, A. D., EVANS, D. J. & SCANLAN, D. J. 2014. Shedding new light on viral photosynthesis. *Photosynthesis Research,* 126**,** 71-97.

QIAO, J., WANG, J., CHEN, L., TIAN, X., HUANG, S., REN, X. & ZHANG, W. 2012. Quantitative iTRAQ LC-MS/MS proteomics reveals metabolic responses to biofuel ethanol in cyanobacterial *Synechocystis* sp. PCC 6803. *J Proteome Res,* 11**,** 5286-300.

QUEVEDO, M. & ANADON, R. 2001. Protist control of phytoplankton growth in the subtropical north-east Atlantic. *Marine Ecology Progress Series,* 221**,** 29-38.

RAIDA, M. 2011. Drug target deconvolution by chemical proteomics. *Curr Opin Chem Biol,* 15**,** 570-5.

RAJALAHTI, T., HUANG, F., KLEMENT, M. R., PISAREVA, T., EDMAN, M., SJOSTROM, M., WIESLANDER, A. & NORLING, B. 2007. Proteins in Different *Synechocystis* Compartments Have Distinguishing N-Terminal Features- A Combined Proteomics and Multivariate Sequence Analysis. *J Proteome Res,* 6**,** 2420-2434.

RAPPAPORT, F. & DINER, B. 2008. Primary photochemistry and energetics leading to the oxidation of the $(Mn)_4Ca$ cluster and to the evolution of molecular oxygen in Photosystem II. *Coordination Chemistry Reviews,* 252**,** 259-272.

RASTOGI, R. P., SONANI, R. R. & MADAMWAR, D. 2015. Cyanobacterial Sunscreen Scytonemin: Role in Photoprotection and Biomedical Research. *Applied Biochemistry and Biotechnology,* 176**,** 1551-1563.

RAUNIYAR, N., GAO, B., MCCLATCHY, D. B. & YATES, J. R. 2012. Comparison of Protein Expression Ratios Observed by Sixplex and Duplex TMT Labeling Method. *J Proteome Res,* 12**,** 1031-1039.

REDFIELD, A. C. 1958. The biological control of chemical factors in the environment. . *Am. Sci.,* 46**,** 205-221.

REES, D. C., JOHNSON, E. & LEWINSON, O. 2009. ABC transporters: the power to change. *Nat Rev Mol Cell Biol,* 10**,** 218-27.

REUMANN, S., DAVILA-APONTE, J. & KEEGSTRA, K. 1999. The evolutionary origin of the protein-translocating channel of chloroplastic envelope membranes- Identification of a cyanobacterial homolog. *Proc Natl Acad Sci U S A,* 96**,** 784-789.

ROBINSON, C. & BOLHUIS, A. 2004. Tat-dependent protein targeting in prokaryotes and chloroplasts. *Biochim Biophys Acta,* 1694**,** 135-47.

ROCAP, G., DISTEL, D. L., WATERBURY, J. B. & CHISHOLM, S. W. 2002. Resolution of *Prochlorococcus* and *Synechococcus* Ecotypes by Using 16S-23S Ribosomal DNA Internal Transcribed Spacer Sequences. *Appl Environ Microbiol,* 68**,** 1180-1191.

ROMAY, C., GONZÁLEZ, R., LEDÓN, N., REMIREZ, D. & RIMBAU, V. 2003. C-Phycocyanin- A Biliprotein with Antioxidant, Anti-Inflammatory and Neuroprotective Effects. *Current Protein and Peptide Science,* 4**,** 207-216.

ROOSE, J. L., WEGENER, K. M. & PAKRASI, H. B. 2007. The extrinsic proteins of Photosystem II. *Photosynthesis Research,* 92**,** 369-387.

ROQUE, A. C. & LOWE, C. R. 2008. Affinity chromatography: history, perspectives, lmitations and prospects. *In:* ZACHARIOU, M. (ed.) *Affinity Chromatography: Methods and Protocols 2nd Edition.* Human Press.

RUBIO, L. M., HERRERO, A. & FLORES, E. 1996. A cyanobacterial narB gene encodes a ferredoxin-dependent nitrate reductase. *Plant Molecular Biology,* 30**,** 845-850.

RUBIO, V., SHEN, Y., SAIJO, Y., LIU, Y., GUSMAROLI, G., DINESH-KUMAR, S. P. & DENG, X. W. 2005. An alternative tandem affinity purification strategy applied to *Arabidopsis* protein complex isolation. . *Plant J,* 41**,** 767-78.

SAITO, M., MOFFETT, J. W., CHISHOLM, S. W. & WATERBURY, J. B. 2002. Cobalt limitation and uptake in *Prochlorococcus*. *Limnol. Oceanogr.,* 47**,** 1629-1636.

SAITO, M. A., MOFFETT, J. W. & DITULLIO, G. R. 2004. Cobalt and nickel in the Peru upwelling region: A major flux of labile cobalt utilized as a micronutrient. *Global Biogeochemical Cycles,* 18.

SAITO, M. A., ROCAP, G. & MOFFETT, J. W. 2005. Production of cobalt binding ligands in a *Synechococcus* feature at the Costa Rica upwelling dome. *Limnol. Oceanogr.,* 50**,** 279-290.

SAITO, M. A., SIGMAN, D. M. & MOREL, F. M. M. 2003. The bioinorganic chemistry of the ancient ocean: the co-evolution of cyanobacterial metal requirements and biogeochemical cycles at the Archean–Proterozoic boundary? *Inorganica Chimica Acta,* 356**,** 308-318.

SAKAMOTO, T. INOUE-SAKAMOTO, K. & BRYANT, D.A. 1999. A novel nitrate/nitrite permease in the marine cyanobacterium *Synechococcus* sp. Strain PCC7002. *J Bacteriol,* 181, 7363-7372.

SALOMON, E. & KEREN, N. 2011. Manganese Limitation Induces Changes in the Activity and in the Organization of Photosynthetic Complexes in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803. *Plant Physiol,* 155**,** 571-579.

SANDER, J., NOWACZYK, M., BUCHTA, J., DAU, H., VASS, I., DEAK, Z., DOROGI, M., IWAI, M. & ROGNER, M. 2010. Functional characterization and quantification of the alternative *Psb*A copies in *Thermosynechococcus elongatus* and their role in photoprotection. *J Biol Chem,* 285**,** 29851-6.

SANDERS, C. E., MELIS, A. & ALLEN, J. F. 1989. In vivo phosphorylation of proteins in the cyanobacterium *Synechococcus* 6301 after chromatic acclimation to photosystem I or photosystem II light. *Biochimica et Biophysica Acta***,** 168-172.

SANKARAN, K. & WU, H. C. 1993. Bacterial lipoproteins. *In:* SCHLESINGER, M. J. (ed.) *Lipid modifications of proteins. .* Boca Raton: CRC Press.

SATO, M., YOSHIKAWA, T., TAKEDA, S. & FURUYA, K. 2007. Application of the size-fractionation method to simultaneous estimation of clearance rates by heterotrophic flagellates and ciliates of pico- and nanophytoplankton. *Journal of Experimental Marine Biology and Ecology,* 349**,** 334-343.

SAUER, J., SCHREIBER, U., SCHMID, R., VOLKER, U. & FORCHHAMMER, K. 2001. Nitrogen starvation-induced chlorosis in *Synechococcus* PCC 7942. Low-level photosynthesis as a mechanism of long-term survival. *Plant Physiol,* 126**,** 233-243.

SCANLAN, D. J., OSTROWSKI, M., MAZARD, S., DUFRESNE, A., GARCZAREK, L., HESS, W. R., POST, A. F., HAGEMANN, M., PAULSEN, I. & PARTENSKY, F. 2009. Ecological genomics of marine picocyanobacteria. *Microbiol Mol Biol Rev,* 73**,** 249-99.

SCHAERLAEKENS, K., SCHIEROVA, M., LAMMERTYN, E., GEUKENS, N., ANNE, J. & VAN MELLAERT, L. 2001. Twin-arginine translocation pathway in *Streptomyces lividans*. *J Bacteriol,* 183**,** 6727-32.

SCHMITT, L. & TAMPE, R. 2002. Structure and mechanism of ABC transporters. *Current Opinion in Structural Biology,* 12**,** 754-760.

SCHRADER, P. S., MILLIGAN, A. J. & BEHRENFELD, M. J. 2011. Surplus Photosynthetic Antennae Complexes Underlie Diagnostics of Iron Limitation in a Cyanobacterium. *PLoS ONE,* 6.

SCHULZ, G. E. 1993. Bacterial porins- structure and function. *Current opinion in Cell Biology,* 5**,** 701-707.

SCHWARZ, R. & FORCHHAMMER, K. 2005. Acclimation of unicellular cyanobacteria to macronutrient deficiency: emergence of a complex network of cellular responses. *Microbiology,* 151**,** 2503-14.

SEN, K. & NIKAIDO, H. 1991. Lipopolysaccharide Structure Required for In Vitro Trimerization of *Escherichia coli* OmpF Porin.pdf. *J Bacteriol ,* 173**,** 926-928.

SEN, S., AGRAWAL, C., MISHRA, Y., RAI, S., CHATTERJEE, A., YADAV, S., SINGH, S. & RAI, L. C. 2015. Exploring the membrane proteome of the diazotropic cyanobacterium *Anabaena* PCC7120 through gel-based proteomics and in silico approaches. *J Proteomics,* 127**,** 161-8.

SETHURAMAN, M., MCCOMB, M. E., HUANG, H., HUANG, S., HEIBECK, T., COSTELLO, C. E. & COHEN, R. A. 2004. Isotope-Coded Affinity Tag (ICAT) Approach to Redox Proteomics- Identification and Quantitation of Oxidant-Sensitive Cysteine Thiols in Complex Protein Mixtures. *J Proteome Res,* 3**,** 1228-1233.

SHAKED, Y., XU, Y., LEBLANC, K. & MOREL, F. M. M. 2006. Zinc availability and alkaline phosphatase activity in *Emiliania huxleyi-* Implications for Zn–P co-limitation in the ocean. *Limnol. Oceanogr.,* 51**,** 299-309.

SHARFF, A., FANUTTI, C., SHI, J., CALLADINE, C. & LUISI, B. 2001. The role of the TolC family in protein transport and multidrug efflux. From stereochemical certainty to mechanistic hypothesis. *Eur. J. Biochem.,* 268**,** 5011-5026.

SHIH, P.M., WU, D.Y., LATIFI, A., AXEN, S.D., FEWER, D.P., TALLA, E., CALTEAU, A., CAI, F., MARSAC, N.T.D., RIPPKA, R., HERDMAN, M., SIVONEN, K., COURSIN, T., LAURENT, T., GOODWIN, L., NOLAN, M., DAVENPORT, K.W., HAN, C.S., RUBIN, E.M., EISEN, J.A., WOYKE, T., GUGGER, M. & KERFELD, C.A. 2013. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. PNAS, 110, 1053-1058.

SHVAREV, D. & MALDENER, I. 2018. ATP-binding cassette transporters of the multicellular cyanobacterium *Anabaena* sp. PCC 7120: a wide variety for a complex lifestyle. *FEMS Microbiology Letters,* 365.

SIDLER, W.A. 1994. Phycobilisome and phycobiliprotein structures. *In:* BRYANT, D.A. (ed.) *The Molecular Biology of Cyanobacteria*. The Netherlands: Kluwe Academic Publishers.

SIX, C., THOMAS, J. C., BRAHAMSHA, B., LEMOINE, Y. & PARTENSKY, F. 2004. Photophysiology of the marine cyanobacterium *Synechococcus* sp. WH8102, a new model organism. *Aquat Microb Ecol,* 35**,** 17-29.

SIX, C., THOMAS, J. C., GARCZAREK, L., OSTROWSKI, M., DUFRESNE, A., BLOT, N., SCANLAN, D. J. & PARTENSKY, F. 2007. Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. *Genome Biol,* 8**,** R259.

SLOBODIN, B. & GERST, J. E. 2011. RaPID: an aptamer-based mRNA affinity purification technique for the identification of RNA and protein factors present in

ribonucleoprotein complexes. *In:* GERST, J. E. (ed.) *RNA Detection and Visualization methods and Protocols* New York: Humana Press.

SOHM, J. A., AHLGREN, N. A., THOMSON, Z. J., WILLIAMS, C., MOFFETT, J. W., SAITO, M. A., WEBB, E. A. & ROCAP, G. 2016. Co-occurring *Synechococcus* ecotypes occupy four major oceanic regimes defined by temperature, macronutrients and iron. *ISME J,* 10**,** 333-45.

SRIVASTAVA, R., PISAREVA, T. & NORLING, B. 2005. Proteomic studies of the thylakoid membrane of *Synechocystis* sp. PCC 6803. *Proteomics,* 5**,** 4905-16.

STAINER, R. Y. & COHEN-BAZIRE, G. 1977. Phototrophic prokaryotes- the cyanobacteria. *Annu Rev Microbiol,* 31**,** 225-274.

STANISŁAWSKA-SACHADYN, A., SACHADYN, P., IHLE, K., SYDORCZUK, C., WIEJACHA, K. & KUR, J. 2006. The construction of bifunctional fusion proteins consisting of MutS and GFP. *Journal of Biotechnology,* 121**,** 134-143.

STAUBER, J. L. & FLORENCE, T. M. 1985. The influence of iron on copper toxicity to the marine diatom, *Nitzschia closterium* (ehrenberg) W. Smith. *Aquatic Toxicology,* 6**,** 297-305.

STEINBERG, D. K., CARLSON, C. A., BATES, N. R., JOHNSON, R. J., MICHAELS, A. F. & KNAP, A. H. 2001. Overview of the US JGOFS Bermuda Atlantic Time-Series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. *Deep Sea Research II,* 48**,** 1405-1447.

STENSJO, K., OW, S. Y., BARRIOS-LLERENA, M. E., LINDBLAD, P. & WRIGHT, P. C. 2007. An iTRAQ-Based Quantitative Analysis To Elaborate the Proteomic Response of *Nostoc* sp. PCC 7120 under N_2 Fixing Conditions. *J Proteome Res,* 6**,** 621-635.

STOCK, A. M., ROBINSON, V., L. & GOUDREAU, P., N. 2000. Two-Component Signal Transduction. *Annu Rev Biochem,* 69**,** 183-215.

STOCKNER, J. G., CALLIERI, C. & CRONBERG, C. 2000. Picoplankton and other non-bloom-forming cyanobacteria in lakes. *In:* WHILLON, B. A. & POTTS, M. (eds.) *In The Ecology of Cyanobacteria. .* Dordrecht:Kluwer.

STOMP, M., HUISMAN, J., JONGH, F., VERAART, A. J., GERLA, D., RIJKEBOER, M., IBELINGS, B. W., WOLLENZIEN, U. I. A. & STAL, L. J. 2004. Adaptive divergence in pigment composition promotes phytoplankton biodiversity. *Nature,* 432**,** 104-107.

STOMP, M., HUISMAN, J., VOROS, L., PICK, F. R., LAAMANEN, M., HAVERKAMP, T. & STAL, L. J. 2007. Colourful coexistence of red and green picocyanobacteria in lakes and seas. . *Ecol Lett,* 10**,** 290-8.

STREHL, B., HOLTZENDORFF, J., PARTENSKY, F. & HESS, W. R. 1999. A small and compact genome in the marine cyanobacterium *Prochlorococcus marinus* CCMP 1375: lack of an intron in the gene for tRNA(Leu)^{UAA} and a single copy of the rRNA operon. *FEMS Microbiology Letters,* 181**,** 261-266.

STUART, R. K., BRAHAMSHA, B. & PALENIK, B. 2014. A Deg-protease family protein in marine *Synechococcus* is involved in outer membrane protein organization. . *Frontiers in Marine Science,* 1.

STUART, R. K., DUPONT, C. L., JOHNSON, D. A., PAULSEN, I. T. & PALENIK, B. 2009. Coastal strains of marine *Synechococcus* species exhibit increased tolerance to copper shock and a distinctive transcriptional response relative to those of open-ocean strains. *Appl Environ Microbiol,* 75**,** 5047-57.

SUDESH, K., TAGUCHI, K. & DOI, Y. 2002. Effect of increased PHA synthase activity on polyhydroxyalkanoates biosynthesis in *Synechocystis* sp. PCC6803. *International Journal of Biological Macromolecules* 30**,** 97-104.

SUGA, M., AKITA, F., HIRATA, K., UENO, G., MURAKAMI, H., NAKAJIMA, Y., SHIMIZU, T., YAMASHITA, K., YAMAMOTO, M., AGO, H. & SHEN, J. R. 2015. Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature,* 517**,** 99-103.

SUGIURA, M., KATO, Y., TAKAHASHI, R., SUZUKI, H., WATANABE, T., NOGUCHI, T., RAPPAPORT, F. & BOUSSAC, A. 2010. Energetics in Photosystem II from *Thermosynechococcus elongatus* with a D1 protein encoded by either the *psb*A$_1$ or *psb*A$_3$ gene. *BBA-Bioenergetics,* 1797**,** 1491-1499.

SULLIVAN, M.B., LINDELL, D., LEE, J.A., THOMPSON, L.R., BIELAWSKI, J.P. & CHISHOLM, S.W. 2006. Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLOS Biol,* 4, 1344-1357.

SUNDA, W. G. 1989. Trace Metal Interactions with Marine Phytoplankton. *Biological Oceanography,* 6**,** 411-442.

SUNDA, W. G. 2012. Feedback Interactions between Trace Metal Nutrients and Phytoplankton in the Ocean. *Front Microbiol,* 3**,** (Article 204).

SUNDA, W. G. & HUNSTMAN, S. A. 1995. Cobalt and zinc interreplacement in marine phytoplankton biological and geochemical implications. *Limnol. Oceanogr.,* 40**,** 1404-1417.

SUNDA, W. G. & HUNSTMAN, S. A. 1998. Interactive effects of external manganese, the toxic metals copper and zinc, and light in controlling cellular manganese and growth in a coastal diatom. *Limnol. Oceanogr.,* 43**,** 1467-1475.

SUTTLE, C. A. 2000. Cyanophages and their role in the ecology of cyanobacteria. *In:* WHITTON, B. A. & POTTS, M. (eds.) *The ecology of cyanobacteria: their diversity in time and space.* Boston, MA.: Kulwer Academic Publishers.

SUTTLE, C. A. & CHAN, A. M. 1993. Marine cyanophages infecting oceanic and coastal strains of *Synechococcus*: abundance, morphology, cross-infectivity and growth characteristics. . *Marine Ecology Progress Series,* 92**,** 99-109.

SUTTLE, C. A. & CHAN, A. M. 1994. Dynamics and distribution of cyanophages and their effect on marine *Synechococcus* spp. *Appl Environ Microbiol,* 60**,** 3167-3174.

TABB, D. L., FERNANDO, C. G. & CHAMBERS, M. C. 2007. MyriMatch: highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J Proteome Res,* 6**,** 654-61.

TABB, D. L., NARASIMHAN, C., STRADER, M. B. & HETTICH, R. L. 2005. DBDigger- Reorganized Proteomic Database Identification That Improves Flexibility and Speed. *Anal. Chem.,* 77**,** 2464-2474.

TAI, V. & PALENIK, B. 2009. Temporal variation of *Synechococcus* clades at a coastal Pacific Ocean monitoring site. *The ISME Journal,* 3**,** 903-915.

TAKAHASHI, H., KOPRIVA, S., GIORDANO, M., SAITO, K. & HELL, R. 2011. Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. *Annu Rev Plant Biol,* 62**,** 157-84.

TANG, D., WARNKEN, K. W. & SANTSCHI, P. H. 2001. Organic complexation of copper in surface waters of Galveston Bay. *Limnol. Oceanogr.,* 46**,** 321-330.

TANG, K., JIAO, N., LIU, K., ZHANG, Y. & LI, S. 2012. Distribution and functions of TonB-dependent transporters in marine bacteria and environments: implications for dissolved organic matter utilization. *PLoS One,* 7**,** e41204.

TASHIAN, R. E. 1989. The carbonic anhydras: widening perspectives on their evolution, expression and function. *Bioessays,* 10**,** 186-192.

TETU, S. G., BRAHAMSHA, B., JOHNSON, D. A., TAI, V., PHILLIPPY, K., PALENIK, B. & PAULSEN, I. T. 2009. Microarray analysis of phosphate regulation in the marine cyanobacterium *Synechococcus* sp. WH8102. *ISME J,* 3**,** 835-49.

THINGSTAD, T. F., KROM, M. D., MANTOURA, R. F. C., FLATEN, G. A. F., GROOM, S., HERUT, B., KRESS, N., LAW, C. S., PASTERNAK, A., PITTA, P., PSARRA, P., RASSOULZADEGAN, F., TANAKA, T., TSELEPIDES, A., WASSMANN, P., WOODWARD, E. M. S., RISER, C. W., ZODIATIS, G. & ZOHARY, T. 2005. Nature of phosphorus limitation in the ultraoligotrophic eastern Mediterranean. *Science,* 309**,** 1068-1071.

THOMPSON, L. R., ZENG, Q., KELLY, L., HUANG, K. H., SINGER, A. U., STUBBE, J. & CHISHOLM, S. W. 2011. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc Natl Acad Sci U S A,* 108**,** E757-64.

THURSTON, T. L., RYZHAKOV, G., BLOOR, S., VON MUHLINEN, N. & RANDOW, F. 2009. The TBK1 adaptor and autophagy receptor NDP52 restricts the proliferation of ubiquitin-coated bacteria. *Nat Immunol,* 10**,** 1215-21.

TIAN, X., CHEN, L., WANG, J., QIAO, J. & ZHANG, W. 2013. Quantitative proteomics reveals dynamic responses of *Synechocystis* sp. PCC 6803 to next-generation biofuel butanol. *J Proteomics,* 78**,** 326-45.

TOLEDO, G. & PALENIK, B. 1997. *Synechococcus* Diversity in the California Current as Seen by RNA Polymerase (rpoC1) Gene Sequences of Isolated Strains. *Appl Environ Microbiol,* 63**,** 4298-4303.

TOLEDO, G. & PALENIK, B. 2003. A *Synechococcus* serotype is found preferentially in surface marine waters. *Limnol. Oceanogr.,* 48**,** 1744-1755.

TOMBOLINI, R., UNGE, A., DAVEY, M. E., DE BRUIJN, F. J. & JANSSON, J. K. 1997. Flow cytometric and microscopic analysis of GFP-tagged *Pseudomonas fluorescens* bacteria. *FEMS Microbiol Ecology,* 22**,** 17-28.

TOTH, T. N., CHUKHUTSINA, V., DOMONKOS, I., KNOPPOVA, J., KOMENDA, J., KIS, M., LENART, Z., GARAB, G., KOVACS, L., GOMBOS, Z. & VAN AMERONGEN, H. 2015. Carotenoids are essential for the assembly of cyanobacterial photosynthetic complexes. *Biochim Biophys Acta,* 1847**,** 1153-65.

TRAUTMANN, D., VOSS, B., WILDE, A., AL-BABILI, S. & HESS, W. R. 2012. Microevolution in cyanobacteria: re-sequencing a motile substrain of *Synechocystis* sp. PCC 6803. *DNA Res,* 19**,** 435-48.

TSAI, B. P., WANG, X., HUANG, L. & WATERMAN, M. L. 2011. Quantitative Profiling of In Vivo-assembled RNA-Protein Complexes Using a Novel Integrated Proteomic Approach. *Molecular & Cellular Proteomics,* 10**,** 1-15.

TULK, B. M., SCHLESINGER, P. H., KAPADIA, S. A. & EDWARDS, J. C. 2000. CLIC-1 functions as a chloride channel when expressed and purified from bacteria. *J Biol Chem,* 275**,** 26986-93.

TWINING, B. S., NUNEZ-MILLAND, D., VOGT, S., JOHNSON, R. S. & SEDWICK, P. N. 2010. Variations in *Synechococcus* cell quotas of phosphorus, sulfur, manganese, iron, nickel, and zinc within mesoscale eddies in the Sargasso Sea. . *Limnol. Oceanogr.,* 55**,** 492-506.

UMENA, Y., KAWAKAMI, K., SHEN, J. R. & KAMIYA, N. 2011. Crystal structure of oxygen-evolving photosystem II at 1.9 Å resolution. *Nature,* 473**,** 55-60.

VALLADARES, A., MONTESINOS, M. L., HERRERO, A. & FLORES, E. 2002. An ABC-type, high-affinity urea permease identified in cyanobacteria. *Mol Microbiol,* 43**,** 703-715.

VAN DER HENST, C., CHARLIER, C., DEGHELT, M., WOUTERS, J., MATROULE, J.-Y., LETESSON, J.-J. & DE BOLLE, X. 2010. Overproduced *Brucella abortus* PdhS-mCherry forms soluble aggregates in *Escherichia coli*, partially associating with mobile foci of IbpA-YFP. *BMC Microbiology,* 10**,** 248.

VARKEY, D., MAZARD, S., JEFFRIES, T. C., HUGHES, D. J., SEYMOUR, J., PAULSEN, I. T. & OSTROWSKI, M. 2018. Stormwater influences phytoplankton assemblages within the diverse, but impacted Sydney Harbour estuary. *PLoS One,* 13**,** e0209857.

VARKEY, D., MAZARD, S., OSTROWSKI, M., TETU, S. G., HAYNES, P. & PAULSEN, I. T. 2016. Effects of low temperature on tropical and temperate isolates of marine *Synechococcus*. *ISME J,* 10**,** 1252-63.

VARKONYI, Z., MASAMOTO, K., DEBRECZENY, M., ZSIROS, O., UGHY, B., GOMBOS, Z., DOMONKOS, I., FARKAS, T., WADA, H. & SZALONTAI, B. 2002. Low-temperature-induced accumulation of xanthophylls and its structural consequences in the photosynthetic membranes of the cyanobacterium *Cylindrospermopsis raciborskii-* An FTIR spectroscopic study. *PNAS,* 99**,** 2410-2415.

VERAKSA, A., BAUER, A. & ARTAVANIS-TSAKONAS, S. 2005. Analyzing protein complexes in *Drosophila* with tandem affinity purification-mass spectrometry. *Dev Dyn,* 232**,** 827-34.

VINCENT, W. 2000. Cyanobacterial dominance in the polar regions. *In B.A. Whitton and M. Potts [eds.], The ecology of cyanobacteria. Kluwer Academic***,** 321-340.

VINYARD, D. J., GIMPEL, J., ANANYEV, G. M., CORNEJO, M. A., GOLDEN, S. S., MAYFIELD, S. P. & DISMUKES, G. C. 2013. Natural variants of photosystem II subunit D1 tune photochemical fitness to solar intensity. *J Biol Chem,* 288**,** 5451-62.

VON HEIJNE, G. 1985. Signal sequences. The limits of variation. . *J. Mol. Biol.,* 184**,** 99-105.

VON HEIJNE, G. 1989. The structure of signal peptides from bacterial lipoproteins. . *Protein Eng. ,* 2**,** 531-534.

VOULHOUX, R., BOS, M. P., GEURTSEN, J., MOLS, M. & TOMMASSEN, J. 2003. Role of a highly conserved bacterial protein in outer membrane protein assembly. *Science,* 299**,** 262-265.

WADA, H., SCHMIDT, H., HEINZ, E. & MURATA, N. 1993. In Vitro Ferredoxin-Dependent Desaturation of Fatty Acids in Cyanobacterial Thylakoid Membranes. *J Bacteriol ,* 175**,** 544-547.

WANG, K., WOMMACK, K. E. & CHEN, F. 2011. Abundance and Distribution of *Synechococcus* spp. and Cyanophages in the Chesapeake Bay. *Appl Environ Microbiol,* 77**,** 7459-7468.

WANG, W., ZHOU, H., LIN, H., ROY, S., SHALER, T. A., HILL, L. R., NORTON, S., KUMAR, P., ANDERLE, M. & BECKER, C. H. 2003. Quantification of Proteins and Metabolites by Mass Spectrometry without Isotopic Labeling or Spiked Standards. *Anal. Chem.,* 75**,** 4818-4826.

WATERBURY, J. B. & RIPPKA, R. 1989. Order Chroococcales. *In:* STALEY, J. T., BRYANT, M. P., PFENNING, N. & HOLT, J. G. (eds.) *Bergey's manual of systematic bacteriology.* Williams Wilkins, Baltimore, MD.

WATERBURY, J. B. & VALOIS, F. W. 1993. Resistance to co-occurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. *Appl Environ Microbiol,* 59**,** 3393-3399.

WATERBURY, J. B., WATSON, S. W., GUILLARD, R. R. L. & BRAND, L. E. 1979. Widespread occurrence of a unicellular, marine planktonic, cyanobacterium. *Nature,* 277**,** 293-294.

WATERBURY, J. B., WATSON, S. W., VALOIS, F. W. & FRANKS, D. G. 1986. Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*. . *Can Bull Fish Aquat Sci,* 472**,** 71-120.

WEGENER, K. M., SINGH, A. K., JACOBS, J. M., ELVITIGALA, T., WELSH, E. A., KEREN, N., GRITSENKO, M. A., GHOSH, B. K., CAMP, D. G., 2ND, SMITH, R. D. & PAKRASI, H. B. 2010. Global proteomics reveal an atypical strategy for carbon/nitrogen assimilation by a cyanobacterium under diverse environmental perturbations. *Mol Cell Proteomics,* 9**,** 2678-89.

WEI, X., HERBST, A., MA, D., AIKEN, J. & LI, L. 2011. A quantitative proteomic approach to prion disease biomarker research: delving into the glycoproteome. *J Proteome Res,* 10**,** 2687-702.

WEITZ, J. & WILHELM, S. 2012. Ocean viruses and their effects on microbial communities and biogeochemical cycles. *F1000 Biology Reports,* 4.

WHITTON, B. A. 2012. Ecology of Cyanobacteria II:their diversity in space and time. *Springer, Dordrecht***,** 1-753.

WICKNER, S., MAURIZI, M. R. & GOTTESMAN, S. 1999. Posttranslational Quality Control: Folding, Refolding, and Degrading Proteins. *Science,* 286**,** 1888-1893.

WIESE, S., REIDEGELD, K. A., MEYER, H. E. & WARSCHEID, B. 2007. Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. *Proteomics,* 7**,** 340-350.

WILKENS, S. 2015. Structure and mechanism of ABC transporters. *F1000Prime Reports,* 7.

WILKINS, M. R., SANCHEZ, J. C., GOOLEY, A. A., APPEL, R. D., HUMPHREY-SMITH, I., HOCHSTRASSER, D. F. & WILLIAMS, K. L. 1996. From proteins to proteomes- large scale protein identification by two-dimensional electrophoresis and arnino acid analysis. *Nature Biotechnology,* 14**,** 61-65.

WILLIAMS, J. G. K. 1988. Construction of specific mutations in photosystem II photosynthetic reaction center by genetic-engineering methods in *Synechocystis* 6803. *Methods in enzymology,* 167**,** 766-778.

WOOD, A. M. 1985. Adaptation of photosynthetic apparatus of marine ultraphytoplankton to natural light fields. *Nature,* 316**,** 253-255.

WOOD, A.M., LIPSEN, M. & COBLE, P. 1999. Fluorescence-based characterization of phycoerythrin-containing cyanobacterial communities in the Arabian Sea during the Northeast and early Southwest Monsoon (1994-1995). *Deep-Sea Res II*, 46, 1769-1790.

WOOD, A. M., PHINNEY, D. A. & YENTSCH, C. S. 1998. Water column transparency and the distribution of spectrally distinct forms of phycoerythrin-containing organisms. *Marine Ecology Progress Series,* 162**,** 25-31.

WU, J., SUNDA, W., BOYLE, E. A. & KARL, D. M. 2000. Phosphate depletion in the Western North Atlantic Ocean. *Science,* 289**,** 759-762.

WYMAN, M., GREGORY, R. P. F. & CARR, N. G. 1985. Novel role for phycoerythrin in a marine cyanobacterium *Synechococcus* strain DC-2. *Science,* 230**,** 818-820.

XIA, X., VIDYARATHNA, N. K., PALENIK, B., LEE, P. & LIU, H. 2015. Comparison of the seasonal variations of *Synechococcus* assemblage structures in estuarine waters and coastal waters of Hong Kong. *Appl Environ Microbiol,* 81**,** 7644-55.

XIONG, Q., FENG, J., LI, S.-T., ZHANG, G.-Y., QIAO, Z.-X., CHEN, Z., WU, Y., LIN, Y., LI, T., GE, F. & ZHAO, J.-D. 2015. Integrated Transcriptomic and Proteomic Analysis of the Global Response of *Synechococcus* to High Light Stress. *Molecular & Cellular Proteomics,* 14**,** 1038-1053.

XU, W., TANG, H., WANG, Y. & CHITNIS, P. R. 2001. Proteins of the cyanobacterial photosystem I. *Biochimica et Biophysica Acta,* 1507**,** 32-40.

YAMAMOTO, Y. 2001. Quality Control of Photosystem II. *Plant and Cell Physiology,* 42**,** 121-128.

YANG, M. K., QIAO, Z. X., ZHANG, W. Y., XIONG, Q., ZHANG, J., LI, T., GE, F. & ZHAO, J. D. 2013. Global phosphoproteomic analysis reveals diverse functions of serine/threonine/tyrosine phosphorylation in the model cyanobacterium *Synechococcus* sp. strain PCC 7002. *J Proteome Res,* 12**,** 1909-23.

YANG, T. T. C. & CHOW, C.-W. 2011. Elucidating Protein: DNA complex by Oligonucleotide DNA Affinity Purification *In:* VANCURA, A. (ed.) *Transcriptional Regulation.* New York: Humana Press.

YATES III, J.R. 2011. A century of mass spectrometry: from atoms to proteomes. *Nat Methods,* 8,633-637.

YONA, D., PARK, M.O., OH, S.J. & SHIN, W.S. 2014. Distribution of *Synechococcus* and its phycoerythrin pigment in relation to environmental factors in the East Sea, Korea. *Ocean Sci J,* 49, 367-382.

ZAK, E., NORLING, B., MAITRA, R., HUANG, F., ANDERSSON, B. & PAKRASI, H. B. 2001. The initial steps of biogenesis of cyanobacterial photosystems occur in plasma membranes. *PNAS,* 98**,** 13443-13448.

ZAKAR, T., LACZKO-DOBOS, H., TOTH, T. N. & GOMBOS, Z. 2016. Carotenoids Assist in Cyanobacterial Photosystem II Assembly and Function. *Front Plant Sci,* 7**,** 295.

ZHENG, Q., WANG, Y., XIE, R., LANG, A. S., LIU, Y., LU, J., ZHANG, X., SUN, J., SUTTLE, C. A. & JIAO, N. 2018. Dynamics of Heterotrophic Bacterial Assemblages within *Synechococcus* Cultures. *Appl Environ Microbiol,* 84.

ZHOU, F., LU, Y., FICARRO, S. B., ADELMANT, G., JIANG, W., LUCKEY, C. J. & MARTO, J. A. 2013. Genome-scale proteome quantification by DEEP SEQ mass spectrometry. *Nat Commun,* 4**,** 2171.

ZUCKERT, W. R. 2014. Secretion of bacterial lipoproteins: through the cytoplasmic membrane, the periplasm and beyond. *Biochim Biophys Acta,* 1843**,** 1509-16.

ZWIRGLMAIER, K., HEYWOOD, J. L., CHAMBERLAIN, K., WOODWARD, E. M., ZUBKOV, M. V. & SCANLAN, D. J. 2007. Basin-scale distribution patterns of picocyanobacterial lineages in the Atlantic Ocean. *Environ Microbiol,* 9**,** 1278-90.

ZWIRGLMAIER, K., JARDILLIER, L., OSTROWSKI, M., MAZARD, S., GARCZAREK, L., VAULOT, D., NOT, F., MASSANA, R., ULLOA, O. & SCANLAN, D. J. 2008. Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. *Environ Microbiol,* 10**,** 147-161.

Chapter 2 Preface

Membrane and membrane-associated proteins make up more than one-third of the proteins in bacteria. The cell membrane is the primary barrier through which ions, molecules and/or substrates move in/out of cells. The following chapter is the first proteomic investigation of marine cyanobacteria *Synechococcus* cell membranes. We have applied this technique to four strains, representing divergent taxonomy and ecological niches, in order to represent the diversity of marine *Synechococcus*.

Objectives:

1. To describe the membrane proteomes of marine cyanobacteria genus *Synechococcus*.
2. To compare the membrane protein expression of four marine *Synechococcus* strains, representative of different ecotypes.
3. To infer the nutrients acquisition strategies of different marine *Synechococcus*.

Comparative membrane proteomics reveal contrasting adaptation strategies for coastal and oceanic marine *Synechococcus* cyanobacteria

Fallen Teoh[1], Bhumika Shah[1], Martin Ostrowski[1,2], Ian Paulsen[1*]

[1]Department of Molecular Sciences, Macquarie University, Sydney, Australia.

Present address: [2]Climate Change Cluster, University of Technology Sydney, Australia

[*]Author for correspondence

Mailing address: 4 Wally's Walk, Department of Molecular Sciences, Macquarie University NSW 2109.

Tel: (61) 2 98508152.

E-mail: ian.paulsen@mq.edu.au

ABSTRACT

Marine cyanobacteria genus *Synechococcus* are among the most abundant and widespread primary producers in the open ocean. *Synechococcus* strains belonging to different clades have adapted distinct strategies for growth and survival across a range of marine conditions. Clades I and IV are prevalent in colder, mesotrophic, coastal waters, while clades II and III prefer warm, oligotrophic open oceans. To gain insight into the cellular resources these unicellular organisms invest into adaptation strategies we performed gel-based membrane proteomics of four *Synechococcus* spp. strains namely CC9311 (clade I), CC9605 (clade II), WH8102 (clade III) and CC9902 (clade IV). Comparative membrane proteomes analysis demonstrated that CC9902 and WH8102 showed high resource

allocation for phosphate uptake, accounting for 44% and 38% of overall transporter protein expression of the species. WH8102 showed high expression of the iron uptake ABC binding protein FutA, suggesting that high binding affinity for iron is possibly a key adaptation strategy for some strains in oligotrophic ocean environments. One protein annotated as a phosphatase 2c (Sync_2505 and Syncc9902_0387) was highly expressed in the coastal mesotrophic strains CC9311 and CC9902, constituting 14-16% of total membrane protein, indicating a vital, but undefined function, for strains living in temperate mesotrophic environments.

INTRODUCTION

Marine cyanobacteria from the genus *Synechococcus* are essentially ubiquitous in the open ocean. Their distribution ranges from coastal to open ocean habitats that extend deep into temperate high-latitudes across 120^0 of latitude (Zwirglmaier et al., 2007, Zwirglmaier et al., 2008, Flombaum et al., 2013, Hunter-Cevera et al., 2016). They play an important role as primary producers, and make a major contribution to the global carbon (Glover et al., 1985, Li, 1994, Agawin et al., 2000, Richardson and Jackson, 2007, Wegener et al., 2010) and nitrogen cycles (Wegener et al., 2010). *Synechococcus* and *Prochlorococcus* together are estimated to account for ~25% of net marine primary productivity in the world (Flombaum et al., 2013). *Synechococcus* spp. are genetically diverse and have been divided into three main subclusters: 5.1, 5.2 and 5.3 based on rRNA phylogeny (Scanlan et al. 2009). In addition, they are further classified into more than 20 clades which display distinct geographical or seasonal distributions (Ferris and Palenik, 1998, Rocap et al., 2002, Fuller et al., 2003, Ahlgren and Rocap, 2006, Penno et al., 2006, Farrant et al., 2016, Sohm et al., 2016). *Synechococcus* spp. are most abundant in the surface mixed layer and

gradually decline at depth as light intensity decreases (Partensky et al., 1999b, DuRand et al., 2001, Rocap et al., 2003, Mackey et al., 2009, Scanlan et al., 2009, Mackey et al., 2011).

Synechococcus spp. strains from clades I, II, III and IV are the most common genotypes found throughout the world's oceans ((Zwirglmaier et al., 2008, Scanlan et al., 2009). Clades I and IV are prevalent in colder, mesotrophic coastal environments (Brown and Fuhrman, 2005, Brown et al., 2005, Ahlgren and Rocap, 2006, Zwirglmaier et al., 2007, Zwirglmaier et al., 2008, Tai and Palenik, 2009, Mazard et al., 2012) while clade II and III is found abundantly in subtropical/tropical (Fuller et al., 2006, Zwirglmaier et al., 2008, Huang et al., 2012, Farrant et al., 2016, Sohm et al., 2016) permanently stratified oceanic waters (Ferris and Palenik, 1998, Toledo and Palenik, 2003, Ahlgren and Rocap, 2006).

Comparative genomics have revealed several hypotheses related to how each lineage has adapted to the distinctly different environmental conditions: For example, *Synechococcus* sp. strains prevalent in open ocean waters such as WH8102 has a limited repertoire of two component signal transduction systems, comprising only five histidine kinase sensors and nine response regulators, consistent with adaptation to a relatively constant ecosystem. In contrast, *Synechococcus* strains common in more variable coastal environments, such as CC9311 has a much greater complement of signal transduction systems with 11 histidine kinases and 17 response regulators (Palenik et al., 2006, Dufresne et al., 2008). In another example, *Synechococcus* open ocean strains such as CC9605 and WH8102, prevalent in P-limited environment, harbour multiple gene copies encoding periplasmic phosphate binding protein (PstS), e.g. up to four PstS homologues in *Synechococcus* sp. strain WH8102. In contrast, coastal *Synechococcus* strains, CC9311 and CC9902, have only a single copy of the gene for PstS (Scanlan et al., 2009).

Several studies have characterized changes in gene expression in response to different environmental conditions in *Synechococcus* (Su et al., 2006, Rivers et al., 2009, Stuart et al., 2009, Tetu et al., 2009, Varkey et al., 2016). Physiological studies have also shown varied adaptability between *Synechococcus* clades to different environmental conditions including light intensity (Palenik, 2001, Ahlgren and Rocap, 2006, Six et al., 2007), nitrogen availability (Moore et al., 2002, Ahlgren and Rocap, 2006, Fuller et al., 2006), metal ion concentrations such as copper (Stuart et al. 2009) and iron (Rivers et al., 2009). For example, coastal marine *Synechococcus* spp. have been found to be more tolerant to copper toxicity than open ocean strains (Stuart et al. 2009). The coastal strains activated more transporters and regulatory elements which are not conserved in open ocean strains in response to copper stress (Stuart et al. 2009).

Membrane proteins are key to adaptation for bacteria to different environments playing critical roles in nutrient transport, ion homeostasis, environmental sensing and regulation. Based on bioinformatic analyses, typically about one third of bacterial proteins are tightly associated with membranes (Wallin and Von Heijne, 1998, Krogh et al., 2001). Little is known about *Synechococcus* membrane protein composition, and there have been no focused studies on comparison of membrane protein expression between different marine *Synechococcus* clades, and how these may impact their life strategies and ecological niches. Hence, we conducted membrane proteomic analysis of four representative marine *Synechococcus* sp. CC9311 (clade I), CC9605 (clade II), WH8102 (clade III) and CC9902 (clade IV), to examine their membrane protein expression profiles and understand their relative resource allocation strategies.

METHODOLOGY

Experimental Procedures

Growth conditions of *Synechococcus* strains

Four *Synechococcus* strains WH8102, CC9311, CC9605 and CC9902 were grown separately in 1L synthetic ocean water (SOW) culture media (Morel et al., 1979, Su et al., 2006) (Supplementary Info 2.1) and incubated at 22^0C in continuous white light intensity of 65-75 μmol photons $m^{-2}s^{-1}$ with agitation at 100 r.p.m. The cyanobacterial cells were harvested during their exponential phase at OD_{750}=0.15-0.2 (approximately 10^7-10^8 cells per ml) measured by spectrophotometer (Beckman DU 640 Spectrophotometer, Beckman Instruments) and cell density were further estimated by BD Influx flow cytometry (BD Biosciences).

Preparation of membranes for Nano LC-MS/MS analysis

Membranes of *Synechococcus* were purified with the modified protocol of Rexroth et al. (2011) (Rexroth et al., 2011). Cells were harvested at the exponential growth phase, by centrifugation at 7,500g at 4^0C for 15 min. The supernatant was disposed and the pellet was mixed with 3ml of BEK buffer (10mM boric acid, 10mM EDTA and 10mM potassium chloride, adjusted to pH 9.5 with potassium hydroxide). Protease inhibitor cocktail (P8465 from Sigma) and DNAseI (11284932001 from Roche) were added per the manufacturer's protocol before cell lysis. Cell disruption was carried out using a French press and membrane proteins were separated by density gradient centrifugation as described previously (Omata and Murata, 1983) with modifications. For the cell disruption, a working pressure of 16,000 psi was applied to a cell suspension and the process was repeated 12 times, on ice. Unbroken cells were removed by centrifugation

(25,000g 10min, 4^0C). Subsequently, *Synechococcus* cell membranes were sedimented by differential centrifugation (150,000g, 45min, 4^0C).

Whole membrane protein concentration was determined using a bicinchoninic acid (BCA) protein assay (Thermo Fisher Scientific) with Bovine Serum Albumin as standard (Bradford, 1976). Membrane protein samples were mixed with sample loading buffer, denatured by boiling (95^0C, 10 min) and separated using SDS-PAGE (4-15% precast gel; Bio-Rad). Sixty micrograms of membrane proteins were loaded on the gel with three replicates for each species. After electrophoresis, the gel was stained with Coomassie fixing solution and agitated for 60 minutes. Subsequently, the gel was stained with colloidal Coomassie concentrate and Coomassie diluent solution with the ratio 4:1 for 3 hours followed by overnight treatment with a destaining solution. Each gel lane was cut into 10 fractions and the peptides were digested using trypsin (V5111 from Promega) as described in Mirzaei et al. 2012. Peptides were analysed on a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-nLC1000 (Thermo Fisher Scientific) instrument (Details provided in Supplementary Info 2.2).

Database search, protein identification and quantitative analysis

Peptide spectra were searched against the *Synechococcus* Cyanorak protein sequence database (extracted from http://application.sb-roscoff.fr/cyanorak/) using Byonic v2.6 (Protein Metrics Inc.) software (Bern et al., 2012). The search criteria were set as following: carbamidomethylation was set as a fixed modification and methionine oxidation as a variable modification. The precursor mass tolerance was set at 10ppm, fragment mass tolerance at 20ppm, up to two missed tryptic cleavages were allowed and a 1% protein false discovery rate cut off was set up in the protein identification search. Proteins found

were only counted if they are found with at least two unique peptides and present in at least 2 of the 3 technical replicates. The proteomics data have been deposited to the ProteomeXchange consortium via the PRIDE (Vizcaino et al. 2016) partner repository with the dataset identifier PXD009805. Protein expression was quantified based on the label-free normalized spectral abundance factor (NSAF) method (Zybailov et al., 2006). The NSAF for each protein (represented by k) in an experiment is calculated as the number of spectral counts (SpC) (average value of the three technical replicates) identifying k divided by its length (L), divided by sum of SpC/L of all identified proteins (Zybailov et al., 2006).

Protein *in silico* analysis

Protein sub-cellular localization was predicted using PSORTb program version 3.0.2 ([www.psort.org/psortb/index.html)](www.psort.org/psortb/index.html) (Yu et al., 2010). The number of transmembrane regions in each protein was predicted using SOSUI 1.11 program ([http://harrier.nagahama-i-bio.ac.jp/sosui/)](http://harrier.nagahama-i-bio.ac.jp/sosui/) (Hirokawa et al., 1998). Functional categories for all proteins were assigned from EggNOG 4.5.1 (Huerta-Cepas et al., 2016).

RESULTS AND DISCUSSION

Overview of comparative membrane proteomic analyses.

Membrane proteomic analyses of four marine *Synechococcus*, namely CC9311 (clade I), CC9605 (clade II), WH8102 (clade III) and CC9902 (clade IV) were conducted using a GeLC-MS/MS (one-dimensional sodium dodecyl sulfate-polyacrylamide gel electrophoresis followed by liquid chromatography-tandem mass spectrometry) proteomic approach. The growth media (synthetic ocean water) used in this study resembles

mesotrophic coastal water conditions. All four *Synechococcus* strains are grown in nutrient replete conditions and harvested in the active mid-log/ exponential phase. This study is thus examining the relative basal membrane protein levels in *Synechococcus* cells during balanced growth, and is presumably not stressed by nutrient deprivation, and hence regulatory systems for scavenging nutrients are likely not induced in the respective *Synechococcus* strains.

The 1D SDS-PAGE gel of the membrane lysates showed variation in membrane protein composition between the *Synechococcus* strains (Figure 1). Membrane proteomic analyses identified 12,560-21,377 unique tryptic peptides with a confidence criterion of 99% from each of the *Synechococcus* strains. Overall, we identified 619, 887, 875, and 757 proteins for CC9311, CC9605, WH8102 and CC9902, respectively, from our membrane extracts, representing 25-33% of their respective predicted proteomes (Supplementary Info 2.3). Among these, >200 proteins from each strain were predicted to contain transmembrane domain regions (Figure 2).
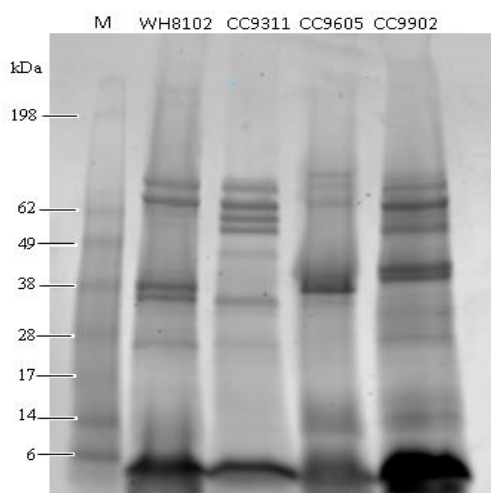
Figure 2.1. 1D SDS-PAGE gel of *Synechococcus* sp. membrane extracts.

Membrane samples (60 μg) from each of the four *Synechococcus* sp. strains WH8102, CC9311, CC9605 and CC9902 were loaded on the 1D SDS-PAGE gel. The lane labelled M represents the protein size markers, with the sizes in kDa provided on the left axis.
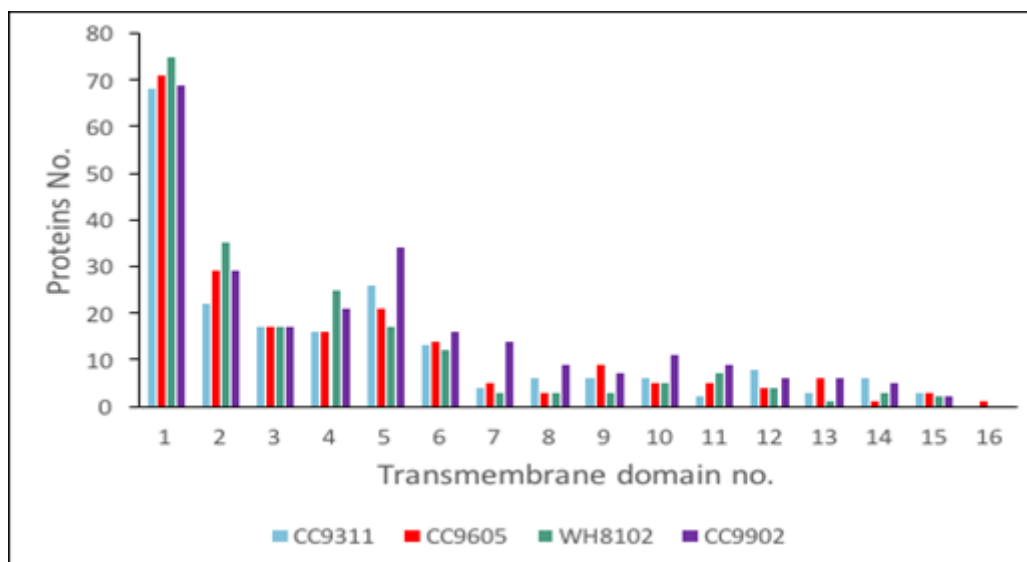
Figure 2.2. Predicted number of transmembrane domains in membrane extract proteins from the four marine *Synechococcus* strains.

All of the proteins detected from the proteomic analysis of the membrane fractions of each of the four *Synechococcus* strains were analyzed using SOSUI to predict their number of transmembrane domains.

Across the four strains, we assigned functional categories to our identified protein set using EggNOG (v4.5) and compared the relative abundance of the proteins using label-free quantitation of NSAF (normalized spectral abundance factor) values in each of these categories. Overall, we observed the relative abundance of proteins varied between the strains based on the protein's functional category (Figure 3). In all four strains, the 'unknown' functional category accounted for the highest relative protein abundance in each strain. Proteins with unknow functions account for about half of the relative protein abundance in CC9311. Meanwhile, this accounts for 26%, 22%, 34% of the relative protein abundance for CC9605, WH8102 and CC9902 strains, respectively (Figure 3). Clearly, there remain significant biological functions within the cyanobacterial cell membrane that have not yet been experimentally characterized. Based on relative protein abundances, the second most common protein functional category in CC9902 is cell wall/membrane/envelope biogenesis, comprising 24% of total membrane proteins. This is almost double the NSAF values observed for this category in the other three strains. We found that proteins under the functional category of inorganic ion transport and metabolism contribute to about 10% of the NSAF value of WH8102 and CC9902 but only about 2.5% of CC9311 and CC9605 strains. As discussed below this is largely due to differences in the relative abundances of phosphate acquisition proteins. The functional category energy production and conversion contributes to 8%, 15%, 11% and 11% of overall NSAF value of CC9311, CC9605, WH8102 and CC9902 strains, respectively, reflecting their importance in the phototrophic lifestyle of these strains. We found a major difference for expression of proteins under the functional group of translation, ribosomal structure and biogenesis. These proteins' expression only contributes to 4.5% of NSAF value for CC9902, however it contributes to about 11%, 20% and 16% of NSAF value for CC9311, CC9605 and WH8102, respectively. The genome sizes of the four strains studied

differ by only 3-17%, with *Synechococcus* strain CC9311 having the largest genome (2.60 Mbp; 2,931 CDS) and strain CC9902 showing the smallest genome (2.23 Mbp; 2,387 CDS). As evident from this section and throughout the manuscript, the observed difference in membrane protein expression levels between *Synechococcus* strains studied is reflective of their adaptation strategies in distinct environments, and is not significantly impacted by the genome size.
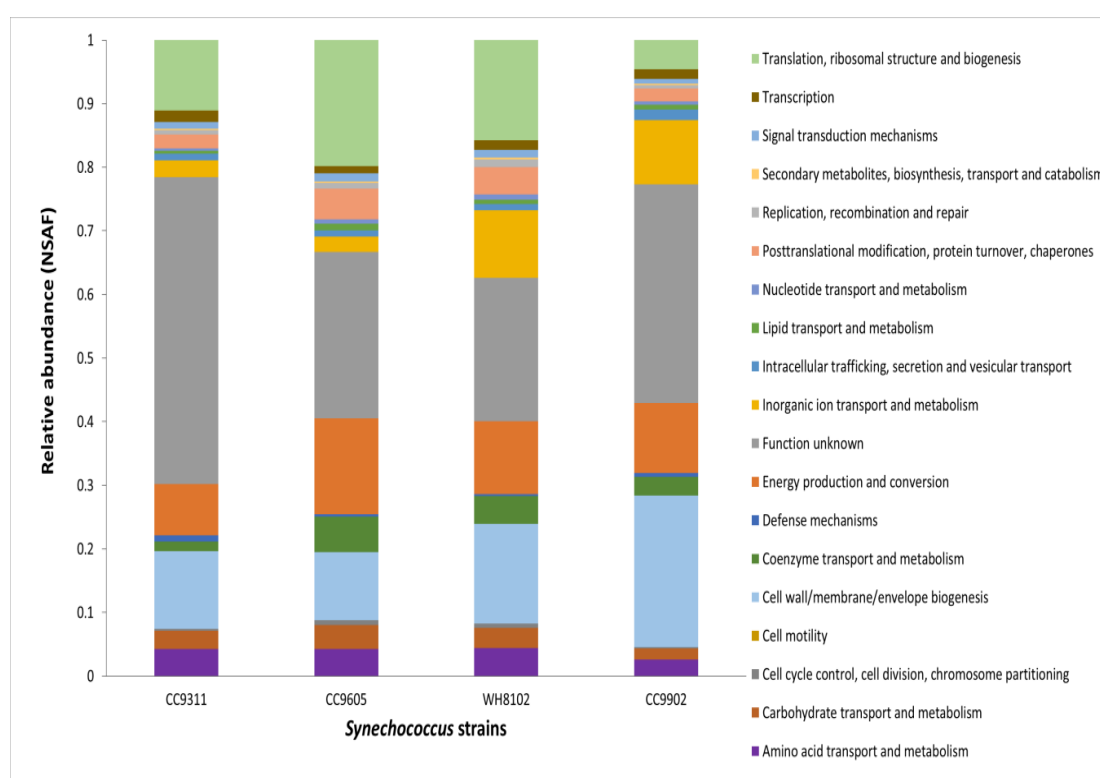
Figure 2.3. Relative abundance of membrane extract proteins of marine *Synechococcus* based on their functional categories.

Functional categories for all of the proteins detected from the proteomic analysis of the membrane fractions of each of the four *Synechococcus* strains were predicted using EggNOG.

Transporter proteins

Transporters are important membrane proteins that facilitate movement of nutrients, ions or substrates in or out from the cells. We detected a higher relative abundance of transport

proteins in WH8102 and CC9902 strains, accounting 16% and 17.2% of overall protein expression NSAF value, respectively. In contrast, transport proteins only accounted for 8.6% and 8.2% for CC9311 and CC9605 strains, respectively (Table 1). There are more than 65 transport proteins comprising 38 orthologous groups identified in all four *Synechococcus* strains from our membrane proteomic data (Supplementary Info 2.4). These include components of the ATP-binding cassette (ABC) transporter systems, ion channels, as well as secondary transporters such as proteins from the RND superfamily, SulP family and MFS superfamily (Supplementary Info 2.4). Notably, a very high number of the transport proteins (>50%) in all four *Synechococcus* strains were components of ABC transport systems (Supplementary Info 2.4), which are known to be a major class of cellular translocation machinery in all bacterial species (Tomii and Kenehisa, 1998), including marine cyanobacteria (Scanlan et al., 2009).

Synechococcus strains / Elements	CC9311	CC9605	WH8102	CC9902
Fe	0.000912	0.001576	0.022915	0.005881
P	0.007367	0.00255	0.061564	0.075047
S	0.001524	0.004001	0.004197	0.000911
N	0.011604	0.010937	0.01569	0.006728
Total transporter protein expression	0.085567	0.082367	0.160091	0.171779

Table 2.1. Relative protein abundance based on cellular resources allocations based on NSAF values. The NSAF values for all detected transporter proteins predicted to be involved in Fe, P, S and N acquisition were totalled to provide an estimate of the cellular resources that are devoted to these processes.

Nutrient acquisition

The succession of marine cyanobacteria in marine waters is highly correlated to the nutrient availability in the habitat (Lindell and Post, 1995, Post et al., 2002). Our comparative proteomics analyses reveal each of the four *Synechococcus* strains displayed distinctive priorities in acquisition and utilization of important macro- and micro-nutrients such as iron (Fe), phosphorus (P), nitrogen (N) and sulfur (S) (Figure 4). These nutrients are known to play a vital role in the growth of marine *Synecchococcus* spp. (Scanlan et al., 2009, Sato et al., 2017).
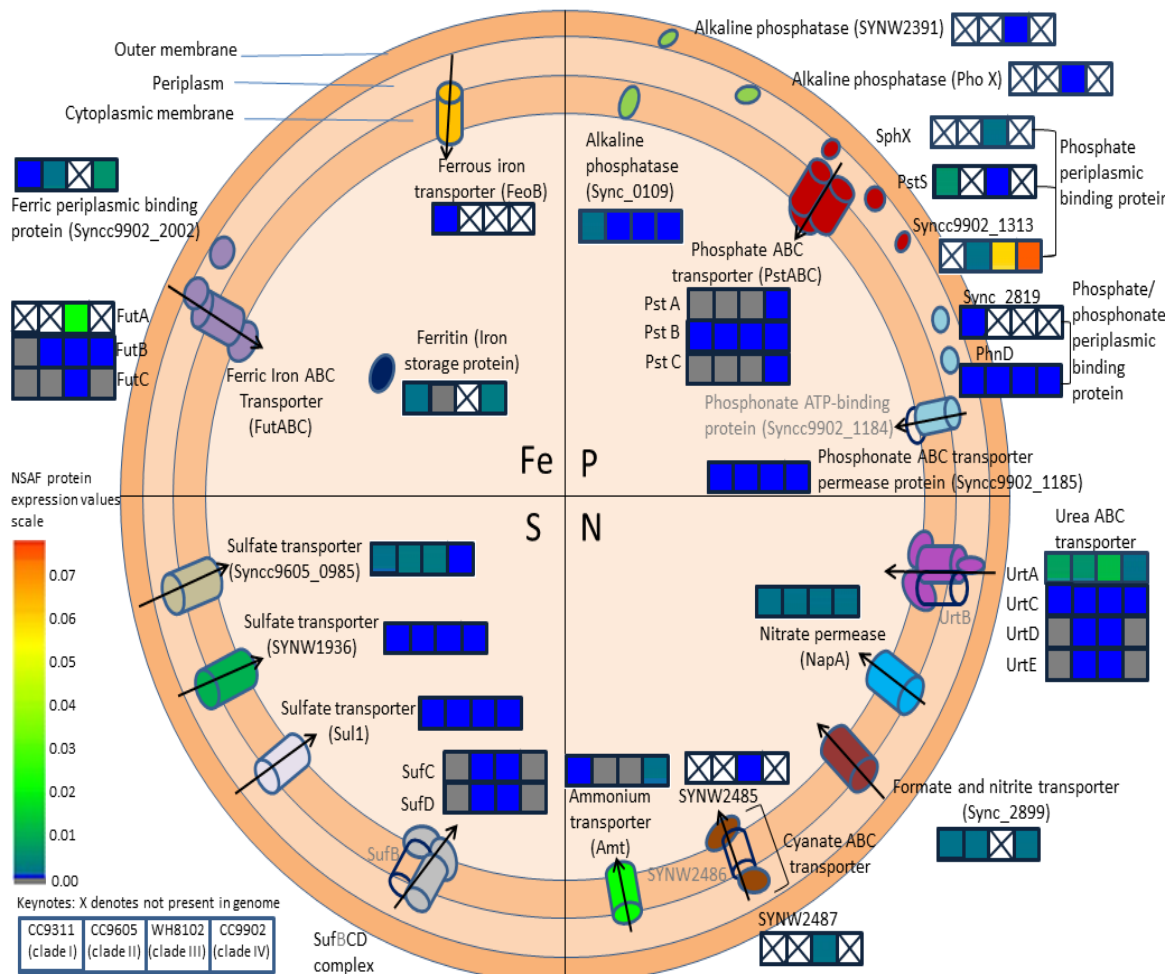
Figure 2.4. Schematic of the membrane transporters and relevant proteins involved in acquisition of iron, phosphorus, nitrogen and sulfur in marine *Synechococcus*.

The schematic shows proteins involved in the acquisition and storage of iron (Fe), phosphorus (P), sulfur (S) and nitrogen (N) in the four *Synechococcus* strains studied. The relative expression levels of these proteins in each of the four strains is indicated in the four boxes, i.e. ▢▢▢▢ , representing in order: CC9311 (clade I), CC9605 (clade II), WH8102 (clade III) and CC9902 (clade IV), respectively. The colour in the boxes represents the NSAF protein expression values, **X** denotes the gene is absent in the genome of the respective strain. The naming of the proteins in the figure is based on the typical name of the proteins from the genome annotations. If no common names were available, then the naming of the protein was based on the protein designation from the

strain which showed the highest NSAF expression value. Transporter subunits not detected in this study are represented in gray scale fonts, and appear as objects with no filled colour in the graphics. The colour of the protein cartoons represents either the subunits of a particular transporter, e.g., PstABC in red; or homologues of a particular enzyme, e.g., alkaline phosphatase in lime green.

Iron acquisition

Iron functions as a cofactor in a plethora of cellular processes such as photosynthesis, respiration, and nitrogen metabolism (Ferreira and Straus, 1994). Iron is depleted in vast expanses of the open ocean and some coastal/continental shelf waters (Behrenfeld and Milligan, 2013). Oxidation of ferrous iron to ferric iron in seawater significantly reduces the biological availability of iron (Boyer et al., 1987), thus organisms have developed several iron acquisition systems.

We identified several components of the ferric iron ABC transporter (FutABC) in all four *Synechococcus* strains studied (Figure 4). Notably, a periplasmic ferric iron binding protein FutA (cluster CK_00057079 in Cyanorak database), is only seen in strain WH8102 and not in any other strains studied. Instead, *Synechococcus* strains CC9311, CC9605 and CC9902 expressed a second periplasmic ferric iron binding protein (cluster CK_00000068) encoded by genes *sync_1545*, *syncc9605_1578* and *syncc9902_2002*, respectively for inorganic iron uptake. The relative abundance of the FutA ferric binding protein in WH8102 is much higher (NSAF value of 0.0225) compared to the second ferric binding protein found in other strains (Figure 4), likely reflecting a higher cellular demand for iron, lower availability of iron in its environmental niche, or both, in WH8102.

As shown in Figure 4, the ferritin iron storage protein expressed in strains CC9311 (Sync_1539) and CC9902 (Syncc9902_2003) is absent in the genome of strain WH8102. This suggests that WH8102 has a distinct adaptation to iron-depleted oligotrophic waters whereby it highly expresses the FutA protein for inorganic iron acquisition, and lacks the iron storage capacity of other *Synechococcus* strains. The distinct FutA protein of WH8102 may potentially bind ferric iron with a higher affinity than the alternate ferric binding proteins found in the other strains.

In addition to the ferric iron transporter, the mesotrophic coastal isolate CC9311 also expressed a ferrous iron transporter FeoB. The *feoB* gene is rarely found in marine *Synechococcus* genomes, but is often seen in freshwater cyanobacteria genomes (Marchetti and Maldonado 2016). The expression of FeoB in CC9311 may reflect environmental exposure to periodic fluxes of freshwater, where ferrous iron would be potentially accessible.

<u>Phosphorus acquisition</u>

Phosphorus is an important building block for DNA, RNA, lipids and intracellular energy (ATP) in cells and has diverse roles in cellular metabolic pathways. The low-nanomolar concentrations of P in marine waters are believed to limit phytoplankton growth, thus marine picocyanobacteria have adopted several approaches such as scavenging or storage to deal with P limitation (Scanlan et al. 2009). Within our study phosphate ABC transporter proteins are highly expressed in all four *Synechococcus* strains (Figure 4). Notably, a periplasmic phosphate binding protein PstS (cluster CK_00000026 in Cyanorak database), is only seen in strains CC9311 and WH8102. Instead, *Synechococcus* strains CC9605 and CC9902 expressed a second periplasmic phosphate binding protein (cluster CK_00043821) encoding genes *syncc9605_1144*, *syncc9605_1145* and *syncc9902_1313*, respectively. This second periplasmic phosphate uptake protein is also expressed in strain

WH8102 (SYNW1018, SYNW1815). Additionally, strain WH8102 expressed a third phosphate binding periplasmic protein, SphX (SYNW1286), not observed in any other strains studied. As depicted in Figure 4, among the four *Synechococcus* strains studied, WH8102 and CC9902 strains showed a relatively higher NSAF value for phosphate periplasmic binding proteins in contrast to CC9311 and CC9605 strains.

Synechococcus clade III strains such as WH8102 are generally the most abundant *Synechococcus* genotypes present in P limited, oligotrophic open water such as the Sargasso Sea (Farrant et al., 2016), North Atlantic Gyre (Mazard et al., 2012) and Eastern Mediterranean (Mella-Flores et al., 2011). Therefore, high expression of phosphate acquisition proteins in clade III strains such as WH8102 is expected. Meanwhile, clade IV strains represented by CC9902 is found predominantly in the nutrient rich, mesotrophic, coastal waters such as the California current (Brown et al., 2005). Therefore, among the four *Synechococcus* strains studied, the highest expression of phosphorus-related periplasmic binding protein (Syncc9902_1313) seen in strain CC9902 is unexpected (Figure 4).

In addition to the phosphate uptake systems, all four *Synechococcus* strains expressed components of a putative phosphonate ABC transporter system, including a phosphonate periplasmic binding protein PhnD (cluster CK_00000860) and phosphonate ABC transporter permease protein (cluster CK_00000861, represented as Syncc9902_1185 in Figure 4). Notably, strain CC9311 expressed an additional putative phosphate/phosphonate binding periplasmic protein (Sync_2819). The higher relative abundance of the PstS protein compared to PhnD proteins in these four strains suggests a preference to utilize phosphate ABC transport machinery for P acquisition.

In addition, a putative complex P_i hydrolyzing enzyme alkaline phosphatase (Sync_0109) is detected in all four *Synechococcus* strains (Cluster CK_00000302). Strain WH8102

expressed two additional putative alkaline phosphatase proteins, PhoX (SYNW1799) and SYNW2391, not present in the genome of other strains studied. Alkaline phosphatase allows phytoplankton to scavenge phosphorus (P) from dissolved organic phosphorus (DOP) when inorganic phosphate is scarce in the ocean (Lin et al., 2015). During phosphorus stress, *Synechococcus* strains have shown increased production of alkaline phosphatases for phosphate uptake (Moore et al., 2005, Tetu et al., 2009). These strategies make them well adapted to live in the oligotrophic open ocean. Expression of several alkaline phosphatase proteins in WH8102 strain may allude to their survival strategy during phosphorus stress. These results correlate with very high expression of zinc binding periplasmic protein, ZnuA, in WH8102 (Figure 2.5). Since zinc is needed by metalloproteins such as alkaline phosphatases for phosphate acquisition (Ostrowski et al., 2010). Therefore, the high expression of ZnuA in WH8102 may be crucial for the acquisition of phosphate (Cox and Saito, 2013).

Previous microarray analyses (Tetu et al., 2009) and knockout mutant studies (Ostrowski et al., 2010) have respectively shown key roles of phosphate regulatory systems, PhoBR and PtrA, in the regulation of P transport and metabolism during phosphate stress in *Synechococcus* strain WH8102. These P-response regulators are also found in the genome of clade II strain CC9605, but are both absent in the genomes of clade I and clade IV strains, CC9311 and CC9902, respectively. Significant basal levels of expression of all 4 PstS homologues in WH8102 (SYNW1018, SYNW1815, SYNW2507, SYNW1286) and 2 PstS homologues in CC9605 (syncc9605_1144, syncc9605_1145) as well as putative alkaline phosphatases (SYNW0120, SYNW1799, SYNW2391) is interesting given the P-stress response regulators in these strains. This implies that P-acquisition systems are relatively highly expressed in these two strains even before any induction of their P-regulons.

Nitrogen acquisition

The availability of biologically utilizable nitrogen is a key environmental factor governing photosynthetic activity in diverse open ocean and coastal marine ecosystems. The light harvesting antenna – PBS, is a N-rich complex, representing major cellular investment which can constitute up to 50% of the soluble protein of the cell (Grossman et al., 1993, Apt et al., 1995). Therefore, the distribution of marine *Synechococcus* spp. is also linked to N availability (Olson et al., 1990, Partensky et al., 1999a).

Marine *Synechococcus* spp. have been reported to utilize diverse N sources including nitrate (NO_3^-), nitrite (NO_2^-), ammonium (NH_4^+), urea and amino acids (Gilbert et al., 1986, Paerl, 1991, Lindell and Post, 1995, Collier et al., 1999). Consistent with these studies, we found a broad range of N uptake transporters in our study (Figure 4). These include N acquisition proteins found in all four *Synechococcus* strains such as those associated with a urea transporter (UrtABCDE), nitrate and nitrite transporter (NapA) and ammonium transporter (Amt). Our proteomic study found that the putative urea ABC transporter (represented as UrtABCDE in Figure 4) is the highest expressed nitrogen transporter for all 4 *Synechococcus* strains. This is intriguing as nitrate is the sole N source in synthetic ocean water (SOW) medium that we used for this study. These results are consistent with previous proteomics study, where urea ABC transporter was the dominant N transporter when *Synechococcus* cells are grown on nitrate as a sole N source (Christie-Oleza et al., 2015).

In addition to N acquisition proteins expressed in all four strains, we also found a predicted formate and nitrite transport protein (cluster CK_00001669) in strains CC9311 (clade I), CC9605 (clade II) and CC9902 (clade IV) encoding genes *sync_2899*, *syncc9605_2657* and *syncc9902_2285*, respectively. This protein is not detected in WH8102 (clade III), instead we identified a very distinct alternate nitrogen source acquisition strategy in

WH8102, expressing components of a putative cyanate ABC transporter (SYNW2485 and SYNW2487) not seen in any other strains we studied.

Cyanate is a N substrate that receives less attention in oceanographic studies (Scanlan et al., 2009) as it is a product of spontaneous urea degradation in aqueous solutions (Hargel et al., 1971). Therefore, cyanate transport is not thought to be a dominant type of N transport in marine organisms, although this is yet to be demonstrated *in situ*. The cyanate transporter is only found in *Synechococcus* clade III strains to date, suggesting this is a unique N acquisition strategy in strains from this clade.

Sulfur acquisition

Sulfur is an essential component of proteins and sulfolipids in cyanobacteria (Takahashi et al., 2011). Compared to other key elements, less attention has been focused on the requirements of marine picocyanobacteria for sulfur, hence the effect of sulfur availability in the oceanic water on marine picocyanobacteria remains less well understood.

Within our study, we detected several S acquisition proteins in all four strains (Figure 4) including sulfate secondary transporters belonging to the SulP family (cluster CK_00056721, represented as SYNW1936 and CK_00008045, represented as Syncc9605_0985 in Figure 4), Sul1, as well as the SufBCD complex that is known to be involved in the biogenesis of Fe-S clusters (Hirabayashi et al., 2015).

Strains CC9605 and WH8102 showed marginally higher protein expression for S transporters, in particular for the sulfate transporter Syncc9605_0985 homologue (Figure 4). Higher expression of sulfur-related acquisition proteins in CC9605 and WH8102 in comparison to the coastal strains (CC9311 and CC9902) may possibly be linked to their higher expression of PsaC, photosystem I iron-sulfur protein (Figure 5) which is essential for photochemical activity.

Highly expressed proteins

The top 50 highest expressed proteins found in each *Synechococcus* strains quantitated using label-free NSAF values vary among the four analysed strains (Figure 5). Overall, porins are the highest expressed proteins in CC9605, WH8102 and CC9902 strains and second highest expressed protein in CC9311 (Figure 5). In addition, as expected a broad range of photosynthesis-associated protein functions showed high expression consistent with their critical role in *Synechococcus* as an autotrophic organism. We also found one hypothetical protein (cluster CK_00000544), encoding genes Sync_0450 in CC9311, Syncc9605 in CC9605, SYNW2055 in WH8102 and Syncc9902_1942 in CC9902 strains, highly expressed in all four *Synechococcus* strains (Figure 5). This hypothetical protein is predicted to have CAAD domain of cyanobacterial aminoacyl-tRNA synthetase, whose function is suggested to mediate the membrane anchorage of the proteins in the thylakoid (Olmedo-Verd et al., 2011). Even so, the exact functional role of this protein remains unexplored.

We detected a very high expression of a protein phosphatase 2c protein in both coastal strains – CC9311 and CC9902 (Figure 5) in contrast to the open ocean strain WH8102 and offshore strain CC9605, suggesting the expression of this protein could be highly environmentally relevant. Nonetheless, the precise role of protein phosphatase 2C remains elusive. We propose phosphatase 2C is involved in the signal transduction of cyanobacteria by dephosphorylating substrate proteins, performing an antagonistic role with a kinase.

In addition, in strain CC9902 (Clade IV), we detected high expression of carbamoyl phosphate synthetase, that catalyzes the formation of carbomyl phosphate from bicarbonate, two molecules of MgATP and ammonia (Kasahara and Ohmori, 1997). Since ammonia is required for formation of carbamoyl phosphate, it could drive the expression

of ammonia transporter (Figure 5). Also, carbamoyl phosphate is a common intermediate of arginine biosynthesis (Nicoloff et al., 2001) and therefore may have an important role in urea production. We also detected high relative expression of the ATP synthase in CC9902 in contrast to other strains in the study (Figure 5). ATP synthase is an enzyme that catalyzes the synthesis of ATP from ADP and phosphate and can possibly contribute to the high demand of P resources in CC9902. Other highly expressed proteins in CC9902 include a periplasmic phosphate binding protein (Syncc9902_1313), putative RND family outer membrane efflux protein (homologue of SYNW2187), periplasmic iron binding protein (Syncc9902_2002) and extracellular solute-binding protein (Syncc9902_0847) (Figure 5).

In strain CC9311 (Clade I), we found high expression of two mechanosensitive channel proteins (Sync_2685 and MscL) (Figure 5). These proteins are suggested to act as emergency release valves in response to osmotic shock (Booth and Louis, 1999). Genome analysis of CC9311 strain showed an expansion of genes encoding mechanosensitive channels by having 7 copies of such genes (Palenik et al., 2006), suggesting that it has special adaptation to cope with the osmotic stress in the coastal environment. Other highly expressed proteins in CC9311 include PstS, UrtA, stomatin/band7 family transporter (Sync_1798), solute binding protein ABC transporter predicted to transport amino acid (Sync_2366) as well as hypothetical proteins including Sync_1972 (a possible lipoprotein), Sync_1845 (possible outer membrane protein) and Sync_1098. These hypothetical proteins are unique to the CC9311 strain and completely absent in the genome of the other 3 strains (Figure 5).

The subtropical strain WH8102 (Clade III) showed high expression of various transporters involved in nutrient, substrate or ion transport. These include the periplasmic phosphate binding protein (SYNW1018 and SYNW1815 (homologue of Syncc9902_1313)), FutA,

UrtA, ZnuA, stomatin/band 7 family transporter (SYNW1909) and RND family outer membrane efflux protein (SYNW2187) (Figure 5). The high expression of the transporters involved in P, N, Fe uptake is consistent with the fact that WH8102 was isolated from oligotrophic waters, and likely has evolved efficient nutrient scavenging strategies to maintain its cellular requirements (Tetu et al., 2009, Mackey et al., 2015). Besides that, we found a high expression of SYNW0406 which is unique to WH8102 (Figure 5). SYNW0406 is a heavily glycosylated integral outer membrane protein whose role has not been fully characterised (Brahamsha 1996, Palenik 2011).

Synechococcus strain CC9605 (clade II) displays different membrane protein adaptations in contrast to WH8102 (clade III) strain, even though they share some common ecological niches (Ferris and Palenik, 1998, Toledo and Palenik, 2003). Clade II is the marine *Synechococcus* group which has colonized the broadest geographical area in the ocean. Protein functions associated with photosynthesis such as phycobilisomes, photosystem I and photosystem II are highly expressed in strain CC9605 (Figure 5). Overall, CC9605 strain shows moderate expression of transport proteins. Only two transporters – UrtA and RND family outer membrane efflux protein (homologue of SYNW2187) are found in the top 50 highest expression protein of CC9605 strain (Figure 5). The overall transporter protein expression in CC9605 strain is the lowest among four *Synechococcus* strains (Table 1). This may reflect a strategy where energy production is emphasized while nutrient scavenging is not as high a priority as seen in WH8102 that is adapted to more oligotrophic settings.
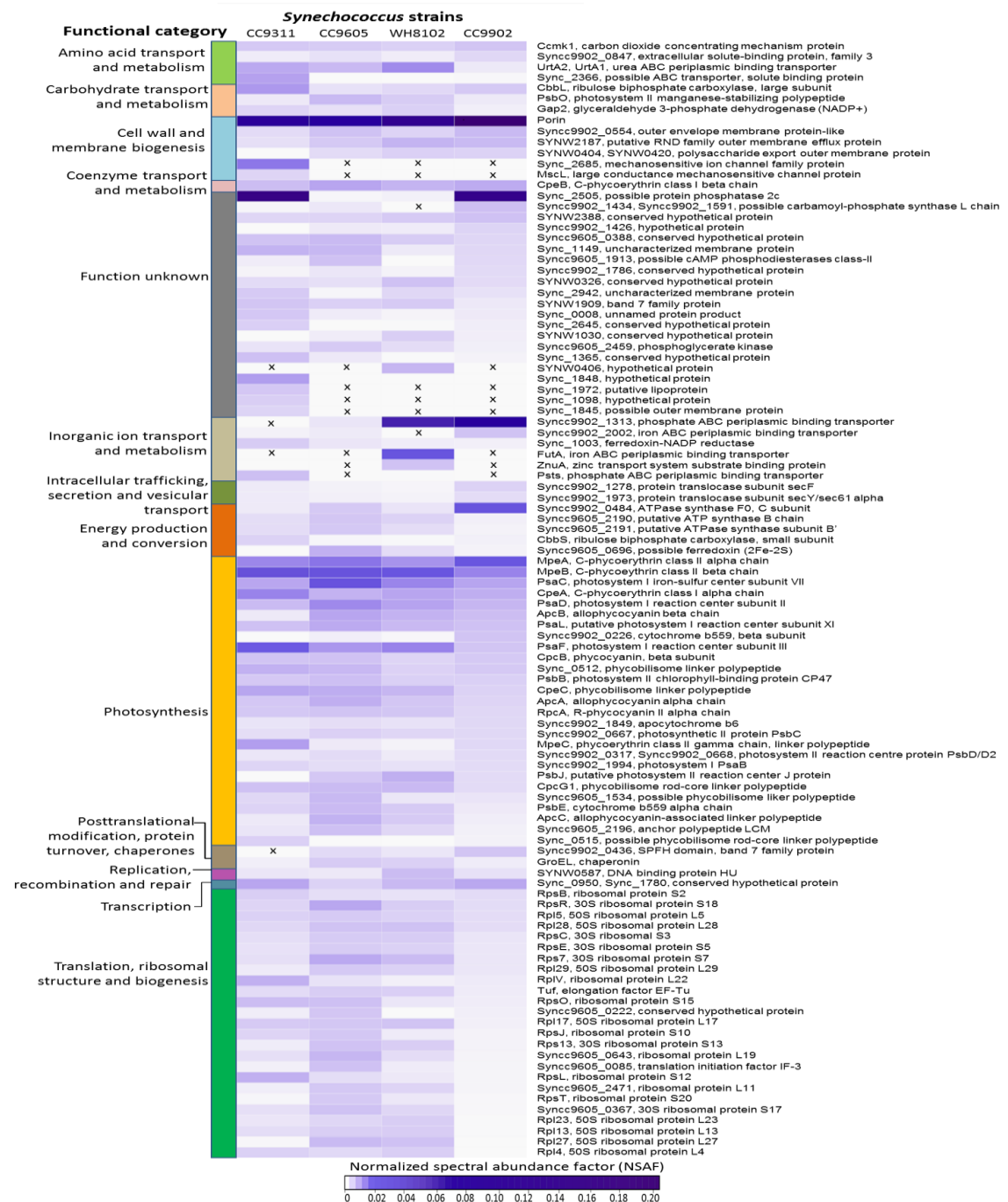
Figure 2.5. Highly expressed proteins from the membrane extract proteomes of 4 marine *Synechococcus* strains. The heatmap depicts the top 50 highest expressed proteins based on NSAF value for each of the *Synechococcus* strains (CC9311, CC9605, WH8102 and CC9902). X denotes the gene is not present in the genome of the respective *Synechococcus* strain. Left lane represents functional category of proteins. The functional category of the proteins is manually assigned with reference from EggNOG.

CONCLUSIONS

The goal of this study was to gain a global membrane proteomic perspective on 4 strains of *Synechococcus* representative of clades I-IV, particularly looking at the basal membrane protein expression levels during exponential growth phase, thus providing interesting insights into common and distinct adaptation strategies. Overall, we observed high expression of transporters involved in iron, phosphorus and nitrogen acquisition in *Synechococcus* strain WH8102. This fits to its adaptation strategy by scavenging nutrients or substrates found in their oligotrophic living environment. Coastal strains CC9311 and CC9902 showed different adaptation strategies though they share similar ecological niches: CC9311 showed high expression of transporters involved in nitrogen acquisition, and mechanosensitive ion channels for osmotic regulation of cells, while CC9902 showed high expression of transporters involved in phosphorus and iron. The highest expression of transporters involved in phosphorus acquisition in CC9902, among four *Synechococcus* strains, suggests a high demand for phosphorus for reasons that remain puzzling. In contrast, CC9605 a clade II *Synechococcus* strain generally showed low or moderate expression of transporters. We do not understand why this may be a successful adaptation strategy, it may need to be verified by looking at additional strains from clade II. While variation of transporter expression among different *Synechococcus* spp. is the main finding of this study, we found that the high expression of a putative phosphatase 2C enzyme in CC9311 and CC9902 may be a crucial signalling or regulatory adaptation to coastal environments.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Supplementary information

Supplementary Info 2.1. SOW culture media composition

a)

Chemical Reagent Components	Volume (mL)
Disodium ethylenediaminetetraacetate dihydrate (from 2.68mM stock)	5
Sodium nitrate (from 1M stock)	9
Dipotassium phosphate (from 35.02mM stock)	2.5
Sodium carbonate (from 32.26mM stock)	3
Trace metal mixture	1
Vitamins	Refer vitamin table
1M HEPES, pH8	1.5

(Notes: Salt base was mixed according to table (b) and deionized water was summed up to 950mL, stirred well and autoclaved. Chemical reagent components in table (a) was filter sterile and summed up to 50mL with deionized water before added to the autoclaved solution (b))

b)

Salts base	Volume (g)
Sodium chloride	24.35
Magnesium chloride hexahydrate	11.1
Sodium sulfate	4.09
Calcium chloride dihydrate	1.54
Potassium chloride	0.7
Sodium bicarbonate	0.2
Potassium bromide	0.1
Boric acid	0.03
Strontium chloride hexahydrate	0.017
Sodium fluoride	0.003

c)

Trace metals	Concentration $(g.l^{-1})$
Zinc sulfate heptahydrate	0.222
Manganese (II) chloride tetrahydrate	1.4
Cobalt (II) chloride hexahydrate	0.0204
Sodium molybdate dehydrate	0.39
Citric acid hydrate	6.25
Nickel (II) chloride hexahydrate	0.0137
Sodium selenite	0.0132
Iron (III) chloride	0.2

d)

Vitamins	Volume
Thiamine hydrochloride	10mg
Biotin (from 0.1 $mg.ml^{-1}$ stock)	500µl
Cobalamin (from 1$mg.ml^{-1}$ stock)	50µl

Supplementary Info 2.2. Summary of mass spectrometry settings.

Reversed-phase chromatographic separation was conducted on a 75µm i.d. x100mm, C18 HALO column, 2.7µm bead size, 160Å pore size. A linear gradient of 1-50% solvent B (99.9% ACN/0.1% FA) was run over 50 minutes, followed by a slope gradient of solvent B from 50% to 85% for 2 minutes and maintain at 85% for 8 minutes. The mass spectrometer was operated in the data-dependent mode to automatically switch between Orbitrap MS and ion trap MS/MS acquisition. Survey full scan MS spectra (from *m/z* 350 to 2000) were acquired with a resolution of 35,000 and an AGC (Automatic Gain Control) target value of $1\text{x}10^6$ ions. Ten most abundant ions were selected for higher energy collisional dissociation (HCD) fragmentation. HCD normalised collision energy was set to 35% and fragmentation ions were detected in the Orbitrap at a resolution of 17,500. Target ions that had been selected for MS/MS were dynamically excluded for 20s.

Supplementary Info 2.3. Complete list of proteins identified from the membrane extracts.

Available in the attached file named Suppl 2.2_Complete list of proteins identified from the membrane extracts.

Supplementary Info 2.4. Table of membrane transporters identified in the membrane extract proteomes of the four *Synechococcus* strains.

Synechococcus strains / Transporters	CC9311	CC9605	WH8102	CC9902
ATP-dependent Transporter				
ABC	28	42	45	44
P-ATPase	1	1	1	1
F-ATPase	7	7	7	7
Total	36	50	53	52
Secondary Transporter				
Resistance-Nodulation-Cell Division (RND) Superfamily	7	3	4	4
Sulfate Permease (SulP) Family	3	3	3	3
Major Facilitator Superfamily (MFS)	3	4	3	4
Divalent Anion:Na+ Symporter (DASS) Family	0	1	1	1
Monovalent Cation: Proton Antiporter-1 (CPA1) Family	0	1	0	2
Monovalent Cation: Proton Antiporter-2 (CPA2) Family	1	1	2	2
K+ Transporter (Trk) Family	0	1	0	1
Ca2+: Cation Antiporter (CaCA) Family	0	0	1	0
Neurotransmitter: Sodium Symporter (NSS) Family	1	1	2	2
Solute: Sodium Symporter (SSS) Family	2	1	0	1
Betaine/Carnitine/Choline Transporter (BCCT) Family	0	0	0	1
Formate-Nitrite Transporter (FNT) Family	1	1	0	1
Ni2+-CO2+ Transporter (NiCoT) Family	1	0	1	1
Amino Acid-Polyamine-Organocation (APC) Family	1	0	0	1
Chloride Carrier/Channel (CIC) Family	1	0	1	2
Tripartite ATP-independent Periplasmic Transporter (TRAP-T) Family	1	0	0	0
Auxin Efflux Carrier (AEC) Family	1	0	0	0
Alanine or Glycine: Cation Symporter (AGCS) Family	1	1	1	1
Cation Diffusion Facilitator (CDF) Family	0	0	0	1
Cytochrome Oxidase Biogenesis (Oxa1) Family	1	1	1	1
Autoinducer-2 Exporter (AI-2E) Family (Formerly PerM Family)	0	2	1	1
Drug/Metabolite Transporter (DMT) Superfamily	0	0	0	1
Twin Arginine Targeting (Tat) Family	0	0	1	0

Total	25	21	22	31
Ion Channels				
Voltage-gated Ion Channel (VIC) Superfamily	0	3	1	1
Ammonia Transporter Channel (Amt) Family	1	0	0	1
Small Conductance Mechanosensitive Ion Channel (MscS) Family	2	0	1	3
Glutamate-gated Ion Channel (GIC) Family of Neurotransmitter Receptors	0	0	1	1
Large Conductance Mechanosensitive Ion Channel (MscL) Family	1	1	0	0
Stomatin/podocin/band 7/nephrosis.2/SPFH (Stomatin) family protein	1	1	1	1
Total	5	5	4	7
Unclassified				
Mg2+ Transporter-E (MgtE) Family	0	1	1	1
HlyC/CorC (HCC) Family	1	2	2	2
YggT or Fanciful K+ Uptake-B (FkuB; YggT) Family	0	0	0	1
Tellurium Ion Resistance (TerC) Family	0	1	0	0
Ferrous Iron Uptake (FeoB) Family	1	0	0	0
Capsular Polysaccharide Exporter (CPS-E) Family	0	1	1	1
Total	2	5	4	5
SUM of all transporters	68	81	83	95

Transporter proteins are classified into families and energy coupling mechanisms as per TCDB (Saier et al. 2006).

REFERENCES

AGAWIN, N. S. R., DUARTE, C. M. & AGUSTI, S. 2000. Nutrient and temperature control of the contribution of picoplankton to phytoplankton biomass and production. *Limnol Oceanogr,* 45**,** 591-600.

AHLGREN, N. A. & ROCAP, G. 2006. Culture isolation and culture-independent clone libraries reveal new marine *Synechococcus* ecotypes with distinctive light and N physiologies. *Appl Environ Microbiol,* 72**,** 7193-204.

APT, K. E., COLLIER, J. L. & GROSSMAN, A. R. 1995. Evolution of the Phycobiliproteins. *J Mol Biol,* 248**,** 79-96.

BEHRENFELD, M. J. & MILLIGAN, A. J. 2013. Photophysiological expressions of iron stress in phytoplankton. *Ann Rev Mar Sci,* 5**,** 217-46.

BERN, M., KIL, Y. J. & BECKER, C. 2012. Byonic: Advanced Peptide and Protein Identification Software. *Curr Protoc Bioinformatics (Chapter 13, Unit13 20).*

BOOTH, I. R. & LOUIS, P. 1999. Managing hypoosmotic stress: aquaporins and mechanosensitive channels in *Escherichia coli. Curr Opin Microbiol,* 2**,** 166-169.

BOYER, G. L., GILLAM, A. H. & TRICK, C. 1987. Iron chelation and uptake. *In:* FAY, P. & VAN BAALEN, C. (eds.) *The cyanobacteria.* Armsterdam, The Netherlands: Elsevier.

BRADFORD, M. M. 1976. Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem,* 72**,** 248-254.

BRAHAMSHA, B. 1996. An abundant cell-surface polypeptide is required for swimming by the nonflagellated marine cyanobacterium *Synechococcus. Proc Natl Acad Sci USA,* 93, 6504-6509.

BROWN, M. V. & FUHRMAN, J. A. 2005. Marine bacterial microdiversity as revealed by internal transcribed spacer analysis. *Aquat Microb Ecol,* 41**,** 15-23.

BROWN, M. V., SCHWALBACH, M. S., HEWSON, I. & FUHRMAN, J. A. 2005. Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity: development and application to a time series. *Environ Microbiol,* 7**,** 1466-79.

CHRISTIE-OLEZA, J. A., ARMENGAUD, J., GUERIN, P. & SCANLAN, D. J. 2015. Functional distinctness in the exoproteomes of marine *Synechococcus*. *Environ Microbiol,* 17**,** 3781-3794.

COLLIER, J. L., BRAHAMSHA, B. & PALENIK, B. 1999. The marine cyanobacterium *Synechococcus* sp. WH7805 requires urease (urea amiohydrolase, EC 3.5.1.5) to utilize urea as a nitrogen source- molecular-genetic and biochemical analysis of the enzyme. *Microbiol Mol Biol Rev,* 145**,** 447-459.

COX, A. D. & SAITO, M. A. 2013. Proteomic responses of oceanic *Synechococcus* WH8102 to phosphate and zinc scarcity and cadmium additions. *Front Microbiol,* 4**,** 387.

DUFRESNE, A., OSTROWSKI, M., SCANLAN, D. J., GARCZAREK, L., MAZARD, S., PALENIK, B. P. et al. 2008. Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol,* 9**,** R90.

DURAND, M. D., OLSON, R. J. & CHISHOLM, S. W. 2001. Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. *Deep Sea Res Part II: Trop Stud Oceanogr,* 48**,** 1983-2003.

FARRANT, G. K., DORE, H., CORNEJO-CASTILLO, F. M., PARTENSKY, F., RATIN, M., OSTROWSKI, M. et al. 2016. Delineating ecologically significant taxonomic

units from global patterns of marine picocyanobacteria. *Proc Natl Acad Sci USA,* 113**,** E3365-74.

FERREIRA, F. & STRAUS, N. A. 1994. Iron deprivation in cyanobacteria. *J Appl Phycol,* 6**,** 199-210.

FERRIS, M. J. & PALENIK, B. 1998. Niche adaptation in ocean cyanobacteria. *Nature,* 396**,** 226-228.

FLOMBAUM, P., GALLEGOS, J. L., GORDILLO, R. A., RINCON, J., ZABALA, L. L., JIAO, N. et al. 2013. Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus. Proc Natl Acad Sci USA,* 110**,** 9824-9.

FULLER, N. J., MARIE, D., PARTENSKY, F., VAULOT, D., POST, A. F. & SCANLAN, D. J. 2003. Clade-Specific 16S Ribosomal DNA Oligonucleotides Reveal the Predominance of a Single Marine *Synechococcus* Clade throughout a Stratified Water Column in the Red Sea. *Appl Environ Microbiol,* 69**,** 2430-2443.

FULLER, N. J., TARRAN, G. A., YALLOP, M., ORCUTT, K. M. & SCANLAN, D. J. 2006. Molecular analysis of picocyanobacterial community structure along an Arabian Sea transect reveals distinct spatial separation of lineages. *Limnol Oceanogr,* 51**,** 2515-2526.

GILBERT, P. M., KANA, T. M., OLSON, R. J., KIRCHMAN, D. L. & ALBERTE, R. S. 1986. Clonal comparisons of growth and photosynthetic responses to nitrogen availability in marine *Synechococcus* spp. *J Exp Mar Bio Ecol,* 101**,** 199-208.

GLOVER, H. E., SMITH, A. E. & SHAPIRO, L. 1985. Diurnal variations in photosynthetic rates- comparisons of ultraphytoplankton with a larger phytoplankton size fraction. *J Plankton Res,* 7**,** 519-535.

GROSSMAN, A. R., SCHAEFER, M. R., CHIANG, G. G. & COLLIER, J. L. 1993. The Phycobilisome, a Light-Harvesting Complex Responsive to Environmental Conditions. *Microbiol Rev,* 57**,** 725-749.

HARGEL, P., GERDING, J. J. T., FIEGGEN, W. & BLOEMENDAL, H. 1971. Cyanate formation in solutions of urea. Calculations of cyanate concentrations at different temperature and pH. . *Biochim Biophys Acta,* 243**,** 366-373.

HIRABAYASHI, K., YUDA, E., TANAKA, N., KATAYAMA, S., IWASAKI, K., MATSUMOTO, T. et al. 2015. Functional Dynamics Revealed by the Structure of the SufBCD Complex, a Novel ATP-binding Cassette (ABC) Protein That Serves as a Scaffold for Iron-Sulfur Cluster Biogenesis. *J Biol Chem,* 290**,** 29717-29731.

HIROKAWA, T., SEAH, B.-C. & MITAKU, S. 1998. SOSUI- classification and secondary structure prediction system for membrane proteins. *Bioinformatics,* 14**,** 378-379.

HUANG, S., WILHELM, S. W., HARVEY, H. R., TAYLOR, K., JIAO, N. & CHEN, F. 2012. Novel lineages of *Prochlorococcus* and *Synechococcus* in the global oceans. *ISME J,* 6**,** 285-97.

HUERTA-CEPAS, J., SZKLARCZYK, D., FORSLUND, K., COOK, H., HELLER, D., WALTER, M. C. et al. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res,* 44**,** D286-93.

HUNTER-CEVERA, K. R., NEUBERT, M. G., OLSON, R. J., SOLOW, A. R., SHALAPYONOK, A. & SOSIK, H. M. 2016. Physiological and ecological drivers of early spring blooms of a coastal phytoplankter. *Science,* 354**,** 326-329.

KASAHARA, M. & OHMORI, M. 1997. Carbamoyl Phosphate Synthetase of the Cyanobacteria *Anabaena Cylindrica*. *Plant Cell Physiol,* 38**,** 73-739.

KROGH, A., LARSSON, B., VON HEIJNE, G. & SONNHAMMER, E. L. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol,* 305**,** 567-80.

LI, W. K. W. 1994. Primary production of prochlorophytes, cyanobacteria and eucaryotic ultraphytoplankton- Measurements from flow cytometric sorting. *Limnol Oceanogr,* 39**,** 169-175.

LIN, X., WANG, L., SHI, X. & LIN, S. 2015. Rapidly diverging evolution of an atypical alkaline phosphatase (PhoA(aty)) in marine phytoplankton: insights from dinoflagellate alkaline phosphatases. *Front Microbiol,* 6**,** 868.

LINDELL, D. & POST, A. F. 1995. Ultraphytoplankton succession is triggered by deep winter mixing in the Gulf of Aqaba (Eilat), Red Sea. *Limnol and Oceanogr,* 40**,** 1130-1141.

MACKEY, K. R., POST, A. F., MCILVIN, M. R., CUTTER, G. A., JOHN, S. G. & SAITO, M. A. 2015. Divergent responses of Atlantic coastal and oceanic *Synechococcus* to iron limitation. *Proc Natl Acad Sci USA,* 112**,** 9944-9.

MACKEY, K. R. M., BRISTOW, L., PARKS, D. R., ALTABET, M. A., POST, A. F. & PAYTAN, A. 2011. The influence of light on nitrogen cycling and the primary nitrite maximum in a seasonally stratified sea. *Prog Oceanogr,* 91**,** 545-560.

MACKEY, K. R. M., RIVLIN, T., GROSSMAN, A. R., POST, A. F. & PAYTAN, A. 2009. Picophytoplankton responses to changing nutrient and light regimes during a bloom. *Mar Biol,* 156**,** 1531-1546.

MARCHETTI, A. & MALDONADO, M.T. 2016. Iron. *In* BOROWITZKA, M.A., BEARDALL, J. & RAVEN, J.A. (eds.) *The Physiology of Microalgae.* Cham: Springer International Publishing.

MAZARD, S., OSTROWSKI, M., PARTENSKY, F. & SCANLAN, D. J. 2012. Multi-locus sequence analysis, taxonomic resolution and biogeography of marine *Synechococcus*. *Environ Microbiol,* 14**,** 372-386.

MELLA-FLORES, D., MAZARD, S., HUMILY, F., PARTENSKY, F., MAHE, F., BARIAT, L., COURTIES, C., MARIE, D., RAS, J., MAURIAC, R., JEANTHON, C., BENDIF, E.M., OSTROWSKI, M., SCANLAN, D.J. & GARCZAREK, L. 2011. Is the distribution of *Prochlorococcus* and *Synechococcus* ecotypes in the Mediterranean Sea affected by global warming? *Biogeosciences,* 8, 2785-2804.

MIRZAEI, M., SOLTANI, N., SARHADI, E., PASCOVICI, D., KEIGHLEY, T., SALEKDEH, G.H., HAYNES, P.A., ATWELL, B.J. 2012. Shotgun proteomic analysis of long-distance drought signaling in rice roots. *J. Proteome Res.,* 11(1), 348-358.

MOORE, L.R., OSTROWSKI, M., SCANLAN, D.J., FEREN, K., SWEETSIR, T. 2005. Ecotypic variation in phosphorus-acquisition mechanisms within marine picocyanobacteria. *Aquat Microb Ecol,* 39, 257-269.

MOORE, L. R., POST, A. F., ROCAP, G. & CHISHOLM, S. W. 2002. Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*.pdf. *Limnol Oceanogr,* 47**,** 989-996.

MOREL, F. M. M., RUETER, J. G., ANDERSON, D. M. & GUILLARD, R. R. L. 1979. Aquil: a chemically defined phytoplankton culture medium for trace metal studies. *J Phycol,* 15**,** 135-141.

NICOLOFF, H., HUBERT, J.-C. & BRINGEL, F. 2001. Carbamoyl-phosphate synthetase (CPS) in lactic acid bacteria and other Gram-positive bacteria. *Le Lait,* 81**,** 151-159.

OLMEDO-VERD, E., SANTAMARIA-GOMEZ, J., OCHOA DE ALDA, J. A., RIBAS DE POUPLANA, L. & LUQUE, I. 2011. Membrane anchoring of aminoacyl-tRNA synthetases by convergent acquisition of a novel protein domain. *J Biol Chem,* 286**,** 41057-68.

OLSON, R. J., CHISHOLM, S., W., ZETTLER, E., R. & ARMBRUST, E. V. 1990. Pigments, size, and distribution of *Synechococcus* in the North Atlantic and Pacific Oceans. *Limnol Oceanogr,* 35**,** 45-58.

OMATA, T. & MURATA, N. 1983. Isolation and Characterization of the Cytoplasmic Membranes from the Blue-Green Alga (Cyanobacterium) *Anacystis nidulans. Plant Cell Physiol,* 24**,** 1101-1112.

OSTROWSKI, M., MAZARD, S., TETU, S. G., PHILLIPPY, K., JOHNSON, A., PALENIK, B. et al. 2010. PtrA is required for coordinate regulation of gene expression during phosphate stress in a marine *Synechococcus. ISME J,* 4**,** 908-21.

PAERL, H. W. 1991. Ecophysiological and Trophic Implications of Light-Stimulated Amino Acid Utilization in Marine Picoplankton. *MBio,* 57**,** 473-479.

PALENIK, B. 2001. Chromatic adaptation in marine *Synechococcus* strains. *Appl Environ Microbiol,* 67**,** 991-4.

Palenik, B. 2011. Recent functional genomics studies in marine *Synechococcus. In:* BURNAP, R.L. & VERMAAS, W.F.J. (eds.) *Functional Genomics and Evolution of Photosynthetic Systems.*: Springer Netherlands, Dordrecht.

PALENIK, B., REN, Q., DUPONT, C. L., MYERS, G. S., HEIDELBERG, J. F., BADGER, J. H. et al. 2006. Genome sequence of *Synechococcus* CC9311- Insights into adaptation to a coastal environment. *PNAS,* 103**,** 13555-13559.

PARTENSKY, F., BLANCHOT, J. & VAULOT, D. 1999a. Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters- a review. *Bull Inst Oceanogr Monaco-Numero Special,* 19**,** 457-476.

PARTENSKY, F., HESS, W. R. & VAULOT, D. 1999b. *Prochlorococcus*, a Marine Photosynthetic Prokaryote of Global Significance. *Microbiol Mol Biol Rev,* 63**,** 106-127.

PENNO, S., LINDELL, D. & POST, A. F. 2006. Diversity of *Synechococcus* and *Prochlorococcus* populations determined from DNA sequences of the N-regulatory gene *ntcA*. *Environ Microbiol,* 8**,** 1200-11.

POST, A. F., DEDEJ, Z., GOTTLIEB, R., LI, H., THOMAS, D. N., EL-ABSAWI, M., EL-NAGGAR, A. et al. 2002. Spatial and temporal distribution of *Trichodesmium* spp. in the stratified waters of the Gulf of Aqaba. *Mar Ecol Prog Ser,* 239**,** 241-250.

REXROTH, S., MULLINEAUX, C. W., ELLINGER, D., SENDTKO, E., ROGNER, M. & KOENIG, F. 2011. The plasma membrane of the cyanobacterium *Gloeobacter violaceus* contains segregated bioenergetic domains. *Plant Cell,* 23**,** 2379-90.

RICHARDSON, T. L. & JACKSON, G. A. 2007. Small phytoplankton and carbon export from the surface ocean. *Science,* 315**,** 838-40.

RIVERS, A. R., JAKUBA, R. W. & WEBB, E. A. 2009. Iron stress genes in marine *Synechococcus* and the development of a flow cytometric iron stress assay. *Environ Microbiol,* 11**,** 382-96.

ROCAP, G., DISTEL, D. L., WATERBURY, J. B. & CHISHOLM, S. W. 2002. Resolution of *Prochlorococcus* and *Synechococcus* Ecotypes by Using 16S-23S Ribosomal DNA Internal Transcribed Spacer Sequences. *Appl Environ Microbiol,* 68**,** 1180-1191.

ROCAP, G., LARIMER, F. W., LAMERDIN, J., ARELLANO, A., COLEMAN, M., HAUSER, L., et al. 2003. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature,* 424**,** 1042-1047.

SAIER, M. H., JR., TRAN, C. V. & BARABOTE, R. D. 2006. TCDB: the Transporter Classification Database for membrane transport protein analyses and information. *Nucleic Acids Res,* 34**,** D181-6.

SATO, N., KAMIMURA, R., KANETA, K., YOSHIKAWA, M. & TSUZUKI, M. 2017. Species-specific roles of sulfolipid metabolism in acclimation of photosynthetic microbes to sulfur-starvation stress. *Plos One,* 12.

SCANLAN, D. J., OSTROWSKI, M., MAZARD, S., DUFRESNE, A., GARCZAREK, L., HESS, W. R. et al. 2009. Ecological genomics of marine picocyanobacteria. *Microbiol Mol Biol Rev,* 73**,** 249-99.

SIX, C., THOMAS, J. C., GARCZAREK, L., OSTROWSKI, M., DUFRESNE, A., BLOT, N., SCANLAN, D. J. & PARTENSKY, F. 2007. Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. *Genome Biol,* 8**,** R259.

SOHM, J. A., AHLGREN, N. A., THOMSON, Z. J., WILLIAMS, C., MOFFETT, J. W., SAITO, M. A. et al. 2016. Co-occurring *Synechococcus* ecotypes occupy four major oceanic regimes defined by temperature, macronutrients and iron. *ISME J,* 10**,** 333-45.

STUART, R. K., DUPONT, C. L., JOHNSON, D. A., PAULSEN, I. T. & PALENIK, B. 2009. Coastal strains of marine *Synechococcus* species exhibit increased tolerance to copper shock and a distinctive transcriptional response relative to those of open-ocean strains. *Appl Environ Microbiol,* 75**,** 5047-57.

SU, Z., MAO, F., DAM, P., WU, H., OLMAN, V., PAULSEN, I. T. et al. 2006. Computational inference and experimental validation of the nitrogen assimilation regulatory network in cyanobacterium *Synechococcus* sp. WH 8102. *Nucleic Acids Res,* 34**,** 1050-65.

TAI, V. & PALENIK, B. 2009. Temporal variation of *Synechococcus* clades at a coastal Pacific Ocean monitoring site. *ISME J,* 3**,** 903-915.

TAKAHASHI, H., KOPRIVA, S., GIORDANO, M., SAITO, K. & HELL, R. 2011. Sulfur assimilation in photosynthetic organisms: molecular functions and regulations of transporters and assimilatory enzymes. *Annu Rev Plant Biol,* 62**,** 157-84.

TETU, S. G., BRAHAMSHA, B., JOHNSON, D. A., TAI, V., PHILLIPPY, K., PALENIK, B. & PAULSEN, I. T. 2009. Microarray analysis of phosphate regulation in the marine cyanobacterium *Synechococcus* sp. WH8102. *ISME J,* 3**,** 835-49.

TOLEDO, G. & PALENIK, B. 2003. A *Synechococcus* serotype is found preferentially in surface marine waters. *Limnol Oceanogr,* 48**,** 1744-1755.

TOMII, K. & KENEHISA, M. 1998. A Comparative Analysis of ABC Transporters in Complete Microbial Genomes. *Genome Res,* 8**,** 1048-1059.

VARKEY, D., MAZARD, S., OSTROWSKI, M., TETU, S.G., HAYNES, P. & PAULSEN, I.T. 2016. Effects of low temperature on tropical and temperate isolates of marine *Synechococcus. ISME J,* 10, 1252-1263

VIZCAINO, J.A. CSORDAS, A., DEL-TORO, N., DIANES, J.A., GRISS, J., LAVIDAS, L. et al. 2016. Update of the PRIDE database and its related tools. *Nucleic Acids Res*, 44 (D1), D447-D456.

WALLIN, E. & VON HEIJNE, G. 1998. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci,* 7**,** 1029-1038.

WEGENER, K. M., SINGH, A. K., JACOBS, J. M., ELVITIGALA, T., WELSH, E. A., KEREN, N. et al. 2010. Global proteomics reveal an atypical strategy for carbon/nitrogen assimilation by a cyanobacterium under diverse environmental perturbations. *Mol Cell Proteomics,* 9**,** 2678-89.

YU, N. Y., WAGNER, J. R., LAIRD, M. R., MELLI, G., REY, S., LO, R., et al. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics,* 26**,** 1608-15.

ZWIRGLMAIER, K., HEYWOOD, J. L., CHAMBERLAIN, K., WOODWARD, E. M., ZUBKOV, M. V. & SCANLAN, D. J. 2007. Basin-scale distribution patterns of picocyanobacterial lineages in the Atlantic Ocean. *Environ Microbiol,* 9**,** 1278-90.

ZWIRGLMAIER, K., JARDILLIER, L., OSTROWSKI, M., MAZARD, S., GARCZAREK, L., VAULOT, D. et al. 2008. Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. *Environ Microbiol,* 10**,** 147-161.

ZYBAILOV, B., MOSLEY, A. L., SARDIU, M. E., COLEMAN, M. K., FLORENS, L. & WASHBURN, M. P. 2006. Statistical Analysis of Membrane Proteome Expression Changes in *Saccharomyces cerevisiae*. *J Proteome Res,* 5**,** 2339-2347.

Chapter 3 Preface

The roles of phosphatases in bacteria remain poorly understood. Our previous investigation into the membrane proteome of marine *Synechococcus* (Chapter 2) shows that the phosphatase Sync_2505 is highly expressed in marine *Synechococcus* sp. CC9311. This finding suggests that uncharacterised phosphatase in *Synechococcus* may play a role in ecological niche adaptation. Therefore, we conducted a global protein-protein interactions (PPIs) study of the set of putative protein phosphatases encoded by *Synechococcus* sp. CC9311 in Chapter 3 and Chapter 4. Chapter 3 focuses on the PPIs of phosphatases and response regulators in *Synechococcus* sp. CC9311 and Chapter 4 investigates the interaction of phosphatases with the cytoplasmic lysate of *Synechococcus* sp. CC9311. Chapter 3 and 4 were conducted simultaneously as the successful purified phosphatases would be used for both studies.

The protein phosphatases found in the genome of marine *Synechococcus* sp. CC9311 include a histidine phosphatase, serine/threonine phosphatases and tyrosine phosphatases. Even though the reversible phosphorylation on serine/threonine/tyrosine protein residues has been identified in prokaryotes, their physiological roles are largely unknown. We hypothesized serine/threonine/tyrosine phosphatases may be involved in the two-component system bacterial signal transduction mechanism in cyanobacteria. Hence, we set up an experiment to investigate whether all these phosphatases interact with response regulators in marine cyanobacterium *Synechococcus* sp. CC9311.

We generated a list of phosphatases (Table 3.1) and response regulators (Table 3.2) in marine cyanobacteria *Synechococcus* sp. CC9311. My work involved recombinant cloning and expression of 9 protein phosphatases and 17 response regulators found in *Synechococcus* sp. CC9311. Successful purified protein products were subsequently utilised for pull-down experiments coupled with mass spectrometry analysis to study PPIs between putative phosphatases and response regulators.

Objectives:

i. To investigate whether response regulators interact with phosphatases. If yes, which phosphatases interact with which response regulators.

ii. To assess the applicability of affinity-purification mass spectrometry (AP-MS) to *in vitro* phosphatase protein-protein interactions study.

Protein-protein interactions of marine *Synechococcus* phosphatases and response regulators

3.1 INTRODUCTION

3.1.1 Phosphatases

Phosphorylation is a covalent modification that proteins frequently undergo following protein biosynthesis in living cells from prokaryotes to eukaryotes (Shi, 2009). The post-translational modifications (PTMs) of proteins by phosphorylation may take a variety of forms, but in essence it is a fundamental mechanism in signal transduction for enzyme activity and plays a vital role for regulating a broad range of physiological processes in all living organisms. The level of cellular phosphorylation is regulated by the counteracting activities of kinases and phosphatases. A protein kinase catalyses the transfer of γ-phosphate from ATP (or GTP (guanosine-5'-triphosphate)) to its protein substrates while a protein phosphatase catalyses the removal of the phosphate group from a phosphorylated substrate (Figure 3.1). Generally, proteins can be phosphorylated on histidine (His), aspartate (Asp), serine (Ser), threonine (Thr) and tyrosine (Tyr) residues (Swanson et al., 1994, Cozzone et al., 2004, Macek et al., 2007, Macek et al., 2008, Yang et al., 2013).
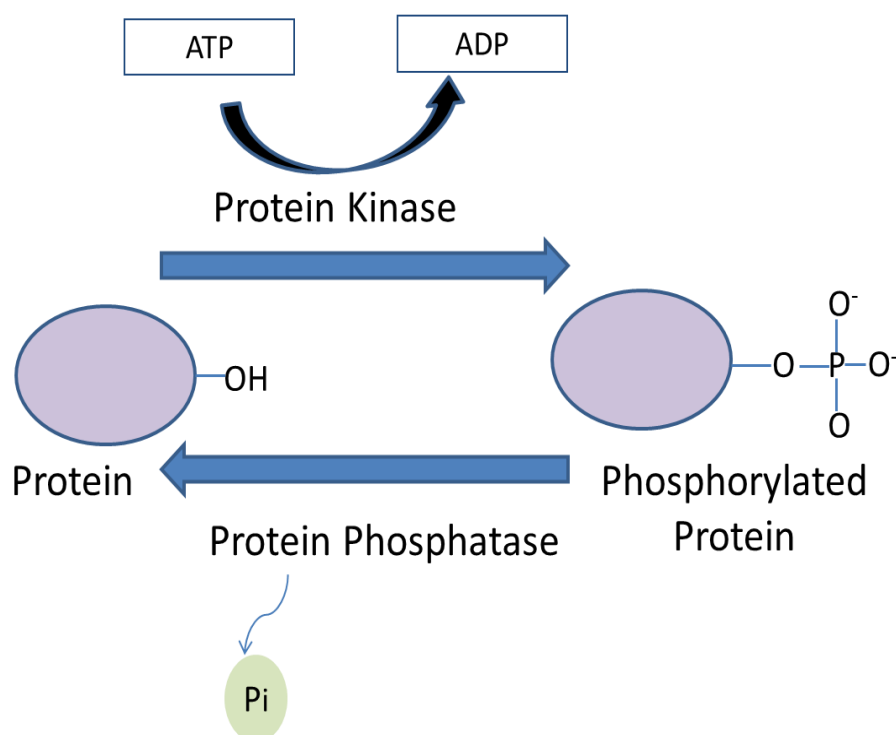
Figure 3.1. A reversible phosphorylation protein mechanisms in bacteria

Protein kinase and phosphatase are enzymes which function to catalyse the transfer of phosphate between their substrates. A protein kinase catalyses the transfer of γ-phosphate from ATP (or GTP) to its protein substrates while a protein phosphatase catalyses the removal of the phosphate group from the phosphorylated protein.

In bacteria, signalling phosphorylation is thought to occur primarily on His and Asp residues (Mizuno 1998). Bacteria including cyanobacteria (Mizuno et al., 1996, Hsiao et al., 2004, Cadoret et al., 2005,) generally use two-component regulatory systems (TCRS) which are composed of histidine kinases to sense and cognate response regulators to respond to environmental stimuli (Hoch, 2000). However, the need for dedicated phosphatases was not initially appreciated in the context of two-component systems and phosphorelay signal transduction (Hoch, 2000), since both phosphohistidine and aspartyl-phosphate residues undergo relatively rapid hydrolysis (Zhang, 1996, Sickmann and Meyer, 2001) and some histidine kinases exhibit phosphatase activity (Egger et al., 1997, Liu et al. 2017). Hence phosphatases were regarded as being mere housekeepers responsible for restoring the cellular ground state, and received much less attention in the past. Nonetheless, recent phosphoproteomic studies on bacteria show numerous proteins are phosphorylated on serine (Ser) and/or threonine (Thr) and/or tyrosine (Tyr) residues as

seen in eukaryotes and a variety of phosphatase family proteins have been identified (Shi et al., 1998, Kennelly, 2002). Unlike the kinase superfamily proteins with very similiar 3D protein structures and active site mechanism, phosphatases are differentiated into several enzyme superfamilies with variable 3D structures and diverse biochemistry mechanisms with different active sites and mechanisms of hydrolysis (Brautigan, 2013). In marine cyanobacteria, the recent identification of high expression of phosphatase proteins (Christie-Oleza et al., 2015) (Chapter 2) suggests they may play important, yet undefined roles in these key marine microbes.

The histidine phosphatase family is the first type of phosphatase family to be characterised, and includes both protein phosphatases and metabolite phosphatases. Research on this protein superfamily dates back to 1935 in yeast capable of interconverting 2-phosphoglycerate and 3-phosphoglycerate (Meyerhof and Kiessling, 1935). Generally, it consists of two branches sharing very limited sequence similarity: the larger branch comprises of proteins which function in metabolic regulation, intermediary metabolism and developmental processes, for example, phosphohistidine SixA, cofactor dependent phosphoglycerate mutase, alpha-ribazole phosphatase (CobC), mannitol-1-phosphatase, fructose-2,6-biphosphatase and acid phosphatase (PhoE) (Fothergill and Harkins, 1982, Pilkis et al., 1987, Rigden, 2008). The smaller branch comprises mainly acid phosphatases such as glucose-1-phosphatase (Pradel and Boquet, 1988) and phytases (Irving and D.J., 1972) with functions ranging from extracellular metabolism to developmental processes (Rigden, 2008). The functions of enzymes in this superfamily are based on a conserved catalytic histidine residue in the motif 'RHG' at the N terminal of the protein, which phosphorylates during the reaction (Rigden, 2008).

Phosphatases that react on Ser/Thr residues (PSPs) comprise three major families: phosphoprotein phosphatases (PPPs), metal-dependent protein phosphatases (PPMs), and the aspartate-based phosphatases. PPP family members contain regulatory subunits for interaction with substrates and their member include protein phosphatase 1 (PP1), PP2A and PP2B (commonly known as calcineurin) (Shi et al., 1998, Kennelly, 2002, Shi, 2009,). The PPM family members include protein phosphatases dependent on manganese/magnesium ions (Mn^{2+}/Mg^{2+}), such as PP2C and pyruvate dehydrogenase phosphatase. PPM proteins do not have regulatory subunits, instead contains additional domains and conserved sequence motifs that may help determine substrate specificity. Metal ions are important cofactors for both PPP and PPM to activate water molecules for

the dephosphorylation reaction. Aspartate-based phosphatases are very poorly understood, however one member protein FCP/SCP (TFIIF-associating component of RNA polymerase II CTD phosphatase/small CTD phosphatase) is known to react with C-terminal domain (CTD) of RNA polymerase II, which contains tandem repeats of a serine-rich heptapeptide.

Bacterial protein tyrosine phosphatases (PTPs) can be divided into three main groups: eukaryotic-like and dual-specificity phosphatases (Shi et al., 1998), low-molecular-weight protein tyrosine phosphatases (LMW PTP) (Shi et al., 1998, Grangeasse et al., 2003) and the polymerase and histidinol family of phosphoesterases (PHP) (Aravind and Koonin, 1998, Morona et al., 2002, Standish and Morona, 2014). All are specific for phosphorylated tyrosine, except eukaryotic-like and dual-specificity phosphatases, which can also hydrolyze protein-bound phosphoserine and/or phosphothreonine residues (Cozzone et al., 2004). The first two groups of PTPs are ubiquitously found in both prokaryotes and eukaryotes and contain an active site sequence motif $C(X)_5R$ which also recognized as PTP signature motif in general. Unlike the other groups, PHPs are found predominantly in Gram-positive bacteria and lack the characteristic $C(X)_5R$ motif (Cozzone et al., 2004). PHPs show optimal activity in basic pH and are metal dependent (Morona et al., 2002, Mijakovic et al., 2005).

3.1.2 Response regulators

Response regulators (RRs) are one of the main components of TCRS. Response regulator (RR) genes are identified in a wide range of bacteria and archaeal species (Galperin, 2010, Jenal and Galperin, 2009). The TCRS are activated upon detection of an internal or external stimulus that results in the histidine kinase undergoing auto-phosphorylation on a histidine residue. The phosphate group is then transferred to an aspartate residue of the receiver cognate RR. RR typically consists of two domains: a receiver (REC) domain and an output domain. Receiver domains contain the phosphoacceptor aspartate and several other highly conserved amino acids that catalyse phosphotransfer from a histidine kinase. Most RRs contain a DNA-binding output domain, and serve as transcriptional regulators when activated by phosphorylation of the receiver domain (Galperin, 2010). Meanwhile some RRs also contain RNA-binding, ligand-binding, protein binding, enzymatic and transporter output domains and are involved in regulation at the transcriptional, post-transcriptional or post-translational levels (Galperin, 2010). Commonly, the regulator

domain is found at the amino terminus of a composite protein that also encodes a DNA-binding domain. The sheer number and diversity of RRs suggest the diversity of functions that they perform. Generally, RRs can be divided into 4 main classes: RR class I, II, III, and IV and discussed in the following paragraphs

RR class I (RRI) consists of proteins with an RR domain within a polypeptide less than 200 amino acids (Ashby, 2004). RRI proteins lack an identifiable output domain and raises question of their mechanisms involved in cell signalling (Ashby, 2004). It is suspected that RRI proteins interact with no more than one partner besides its cognate kinase. For example, in *Escherichia coli*, after autophosphorylation of the CheA histidine kinase, the phosphoyl group is transferred to the CheY (RRI), which then binds with flagellar motor proteins (Bilwes et al., 1999, Webre et al., 2003).

RR class II comprises of an N-terminal RR domain fused to an output DNA-binding domain, either a T_reg (OmpR subclass) (Martinez-Hackert and Stock, 1997), HTH_LuxR (NarL subclass) (Schwartz et al., 1998), or AraC (AraC subclass) (Tobes and Ramos, 2002). RR class II proteins are proposed to function as transcriptional regulators and this is the most common type of RR found in cyanobacteria (Ashby, 2004). Examples of the cyanobacterial RR class II proteins include the sucrose synthesis regulator OrrA (Schwartz et al., 1998), manganese regulator ManR (Ogawa et al., 2002, Yamaguchi et al., 2002) and phosphate regulator PhoB (Aiba et al., 1993, Suzuki et al., 2004).

Some cyanobacterial response regulators have two or three RR domains, together with T_reg-Hpt domains (for histidine phosphotranfer) grouped in RR class III. Examples of RR class III in cyanobacteria include Crr93 from heterocystous N_2 fixers, and RcaC from *Fremyella displosiphon* (Chiang et al., 1992). Crr93 has a similar domain organization as RcaC and have been found to involve in complementary chromatic adaptation (Chiang et al., 1992). Both the N-terminal RR and Hpt domains play a crucial role in regulation of phycocyanin gene expression (Li and Kehoe, 2005).

The majority of the proteins in RR class IV do not have any DNA-binding domains, but a number of them have output domains with putative catalytic activities. More than 40% of these polypeptides possess a GGDEF domain, also called DUF1. Expression of recombinant GGDEF domains from different bacteria including freshwater cyanobacteria *Synechocystis* sp. strain PCC6803 showed that they all exhibit diguanylate cyclase activity

(Ryjenkov et al., 2005). Phosphorylation of the RR is pre-requisite for signal transduction of GGDEF domain and production of a secondary messenger, cyclic diguanylic acid.

Recently, there are a number of metabolic enzymes been found to bind to the REC domain of response regulators in bacteria (Galperin, 2010). These enzymes include P-loop-type ATPases of the MinD/ParA and PilB families, nucleotide phosphorylase, sugar transferase, dolichyl-phosphate glucosyltransferase, NAD(P)-dependent glutamate dehydrogenase, threonine synthase and potential metal-dependent hydrolase (Galperin, 2010). This suggests a very diverse range of roles of response regulators in the regulation of cell signalling in bacteria.

3.2 METHODOLOGY

The workflow implemented in this work is represented in Figure 3.2.
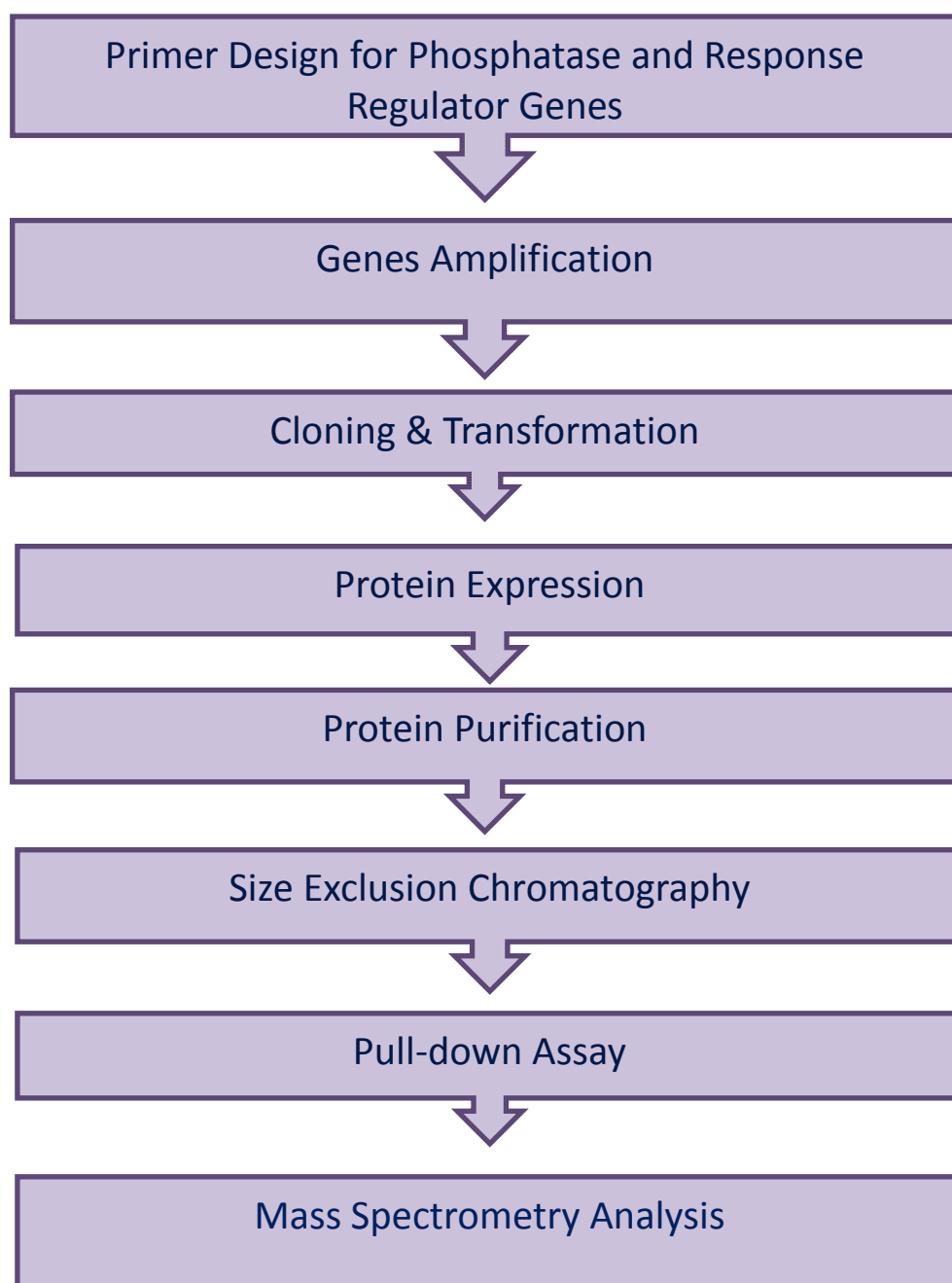
Figure 3.2. Experimental workflow of Chapter 3.

3.2.1. Primer design, PCR amplification and cloning of target genes

Nine protein phosphatase and seventeen response regulator genes were identified from the *Synechococcus* sp. CC9311 genome as shown in Table 3.1 and Table 3.2, which were used for the protein-protein interactions study of phosphatases and response regulators in this chapter.

Synechococcus sp. CC9311 genomic DNA was obtained from postdoctoral fellow Dr. Bhumika Shah in the Paulsen laboratory. Polymerase chain reaction (PCR) was conducted for every target gene using a high-fidelity Qiagen Taq DNA polymerase (Qiagen). Primers were designed to introduce ±15bp extensions homologous to the ends of the N-terminal His-tag plasmid pOPINF (for the cloning of the response regulators) and N-terminal His and GST tag plasmid pOPINJ (primarily for the cloning of the protein phosphatases) (OPPF, UK) (see Table 3.3 for the plasmids used, Supplementary Info 3.1 for the plasmid maps). pOPINJ and pOPINF plasmid constructs were made by In-Fusion™ cloning (Clontech) full target genes into the pOPINF (for response regulators) and pOPINJ (for phosphatases) plasmids. For the response regulators, primers for cloning a partial length gene encoding the response regulator domain were designed as well for In-Fusion™ cloning. The designed primers were analysed using Oligocalc (Kibbe, 2007) with the melting temperature adjusted to 65 ± 5^{0}C for most sequences and manufactured by Integrated DNA Technologies (IDT) Australia (see Supplementary Info 3.2, 3.3, 3.4 for primer sequences). The components and conditions of the PCR reactions are outlined in Table 3.4 and Table 3.5. Amplified PCR products were purified using a commercial purification kit (Wizard ® SV Gel and PCR Clean-Up System, Promega) and visualized on a 1% agarose gel.

pOPINJ and pOPINF plasmids were linearized by digestion with *Kpn*I and *Hin*dIII restriction enzymes. The digestion reaction components mixture is summarized in Table 3.6. The digestion mixture was incubated at 37^{0}C for 1 hour. Linearized vector was purified using a commercial purification kit (Wizard ® SV Gel and PCR Clean-Up System, Promega) and visualized on a 0.8% agarose gel.

Table 3.1. Protein phosphatases identified in *Synechococcus* sp. CC9311

Gene	Gene product	Protein domain	Protein family	Protein size (kDa)	pI	Signal peptides (Y/N)	Cysteine no.
sync_0012	SpoII domain protein	Stage II sporulation protein E (SpoIIE)	RsbU	52	8.71	N	4
sync_0151	Putative phosphatase	Dual specificity phosphatase, catalytic domain	Protein tyrosine phosphatase (PTP)	18	6.75	N	4
sync_0484	Low molecular weight phosphotyrosine protein	Low molecular weight phosphotyrosine protein phosphatase	Low molecular weight phosphatase (LMWP)	18	5.57	N	3
sync_1056	Serine/threonine specific protein phosphatase	Calcineurin-like phosphoesterase	Metallophosphatase superfamily/Calcineurin-like phosphoesterase	22	6.58	N	6
sync_1732	Serine/threonine specific protein phosphatase	Calcineurin-like phosphoesterase domain, apaH type; metallophosphoesterase, DNA ligase-associated	Metallophosphatase	31	7.34	N	7
sync_1857	Serine/threonine specific protein phosphatase	Twin arginine translocation (Tat) signal profile; calcineurin-like phosphoesterase domain, apaH type	CpdA superfamily	39	8.32	Y	4
sync_2505	Possible protein phosphatase 2C	-	DUF1400 superfamily	23	5	Y	2
sync_2627	Phosphohistidine phosphatase SixA	Histidine phosphatase superfamily	Histidine phosphatase (HP) superfamily	19	9.22	N	8
sync_2828	Possible protein phosphatase 2A regulatory B subunit	-	-	19	4.77	N	1

Table 3.2. Response regulators identified in *Synechococcus* sp. CC9311

Gene	Gene product	Protein domain	Protein family	Protein size (kDa)	pI	RR class
sync_0115	Two-component response regulator	REC, Trans_reg_C domains	OmpR family	28	5.31	-
sync_0265	Two-component response regulator	REC domain	CheY family	15	4.27	RRI
sync_0574	Two-component response regulator	REC, HTH_LUXR domains	NarL family	21	5.52	RRII
sync_0586	Possible DNA-binding response regulator	REC, HTH_LUXR domains	NarL family	29	7.09	RRII
sync_0669	Two-component response regulator	REC, Trans_reg_C domains	OmpR family	28	6.63	RRII
sync_0707	DNA-binding response regulator	REC, Trans_reg_C domains	OmpR family	27	4.95	RRII
sync_0807	Two-component response regulator	Trans_reg_C domain	OmpR family	25	6.45	RII
sync_1079	Sensory box histidine kinase/response regulator	PAS, HisKA, HATPase, REC domains	Hybrid HisK	58	5.72	-
sync_1116	DNA-binding response regulator	REC, HTH_LUXR domains	NarL family	25	6	RRII
sync_1134	Response regulator receiver domain	REC domain	CheY family	13	4.36	RRI
sync_1145	Possible Two-component response regulator	REC, Trans_reg_C domains	OmpR family	28	7.07	RRII
sync_1219	Two-component response regulator	REC, Trans_reg_C domains	OmpR family	27	7.92	RRII
sync_1232	Two-component response regulator	REC, Trans_reg_C domains	OmpR family	28	5.37	RRII
sync_2111	DNA-binding response regulator	REC, Trans_reg_C domains	OmpR family	28	4.65	RRI
sync_2591	DNA-binding response regulator, RpaA	REC, Trans_reg_C domains	OmpR family	30	5.48	RRII
sync_2597	DNA-binding response regulator	REC, Trans_reg_C domains	OmpR family	28	4.84	RRII
sync_2642	DNA-binding response regulator	REC, HTH_LUXR domains	NarL family	27	4.85	RRII

Table 3.3. Vectors used for cloning

Vector	Product	Vector size (bp)	Forward primer extension	Reverse primer extension
pOPINF	His6-3C-POI	5531	AAGTTCGTTTCAGGCCG	ATGGTCTAGAAAGCTTTA
pOPINJ	His6-GST-3C-POI	6218	AAGTTCGTTTCAGGCCG	ATGGTCTAGAAAGCTTTA
pOPINM	His6-MBP-3C-POI	6665	AAGTTCGTTTCAGGCCG	ATGGTCTAGAAAGCTTTA
pOPIN3C	His6-SUMO-3C-POI	5824	AAGTTCGTTTCAGGCCG	ATGGTCTAGAAAGCTTTA

(Notes: Cloning and transformation procedures were the same for all vectors unless otherwise mentioned).

Table 3.4. PCR reaction components

Components	In-Fusion™ cloning PCR (μl)	Colony PCR (μl)
25mM $MgCl_2$	2	2.5
Buffer	5 (10x Buffer)	5 (5x Buffer)
10mM dNTP	1	1
Forward primer	5	0.5
Reverse primer	5	0.5
DNA template	5	Single colony
DNA polymerase	0.2 (Qiagen *Taq* Polymerase (5 units/μl))	0.1 (GoTaq DNA Polymerase (5 units/μl))
Sterile water	26.8	15.4
Total volume	50	25

Table 3.5. PCR reaction conditions

In-Fusion™ cloning PCR			Colony PCR		
94^0C	3 mins		95^0C	2 mins	
94^0C	30 s		95^0C	30 s	
55^0C	30 s	30 cycles	55^0C	30 s	30 cycles
72^0C	1 min		72^0C	1.5-2.5 mins	
72^0C	10 mins		72^0C	5 mins	
4^0C	∞		4^0C	∞	

Table 3.6. Enzyme Digestion Components

Contents	Reaction Volume (μl)
Plasmid/DNA	10
10X buffer	2
Kpn I enzyme	1
*Hin*d III enzyme	1
Sterile water	6
Total	**20**
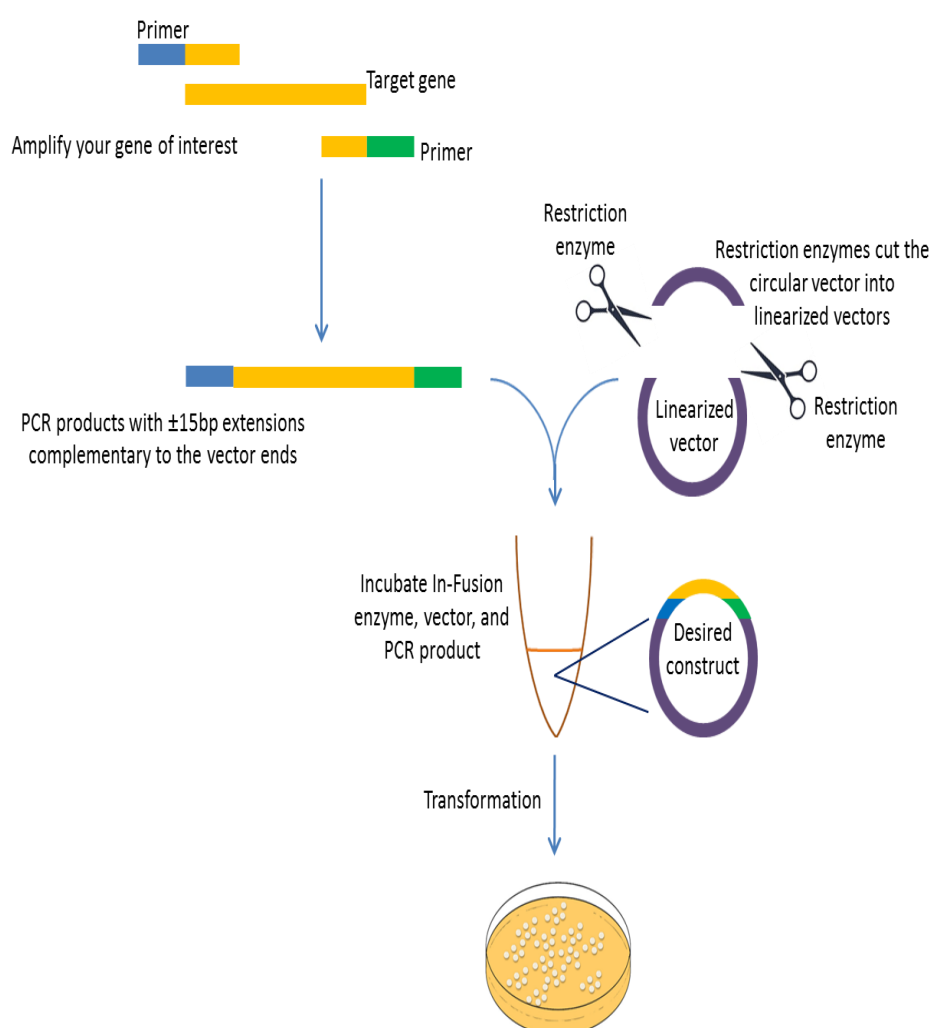
Figure 3.3. Ligation independent In-Fusion™ cloning (modified from In-Fusion™ Clonetech instruction manual).

Ligation-independent cloning (LIC) (Aslanidis and De Jong, 1990, Haun and Moss, 1992) was conducted using the In-Fusion™ system (Clontech) (Figure 3.3). Purified PCR product and digested vector were mixed in 3:1 molar ratio, 0.66μl In-Fusion™ HD Enzyme Premix was added and the final volume was adjusted to 3.3μl with sterile water.

3.2.2 Transformation

3.2.2.1 Transformation into *E. coli* Stellar™ Competent Cells.

The recombinant vector was transformed into Stellar™ Competent Cells (Clontech) as recommended by Clonetech's In-Fusion™ PCR Cloning Kits. The transformation procedures involved adding recombinant vector to Stellar™ Competent Cells and incubated on ice for 30 minutes. It is then followed by heat-shock for 30-45 seconds at 42^0C in the water bath and returned to ice for 2 minutes. After that, 300μl SOC (Super optimal broth with catabolite repression) media (Clontech) was added to the competent cells and incubated in incubator shaker at 37^0C for an hour, agitated at 200rpm. The entire transformation mixture was then plated onto antibiotic-selective LB (Luria-Bertani) agar plates (Ampicillin 50μg/ml) supplemented with 1mM IPTG (isopropyl β-D-1-thiogalactopyranoside) and 0.02% X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galacto-pyranoside) in DMF (dimethyl formamide). The culture plates were incubated at 37^0C for 16-18 hours. All the cell culturing was performed in a laminar flow cabinet.

Successful transformants containing the desired cloned gene were easily picked up with the blue-white screening technique based on expression of the *lacZ* gene, which is interrupted by the cloned insert (Woodman, 2008). Putative clones were subsequently verified by colony PCR (refer Table 3.4 and Table 3.5 for colony PCR conditions) using a gene-specific reverse and a vector-specific T7 forward primer. The colony PCR reactions were visualized with a 1.0% agarose gel. Plasmids were extracted from the cultured colonies using Wizard® Plus SV Minipreps DNA Purification System (Promega) and further verified by DNA sequencing (Macrogen, Inc.).

3.2.2.2 Transformation into expression hosts

a) Rosetta™ 2 (DE3) p*Lac*l *E. coli* competent cells

Rosetta™ 2 (DE3) p*Lac*l *E. coli* competent cells were selected as expression hosts for transformation. Recombinant plasmid (1-2µl) was added to the expression host (25µl) and incubated on ice for 30 minutes, followed by heat-shock at 42^0C for 30-45 seconds and then back to ice incubation for 5 minutes. Cells were recovered at 37^0C under agitation at 200rpm for 90 minutes after addition of SOC media (250µl). An aliquot of 20µl cell cultures was spread on 24-well antibiotic selective LB agar plates and incubated at 37^0C overnight. The culture plates were stored at 4^0C for small-scale protein expression trials and subsequently large scale production.

b) Lemo21 (DE3) *E. coli* competent cells

For proteins that did not express or were expressed as an insoluble product, Lemo21 (DE3) (New England Biolabs Inc.) were used for expression. The Lemo21 (DE3) competent *E. coli* cells were thawed on ice for 10 mins before 3µl of plasmid DNA was added to it. The mixture was incubated on ice for 30 minutes and heat-shocked for 42^0C for exactly 10 seconds and then placed on ice for 5 mins. Next, 400µl of SOB (super optimal broth) solution (New England Biolabs Inc.) was mixed with the mixture at 37^0C under agitation at 200rpm for 60 minutes. Then, 50µl of the mixture was added to the antibiotic selective plate and incubated at 37^0C overnight. The culture plate was stored at 4^0C for small-scale protein expression trials and subsequently large scale production.

3.2.3 Protein expression and purification

3.2.3.1 Small-scale Protein Expression Screening

A 24-deep well plate was used for growing 2ml bacterial cell cultures in LB broth at 37^0C overnight with appropriate antibiotics. One tenth of the overnight culture was transferred to a new 24-deep well block plate supplemented with 3 ml of Power Broth (Molecular Dimensions) with appropriate antibiotics. Bacterial cells were grown at 37^0C under agitation about 3 hours until reaching OD_{600} of ~0.5. The cultures were cooled down to 20^0C for 20 minutes and 1mM IPTG was added and incubated for ~18 hours, shaking at 200 rpm at 20^0C. Cultured cells were harvested by centrifugation at 4000 rpm at 4^0C for

30 minutes. The media was decanted with caution without disturbing the pellet. The cultured plate was sealed and stored in -80^{0}C freezer until ready for protein extraction. Protein expression of Lemo21 (DE3) *E. coli* competent cells was similar to the protein expression methods described above, except L-rhamnose (from 0 to 2mM) was added when the cultures initially grow and a final concentration of 0.4mM IPTG was used for induction.

Protein extraction was initiated by resuspending defrosted cell pellets completely in 210μl of Lysis Buffer A (refer Table 3.7 for all the buffers used for protein extraction) and incubated on ice for 30 minutes before centrifugation at 4000rpm for 30 minutes at 4^{0}C. Before centrifugation finished, 20μl of the Ni-NTA (nickel nitrilotriacetic acid) magnetic beads (GenScript) were added to the wells of 96-well plate and the plate was placed on the 96-well magnetic rack (Qiagen) for 5 minutes. The aqueous solution of magnetic beads was decanted with pipette and the magnetic beads were washed with water two times. After a second wash, the magnetic beads were equilibrated with 200μl Binding Buffer A twice. The supernatant solution (extracted proteins) was transferred to the plate placed on a 96-well magnetic rack for 2 minutes, placed on the magnetic rack for 2 minutes, and then the solution was pipetted off. Next, 200μl Wash Buffer A was mixed with the magnetic beads for 5 minutes, placed on the magnetic rack for 2 minutes and then decanted with pipette. This process was repeated twice. Ultimately, 50μl of Elution Buffer A was mixed with the beads for 1 minute, placed on the magnetic rack for 1 minute and the elution was transferred to a new 1.5ml Eppendorf tubes and prepared for one dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis. In one case, instead of Buffer A, Buffer X was used to help solubilize the proteins (refer Supplementary info 3.5 for Buffer X components).

The reagents required for SDS-PAGE analysis are presented in Table 3.8. Protein samples were boiled in loading dye for 10 minutes and loaded on a 4-15% precast gel (BioRad). The SDS-PAGE running condition was set up at 80V for 10 minutes followed by 165V for 40 minutes. After this, the SDS-PAGE gel was immersed in fixing solution for an hour. Colloidal Coomassie concentrate was mixed with colloidal Coomassie diluent in the ratio of 4:1 for total volume of 50ml and added onto the gel for 2 hours incubation under agitation. After that, the staining solution was removed and the gel was washed with distilled water twice before destaining solution was added to destain the unbound colour dye on the gel, so that the protein bands can be clearly visualized.

Table 3.7. Protein preparation buffers

Notes: Buffer A was used for small scale protein expression and Buffer B for large scale protein expression.

Buffer	Composition
Lysis Buffer A	50mM NaH$_2$PO$_4$, 300mM NaCl, 10mM imidazole, 1% v/v Tween 20, adjust pH to 8.0 using NaOH. Lysozyme (1mg/ml) and 400 Kunits/ml of DNAse I is added when mix with cell pellet.
Lysis Buffer B	HEPES buffer pH 8.0, 300mM NaCl, 5% glycerol, 400 Kunits/ml DNAse I, 1% v/v protease inhibitor cocktail.
Binding Buffer A	Buffer NPI-10-Tween: 50mM NaH$_2$PO$_4$, 300mM NaCl, 10mM imidazole, 1% v/v Tween 20, adjust pH to 8.0 using NaOH.
Binding Buffer B	50mM HEPES buffer pH8.0, 300mM NaCl, 5% glycerol, 20mM Imidazole.
Wash Buffer A	Buffer NPI-20-Tween: 50mM NaH$_2$PO$_4$, 300mM NaCl, 20mM imidazole, 0.05% v/v Tween 20, adjust pH to 8.0 using NaOH.
Wash Buffer B	50mM HEPES buffer pH 8.0, 300mM NaCl, 5% glycerol, 50mM Imidazole, pH is adjusted to pH 8.0 after mixture.
Elution Buffer A	Buffer NPI-250-Tween: 50mM NaH$_2$PO$_4$, 300mM NaCl, 250mM imidazole, 0.05% v/v Tween 20, adjust pH to 8.0 using NaOH.
Elution Buffer B	50mM HEPES buffer pH 8.0, 300mM NaCl, 5% glycerol, 400mM Imidazole, pH is adjusted to pH 8.0 after mixture.
Running Buffer for SEC	50mM HEPES buffer pH 8.0, 300mM NaCl, 5% glycerol

Table 3.8. Solutions used for SDS-PAGE analysis

Solution	Components
Loading Dye	100mM Tris buffer pH 6.8, sodium dodecyl sulfate (SDS) (4%w/v), bromophenol blue (0.2%w/v), glycerol (20% w/v), 200mM dithiothreitol (DTT)
Fixing Solution	40% methanol, 7% acetic acid
Coomassie Staining Solution (concentrate)	Coomassie Brilliant Blue G-250 (0.1875% w/v), 25% Methanol, 85% Phosphoric acid
Colloidal Coomassie Diluent Solution	50% Ammonium sulfate
Destaining Solution	25% Methanol

3.2.3.2 Large-scale protein expression and purification

3.2.3.2a Protein expression

Large-scale protein expression was conducted using ZYP-5052 auto-induction media (Studier, 2005). The components of the media are described in Supplementary Info 3.6 and 3.7. A single colony was inoculated into 10ml of Power Broth with appropriate antibiotics in a 50ml Falcon tube. The culture was incubated at 37^0C, with 200rpm agitation overnight. The overnight culture was then transferred to 8x 2L shake flasks with 500ml of ZYP-5052 media and incubated at 37^0C under 200rpm agitation for 3-7 hours until the OD_{600} reached 0.4-0.6. After bacterial growth reaching mid-log phase, the culture flask was shifted to shaking incubator and grown overnight at 20^0C with agitation at 200rpm. The bacterial cells were harvested with JLA-10.500 centrifuge rotor, centrifuging at 5,000 rpm for 20 minutes at 4^0C. The supernatant was discarded and the cell pellet was resuspended in Lysis Buffer B without DNaseI and stored at -80^{0}C until ready for protein extraction. Two batches of bacterial culture were conducted for recombinant protein phosphatases with the same bacterial cultures condition.

3.2.3.2b Protein purification

The protein extraction and purification of two batches of bacterial culture for recombinant protein phosphatases were identical except every buffer used in the second batch bacterial culture was added with 0.5mM TCEP (Tris(2-carboxyethyl) phosphine). Meanwhile, the conditions of protein extraction and purification of response regulator proteins were same as first batch bacterial culture of recombinant protein phosphatases. Frozen cells were thawed on ice before cell lysis was conducted. Dnase I was added before cells were lysed by sonication. After cell lysis, the cells were centrifuged at 20,000 x g at 4^0C for 20 minutes. Purification of His-tagged proteins for all phosphatase and response regulator proteins were carried out using HisPur™ Ni-NTA Spin Columns (Thermo Fisher Scientific). Solutions required for His-tagged protein purification are represented in Table 3.7. All purification steps were conducted at 4^0C. HisPur ™ Ni-NTA Spin Columns were first equilibrated with 5ml Binding Buffer B twice. Then, the resin in the column was mixed with soluble lysed cells and incubated for 2 hours. After that, columns were spun at 700 x g for 2 minutes. This step was repeated until all soluble lysed cells were spun through the column. Next, 5ml of Wash Buffer B was added and centrifuged at 700 x g for 2 minutes. This process was repeated six times, and then new centrifuge tubes were used to collect the eluted proteins. Elution was conducted by adding 1ml of Elution Buffer B followed by centrifugation at 700 x g for 90 seconds and this process was repeated 5 times. Eluted proteins were analysed by SDS-PAGE and the rest of the purified protein was stored at 4^0C, and the protein purity was further enhanced by size exclusion chromatography.

3.2.4 Size exclusion chromatography

Generally, three solutions were required for the size exclusion chromatography (SEC) assay: 20% ethanol, water and running buffer. The components of the running buffer were represented in Table 3.7. Two different conditions of running buffer was used for SEC: one was added with 0.5mM TCEP for eluted proteins in TCEP buffers (phosphatase proteins only) while another one was without TCEP for eluted proteins without TCEP (phosphatase and response regulator proteins). Superdex™ 200 10/300 GL (column bed volume of 24ml) (GE Healthcare Life Sciences) and HiLoad™ 16/600 Superdex™ 200pg (column bed volume of 120ml) (GE Healthcare Life Sciences) SEC columns were used

with the Akta Pure chromatography system (GE Healthcare Life Sciences) for protein purification and elimination of imidazole which can interfere with the protein analysis. Purified proteins were collected from the chromatography column and their concentrations were estimated spectrometrically, using the extinction coefficients of the proteins. Vivaspin® 6 centrifugal concentrator (GE Healthcare Life Sciences) was used if the concentrations of the protein were low. Purified proteins were stored at -80^0C until ready for the protein-protein interaction study.

3.2.5 Protein-protein interactions of protein phosphatases with response regulators

3.2.5.1 Pull-Down Assay

The mechanisms involved in the pull-down assay are illustrated in Figure 3.4. Purified phosphatase (40μg) was gently mixed with purified response regulator (40μg) and incubated at 4^0C for 90 minutes. During the interval time of protein incubation, 20μl of glutathione magnetic beads (Genscript) were transferred to a 1.5ml Eppendorf tube and placed on a magnetic rack (Qiagen) for 2 minutes. The solution was pipetted off and 100μl of water was added to the beads, placed on the magnetic rack for 2 minutes and then the solution was pipetted off. This process was repeated twice. After discarding the water, 100μl GST Binding/Wash buffer (50mM Tris-HCl, pH 7.5, 300mM NaCl, 5% glycerol) was added to equilibrate the magnetic beads, placed on the magnetic rack for 2 minutes and then pipetted off. This process was repeated twice and then the mixture of proteins were transferred to the magnetic beads and incubated at 4^0C for 40 minutes. The mixture was then placed on the magnetic rack for 2 minutes and the solution was pipetted off. Next, 200μl GST Wash Buffer was added to the protein mixture, placed on the magnetic rack for 2 minutes and then the solution was pipetted off. This process was repeated three times. Finally, 100 μl GST Elution Buffer (20mM reduced glutathione (GSH), 50mM Tris-HCl, pH8.0, 300mM NaCl, 5% glycerol) was added to the protein mixture, incubated at 4^0C for 2 minutes, followed by magnetic stand for 2 minutes before the proteins were eluted. This process was repeated twice. Native gel electrophoresis analysis was conducted to visualize the protein bands from the pull-down assay. Native gel electrophoresis running condition was set up at 80V for 10 minutes followed by 165V for 45 minutes. The reagents used for the native gel electrophoresis are represented in Table 3.9 while gel staining and

destaining reagents used after electrophoresis were same as described in SDS-PAGE. Thirty microliters of the initial protein elution were used for native gel electrophoresis while the rest was used for mass spectrometry analysis for PPIs verification.
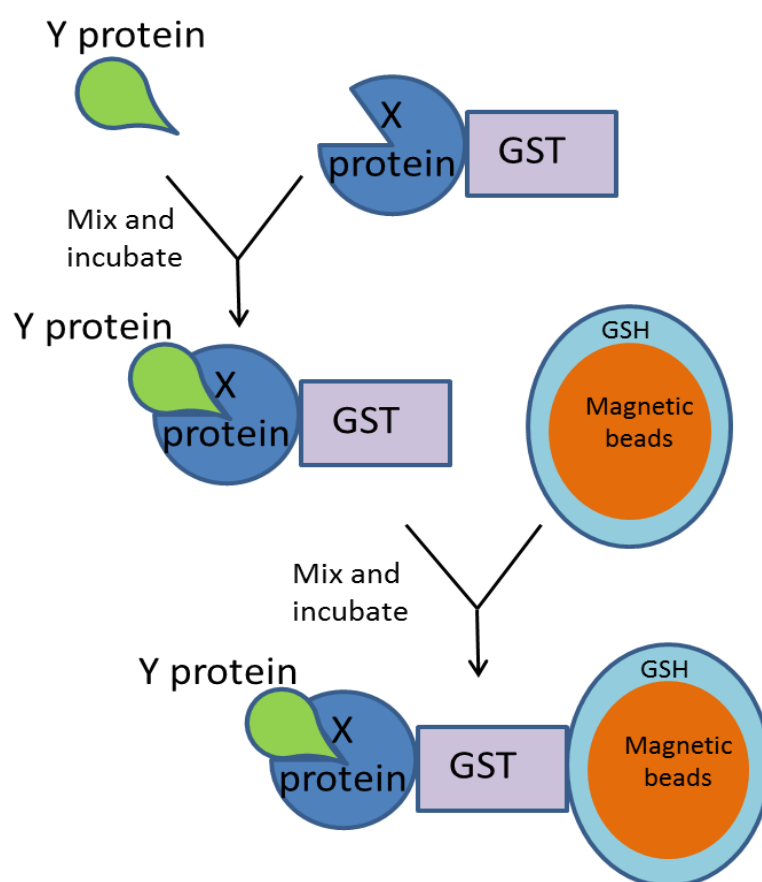
Figure 3.4. Illustration of GST-tagged protein pull-down assay. GST-tagged phosphatase protein (represented by X protein) was mixed with response regulator protein (represented by Y protein) and incubated at 4^0C. The solution was then mixed with the GST-magnetic beads. If X and Y were interacted, they were bound to the GST-magnetic beads.

Table 3.9. Native gel electrophoresis solution

Reagents	Components
Sample buffer (2X)	62.5mM Tris-HCl, pH 6.8, 25% glycerol, 1% bromophenol blue
Running buffer (1X)	25mM Tris, 192mM glycine (Running buffer should be ~pH 8.3)

3.2.5.2 Proteins/Peptides extraction

3.2.5.2a Methanol/Chloroform Protein Precipitation

Protein samples (~40-80µg) were mixed with 400µl Methanol and vortexed. This was followed by 100µl of chloroform and vortexed. Then, 300µl of water was added and vortexed (the samples would look cloudy after the vortexing). Next, the sample was centrifuged at 14,000 x g for 2 minutes. The top aqueous layer of the solution was pipetted off. The protein should be located between the layers and may be visible as a thin wafer. After the top aqueous layer removal, 400µl of methanol was added, vortexed and followed by centrifugation at 14,000 x g for 3 mins. The supernatant was pipetted off as much as possible retaining the pellet without disturbance. The remaining solution was dried out in Speed-vac vacuum concentrator and ready for peptide extractions.

3.2.5.2b In-solution protein digestion

Refer Table 3.10 for all the reagents required for in-solution digestion. Proteins were mixed with 100µl of urea, vortexed and sonicated for 2 minutes. This was then followed by addition of 5µl of dithiothreitol (DTT) reducing reagent, vortexed and incubated at room temperature for 60 minutes. After that, 20µl of iodoacetamide (IAA) alkylating reagent was added, vortexed and incubated at room temperature in the dark for 60 minutes. Then, 20µl DTT was added, vortexed and incubated for another 30 minutes. Last but not least, the urea concentration was diluted with 775µl MilliQ water and vortexed before trypsin was added in an estimated ratio of 1:30 for peptide digestion.

3.2.5.2c Peptide Desalting and Purification

Trypsinated peptides were reconstituted in 30µl 2% formic acid (FA), vortexed for 45 seconds and centrifuged at 3,000 x *g* for 1 minute. Stage Tips were pre-conditioned with 60µl of Solvent A, centrifuged at 450 x *g* for 2 minutes (refer Table 3.11 for all the solvent required for peptide desalting and purification). It was then followed by 60µl of Solvent B and centrifuged at 450 x *g* for 2 minutes. Peptide samples were loaded into Stage Tips, and centrifuged at 1,500 x *g* for 2 minutes. This process was continued until all the peptide samples were finished loading. Stage Tips were washed with 60µl of Wash Solvent A, centrifuged at 1,500 x *g* for 2 minutes. This process was repeated twice. It was then followed by 60µl of Wash Solvent B, centrifuged at 1,500 x *g* for 2 minutes. After this, the Stage Tip assembly was placed onto a new 1.5mL Eppendorf tube and elution buffer was loaded onto it and centrifuged at 1,500 x *g* for 2 minutes. Eluted peptide solutions were dried with a Speed-vac vacuum concentrator and stored at 4^0C until ready for mass spectrometry analysis.

Table 3.10. Reagents for in-solution protein digestion

Reagents
0.4 M Tris Stock, pH 7.8
6M Urea in Tris buffer, pH7.8
Reducing reagents: 200mM DTT in Tris buffer, pH 7.8
Alkylating reagents: 200mM iodoacetamide in Tris buffer, pH 7.8
20 µg Trypsin solution (125µl Tris stock+ 375µl MiliQ water)

Table 3.11. Solvent for peptide desalting and purification

Solution	Mixture components
Sample reconstitution solvent	0.2% formic acid (FA) in ultrapure water
Conditioning Solvent A	acetonitrile
Conditioning Solvent B	ultrapure water
Wash Solvent A	water: acetonitrile: FA in the ratio of 94.8:5:0.2
Wash Solvent B	acetonitrile
Elution buffer	acetonitrile: water: ammonium hydroxide in the ratio of 60:35:5

3.2.5.3 Mass spectrometry analysis

Dried peptides were solubilized in 21μl of 0.1% FA, vortexed and centrifuged at 14,000 rpm for 15 minutes. The process was repeated twice. For mass spectrometry analysis, 10μl of samples was transferred into a MS-vial and loaded into a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-nLC1000 (Thermo Fisher Scientific) instrument for peptide analysis. Reversed-phase chromatographic separation was conducted on a 75μm i.d. x100mm, C18 HALO column, 2.7μm bead size, 160Å pore size. A linear gradient of 1-50% solvent B (99.9% ACN/0.1% FA) was run over 30 minutes, followed by a slope gradient of solvent B from 50% to 85% for 2 minutes and maintained at 85% for 8 minutes. The mass spectrometer was operated in the data-dependent mode to automatically switch between Orbitrap MS and ion trap MS/MS acquisition. Survey full scan MS spectra (from *m/z* 350 to 2000) were acquired with a resolution of 35,000 and an AGC (Automatic Gain Control) target value of $1x10^6$ ions. The ten most abundant ions were selected for higher energy collisional dissociation (HCD) fragmentation. HCD normalised collision energy was set to 35% and fragmentation ions were detected in the Orbitrap at a resolution of 17,500. Target ions that had been selected for MS/MS were dynamically excluded for 10sec.

3.2.6 Bioinformatic analysis

Bioinformatic analyses conducted including phylogenetic analysis of the protein phosphatases and response regulators in marine cyanobacteria *Synechococcus* and protein identification analysis from the pull-down assay coupled with mass spectrometry.

3.2.6.1 Phylogenetic analysis of phosphatases and response regulators

The putative protein phosphatase and response regulator amino acid sequences from marine cyanobacteria *Synechococcus* were extracted from Cyanorak database (http://www.sb-roscoff.fr/cyanorak). The protein sequences were clustered using Clustal X v2.1 software (Larkin et al., 2007) and the phylogenetic tree was visualized using iTOL software (Letunic and Bork, 2016).

3.2.6.2 Protein identification from pull-down assays

Thermo XCalibur raw files from Q-Exactive mass spectrometer were analysed with MaxQuant software (Tyanova et al., 2016) for protein identification. *Synechococcus* sp. CC9311 protein sequence database extracted from Cyanorak was used as the database for peptide searching. The search parameter used for the database searching: trypsin was chosen as a digestive enzyme, the first search peptide tolerance was set at 20ppm, main search peptide tolerance at 4.5ppm and the MS/MS tolerance at 20ppm. Two missed cleavages were allowed and carbamidomethylation of cysteines was set as a fixed modification and methionine oxidation as a variable modification. Only proteins identified with at least two unique peptides were analysed.

3.3 RESULTS

3.3.1 Phylogenetic tree of phosphatases in marine cyanobacteria

To better understand the distribution of protein phosphatases across the major *Synechococcus* clades, representing distinct marine environmental niches, the predicted protein phosphatase protein sequences of marine cyanobacteria *Synechococcus* sp. CC9311 (Clade I), CC9605 (Clade II), WH8102 (Clade III) and CC9902 (Clade IV) were extracted from the Cyanorak database (http://www.sb-roscoff.fr/cyanorak). The phylogenetic trees are shown in Figure 3.5 including a total of 29 phosphatases across the four *Synechococcus* strains within three major families of phosphatases namely, serine/threonine (Ser/Thr) phosphatases, tyrosine (Tyr) phosphatases and histidine (His) phosphatases. Among them, the serine/threonine phosphatases were the largest protein phosphatase family found in marine *Synechococcus*.

Ser/Thr phosphatases in marine *Synechococcus* comprises of PPP and PPM phosphatase superfamilies. The primary member of PPP is phosphatase 2C (Sync_2505, Syncc9605_2305, SYNW2159 and Syncc9902_0387) while the members of PPM include protein phosphatase 2A, protein phosphatase 2B (commonly known as calcineurin) and SpoIIE. The SpoIIE (Sync_0012, Syncc9605_0012, SYNW0012, Syncc9902_0012) protein homologs are shown in Figure 3.5a are the most recently evolved group of proteins in the Ser/Thr phosphatase protein family in *Synechococcus* spp.

Tyrosine phosphatase proteins have been identified in the genome of all four marine *Synechococcus* studied. Tyrosine-specific PTPs that act on phosphotyrosine and dual-specificity protein tyrosine phosphatase (DsPTP), which can dephosphorylate both phosphotyrosine and phosphoserine/phosphothreonine are shown in the phylogenetic tree of tyrosine phosphatases. Interestingly, the dual-specificity phosphatase was only identified in *Synechococcus* sp. CC9311 (Clade I) (protein Sync_0151) among the four representative strains of marine *Synechococcus* clade I-IV (Figure 3.5b).

Unlike the Ser/Thr phosphatase and Tyr phosphatase protein family, only one histidine phosphatase protein (phosphohistidine phosphatase SixA) is seen in each of the 4 marine *Synechococcus* strains studied (Figure 3.5c).

3.3.2 Phylogenetic tree of response regulators in marine cyanobacteria *Synechococcus* sp. CC9311

The phylogenetic tree of 17 response regulators (RRs) found in marine *Synechococcus* sp. CC9311 is depicted in Figure 3.6. These include RRs from the OmpR (10 homologs), NarL (4 homologs), CheY (2 homologs) protein families as well as one hybrid kinase (Table 3.2). Sync_1079 protein is a hybrid kinase, which contains a HisKA (dimerization and phosphoacceptor), HATPase (histidine kinase ATPase) and PAS domains, which make a histidine kinase, and a receiver (REC) domain which make a response regulator.

The evolutionary divergence of the response regulators in *Synechococcus* sp. CC9311 can be divided into three main branches (Figure 3.6). The first cluster is composed of two response regulators Sync_0707 and Sync_2597 belonging to the OmpR family. The second cluster includes two main subgroups with majority of proteins belonging to the OmpR family: the first group includes Sync_1079, Sync_1145 and Sync_1232, in which Sync_1079 (hybrid kinase) is more deeply branching, while, the second group consists of Sync_0115, Sync_2642, Sync_0669, and Sync_2111. Meanwhile the rest of the proteins fall into the third cluster composed of a variety of response regulators from different families including OmpR, NarL and CheY.

a) Serine/Threonine Phosphatases

Tree scale: 0.01

Syncc9605_1654
SYNW0913
Sync_1056
SYNW2159
SYNW0186
Sync_2828
Syncc9902_2217
Syncc9605_2562
SYNW2398
Sync_2505
Syncc9605_2305
Syncc9902_0387
Sync_1732
Syncc9605_0668
Sync_1857
Sync_0012
Syncc9605_0012
SYNW0012
Syncc9902_0012

b) Tyrosine phosphatases

Tree scale: 0.01 ⊢⊣

- Sync_0151
- SYNW0915
- Sync_0484
- Syncc9902_1913
- Syncc9605_0417
- SYNW2026

c) Histidine phosphatases

Tree scale: 0.01 ⊢⊣

- Syncc9605_2414
- SYNW2276
- Sync_2627
- Syncc9902_2004
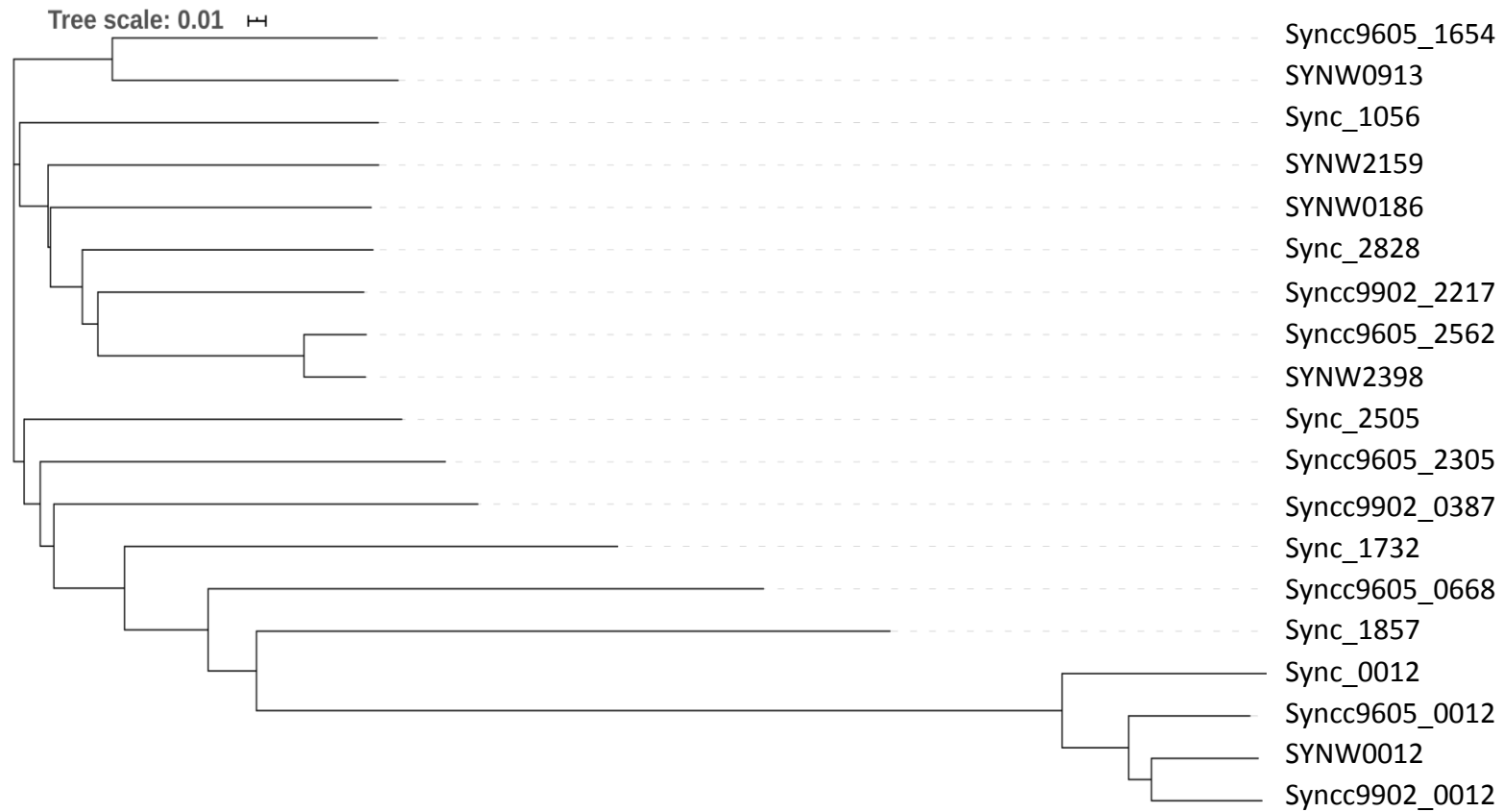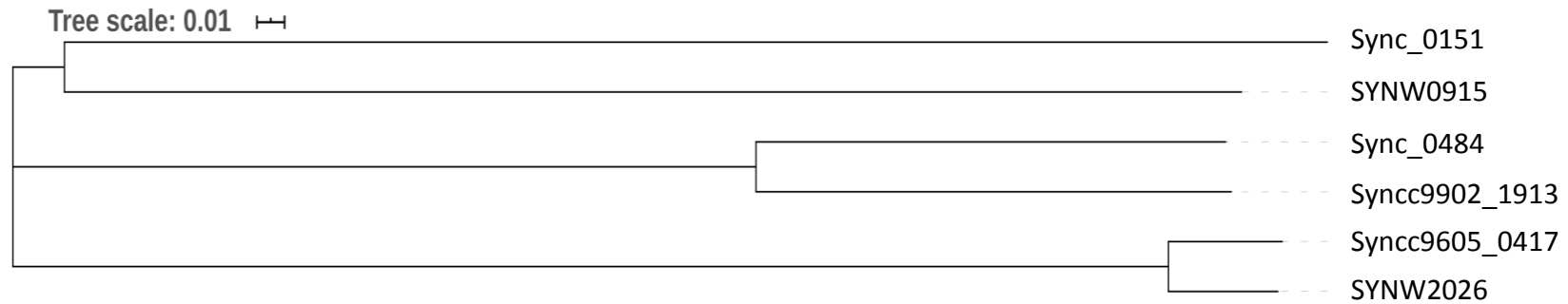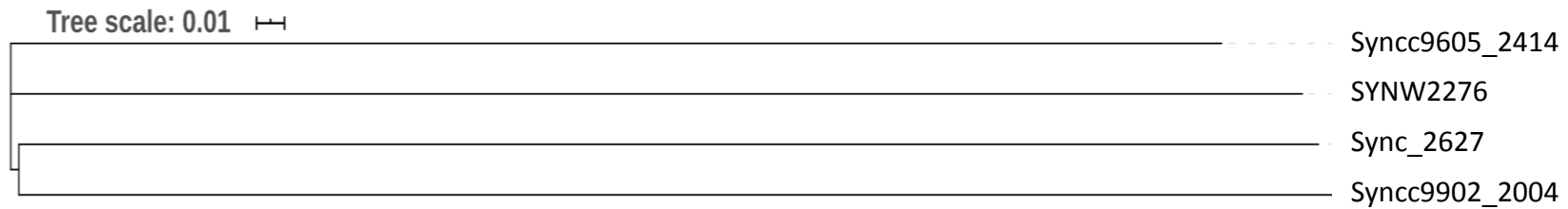
Figure 3.5. Phylogenetic trees of a) serine/threonine phosphatase, b) tyrosine phosphatase and, c) histidine phosphatase family proteins in marine cyanobacteria *Synechococcus* spp.

Tree scale: 0.1

Sync_0707 (OmpR family)
Sync_2597 (OmpR family)
Sync_1079 (Hybrid HisK)
Sync_1145 (OmpR family)
Sync_1232 (OmpR family)
Sync_0115 (OmpR family)
Sync_2642 (NarL family)
Sync_0669 (OmpR family)
Sync_2111 (OmpR family)
Sync_0586 (NarL family)
Sync_1219 (OmpR family)
Sync_2591 (OmpR family)
Sync_1116 (NarL family)
Sync_0807 (OmpR family)
Sync_0574 (NarL family)
Sync_0265 (CheY family)
Sync_1134 (CheY family)

Figure 3.6. Phylogenetic tree of response regulator proteins in marine cyanobacteria *Synechococcus* sp. CC9311.

3.3.3 PCR amplification, cloning and transformation

PCR amplified products of all the protein phosphatase and response regulator genes were visualized on 1% agarose gel as represented in Figure 3.7. All selected protein phosphatase and response regulator genes were amplified successfully from the genomic DNA of *Synechococcus* sp. CC9311, except for the response regulator gene *sync_1079*.

Amplified PCR products of the protein phosphatase and response regulator genes were cloned in pOPINJ and pOPINF vectors, respectively. Positive clones were further verified via DNA sequencing. A frameshift mutation was detected for the protein phosphatase Sync_1732 and this gene was therefore re-amplified using a high-fidelity Platinum Pfx DNA polymerase (Thermo Fisher Scientific) followed by a repeat of the cloning and transformation steps, however, the same frameshift mutation in the gene was still observed, suggesting that our strain of CC9311 contains a chromosomally-encoded frameshift in this gene. Therefore, Sync_1732 was discontinued from the rest of the experimental procedures.

As described in section 3.3.4, not all of the cloned protein phosphatase genes were observed to produce soluble protein expression. In those cases, pOPINM and pOPINS3c vectors were used as alternative cloning vectors. Overall, 8 protein phosphatase genes were successfully cloned into the pOPINJ vector, 3 into the pOPINM vector, and 6 into the pOPINS3c vector. Meanwhile, all response regulators except *sync_1079* which failed in PCR amplification were successfully cloned into the pOPINF vector. Recombinant protein phosphatase-encoding plasmids were transformed into the *E. coli* expression host Rosetta™ 2 (DE3) p*Lac*I and/or Lemo21 (DE3) for protein expression. Meanwhile the recombinant plasmids encoding the response regulators were transformed into Lemo21 (DE3) *E. coli* cells. A summary of the overall cloning results is shown in Table 3.12 and 3.13.
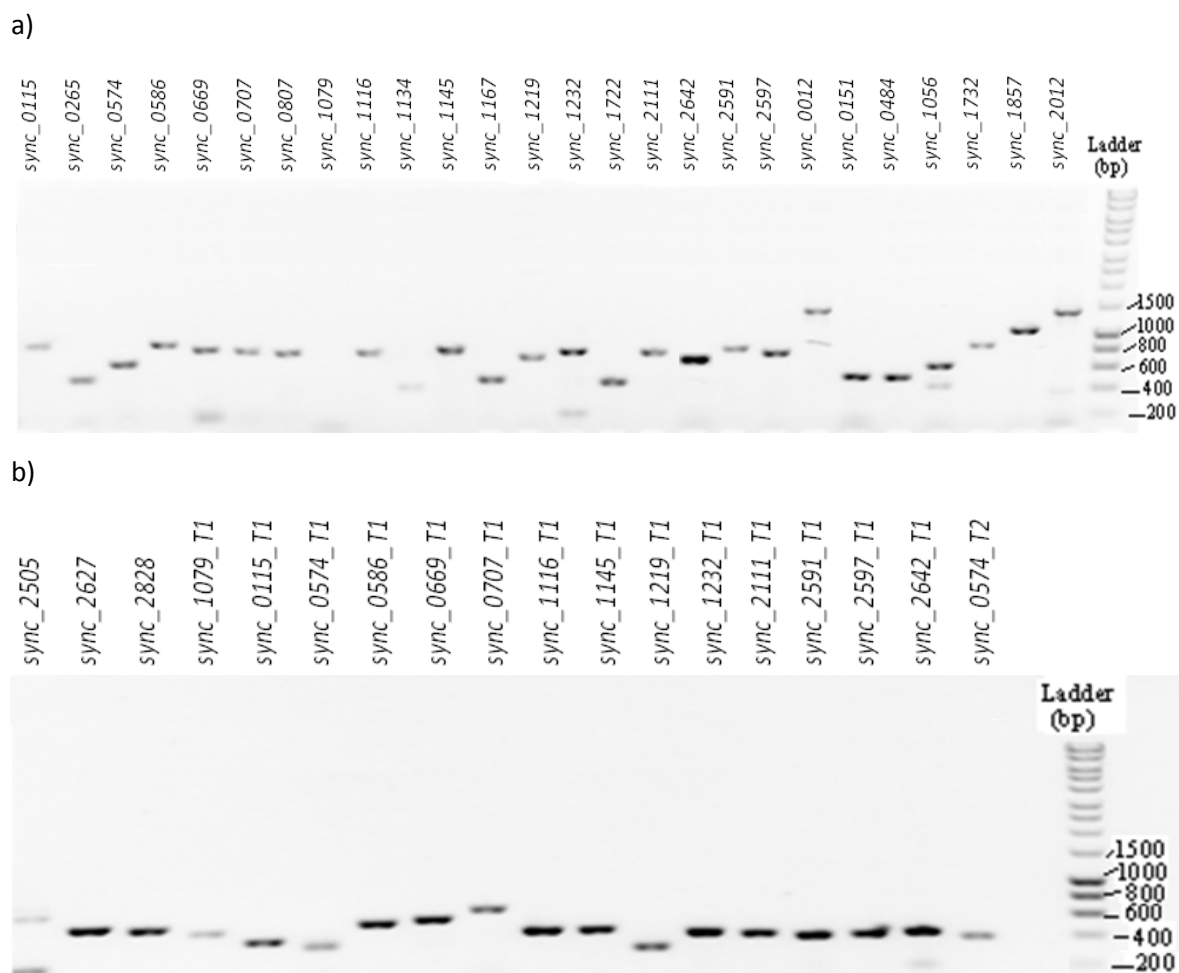
Figure 3.7 a, b. PCR amplification products of *Synechococcus* sp. CC9311 response regulators and phosphatases in agarose 1% gel (refer Table 3.1 and 3.2 for further details of the genes). The DNA ladder is indicated in the right hand lane of both figures with the molecular sizes of the markers indicated. *Sync_1722 and Sync_2012* were used as positive controls. The gene names ending with T1 or T2 were fragments of the response regulator genes that encoded the response regulator domains.

Table 3.12. Summary table of experiment conducted on phosphatases from PCR amplification until protein expression.

Phosphatase genes	PCR amplification	Cloning and transformation						Protein Expression	Solubility
		Rosetta™ 2 (DE3) p*Lac*l *E. coli* competent cells			Lemo21 (DE3) *E. coli* competent cells				
		pOPINJ vector	pOPINM vector	pOPINS3c vector	pOPINJ vector	pOPINM vector	pOPINS3c vector		
sync_0012	✓	✓	F	✓	x	F	N.T.	✓	x
sync_0151	✓	✓	F	✓	x	F	N.T.	✓	x
sync_0484	✓	✓	F	✓	✓	F	N.T.	✓	✓
sync_1056	✓	✓	F	✓	x	F	N.T.	✓	x
sync_1732	✓	*	-	-	-	-	-	-	-
sync_1857	✓	✓	F	✓	✓	F	N.T.	✓	✓
sync_2505	✓	x	✓	✓	x	x	✓	✓	x
sync_2627	✓	✓	✓	N.T.	N.T.	N.T.	N.T.	✓	✓
sync_2828	✓	✓	✓	N.T.	N.T.	N.T.	N.T.	✓	✓

(Notes: ✓ indicates yes/success; x indicates no/not success; - indicates not applicable as failure in former steps; N.T. indicates not tested, * indicates frameshift mutations; F indicates failure in transformation of the plasmid into Stellar™ competent cells in the initial cloning step. Protein expression and solubility results are recorded as yes/success for the successful case of any kind of vector transformed in either Rosetta™ 2 p*Lac*l *E. coli* and Lemo21 (DE3) *E. coli*.

Table 3.13. Summary table of experiment conducted on response regulators from PCR amplification until protein expression.

Response regulator genes	PCR amplification	Cloning and transformation Competent cells		Protein Expression	Solubility
		E. coli Stellar™	*E. coli* Lemo21 (DE3)		
sync_0115	✓	✓	N.T.	N.T.	N.T.
sync_0265	✓	✓	N.T.	N.T.	N.T.
sync_0574	✓	✓	✓	✓	✓
sync_0586	✓	✓	✓	✓	✓
sync_0669	✓	✓	N.T.	N.T.	N.T.
sync_0707	✓	✓	✓	✓	✓
sync_0807	✓	✓	✓	✓	✓
sync_1079	X	-	-	-	-
sync_1116	✓	✓	✓	✓	✓
sync_1134	✓	✓	N.T.	N.T.	N.T.
sync_1145	✓	✓	N.T.	N.T.	N.T.
sync_1219	✓	✓	✓	✓	✓
sync_1232	✓	✓	✓	✓	✓
sync_2111	✓	✓	N.T.	N.T.	N.T.
sync_2591	✓	✓	✓	✓	✓
sync_2597	✓	✓	N.T.	N.T.	N.T.
sync_2642	✓	✓	✓	✓	✓
Response regulator domains					
sync_0115_T1	✓	✓	N.T.	N.T.	N.T.
sync_0574_T1	✓	✓	N.T.	N.T.	N.T.
sync_0574_T2	✓	✓	N.T.	N.T.	N.T.
sync_0586_T1	✓	✓	N.T.	N.T.	N.T.
sync_0669_T1	✓	✓	N.T.	N.T.	N.T.
sync_0707_T1	✓	✓	N.T.	N.T.	N.T.

sync_1079_T1	✓	✓	N.T.	N.T.	N.T.
sync_1116_T1	✓	✓	N.T.	N.T.	N.T.
sync_1145_T1	✓	✓	N.T.	N.T.	N.T.
sync_1219_T1	✓	✓	N.T.	N.T.	N.T.
sync_1232_T1	✓	✓	N.T.	N.T.	N.T.
sync_2111_T1	✓	✓	N.T.	N.T.	N.T.
sync_2591_T1	✓	✓	N.T.	N.T.	N.T.
sync_2597_T1	✓	✓	N.T.	N.T.	N.T.
sync_2642_T1	✓	✓	N.T.	N.T.	N.T.

(Notes: ✓ indicates yes/success; x indicates no/not success; - indicates not applicable as failure in former steps; N.T. indicates not tested). Due to limitations in time, not all of the response regulator gene constructs were transformed into the *E. coli* Lemo21 (DE3) expression system, as it was deemed a higher priority to test a subset of the response regulators in PPI interaction pull-down study, rather than expressing all of the response regulators and not having enough time to undertake the PPIs study.

3.3.4 Protein expression and purification

Small scale analysis of protein expression was conducted for all of the successfully cloned protein phosphatase genes. Eight phosphatase proteins namely, Sync_0012, Sync_0151, Sync_0484, Sync_1056, Sync_1857, Sync_2505, Sync_2627 and Sync_2828 were tested for small scale protein expression (Figure 3.8). All of these protein phosphatase genes were successfully transformed into the expression host Rosetta™ 2 (DE3) p*Lac*l competent *E. coli,* except for Sync_2505. Five out of the remaining seven putative protein phosphatase proteins showed detectable protein expression when visualized by 1D SDS-PAGE gels (Figure 3.8). Two of them (Sync_2627 and Sync_2828) were present in soluble form and were purified by Ni-NTA magnetic beads (Figure 3.8 b, c). Meanwhile the other three (Sync_0151, Sync_0484 and Sync_1056) were found to be expressed in an insoluble form (Figure 3.8d, f).

Several approaches were used to optimize protein expression and solubility: First, Lemo21 (DE3) competent *E. coli* (New England Biolabs Inc.) (Pathak et al., 2016) which has been design for expression of complex proteins was used as an alternative expression host. The plasmid constructs encoding the insoluble recombinant proteins Sync_0012, Sync_0151,

Sync_0484, Sync_1056, Sync_1857 and Sync_2505 were transformed into Lemo21 (DE3) competent *E. coli*. However, among the six of them, only constructs encoding Sync_1857 and Sync_0484 were successfully transformed into the Lemo21 host strain. Both of them were tested for small scale protein expression according to manufacturer's protocol. Varying concentrations of rhamnose (0μM, 100μM, 500μM, 1000μM, 2000μM) were used to optimize the protein expression and solubility of the proteins. A protein band corresponding in size to Sync_1857 was detected in SDS-PAGE gels for all the concentrations of rhamnose tested (Figure 3.9a). Therefore, in large-scale protein expression of phosphatase Sync_1857 (see below), no rhamnose was added since the protein can be purified even without rhamnose in the small-scale protein expression test. Meanwhile, Sync_0484 protein was expressed in insoluble form in the Lemo21 (DE3) *E. coli* host cells (Figure 3.9b, c). Instead of the set of Buffer A solutions in Table 3.7, Buffer X (Supplementary Info 3.5) was used to enhance protein solubility. Buffer X enhanced solubility of the protein phosphatase Sync_0484 enabling this protein to be successfully purified in soluble form (Figure 3.9c).

Another approach used to enhance protein solubility was using other vectors which encode proteins that can assist in protein folding. Vectors pOPINS3c (encoding the SUMO chaperone protein) and pOPINM (which fuses the target gene with maltose binding protein) were also used for the cloning of protein phosphatase genes whose expression had proved recalcitrant in this study. The target genes were cloned into these vectors and transformed into the expression host Rosetta™ 2 (DE3) p*Lac*l competent *E. coli*. The success of the cloning and transformation of these constructs is summarized in Table 3.12. Protein expression from the recombinant genes in vector pOPINM and pOPINS3c are represented in Figure 3.10, 3.11. Recombinant proteins of Sync_2505, Sync_2627 in pOPINM vector showed protein expression but in insoluble form, while Sync_2828 in pOPINM vector exhibited a band the size of the maltose binding protein alone, suggesting that the fused protein phosphatase has been cleaved off (Figure 3.10a). The recombinant proteins in pOPINS3c showed protein expression but in insoluble form, these include Sync_0012, Sync_0151, Sync_0484, Sync_1056 and Sync_1857 (Figure 3.10 a, b). Meanwhile, recombinant protein Sync_2505 showed no protein expression (Figure 3.10 a, b). Sync_2505 was a very highly expressed protein in *Synechococcus* sp. CC9311 (membrane proteomics study in Chapter 2), hence recombinant protein Sync_2505 in pOPINS3c was further tested for protein expression by transforming into Lemo21 (DE3)

E. coli competent. SDS-PAGE showed expression of Sync_2505 recombinant protein using the pOPINS3c vector, but in an insoluble form (Figure 3.11).

For the response regulators, nine response regulator proteins namely Sync_0574, Sync_0586, Sync_0707, Sync_0807, Sync_1116, Sync_1219, Sync_1232, Sync_2591 and Sync_2642 were tested for small scale protein expression. All of them were successfully purified in small scale and subsequently proceed with large-scale purification (see Figure 3.15 j, k). Overall, small-scale protein expression resulted in the purification of 4 putative protein phosphatase proteins, namely Sync_0484, Sync_1857, Sync_2627 and Sync_2828, and 9 response regulator proteins as presented in Table 3.12, 3.13.

Bacteria were cultured in 2L volumes (four 500ml cultures in 2L flask) of media for large-scale protein expression. All protein preparations showing >80% purity in small scale or large scale expression is henceforth referred as pure protein. All successful purified proteins in small-scale protein expression were taken further for large-scale protein purification except Sync_0484. Sync_0484 did not proceed with large-scale protein purification as there was concern with the possibility that the buffer which contained sodium deoxycholate may interfere with the protein-protein interaction between the putative phosphatase and response regulators. A difference from small-scale protein purification which used Ni-NTA magnetic beads, all proteins in large-scale production were purified with HisPur™ Ni-NTA Spin Columns. This is mainly due to the concern that the ferromagnetism of the magnetic beads could influence the binding of the proteins for the pull-down assays, in particular for phosphatase affinity purification mass spectrometry assay in Chapter 4. After protein purification, all the eluted proteins were analysed with SEC.

Figure 3.8. Small-scale protein phosphatase protein expression (recombinant pOPINJ vector in Rosetta™ 2 (DE3) p*Lac*l *E. coli*). 1D SDS-PAGE gel show protein pellets (insoluble fractions), soluble fractions and purified proteins. Arrows point to the targeted protein band. M indicates protein marker in the left hand lane with molecular weight shown in kDa.

a) Protein pellet (insoluble fractions) of putative protein phosphatase proteins. Lane 1, Sync_2627; Lane 2, Sync_2828.

b) Soluble fractions of protein phosphatase proteins. Lane 1, Sync_2627; Lane 2, Sync_2828.

c) Purified protein phosphatase proteins. Lane 1, Sync_2627; Lane 2, Sync_2828. Arrows pointed to the targeted protein band. Both Sync_2627 and Sync_2828 proteins were purified.

d) Protein pellet (insoluble fractions) of the protein phosphatase proteins. Lane 1, Sync_0151; lane 2, Sync_0484, lane 3, Sync_1857; lane 4, Sync_1056 and lane 5, Sync_0012. Sync_0151, Sync_0484 and Sync_1056 showed successful protein expression with arrows indicating the targeted protein band.

e) Soluble fractions of the protein phosphatase proteins. Lane 1, Sync_0151, lane 2, Sync_0484; lane 3, Sync_1857, lane 4, Sync_1056 and lane 5, Sync_0012. None of the targeted proteins are present in soluble form.

f) Purified protein phosphatase proteins. Lane 1, Sync_0151; lane 2, Sync_0484; lane 3, Sync_1857; lane 4, Sync_1056 and lane 5, Sync_0012. None of the targeted proteins were purified.
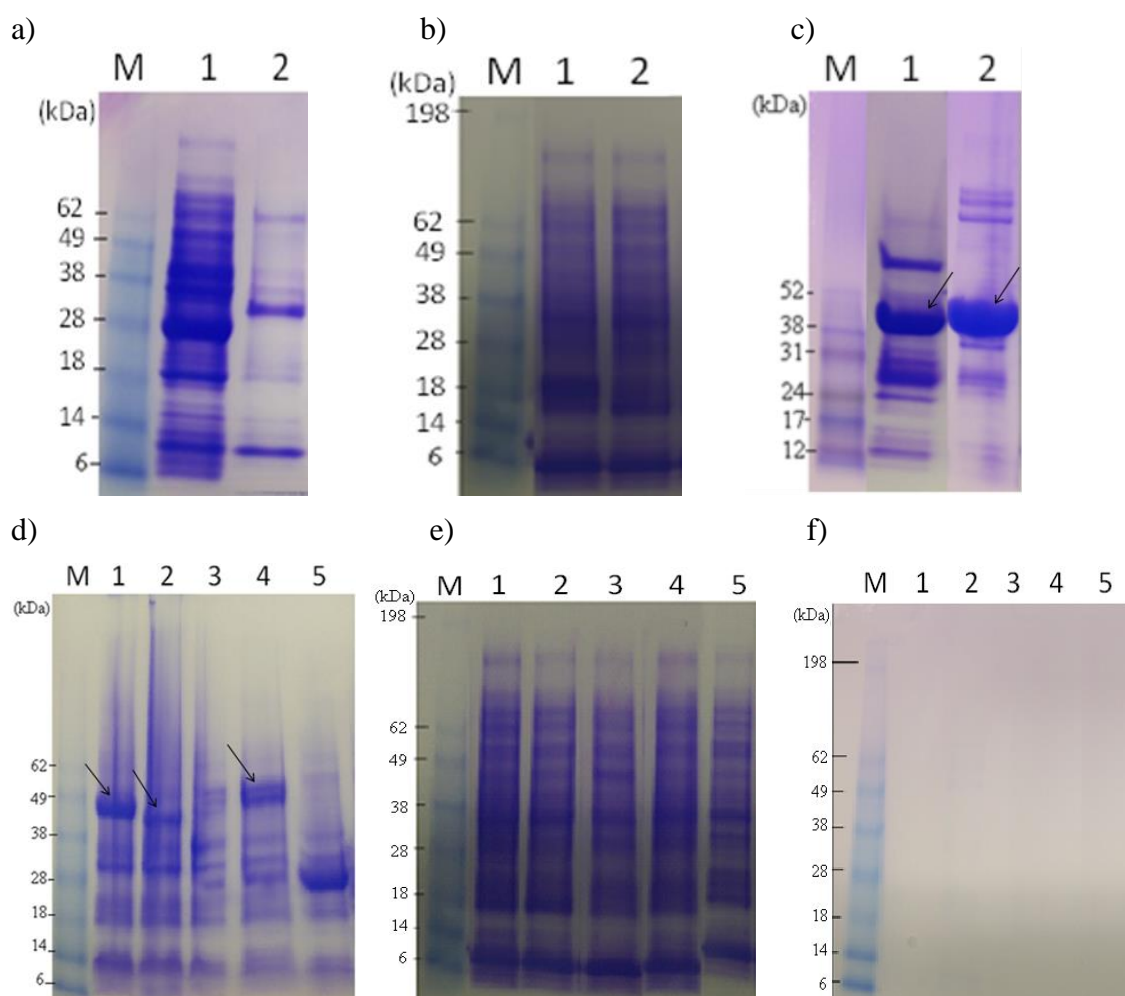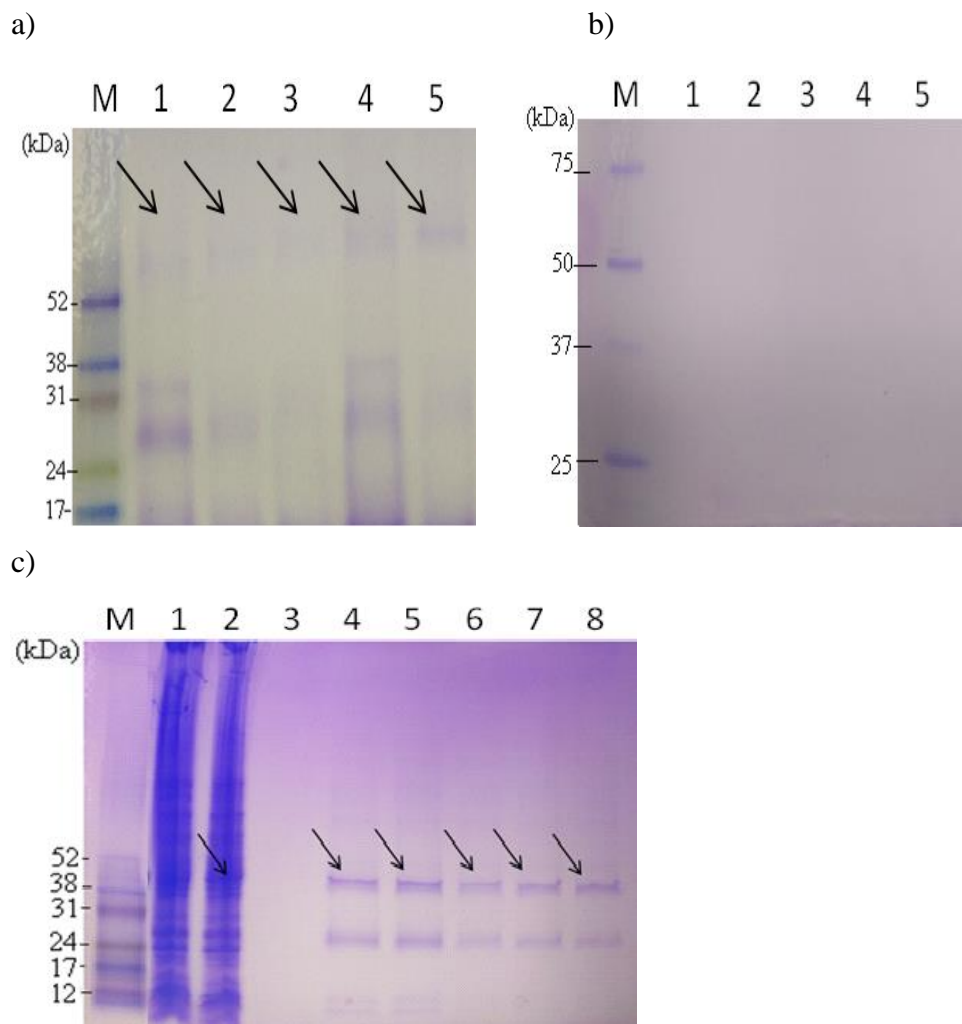

Figure 3.9. Small scale protein phosphatase protein expression (recombinant pOPINJ vector in Lemo21 (DE3) _E. coli_). 1D SDS-PAGE gel show protein expression of the

phosphatase proteins Sync_1857 and Sync_0484. Arrows indicate the target protein band. M indicates protein marker in the left hand lane with molecular weight shown in kDa.

a) Purified Sync_1857 protein following treatment with different concentrations of rhamnose added during bacterial culture: Lane 1-5 (2000µM, 1000µM, 500µM, 100µM and no rhamnose respectively). Purified protein was observed in all rhamnose concentrations tested.

b) Purified Sync_0484 protein following treatment with different concentrations of rhamnose added during bacterial culture. Lane 1-5 (2000µM, 1000µM, 500µM, 100µM and no rhamnose respectively). No purified protein was observed in the gel.

c) Protein expression of Sync_0484 using Buffer X for solubilization. Different concentrations of rhamnose were added during bacterial culture. Lane 1, soluble fraction of Sync_0484 protein, lane 2, total lysate of Sync_0484 protein. Lane 1 and 2 represented Figure 3.9b in which Buffer A was used for protein solubilization. Lane 3 is an empty lane. Lane 4-8 represent purified Sync_0484 protein with the following concentrations of rhamnose: 0µM, 100µM, 500µM, 1000µM and 2000µM, respectively. Purified protein was observed for all rhamnose concentrations tested.
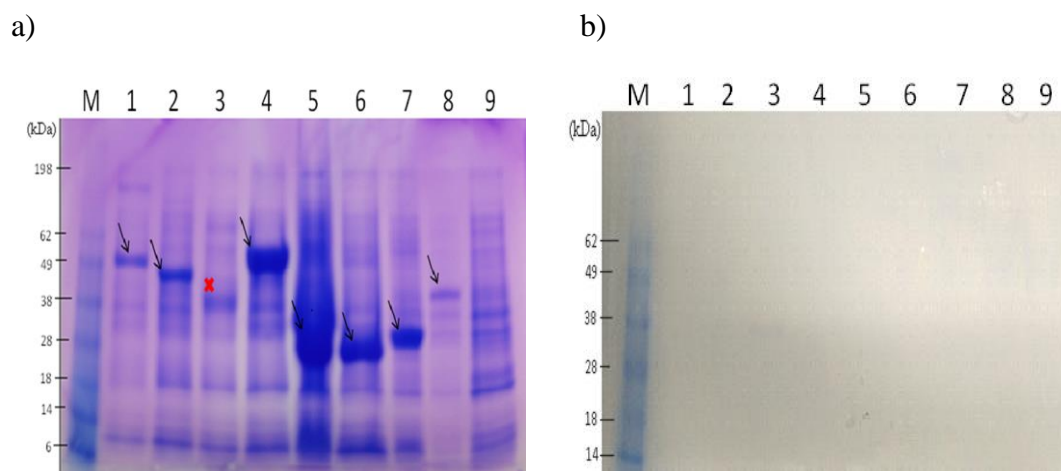
Figure 3.10. Small-scale protein phosphatase protein expression (recombinant pOPINM and pOPINS3c in Rosetta™ 2 (DE3) p*Lac*l *E. coli*. 1D SDS-PAGE gel show Sync_2505, Sync_2627 and Sync_2828 expressed from the pOPINM vector, Sync_0012, Sync_0151, Sync_0484, Sync_1056 and Sync_1857 and Sync_2505 expressed from the pOPINS3c vector. Arrows indicate the target protein band while the red cross indicates a band the size of the maltose binding protein alone, suggesting that the fused protein phosphatase has been cleaved off.

a) Total lysate from cell expressing the recombinant protein phosphatase proteins. Lane 1-3 represent proteins Sync_2505, Sync_2627, Sync_2828 expressed from the pOPINM vector, respectively. Lane 3-8 represent Sync_0012, Sync_0151, Sync_0484, Sync_1056 and Sync_1857 expressed from the pOPINS3c vector respectively.

b) Purified protein phosphatase proteins. Lane 1-3 represent proteins Sync_2505, Sync_2627, Sync_2828 expressed from the pOPINM vector, respectively. Lane 3-8 represent proteins Sync_0012, Sync_0151, Sync_0484, Sync_1056 and Sync_1857 expressed from the pOPINS3c vector, respectively. None of the recombinant proteins were purified.
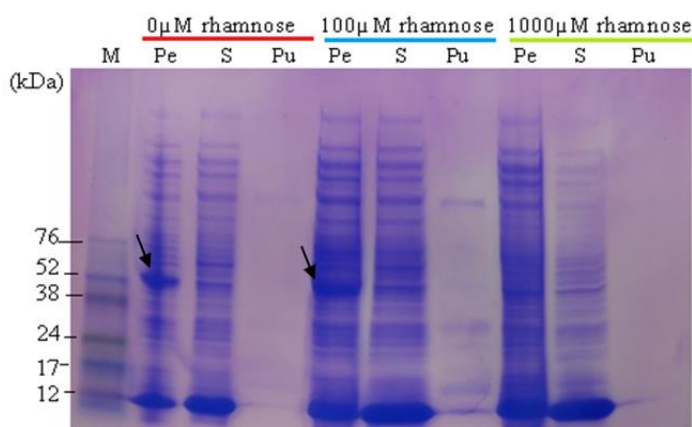
Figure 3.11. Small-scale protein phosphatase Sync_2505 protein expression (expressed from pOPINS3c in Lemo21 (DE3) *E. coli*. 1D SDS-PAGE gel shows Sync_2505 expression following treatment of cells with 3 different concentrations of rhamnose: 0μM, 100μM and 1000μM. Arrows indicate the protein band of Sync_2505. M, protein marker (molecular weight shown in the left lane in kDa); Pe, pellet (insoluble fraction); S, soluble fraction; Pu, purified protein. Protein expression of Sync_2505 was detected, but in insoluble form when cultured in conditions without rhamnose and with 100μM rhamnose.

3.3.5 Size Exclusion Chromatography (SEC)

Size exclusion chromatography analyses were conducted on the purified protein phosphatases Sync_2627, Sync_2828 and Sync_1857 to examine their oligomeric state(s). A HiLoad™ 16/600 Superdex™ 200pg preparatory column with a bed volume of 120ml was used for separation of the proteins based on their molecular weight. A Superdex™ 200 10/300 GL analytical column with a bed volume of 24ml was used for further protein separation. Purified Sync_2627 preparatory SEC shows a broad peak spectrum (Figure 3.12). The various oligomer forms could be due to the 8 cysteine residues in Sync_2627 protein that may form disulfide bridges. Due to the low yield of Sync_2627 making the protein oligomers separation difficult, the Sync_2627 oligomers were not further analysed by analytical SEC column and pull-down assays. SEC analysis was conducted using two running buffers, with and without 0.5mM TCEP, a reducing agent to reduce disulphide bonds. SEC analysis of Sync_2828 showed two peaks representing dimeric (estimated molecular weight of 88kDa) and hexamer forms (estimated molecular weight of 264 kDa) (Figure 3.13, refer Supplementary Info 3.8 for the standard curve of analytical SEC). The

results remained the same with or without TCEP addition. Meanwhile, SEC analysis of Sync_1857 showed four peaks, representing hexameric, tetrameric, dimeric and monomeric forms, with estimated molecular weight of 390kDa, 260kDa, 130kDa, 65kDa respectively, when using the running buffer without addition of TCEP (Figure 3.14). When TCEP was used, there were only two peaks detected, which represented its dimeric and monomeric forms (Figure 3.14). All oligomeric protein forms were collected, stored in -80^0C until ready for the protein-protein interaction study.

The results of size exclusion chromatography of all nine response regulators showed that all of these response regulators were present in monomeric forms (Figure 3.15). Response regulator Sync_2591 showed two peaks in the SEC analysis, representing its dimeric and monomeric forms (Figure 3.15h). The isolated proteins from the chromatography peaks were collected for all response regulators and used for the pull-down assay.
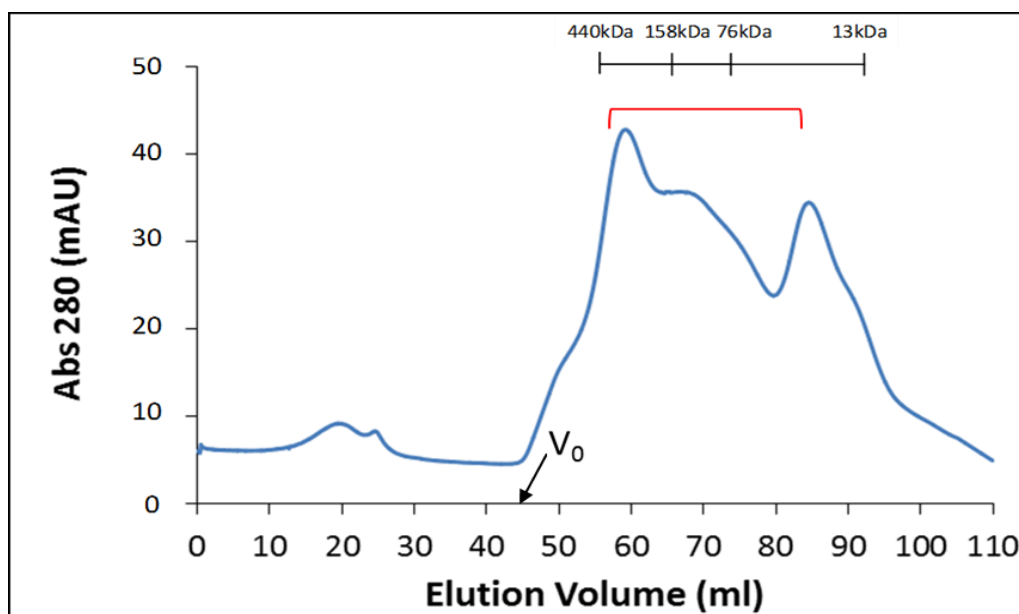
Figure 3.12. SEC of the putative protein phosphatase Sync_2627. The preparatory SEC trace is shown for Sync_2627. The calculated molecular weight according to the amino acid sequence of the recombinant protein Sync_2627 is 45kDa. Corresponding molecular weight standards indicated (horizontal bar). V_0 represents void volume. The red bracket indicates the range of possible oligomerization forms of Sync_2627.
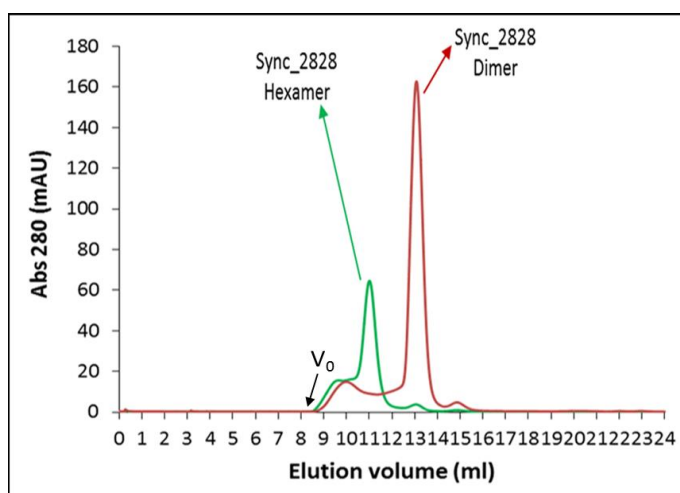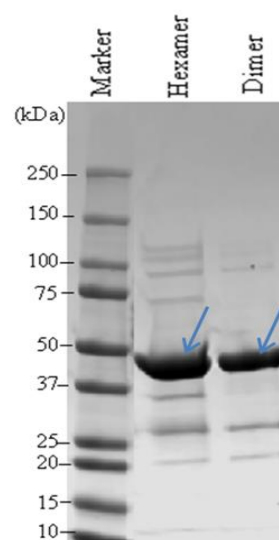
Figure 3.13. SEC and 1D SDS-PAGE gel of the purified putative protein phosphatase Sync_2828.

a) Analytical SEC in running buffer with TCEP shows two peaks representing the hexameric and dimeric forms of Sync_2828. Running buffer without TCEP gave an identical result for the SEC analysis of Sync_2828 (data not shown). V_0 represents void volume.

b) 1D SDS-PAGE gel of the purified Sync_2828 hexamer and dimer. Arrows pointed to targeted protein. Marker: protein markers with their molecular weight indicated in kDa are shown on the left hand lane; hexamer indicates protein Sync_2828 in hexameric form; dimer indicates protein Sync_2828 in dimeric form.
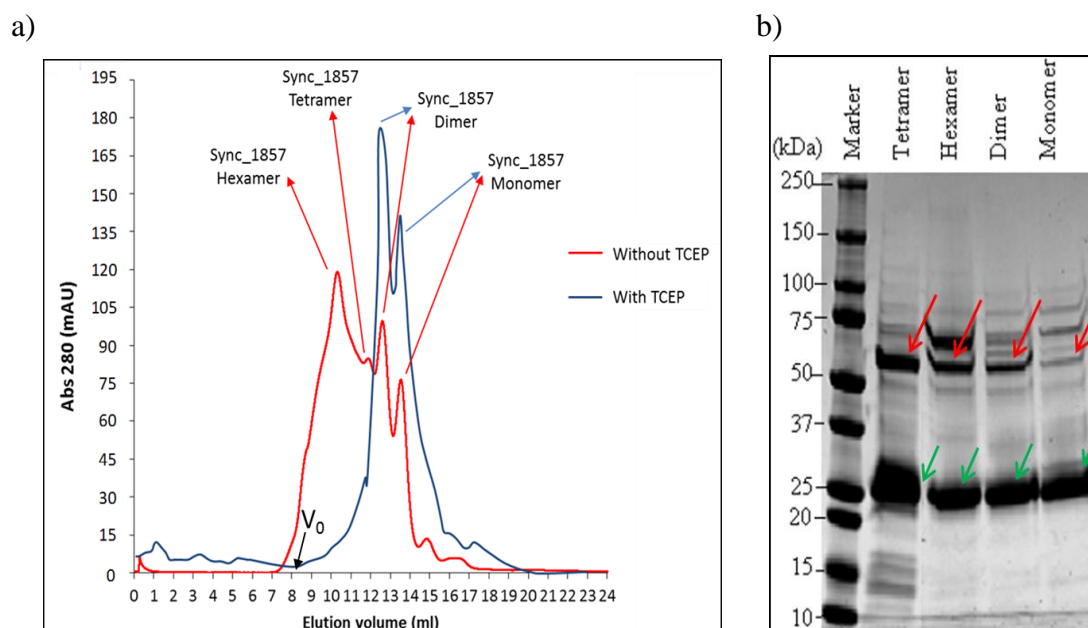
Figure 3.14. SEC and 1D SDS-PAGE gel of the purified putative protein phosphatase Sync_1857.

a) Analytical SEC of Sync_1857 running in buffer with TCEP (blue) and without TCEP (red). V_0 represents void volume. Monomeric, dimeric, tetrameric and hexameric forms are indicated with arrows.

b) 1D SDS-PAGE gel of the monomeric, dimeric, tetramer and hexamer forms of Sync_1857. Red arrows pointed to the Sync_1857 protein band, green arrows pointed to the GST protein (GST-tagged recombinant proteins are prone to cleavage when boiling in SDS). Marker: protein markers with their molecular weight indicated in kDa are shown in the left hand lane; Tetramer, hexamer, dimer and monomer indicates oligomeric state of protein Sync_1857.
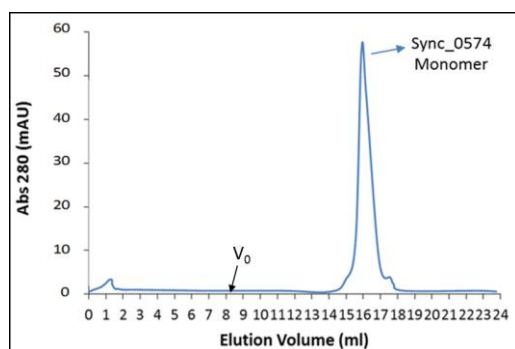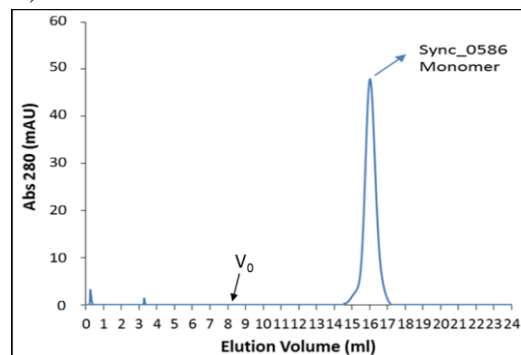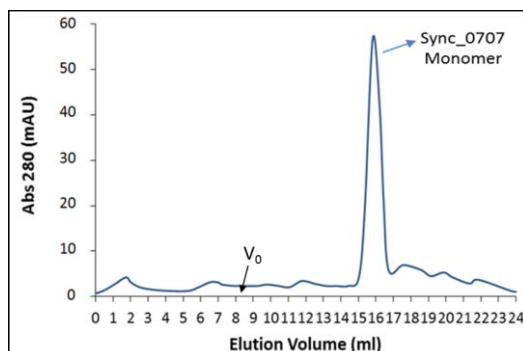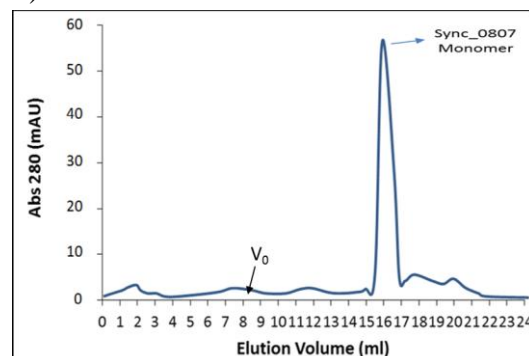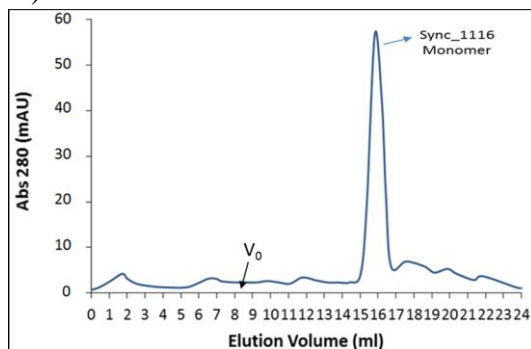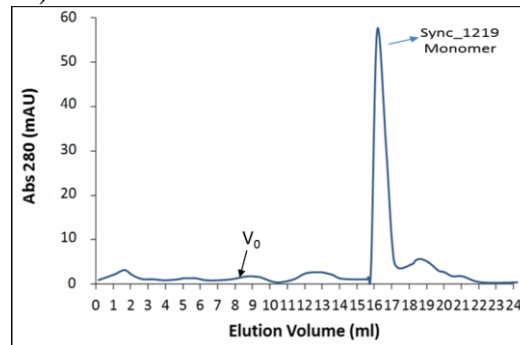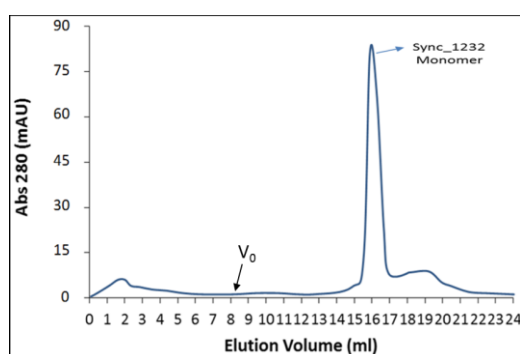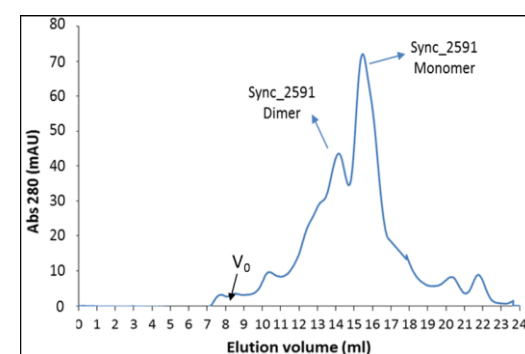
i)
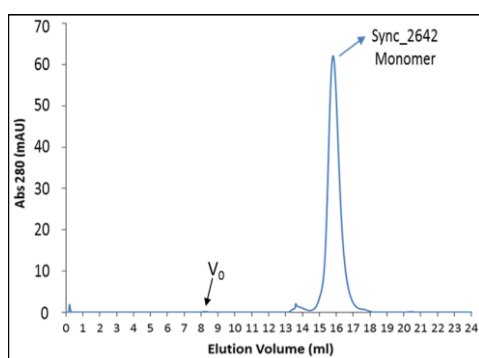
j)
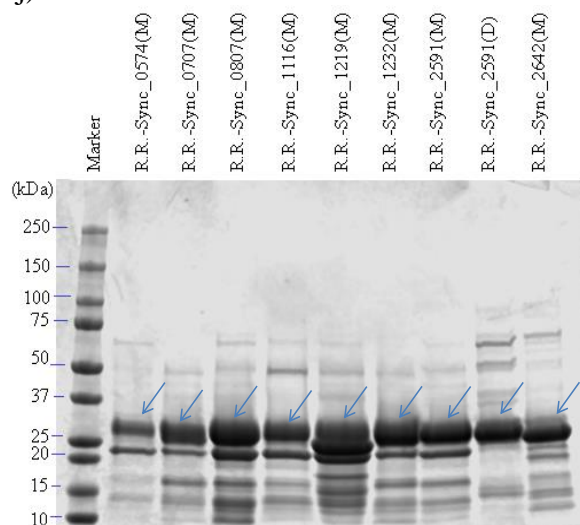
k)
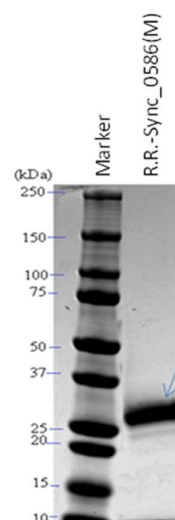
Figure 3.15. SEC and 1D SDS-PAGE gels of the purified response regulator proteins.
Figure 3.15a-i show the analytical size exclusion chromatography of the response regulator proteins Sync_0574, Sync_0707, Sync_0807, Sync_1116, Sync_1219, Sync_1232, Sync_2591, Sync_2642 and Sync_0586 respectively. V_0 represents void volume. Figure 3.15j-k show the 1D SDS-PAGE gels of purified response regulators from a-i. Arrows pointed to the targeted protein. Marker: protein markers with their molecular weight indicated in kDa are shown in the left hand lane. Abbreviations: R.R., response regulator; M, monomer; D, dimer.

3.3.6 Protein-protein interactions between the putative protein phosphatases and response regulators.

A pull-down assay was conducted as described in section 3.2.5.1. GST-tagged phosphatase was used as a bait to capture a response regulator (prey) *in vitro*. This assay was carried out on a one-to-one basis, one protein phosphatase mixed with one response regulator. Native gel analysis was conducted to visualize the protein-protein interactions between the putative protein phosphatase and response regulator. The running buffer for the native gel was represented in Table 3.9. All the conditions of the running buffer remained the same for phosphatase Sync_1857 and Sync_1857 except the running buffer pH was adjusted to 9 for phosphatase Sync_1857 as its isoelectric point (pI) was 8.32 close to the pH 8.3 (the common pH for native gel running buffer). The native gel electrophoresis results are represented in Supplementary Info 3.9. After native gel analysis, the remaining eluted proteins from the pull-down assay were further been processed with in-solution digestion with trypsin before mass spectrometric analysis. Mass spectrometry analysis showed that response regulators Sync_0574 and Sync_0586 did not interact with any of the protein phosphatases tested (Table 3.14). Sync_2828 has a smaller range of oligomerization states, and was thus chosen to test whether these variant forms of oligomerization state showed different PPIs properties. Due to time limitations, only the monomeric form of Sync_1857 was tested. All tested response regulators showed the same results with dimer and hexamer forms of phosphatases except Sync_0707, Sync_0807 and Sync_2591_M. Response regulator Sync_0807 and Sync_2591_M showed positive results only with the dimeric form of protein phosphatase Sync_2828, while the response regulator Sync_0707 showed positive results solely with the hexameric form of the protein phosphatase Syn_2828.

Table 3.14. Bait-prey interactions of phosphatases and response regulators

Phosphatase / Response Regulator	Sync_2828		Sync_1857
	Dimer form	Hexamer form	Monomer form
Sync_0574_M	x	x	x
Sync_0586_M	x	x	x
Sync_0707_M	x	✓	✓
Sync_0807_M	✓	x	✓
Sync_1116_M	✓	✓	✓
Sync_1219_M	✓	✓	✓
Sync_1232_M	✓	✓	✓
Sync_2591_M	✓	x	✓
Sync_2591_D	✓	✓	✓
Sync_2642_M	✓	✓	✓

(Notes: ✓ indicates the response regulator protein was detected in the mass spectrometry analysis; x indicates the response regulator protein was not detected in the mass spectrometry analysis. Abbreviation: M, monomer; D, dimer.)

3.4 DISCUSSION

The pOPIN suite vectors were chosen as the cloning vectors for this study as it contains a variety of vector products with different specificity. The design of the vectors facilitates users to easily change vectors for In-Fusion™ cloning without having to redesign the primers. Therefore, it saves the users time to test vectors that suit to their cloning work and expression of proteins. Previous studies have also demonstrated the successful application of In-Fusion™ cloning for high throughput construction and protein expression screening of pOPIN suite vectors in *E.coli* (Bird 2011; Bird et al. 2013). Therfore, pOPINF, pOPINJ, pOPINM and pOPINS3c vectors were used for the cloning work of phosphatases and response regulators in this study. pOPINF which has a His-tag was used for the protein response regulators, as these proved to be more easily purified. Meanwhile, pOPINJ, pOPINM, pOPINS3c which contain His-tag plus additional tags: GST-tag, MBP-tag and SUMO-tag respectively were used for the protein phosphatase proteins, as these proved to be much more difficult to purify in soluble form. Additionally, a second binding tag was

needed for the recombinant protein phosphatase proteins in order to conduct the pull-down assay study in this chapter and chapter 4.

Two different strains of *E. coli* expression host were utilized for this study. Rosetta™ 2 (DE3) p*Lac*l *E. coli* strain was initially utilized as expression host for the recombinant protein phosphatases. However, the solubility of proteins was low with this expression system hence, Lemo21 (DE3) *E. coli* strain was also used, as this strain has been successfully used to solubilize some recalcitrant proteins in previous studies (Wahlgren et al., 2018, Leclercq et al., 2017, Kuhn et al., 2015, Pathak et al., 2016). Recombinant protein phosphatase proteins transformed in Lemo21 (DE3) such as Sync_1857 and Sync_0484 showed enhanced protein solubility, while Sync_1857 was successfully solubilized in Buffer A and Sync_0484 was only solubilized in Buffer X.

All the purified proteins either phosphatases or response regulators contained His-tag fused to the target protein. Therefore, a Ni-NTA resin was used for affinity purification of all of the recombinant proteins. A high concentration of imidazole was used to elute the recombinant proteins that bound to the Ni-NTA resin. Therefore, SEC chromatography assay was conducted to eliminate imidazole which might interfere with the PPIs assay and to analyse the oligomerization state of the recombinant proteins by separation of the proteins based on their molecular weight. The purified protein phosphatases contain cysteine amino acid(s), hence, it was possible these recombinant proteins can form disulfide bridge, altering the homooligomerization state of the proteins. Owing to this, we used two different buffers for the SEC assay of the protein phosphatase proteins: one with the reducing agent TCEP, while the other with no TCEP. Sync_2828 protein with one cysteine exists in dimer and hexamer forms and showed no difference with both conditions tested. Phosphatase protein Sync_1857 which has 4 cysteines, formed four oligomer states (monomer, dimer, tetramer and hexamer) in buffer without TCEP. Meanwhile, it existed in only monomer and dimer forms when it was in buffer with TCEP. The SEC chromatography shows that Sync_1857 can form disulfide bridge(s) *in vitro* and it seems likely that it forms disulfide bridge(s) *in vivo* as well, as it has a predicted signal peptide, and is likely localized to the oxidizing environment in the periplasm of the cell. In the case of Sync_2627, it was examined only by preparatory SEC using running buffer without TCEP. Due to the low yield of the protein preparation along with the added difficulty of multiple oligomer forms, a decision was made to not proceed with the analytical SEC and

pull-down assay for Sync_2627. Future work could involve a larger scale preparation in fermenter cultures of Sync_2627 to facilitate further work on the protein oligomers.

A pull-down assay was conducted for PPIs study of the *Synechococcus* protein phosphatases and response regulators. Principally, there are three kinds of possible results from the interactions: first, no binding, indicating that there is no interaction between the protein phosphatase and the response regulator. Second, both proteins show evidence of an interaction, which could be either a strong interaction or transient interaction. Forty micrograms of protein for each of the protein phosphatase and the response regulator were used for one pull-down assay. The interactions between phosphatase and response regulator were not clearly been observed in the native gel possibly due to several factors. First, strict conditions were applied for this study for the pull-down assay to avoid false-positive results. Cautious handling in the washing and elution steps for the pull-down assay might reduce the yield of the proteins for visualization in a native gel. Second, the interactions between the protein phosphatase and response regulator could be transient or not stable (possibly resulting in them becoming unbound while the native gel electrophoresis was conducted). However, the protein-protein interactions between the protein phosphatase and the response regulator were detected with mass spectrometry analysis. The mass spectrometry analysis showed the detection of the protein phosphatase with its interacting partner of the response regulator (if they interact) or solely phosphatase (if no interaction). There was no issue of contaminants from the mass spectrometry analysis. Therefore, the protein phosphatases were used for the PPIs study in next chapter as well. The reliability of pull-down assays can be enhanced by bi-directional study using protein A as bait and protein B as prey and vice versa. Therefore, it is suggested that future work of the pull-down assay could encompass using the response regulator as a bait and the protein phosphatase as a prey to enhance validation of the results. Moreover, other oligomeric forms of protein phosphatase of Sync_1857 including dimer, tetramer and hexamer could to be tested in the future to check whether the PPIs results remain the same for all oligomer forms.

3.5 CONCLUSIONS

Nine protein phosphatase and seventeen response regulator genes were bioinformatically found in the genome of marine cyanobacteria *Synechococcus* sp. CC9311. All of the protein phosphatase and response regulator genes were successfully amplified except *sync_1079*. Cloning and transformation were conducted for all the amplified products. One of the cloned protein phosphatase genes, namely *sync_1732* was found to contain a frameshift mutation and was discontinued from further analysis. Four of the protein phosphatases, namely Sync_0484, Sync_1857, Sync_2627 and Sync_2828 were expressed as soluble proteins. Three of them (excluding Sync_0484 which was only solubilized in Buffer X) were produced in large-scale bacterial culture, purified and analysed with SEC. Protein phosphatase Sync_2627 which contains 8 cysteine amino acids, demonstrating a broad range of peaks in the preparatory SEC column analysis. Syc_2627 was not proceeded further with the analytical SEC column analysis and pull-down assay owing to low protein yield and the possible complex isolation of the variety of protein oligomers. Sync_2828 showed two oligomer forms (dimer and hexamer) with or without TCEP in running buffer in the analytical SEC and Sync_1857 showed four oligomer forms (monomer, dimer, tetramer and hexamer) in running buffer without TCEP and two oligomer forms (monomer and dimer) in SEC analysis using running buffer with TCEP, suggesting a role of disulfide bonding in oligomer formation for Sync_1857.

Nine out of the seventeen response regulators, namely Sync_0574, Sync_0586, Sync_0707, Sync_0807, Sync_1116, Sync_1219, Sync_1232, Sync_2591 and Sync_2642 were tested for protein expression and all nine were found to be expressed in a soluble form, suggesting that the response regulator proteins are more easily heterologously expressed than the protein phosphatases. The nine response regulator proteins were purified and analysed with preparatory and analytical SEC. All of them exist as monomers except Sync_2591 which was present in dimer and monomer forms.

PPI analyses between the protein phosphatases and response regulators were conducted using a pull-down assay coupled with mass spectrometry analysis. Two of the protein phosphatases, namely Sync_1857 and Sync_2828 were used as baits for an *in vitro* pull-down assay against prey response regulator proteins. Both the dimer and hexamer forms of Sync_2828 were used for the *in vitro* pull-down assay to test whether different oligomer forms showed different PPIs with the response regulators, while only the monomer of

Sync_1857 was used as a bait for the pull-down assay. Thirty pull-down assays in total were conducted in this study; testing each phosphatase (Sync_1857 and the two oligomeric forms of Sync_2828) react with 10 response regulators (8 + 2 oligomeric forms of Sync_2591). The results from mass spectrometric analyses demonstrated that response regulators Sync_0574 and Sync_0586 did not interact with any of the protein phosphatases tested. Meanwhile, response regulators Sync_0707, Sync_0807 and Sync_2591 (monomer) showed different PPI results with the dimer and hexamer forms of protein phosphatase Sync_2828. Sync_0807 and Sync_2591 (monomer) bound to Sync_2828 in dimer form but not hexamer, and Sync_0707 only bound with the hexameric form of Sync_2828. The biological significance of the differential binding by the different oligomeric forms of Sync_2828 remains to be explored in future work. The findings of this study would be strengthened by future work with pull-down assays using the response regulator proteins as bait and the protein phosphatase as prey to enhance validation of the observed PPIs between the protein phosphatases and the response regulators.

Supplementary Info 3.1. Plasmid maps

pOPINM (6665 bp)

(6415) **XmnI**
(6366 .. 6385) Amp-R
(6296) **ScaI**
(5964) **NmeAIII**
(5886) **BpmI**
(5222 .. 5241) pBR322ori-F
(5189) **DrdI**
(5082) **PciI**
(4938) **AsiSI**
(4482) **SwaI**
(4226) **SphI**
(4079 .. 4098) rbglobpA-R
(3992) **BspHI**
(3960 .. 3979) Bglob-pA-R
(3956) BspDI * - ClaI *
(3924) **MscI**
(3883) **Bsu36I**
(3873) **DraIII**
(3849) **PmeI**
(3831) **HindIII**
(3829) **PluTI**
(3827) **SfoI**
(3826) **NarI**
(3825) **KasI**
(3756) **PvuII**
(3679 .. 3701) M13/pUC Forward
(3670 .. 3687) M13 Forward
(3638 .. 3654) M13 Reverse
(3619 .. 3641) M13/pUC Reverse
(3498) **KpnI**
(3494) **Acc65I**
(3440 .. 3463) MBP-F
BclI * (3408)
(3152) **BlpI**
BsaBI (3041)
BsmI (2986)
BglII (2725)
BsiWI (2656)
BspEI (2485)
NcoI (2331)
PacI (2280)
PpuMI (2208)
T7 (2150 .. 2169)
RsrII (2140)
pCAG-F (2089 .. 2108)
pCAGGS-5 (2005 .. 2027)
SacII (1965)
BmeT110I (1952)
Aval - **BsoBI** (1951)
Apal (1931)
PspOMI (1927)
EagI (1015)
FseI (1017)
MluI (399)
BbvCI - **Bpu10I** (416)

AmpR
baculovirus recombination region (lef2/ORF603)
lef2
ORF603
CMV...
chicken β-actin promoter
ori
loxP
ORF1629
β-globin poly(A) signal
baculovirus reco...
M13 rev
6xHis
lacZα
CAP binding site
T7 promoter
p10 promoter
MBP

pOPINS3C (5824 bp)

(5574) **XmnI**
(5525 .. 5544) Amp-R
(5455) **ScaI**
(5123) **NmeAIII**
(5045) **BpmI**
(4548) **PspFI**
(4544) **BseYI**
(4381 .. 4400) pBR322ori-F
(4348) **DrdI**
(4241) **PciI**
(4097) **AsiSI**
(4055) **BmgBI**
(3641) **SwaI**
(3385) **SphI**
(3312) **BstAPI**
(3238 .. 3257) rbglobpA-R
(3151) **BspHI**
(3119 .. 3138) Bglob-pA-R
(3115) BspDI * - ClaI *
(3083) **MscI**
(3042) **Bsu36I**
(3014) **HindIII**
(3012) **PluTI**
(3010) **SfoI**
(3009) **NarI**
(3008) **KasI**
(2939) **PvuII**
(2862 .. 2884) M13/pUC Forward
M13 Forward (2853 .. 2870)
M13 Reverse (2821 .. 2837)
M13/pUC Reverse (2802 .. 2824)
KpnI (2676)
Acc65I (2672)
NruI (2634)
NcoI (2331)
PacI (2280)
PpuMI (2208)
T7 (2150 .. 2169)
RsrII (2140)
pCAG-F (2089 .. 2108)
pCAGGS-5 (2005 .. 2027)
SacII (1965)
BmeT110I (1952)
Aval - **BsoBI** (1951)
Apal (1931)
PspOMI (1927)
EagI (1015)
FseI (1017)
MluI (399)
BbvCI - **Bpu10I** (416)

AmpR
baculovirus recombination region (lef2/ORF603)
lef2
ORF603
CMV enhancer
chicken β-actin promoter
ori
loxP
ORF1629
polyhedrin 3' UTR
baculovirus recombination region (ORF1629)
β-globin poly(A) signal
M13 fwd
M13 rev
6xHis
p10 promoter
lacZα
lac promoter

170

Supplementary Info 3.2. PCR amplification primers (phosphatases)

Gene	Primer (5'→3')[*]	Length	Tm (^{0}C)	% GC content
sync_0012_F	AAGTTCTGTTTCAGGGCCCGgtgagcagcacaccttc	37	70	57
sync_0012_R	TCTAGAAAGCTTTAtcaggccaaagacctgggtacag	37	66	46
sync_0151_F	GTTCTGTTTCAGGGCCCGatggctcgacaactg	33	68	58
sync_0151_R	CACAAACTGGTCTAGAAAGCTTTAttagctggcccc	36	66	47
sync_0484_F	GTTTCAGGGCCCGatgacaacaaaggtgctgtttg	35	67	51
sync_0484_R	CTGGTCTAGAAAGCTTTAttagcgcttagtggctggagc	39	68	49
sync_1056_F	AAGTTCTGTTTCAGGGCCCGatgccaccgcgctgcaac	38	72	61
sync_1056_R	CTGGTCTAGAAAGCTTTAttagcaaggcactgcctgtagcaatgg	45	69	47
sync_1732_F	GAAGTTCTGTTTCAGGGCCCGgtgataattgcaatc	36	66	47
sync_1732_R	CACAAACTGGTCTAGAAAGCTTTAtcaccggaccagcc	38	68	50
sync_1857_F	AAGTTCTGTTTCAGGGCCCGatgcaaacaagtcg	34	66	50
sync_1857_R	CACAAACTGGTCTAGAAAGCTTTAttaagagccataggac	40	64	40

sync_2505_F	AAGTTCTGTTTCAGGGCCCGatgatgtctatatctccatcg	41	68	46
sync_2505_R	GTCTAGAAAGCTTTAtcaccagaggccgcgaacgg	35	68	54
sync_2627_F	AAGTTCTGTTTCAGGGCCCGgtgccaaccgcttacg	36	70	58
sync_2627_R	CTGGTCTAGAAAGCTTTAttaaacgcagcccggcaac	37	67	49
sync_2828_F	AAGTTCTGTTTCAGGGCCCGttgacctccggcagg	35	70	60
sync_2828_R	CTGGTCTAGAAAGCTTTAtcaggattcagcacgtttccaaggg	43	68	47

(* Uppercase letter represents extension primer, while lowercase letter represents template DNA primer)

Supplementary Info 3.3. PCR amplification primers (response regulators)

Gene	Primer (5'→3')*	Length	Tm (⁰C)	% GC content
sync_0115_F	AAGTTCTGTTTCAGGGCCCGatgcccacagatggg	35	69	57
sync_0115_R	CTGGTCTAGAAAGCTTTAtcaaactgctgcaggtgctc	38	67	47
sync_0265_F	AAGTTCTGTTTCAGGGCCCGttgctttcctgtg	33	66	52
sync_0265_R	CACAAACTGGTCTAGAAAGCTTTActatgcgcttacgaaacag	43	66	42
sync_0574_F	TTCTGTTTCAGGGCCCGatgaactgcctcgtg	32	67	56
sync_0574_R	CTGGTCTAGAAAGCTTTAtcattgatcgatgctttgaatggtg	43	65	40
sync_0586_F	AAGTTCTGTTTCAGGGCCCGatggtgcataaatatcaaatcg	42	66	43
sync_0586_R	CTGGTCTAGAAAGCTTTAttaaatatgtgcttcaacc	37	61	35
sync_0669_F	AAGTTCTGTTTCAGGGCCCGgtgagcaagcacgatc	36	69	56
sync_0669_R	TGGTCTAGAAAGCTTTAtcaggcgctcgttttctccag	38	67	47
sync_0707_F	TTCTGTTTCAGGGCCCGatgcctggaactacttcaacg	38	69	53
sync_0707_R	CTGGTCTAGAAAGCTTTAttattggtcgttcatgcgaagtg	41	66	41
sync_0807_F	TGGAAGTTCTGTTTCAGGGCCCGatgaatgagcaggg	37	69	54

sync_0807_R	CACAAACTGGTCTAGAAAGCTTTAtcaagagctcttc agg	40	66	43
sync_1079_F	AAGTTCTGTTTCAGGGCCCGatgactattgatcgtcatc	39	67	46
sync_1079_R	CTGGTCTAGAAAGCTTTAtcaatgtctcaaaacgttagc caatg	44	65	39
sync_1116_F	AAGTTCTGTTTCAGGGCCCGgtgcggattaacc	33	67	55
sync_1116_R	CACAAACTGGTCTAGAAAGCTTTAtcaaattcctgctt caatcaagc	47	66	38
sync_1134_F	CTGGAAGTTCTGTTTCAGGGCCCGatgaaatacgtag	37	67	49
sync_1134_R	CACAAACTGGTCTAGAAAGCTTTAttagagatatcgct caagc	43	65	40
sync_1145_F	AAGTTCTGTTTCAGGGCCCGgtgacctcctcatcc	35	69	57
sync_1145_R	CTGGTCTAGAAAGCTTTAttagtccctgatgttttcaagg ctg	43	65	40
sync_1219_F	CTGTTTCAGGGCCCGatgcttaattcagatcggcg	35	68	54
sync_1219_R	CTGGTCTAGAAAGCTTTActacttcaagatatagccgac acc	42	66	43
sync_1232_F	AAGTTCTGTTTCAGGGCCCGatggcgatgacccagc	36	70	58
sync_1232_R	CTGGTCTAGAAAGCTTTAtcaacgcgaacgaggttgtcc atc	42	68	48

sync_2111_F	AAGTTCTGTTTCAGGGCCCGatggaccagctgcc	34	69	59
sync_2111_R	CAAACTGGTCTAGAAAGCTTTAttatggggttgtggaagtac	42	65	40
sync_2591_F	AAGTTCTGTTTCAGGGCCCGatgaaaccttccatcc	36	67	50
sync_2591_R	CTGGTCTAGAAAGCTTTAtcaggccgtggcccgatttg	38	69	53
sync_2597_F	AAGTTCTGTTTCAGGGCCCGatgacggccacagccgc	37	72	62
sync_2597_R	CTGGTCTAGAAAGCTTTAtcaggatccttcgggtgcaacgg	41	70	51
sync_2642_F	AAGTTCTGTTTCAGGGCCCGatgacggagactcccag	37	70	57
sync_2642_R	CTGGTCTAGAAAGCTTTAtcaattcacaaggtgatgttggagg	43	66	42

([*] Uppercase letter represents extension primer, while lowercase letter represents template DNA primer)

Supplementary Info 3.4: PCR amplification primers (response regulators (targeting response regulator domain))

Gene	Primer (5'→3')*	Length	Tm (^{0}C)	% GC content
sync_0115_R2	CTGGTCTAGGAAAGCTTTAtctcaggatcgtgaaaatgcg	40	67	45
sync_0574_R2	CTGGTCTAGGAAAGCTTTAcaaagcatcaaaagcatcgg	39	66	44
sync_0574_R3	CTAGGAAAGCTTTAgatcacattgatgcaatgtcgc	36	63	42
sync_0586_R2	CTGGTCTAGGAAAGCTTTAattgttggtcgttcttaaggc	40	66	43
sync_0669_R2	GGTCTAGAAAGCTTTActccccatggtggccaatcag	37	68	51
sync_0707_R2	CTGGTCTAGGAAAGCTTTActtcttgcgcagatatcgc	38	67	47
sync_1079_F2	AAGTTCTGTTTCAGGGCCCGatgcgtgtttctttgc	36	67	50
sync_1116_R2	CTGGTCTAGGAAAGCTTTAattggtagtacttaaggc	37	63	41
sync_1145_R2	CTGGTCTAGGAAAGCTTTAgcgtttaagcaaggtgcgg	38	68	50
sync_1219_R2	CTGGTCTAGGAAAGCTTTActgaacacgtgccaaatagtgatc	43	67	44
sync_1232_R2	CTGGTCTAGGAAAGCTTTAtctcagcactgcttcgatgcg	40	69	50
sync_2591_R2	CTGGTCTAGGAAAGCTTTAacgcctaagcagcg	33	66	52
Sync_2597_R2	CTGGTCTAGGAAAGCTTTAgcgcagaacgcagcgaattc	39	69	51
sync_2642_R2	CTGGTCTAGGAAAGCTTTAgcgcagaacgcagcgaattc	32	67	56

(* Uppercase letter represents extension primer, while lowercase letter represents template DNA primer)

Supplementary Info 3.5 Protein preparation Buffer X

Buffer	Composition
Lysis Buffer X	50mM Tris-HCl (pH8.0), 1% Triton X-100, 10% glycerol, 300mM NaCl, 0.1% sodium deoxycholate, 20mM imidazole. Lysozyme (1mg/ml) and 400 Kunits/ml of DNAse I is added when mix with cell pellet.
Binding Buffer X	50mM Tris-HCl (pH8.0), 1% Triton X-100, 10% glycerol, 300mM NaCl, 0.1% sodium deoxycholate, 20mM imidazole.
Wash Buffer X	50mM Tris-HCl (pH8.0), 1% Triton X-100, 10% glycerol, 300mM NaCl, 0.1% sodium deoxycholate, 50mM imidazole.
Elution Buffer X	50mM Tris-HCl (pH8.0), 1% Triton X-100, 10% glycerol, 300mM NaCl, 0.1% sodium deoxycholate, 400mM imidazole.

Supplementary Info 3.6. ZYP-5052 auto-induction media for 500ml culture

| Components	Volume (ml)
MiliQ Water	424
1000X trace metal mixture[a]	0.5
1M MgSO$_4$	0.5
10x TY[a]	50
GGL[a]	10
20x NPS[a]	25
Ampicillin (100mg/ml)	0.25
Chloramphenicol (50mg/ml)	0.25

[a] See Supplementary Info 3.7 for details.

Supplementary Info 3.7. Stock media preparation for ZYP-5052 autoinduction media

Components	Volume
1000x trace metal mixture (100 ml)	(in ml)
0.1M $FeCl_3 \cdot 6H_2O$	50
1M $CaCl_2$	2
1M $MnCl_2 \cdot 4H_2O$	1
1M $ZnSO_4 \cdot 7H_2O$	1
0.2M $CoCl_2 \cdot 6H_2O$	1
0.1M $CuCl_2 \cdot 2H_2O$	2
0.2M $NiCl_2 \cdot 6H_2O$	1
0.1M $Na_2MoO_4 \cdot 5H_2O$	2
0.1M $Na_2SeO_3 \cdot 5H_2O$	2
0.1M H_3BO_3	2
MiliQ water	sum up to 100
10x TY (500 ml)	(in g)
Tryptone	50
Yeast extract	25
GGL (500 ml)	(in g)
Glycerol	125
Glucose	12.5
α-lactose	50
20x NPS	(in g)
$(NH_4)_2SO_4$	33
KH_2PO_4	68
Na_2HPO_4	71

(Notes: All components were autoclaved except for the antibiotics and trace metal mixture solution before they were used for cell culture. The antibiotics and trace metals mixture were filter sterilized before use.)

Supplementary Info 3.8. Standard curve of analytical SEC
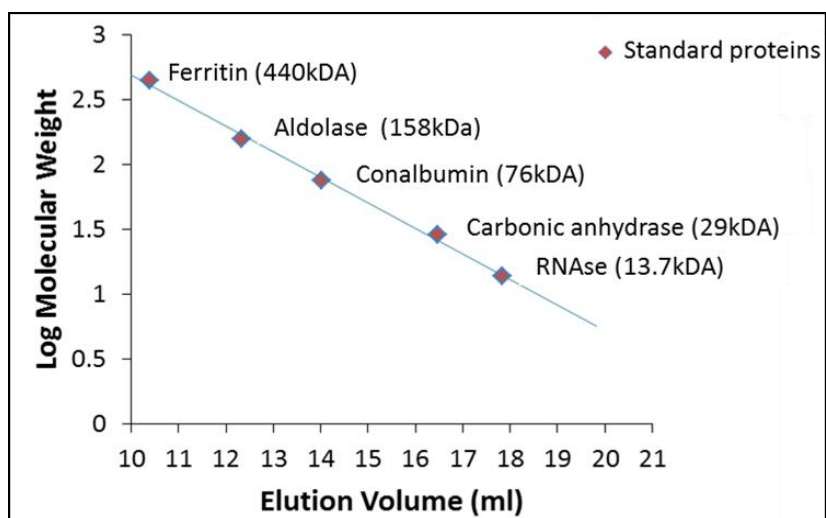
Standard curve generated from a linear fit of the log molecular weight of the protein standards versus elution volume.

Supplementary Info 3.9. Native gel of the pull-down assay with the protein phosphatases and response regulators.
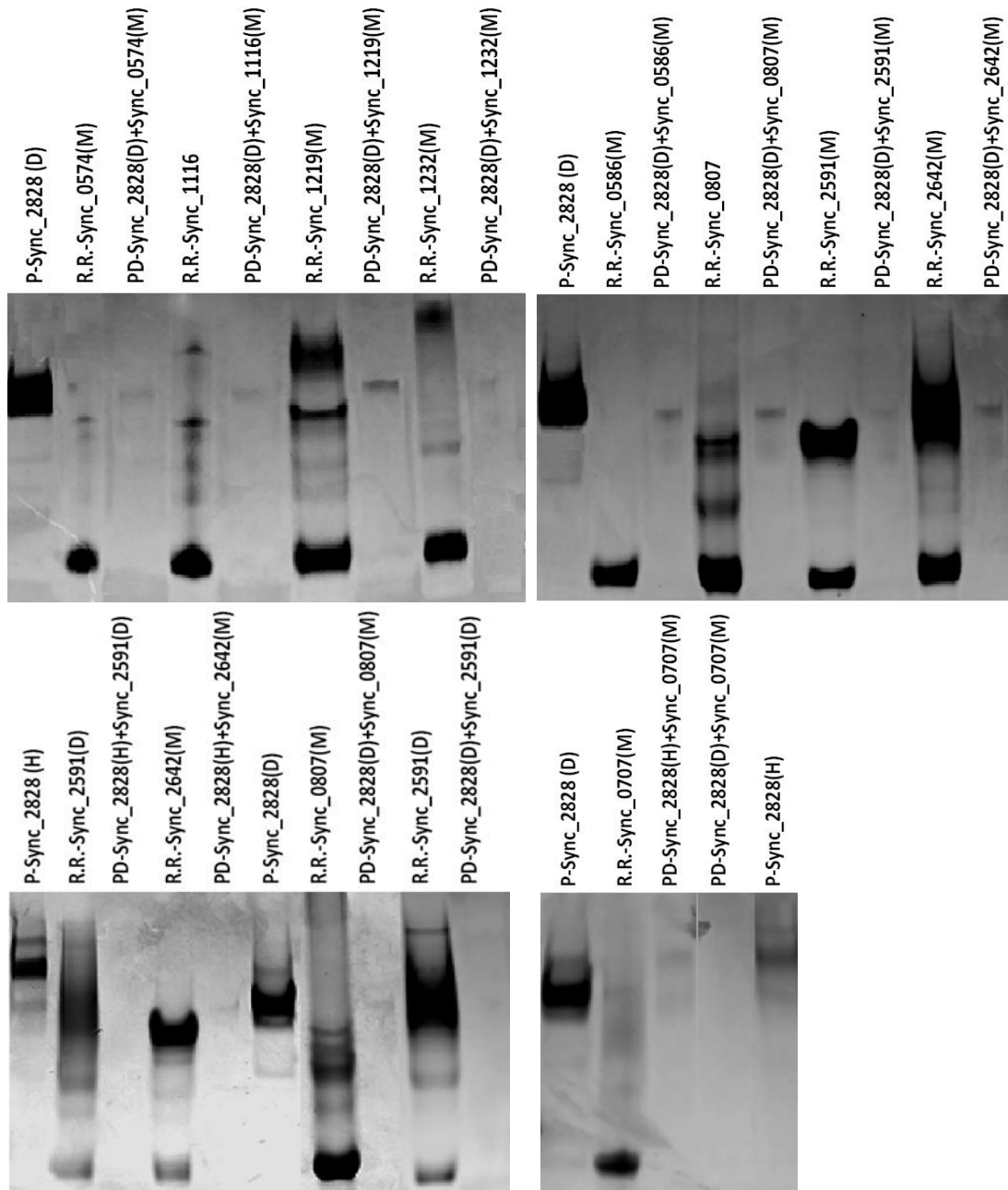
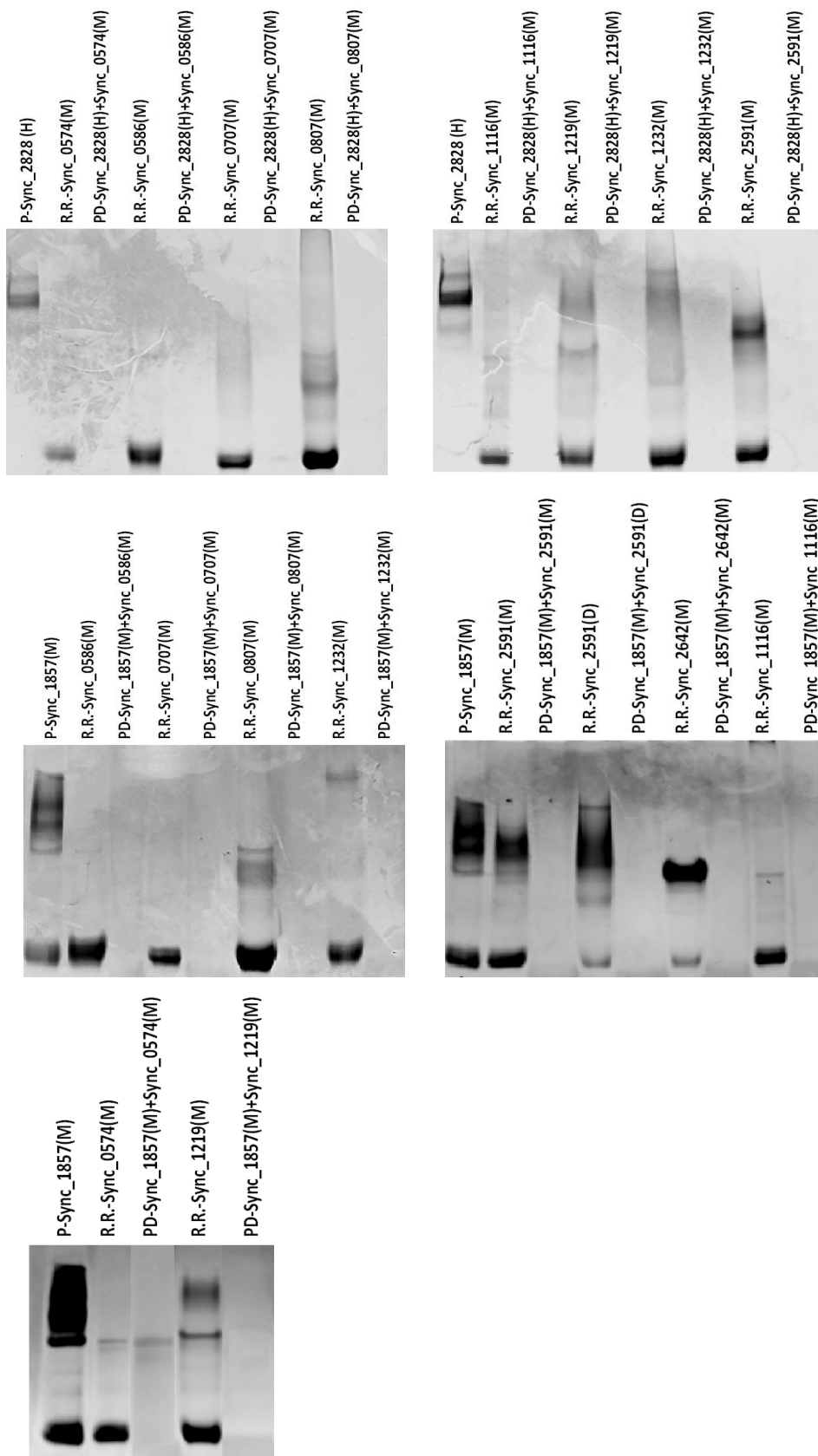(Abbreviations: P, phosphatase; R.R.-response regulator; PD: pull-down; M, monomer, D, dimer; H, hexamer)

REFERENCES

AIBA, H., NAGAYA, M. & MIZUNO, T. 1993. Sensor and regulator proteins from the cyanobacterium *Synechococcus* sp. PCC7942 that belong to the bacterial signal-transduction protein families: implication in the adaptive response to phosphate limitation. *Mol. Microbiol.,* 8**,** 81-91.

ARAVIND, L. & KOONIN, E. V. 1998. Phosphoesterase domains associated with DNA polymerases of diverse origins. *Nucleic Acids Res,* 26**,** 3746-3752.

ASHBY, M. K. 2004. Survey of the number of two-component response regulator genes in the complete and annotated genome sequences of prokaryotes. *FEMS Microbiol Lett,* 231**,** 277-281.

ASLANIDIS, C. & DE JONG, P. J. 1990. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res,* 18**,** 6069.

BILWES, A. M., ALEX, L. A., CRANE, B. R. & SIMON, M. I. 1999. Structure of CheA, a Signal-Transducing Histidine Kinase. *Cell,* 96**,** 131-141.

BIRD, L.E. 2011. High throughput construction and small scale expression screening of muti-tag vectors in *Echerichia coli*. Methods, 55, 29-37.

BIRD, L., RADA, H., FLANAGAN, J., DIPROSE, J. M., GLIBERT, R. J. C. & OWENS, R. J. 2013. Applications of In-Fusion™ cloning for the parallel construction of *E.coli* expression vectors. Methods Mol Biol, 1116, 209-234.

BRAUTIGAN, D. L. 2013. Protein Ser/Thr phosphatases--the ugly ducklings of cell signalling. *FEBS J,* 280**,** 324-45.

CADORET, J. C., ROUSSEAU, B., PEREWOSKA, I., SICORA, C., CHEREGI, O., VASS, I. & HOUMARD, J. 2005. Cyclic nucleotides, the photosynthetic apparatus and response to a UV-B stress in the Cyanobacterium *Synechocystis* sp. PCC 6803. *J Biol Chem,* 280**,** 33935-44.

CHIANG, G. G., SCHAEFER, M. R. & GROSSMAN, A. R. 1992. Complementation of a red-light-indifferent cyanobacterial mutant. *PNAS,* 89**,** 9415-9419.

CHRISTIE-OLEZA, J. A., ARMENGAUD, J., GUERIN, P. & SCANLAN, D. J. 2015. Functional distinctness in the exoproteomes of marine *Synechococcus*. *Environ Microbiol,* 17**,** 3781-3794.

COZZONE, A. J., GRANGEASSE, C., DOUBLET, P. & DUCLOS, B. 2004. Protein phosphorylation on tyrosine in bacteria. *Arch Microbiol,* 181**,** 171-81.

EGGER, L. A., PARK, H. & INOUYE, M. 1997. Signal transduction via the histidyl-aspartyl phosphorelay. *Genes Cells*, 2, 167-84.

FOTHERGILL, L. A. & HARKINS, R. N. 1982. The amino acid sequence of yeast phosphoglycerate mutase. *Proc R Soc Lond B,* 215**,** 19-44.

GALPERIN, M. Y. 2010. Diversity of structure and function of response regulator output domains. *Curr Opin Microbiol,* 13**,** 150-9.

GRANGEASSE, C., OBADIA, B., MIJAKOVIC, I., DEUTSCHER, J., COZZONE, A. J. & DOUBLET, P. 2003. Autophosphorylation of the *Escherichia coli* protein kinase Wzc regulates tyrosine phosphorylation of Ugd, a UDP-glucose dehydrogenase. . *J Biol Chem,* 278**,** 39323-9.

HAUN, R. S. & MOSS, J. 1992. Ligation-independent cloning of glutathione S-transferase fusion genes for expression in *Escherichia coli. Gene,* 112**,** 37-43.

HOCH, J. A. 2000. Two-component and phosphorelay signal transduction. *Curr Opin Microbiol,* 3**,** 165-170.

HSIAO, H. Y., HE, Q., VAN WAASBERGEN, L. G. & GROSSMAN, A. R. 2004. Control of photosynthetic and high-light-responsive genes by the histidine kinase DspA: negative and positive regulation and interactions between signal transduction pathways. *J Bacteriol,* 186**,** 3882-8.

IRVING, G. C. & D.J., C. 1972. Inositol phosphate phosphatases of microbiological origin: the inositol pentaphosphate products of *Aspergillus ficuum* phytases. *J. Bacteriol.,* 112**,** 434-438.

JENAL, U. & GALPERIN, M. Y. 2009. Single domain response regulators: molecular switches with emerging roles in cell organization and dynamics. *Curr Opin Microbiol,* 12**,** 152-60.

KENNELLY, P. J. 2002. Protein kinases and protein phosphatases in prokaryotes- a genomic perspective. *FEMS Microbiol Lett,* 206**,** 1-8.

KIBBE, W. A. 2007. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res,* 35**,** W43-6.

KUHN, P., DRAYCHEVA, A., VOGT, A., PETRIMAN, N. A., STURM, L., DREPPER, F., WARSCHEID, B., WINTERMEYER, W. & KOCH, H. G. 2015. Ribosome binding induces repositioning of the signal recognition particle receptor on the translocon. *J Cell Biol,* 211**,** 91-104.

LARKIN, M. A., BLACKSHIELDS, G., BROWN, N. P., CHENNA, R., MCGETTIGAN, P. A., MCWILLIAM, H., VALENTIN, F., WALLACE, I. M., WILM, A., LOPEZ, R., THOMPSON, J. D., GIBSON, T. J. & HIGGINS, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics,* 23**,** 2947-8.

LECLERCQ, S., DEROUAUX, A., OLATUNJI, S., FRAIPONT, C., EGAN, A. J., VOLLMER, W., BREUKINK, E. & TERRAK, M. 2017. Interplay between Penicillin-binding proteins and SEDS proteins promotes bacterial cell wall synthesis. *Sci Rep,* 7**,** 43306.

LETUNIC, I. & BORK, P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res,* 44**,** W242-5.

LI, L. & KEHOE, D. M. 2005. In vivo analysis of the roles of conserved aspartate and histidine residues within a complex response regulator. *Mol Microbiol,* 55**,** 1538-52.

LIU, Y., ROSE, J., HUANG, S., HU, Y., WU, Q., WANG, D., LI, C., LIU, M., ZHOU, P. & JIANG, L. 2017. A pH-gated conformational switch regulates the phosphatase activity of bifunctional HisKA-family histidine kinases. *Nat Commun,* 8, 2104.

MACEK, B., GNAD, F., SOUFI, B., KUMAR, C., OLSEN, J. V., MIJAKOVIC, I. & MANN, M. 2008. Phosphoproteome Analysis of *E. coli* Reveals Evolutionary Conservation of Bacterial Ser/Thr/Tyr Phosphorylation. *Mol Cell Proteomics,* 7**,** 299-307.

MACEK, B., MIJAKOVIC, I., OLSEN, J. V., GNAD, F., KUMAR, C., JENSEN, P. R. & MANN, M. 2007. The Serine/Threonine/Tyrosine Phosphoproteome of the Model Bacterium *Bacillus subtilis*. *Mol Cell Proteomics,* 6**,** 697-707.

MARTINEZ-HACKERT, E. & STOCK, A. M. 1997. The DNA-binding domain of OmpR-crystal structures of a winged helix transcription factor. *Curr Biol,* 5**,** 109-124.

MEYERHOF, O. & KIESSLING, W. 1935. Uber die isolierung der isomeren phosphoglycerinsauren (glycerinsaure-2-phosphorsaure und glycerinsaure-3-phosphorsaure) aus garansatzen und ihr enzymatisches gleichgewicht. *Biochem. Z.,* 276**,** 239-253.

MIJAKOVIC, I., MUSUMECI, L., TAUTZ, L., PETRANOVIC, D., EDWARDS, R. A., JENSEN, P. R., MUSTELIN, T., DEUTSCHER, J. & BOTTINI, N. 2005. In vitro characterization of the *Bacillus subtilis* protein tyrosine phosphatase YwqE. *J Bacteriol,* 187**,** 3384-90.

Mizuno, T. 1998. His-Asp Phosphotransfer Signal Transduction. *J Biochem,* 123, 555-563.

MIZUNO, T., KANEKO, T. & TABATA, S. 1996. Compilation of All Genes Encoding Bacterial Two-component Signal Transducers in the Genome of the Cyanobacterium, *Synechocystis* sp. Strain PCC 6803. *DNA Res,* 3**,** 407-414.

MORONA, J. K., MORONA, R., MILLER, D. C. & PATON, J. C. 2002. *Streptococcus pneumoniae* Capsule Biosynthesis Protein CpsB Is a Novel Manganese-Dependent Phosphotyrosine-Protein Phosphatase. *J Bacteriol ,* 184**,** 577-583.

OGAWA, T., BAO, D. H., KATOH, H., SHIBATA, M., PAKRASI, H. B. & BHATTACHARYYA-PAKRASI, M. 2002. A two-component signal transduction pathway regulates manganese homeostasis in *Synechocystis* 6803, a photosynthetic organism. *J Biol Chem,* 277**,** 28981-6.

PATHAK, D., BHAT, A. H., SAPEHIA, V., RAI, J. & RAO, A. 2016. Biochemical evidence for relaxed substrate specificity of Nalpha-acetyltransferase (Rv3420c/rimI) of *Mycobacterium tuberculosis*. *Sci Rep,* 6**,** 28892.

PILKIS, S. J., LIVELY, M. O. & EL-MAGHRABI, M. R. 1987. Active site sequence of hepatic fructose-2,6-bisphosphatase. Homology in primary structure with phosphoglycerate mutase. . *The J Biol Chem,* 262**,** 12672-12675.

PRADEL, E. & BOQUET, P. L. 1988. Acid phosphatases of *Escherichia coli*: molecular cloning and analysis of agp, the structural gene for a periplasmic acid glucose phosphatase. . *J. Bacteriol.,* 170**,** 4916-4923.

RIGDEN, D. J. 2008. The histidine phosphatase superfamily: structure and function. *Biochem J,* 409**,** 333-48.

RYJENKOV, D. A., TARUTINA, M., MOSKVIN, O. V. & GOMELSKY, M. 2005. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *J Bacteriol,* 187**,** 1792-8.

SCHWARTZ, S. H., BLACK, T. A., JAGER, K., PANOFF, J.-M. & WOLK, C. P. 1998. Regulation of an Osmoticum-Responsive Gene in *Anabaena* sp. Strain PCC 7120. *J Bacteriol ,* 180**,** 6332-6337.

SHI, L., POTTS, M. & KENNELLY, P. J. 1998. The serine, threonine, and/or tyrosine-specific protein kinases and protein phosphatases of prokaryotic organisms- a family portrait. *FEMS Microbiol Rev,* 22**,** 229-253.

SHI, Y. 2009. Serine/threonine phosphatases: mechanism through structure. *Cell,* 139**,** 468-84.

SICKMANN, A. & MEYER, H. E. 2001. Phosphoamino acid analysis. *Proteomics,* 1**,** 200-206.

STANDISH, A. J. & MORONA, R. 2014. The role of bacterial protein tyrosine phosphatases in the regulation of the biosynthesis of secreted polysaccharides. *Antioxid Redox Signal,* 20**,** 2274-89.

STUDIER, F. W. 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expr Purif ,* 41**,** 207-234.

SUZUKI, S., FERJANI, A., SUZUKI, I. & MURATA, N. 2004. The SphS-SphR two component system is the exclusive sensor for the induction of gene expression in response to phosphate limitation in *Synechocystis. . J. Biol. Chem.,* 279**,** 13234-13240.

SWANSON, R. V., ALEX, L. A. & SIMON, M. I. 1994. Histidine and aspartate phosphorylation- two-component system and the limits of homology. *Trends Biochem Sci,* 19**,** 485-490.

TOBES, R. & RAMOS, J. L. 2002. AraC-XylS database- a family of positive transcriptional regulators in bacteria. *Nucleic Acids Res,* 30**,** 318-321.

TYANOVA, S., TEMU, T. & COX, J. 2016. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc,* 11**,** 2301-2319.

WAHLGREN, W. Y., DUNEVALL, E., NORTH, R. A., PAZ, A., SCALISE, M., BISIGNANO, P., BENGTSSON-PALME, J., GOYAL, P., CLAESSON, E., CAING-CARLSSON, R., ANDERSSON, R., BEIS, K., NILSSON, U. J., FAREWELL, A., POCHINI, L., INDIVERI, C., GRABE, M., DOBSON, R. C. J., ABRAMSON, J., RAMASWAMY, S. & FRIEMANN, R. 2018. Substrate-bound outward-open structure of a Na^+-coupled sialic acid symporter reveals a new Na^+ site. *Nat Commun,* 9.

WEBRE, D. J., WOLANIN, P. M. & STOCK, J. B. 2003. Bacterial chemotaxis. *Curr. Biol. ,* 13**,** R47-R49.

WOODMAN, M. E. 2008. Direct PCR of Intact Bacteria (Colony PCR). *Curr Protoc Microbiol,* 9**,** 1-6.

YAMAGUCHI, A., SUZUKI, I., YAMAMOTO, H., LYUKEVICH, A., BODROVA, I., LOS, D. A., PIVEN, I., ZINCHENKO, V., KANEHISA, M. & MURATA, N. 2002. A Two-Component Mn^{2+}-Sensing System Negatively Regulates Expression of the *mntCAB* Operon in *Synechocystis.Plant Cell,* 14.

YANG, M. K., QIAO, Z. X., ZHANG, W. Y., XIONG, Q., ZHANG, J., LI, T., GE, F. & ZHAO, J. D. 2013. Global phosphoproteomic analysis reveals diverse functions of serine/threonine/tyrosine phosphorylation in the model cyanobacterium *Synechococcus* sp. strain PCC 7002. *J Proteome Res,* 12**,** 1909-23.

ZHANG, C.-C. 1996. Bacterial signalling involving eukaryotic-type protein kinases. *Mol Microbiol,* 20**,** 9-15.

Chapter 4 Preface

In Chapter 3, two phosphatase proteins were successfully purified: Sync_1857 and Sync_2828. In the following chapter, we investigated the protein-protein interactions of each of these phosphatases with *Synechococcus* CC9311 cytoplasmic lysate. Thus, a protein-protein interaction network was created for these candidate phosphatases, and cytoplasmic proteins, in *Synechococcus* sp. CC9311.

Objectives:

1) To identify proteins that interact with the protein phosphatases Sync_1857 and Sync_2828 and shed light on the regulatory role of these phosphatases.

2) To generate a protein-protein interactions network of phosphatases in marine cyanobacteria.

Defining the Protein-protein Interaction Network of Protein Phosphatases in Marine Cyanobacteria

Fallen Teoh[1], Bhumika Shah[1], Ian Paulsen[1*]

[1]Department of Molecular Sciences, Macquarie University, Sydney, Australia.

[*]Author for correspondence

Mailing address: 4 Wally's Walk, Department of Molecular Sciences, Macquarie University NSW 2109. Tel: (61) 2 98508152.

E-mail: ian.paulsen@mq.edu.au

ABSTRACT

Reversible protein phosphorylation post-translational modifications regulate a myriad of cellular processes in prokaryotes to eukaryotes. The level of protein phosphorylation status in cells is manipulated by the concerted reaction of kinases and phosphatases. Kinase enzymes catalyze the transfer of phosphate groups from high energy, phosphate-donating molecules to substrate proteins. Meanwhile, phosphatases act in an antagonistic way, removing the phosphate group from the same residue. Marine picocyanobacteria occupy a broad geographic area and have a significant contribution to the primary production in the ocean. Even though they play a crucial role in the ecosystem of Earth, the regulation of cell signaling in response to intracellular and extracellular stimuli in marine cyanobacteria are not well characterized. We present the first study of protein-protein interactions of phosphatases in marine *Synechococcus* using affinity purification mass spectrometry (AP-MS) approach. We selected two phosphatase proteins: Sync_1857 predicted to be localized in the periplasm and Sync_2828 protein predicted to be localized in the cytoplasm to study their potential protein-protein interactions and investigate their roles in *Synechococcus* cellular metabolism. The affinity purified protein partners of the two phosphatases were analyzed using Orbitrap Q-Exactive mass spectrometry coupled with nano-liquid chromatography. We identified 84 and 26 proteins that bound to phosphatase Sync_1857 and Sync_2828, respectively. Both phosphatase proteins involved in glycolysis, gluconeogenesis and oxidative pentose pathway, suggesting their crucial roles in carbon metabolism. Sync_1857 has a broad range of

interacted proteins, suggesting a diverse role of phosphodiesterase phosphatase in regulation of cellular activities in marine cyanobacteria. Meanwhile, Sync_2828 was found to interact with a range of stress-related proteins, suggesting a potential role in stress response regulation in cyanobacterial cells.

INTRODUCTION

Marine picocyanobacteria genus *Synechococcus* occupy a diverse geographical area in the world oceans. They play a vital role in primary production and make an important contribution to the carbon biogeochemical cycle in the ocean (Richardson and Jackson, 2007, Pittera et al., 2014). *Synechococcus* sp. CC9311 (hereafter CC9311) belongs to clade I of genus *Synechococccus*, which are among the four predominant clades (clade I, II, III and IV) in the world's oceanic water (Zwirglmaier et al., 2008, Scanlan et al., 2009). CC9311 has a greater number of predicted signal transduction systems than other sequenced marine *Synechococcus* strains, and was isolated from temperate coastal waters suggesting that it may have a greater capacity to sense and respond to changes in dynamic coastal waters (Palenik et al., 2006). Hence, we have used CC9311 as a model system for investigating cyanobacterial signal transduction networks.

Reversible protein phosphorylation, a widespread post-translational modification (PTM) from prokaryotes to eukaryotes, mediates cellular responses to growth factors, environmental stimuli, and internal processes by inducing conformational changes or by disruption and creation of protein-protein interaction surfaces (Mann and Jensen, 2003). It is regulated in a coordinated manner by protein kinases and phosphatases. Kinase enzymes catalyze the phosphorylation of proteins commonly at their serine (Ser), threonine (Thr), tyrosine (Tyr), histidine (His) or aspartate (Asp) residues (Paithoonrangsarid et al., 2004, Pereira et al., 2011, Zorina et al., 2011). Meanwhile, phosphatase enzymes cleave the monophosphate esters from the phosphorylated form of the same residues. Therefore, balanced coordination of a multitude of kinases and phosphatases, control a broad range of signaling pathways and physiological processes in bacteria (Pereira et al., 2011)

The functional characterization of phosphatases has lagged behind kinases in prokaryotes. The two-component regulatory systems (TCRS) involving a histidine kinase (HK) and its

cognate response regulator (RR) has been recognized as the primary phosphorylation-based signal transduction in bacteria (Los et al., 2010). The HK catalyzes auto-phosphorylation of a specific histidine residue, generating a high-energy phosphoryl group to transfer to an aspartate residue in the RR. Principally, the phosphorylation level of the RR governs system output. Many bacterial two-component HK are bifunctional, having both kinase activity (acting on histidine) and phosphatase activity (acting on phosphoaspartate) (Klumpp and Krieglstein, 2002). Owing to these, phosphatases receive little attention in the past. However, recent studies have shown that Ser/Thr/Tyr phosphorylation PTM events have high coverage occurrence in bacteria including cyanobacteria (Yang et al., 2013, Spat et al., 2015) as seen in eukaryotes. Moreover, different types of phosphatase family proteins have been revealed, reflecting the diversity of phosphatases and possibly linked to their diverse roles in cells (Shi et al., 1998, Pereira et al., 2011, Standish and Morona, 2014). Moreover, recent studies on cyanobacterial proteomics (Christie-Oleza et al., 2015) (Chapter 2 of this thesis) revealed a very high level of expression of particular protein phosphatases, suggesting that they must be playing key roles in signal transduction in marine cyanobacteria. However, we know little about phosphatases in bacteria, partly due to our poor understanding of the phosphatase protein-protein binding network.

Therefore, we conducted a phosphatase protein-protein interactions (PPIs) study to elucidate the functional role of two phosphatases from marine *Synechococcus* CC9311. Sync_1857 is a putative serine/threonine specific protein phosphatase and includes a predicted twin-arginine translocation (Tat) signal peptide suggesting that it is targeted to the periplasm. Sync_2828 is homologous to the yeast phosphatase 2A regulatory B subunit. Whether Sync_2828 can function as a phosphatase by itself, or interacts with other phosphatase in CC9311, or has functions unrelated to phosphatases is not clear. Sync_2828 does not contain any signal sequences, and is predicted to be localized in CC9311 cytoplasm.

These two proteins were expressed and purified as GST fusions. Affinity purification mass spectrometry (AP-MS) with the GST tag was used to pull-down their potential binding proteins from cytoplasmic lysate of CC9311 and verify their identification using mass spectrometry analysis.

METHODOLOGY

Experimental procedures

Plasmid Constructs and Purification of Recombinant

The phosphatase gene sequences of *sync_1857* and *sync_2828* were PCR-amplified from *Synechococcus* sp. CC9311 genomic DNA using specific primers (Supplementary Info 4.1). The purified PCR fragments were then cloned into linearized pOPINJ vectors (Oxford Protein Production Facility), cut with the restriction enzymes *Kpn*I and *Hin*dIII, using In-Fusion® HD EcoDry™ Cloning Plus kits (Takara). GST and His$_6$-tag sequences in the pOPINJ vector were fused with the target genes *sync_1857* and *sync_2828* corresponding to the N-terminus of the recombinant proteins. All constructs were verified by DNA sequencing (Macrogen). The constructed plasmids, namely 1857J (for sequence *sync_1857* in pOPINJ vector) and 2828J (for sequence *sync_2828* in pOPINJ) were transformed into *Escherichia coli* Lemo21 (DE3) (New England Biolabs Inc.) and Rosetta™ 2 DE3 p*LacI* (Novagen) strains, respectively. The transformed *Escherichia coli* cells for both phosphatases were cultured using ZYP5052 autoinduction media supplemented with 100mg/mL ampicillin and 50mg/mL chloramphenicol. Two batches of 4 liters bacterial culture were conducted for each strain. The bacterial cultures were incubated at 37^0C until the growth rate reached mid-log phase (OD$_{600nm}$ = 0.4-0.6), then, autoinduction of protein expression was induced at 20^0C overnight. The bacterial cultures were harvested by centrifugation at 5000 rpm for 20 minutes at 4^0C. For the first batch culture of each strain, the bacterial lysates were resuspended in Buffer A (50mM HEPES, pH8.0, 300mM NaCl, 5% glycerol, 20mM imidazole) and stored at -80^0C until ready for cell lysis. DNaseI (Sigma) and EDTA-free protease inhibitor cocktail (Bi Make) were added according to manufacturer's protocol upon cell lysis on ice using a sonicator. Cell lysates were clarified by centrifugation at 12,900 rpm for 15 minutes at 4^0C. The resulting supernatants were collected and loaded onto a 1mL HisPur Ni-NTA spin column (Thermo Fisher Scientific), pre-equilibrated with Buffer A. After all sample buffers were spun down by centrifugation, wash buffer (50mM HEPES, pH8.0, 300mM NaCl, 5% glycerol, 50mM imidazole) was added and centrifuged. This process was repeated six times. The His6-tagged proteins were eluted with the 400mM imidazole in Buffer A. The condition for the second batch culture of each strain was same as first batch except 0.5mM TCEP (tris(2-carboxyethyl) phosphine) was added for all solutions mentioned. The recombinant

proteins were fractionated by Superdex™ 200 10/300 GL and HiLoad™ 16/600 and Superdex™ 200pg size exclusion chromatography (SEC) columns (GE Healthcare Life Sciences) using the Akta Pure chromatography system (GE Healthcare Life Sciences). For the first batch culture, the culture Buffer A without imidazole was used as running buffer for SEC analysis while the condition buffer remained the same for the second batch culture except 0.5mM TCEP was added into the buffer. The fractionated proteins from SEC were collected and concentrated with a Vivaspin 6 concentrator (GE Healthcare) and stored at -80^{0}C and until ready for the pull-down *in vivo* assay.

Synechococcus cell culture and cytoplasmic lysate extraction

Synechococcus sp. CC9311 cells were grown in synthetic ocean water (SOW) media (Morel et al., 1979, Su et al., 2006) and incubated at 22^{0}C in continuous white light intensity of 40 μmol photons m^{-2}s^{-1} with agitation at 100 rpm. *Synechococcus* cells were harvested at exponential growth phase at an OD$_{750nm}$ of ~0.2 and centrifuged at 8000rpm, 4^{0}C for 15 minutes. Cell lysates were then mixed in ice-cold lysis buffer (50mM Tris-HCl, pH8.0, 10mM KCl and 5% glycerol) with DNaseI (Sigma) was added according to the manufacturer's protocol before cell lysis. Cell lysis was conducted with French press pressurized cells at 16,000 psi. The samples were then centrifuged at 17,100 rpm, for 10 minutes at 4^{0}C to remove unbroken cells. Subsequently, the supernatant was centrifuged at 41,900 rpm, for 45 minutes at 4^{0}C to collect the cytoplasmic lysate. Cytoplasmic lysate protein concentration was determined using a bicinchoninic acid protein (BCA) assay (Thermo Fisher Scientific) with Bovine Serum Albumin (Bradford, 1976) as a standard.

Affinity Purification of Cytoplasmic Proteome Subset

Affinity purification assays were carried out for 1857J phosphatase monomer, dimer, tetramer and hexamer and 2828J phosphatase dimer and hexamer individually. A milligram of cytoplasmic lysate was mixed with 50μg of each phosphatase protein at 4^{0}C overnight. The solution mixture was then incubated with 50μl of GST agarose beads (GE Healthcare) at 4^{0}C for 2.5 hours. After incubation, the GST agarose beads were washed five times with 1mL

of ice-cold wash buffer (50mM Tris-HCl, pH7.5, 10mM KCl, 5% glycerol). To elute the cytoplasmic proteins that bind with GST-tagged phosphatase, GST agarose beads were reacted with 100µl of 50mM Tris-HCl, pH8 containing 20mM glutathione at 4^0C. Aliquots of eluted proteins were analyzed by 1D SDS-PAGE while the rest of the proteins were proceeded with in-solution digestion.

In-Solution Digestion and MS/MS Mass Spectrometry

Methanol/chloroform protein precipitation (Wessel and Flugge, 1984) was conducted on the affinity purified proteins. 6M urea was used for solubilization of extracted proteins. 5µl DTT (dithiothreitol from 200mM stock) reducing agent was added to the mixture and incubated for an hour at room temperature. It was then followed by 20µl IAA (iodoacetamide from 200mM stock), alkylating agent reaction, in the dark for an hour. Next, 20µl DTT was added and incubated at room temperature for 30 minutes (refer (Kinter and Sherman, 2000) for detail protocols). Prior to trypsin digestion, the solution was diluted to 1M urea concentration. Trypsin enzyme was added to the protein in the ratio of 1:30. The solution was then incubated at 37^0C for 16 hours.

Shotgun Nano-LC-MS/MS Analysis

All samples were analyzed by LC-MS/MS using a Q-Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-nLC1000 (Thermo Fisher Scientific). Reverse-phase chromatographic separation was carried out on a C18 HALO column (75µm i.d. x100 mm, 2.7 µm bead size, 160 Å pore size) (Thermo Fisher Scientific). A linear gradient of 1-50% solvent B (99.9% ACN/0.1% FA) was run over 30 minutes, followed by a slope gradient of solvent B from 50% to 85% for 2 minutes and maintained at 85% for 8 minutes. The mass spectrometer was operated in the data-dependent mode to automatically switch between Orbitrap MS and ion trap MS/MS acquisition. Survey full scan MS spectra (from m/z 350 to 2000) were acquired with a resolution of 35,000 and an AGC (Automatic Gain Control) target value of 1 x 10^6 ions. For identification of peptides, the ten most abundant ions were selected for higher energy collisional energy set to 35% and fragmentation ions were detected in the Orbitrap at a resolution of 17,500. Target ions that

had been selected for MS/MS were dynamically excluded for 10 sec. The lock mass option was enabled in all full scans to improve the mass accuracy of precursor ions.

Database searching for protein/peptide identification

Raw data files produced in Xcalibur software (Thermo Fisher Scientific) were searched against the concatenated forward and reverse Cyanorak *Synechococcus* s.p. CC9311 protein database (http://abims.sb-roscoff.fr/cyanorak) using MaxQuant (version 1.6.3.4) search engine (Tyanova et al., 2016). The following criteria were used: full tryptic specificity was required, 2 missed cleavages were allowed; carbamidomethylation of cysteines was set as a fixed modification and methionine oxidation was set as a variable modification. The first search peptide tolerance was set as 20ppm, main search peptide tolerance at 4.5ppm, and the MS/MS tolerance at 20ppm. Proteins found were only counted if they are found with at least two unique peptides.

Bioinformatic Analysis

Prediction of secreted proteins was analyzed using a series of bioinformatics tools: subcellular localization prediction with PSORTb program version 3.02 (Yu et al., 2010), presence of signal peptides with SignalP v5.0 (Almagro Armenteros et al., 2019), transmembrane helices with TMHMM (Emanuelsson et al., 2007). Functional category of proteins was assigned using EggNOG v4.5 (Huerta-Cepas et al., 2016). Protein-chemical interactions with confidence score of 0.75 were extracted from STITCH (version 5.0) (Kuhn et al., 2008) and associated gene ontology (GO) of biological process and molecular functions of proteins were summarized in Supplementary Info 4.2.

RESULTS AND DISCUSSION

**Construction of affinity tag for phosphatase proteins**

Targeted genes *sync_1857* and *sync_2828* were amplified and cloned into the pOPINJ vector, which contains His_6 and GST-tags at the N-terminus. The constructed vectors were then

transformed into an *E. coli* expression system. The design of the recombinant vectors takes advantage of the His_6-tag for recombinant protein isolation and purification while the GST-tag enables pull-down assays for identification of affinity purified proteins that interact with Sync_1857 and Sync_2828.

Oligomeric states of the phosphatase proteins

In-solution size exclusion chromatography (SEC) analysis of the recombinant phosphatases Sync_1857 and Sync_2828 show they exist in several oligomeric forms. While Sync_1857 oligomerized in monomer, dimer, tetramer and hexamer (Figure 4.1a), Sync_2828 showed dimeric and hexameric forms (Figure 4.1b). As both proteins have cysteine residues in their amino acid sequences, hence they could possibly form disulfide-bridge(s) and change oligomerization state of the proteins. The amino acid sequence of Sync_1857 has 4 cysteine residues while Sync_2828 has one cysteine residue. When a reducing agent, TCEP was added to the running buffer of SEC, Sync_1857 existed as monomeric and dimeric forms while it oligomerized in monomer, dimer, tetramer and hexamer forms without TCEP (Figure 4.1a). This suggests that Sync_1857 forms disulfide-bridges and this is likely to occur *in vivo* as the protein contains signal peptide and likely to be localized in the periplasm, an oxidizing environment of cells. Meanwhile, SEC of Sync_2828 showed no difference in term of oligomerization state with TCEP or without TCEP, consistent with its likely cytoplasmic protein. The study of protein phosphatases can be complicated by their existence in different oligomer forms, for example phosphatase 2A (PP2A) can exist as dimers and trimers (Cohen, 1989, Shenolikar and Naim, 1991). Different oligomers display different enzyme structures, possibly changing the function of the enzymes (Rubiolo et al., 2014). Previous studies have shown that bacterial signal transduction can be regulated by protein oligomerization, however, this has not been well-studied for bacterial phosphatases (Amster-Choder and Wright, 1992, Lazazzera et al., 1993). The native oligomerization state of Sync_1857 and Sync_2828 in *Synechococcus* CC9311 cells remain unknown and whether the oligomerization status changes upon stimulation is also unknown. Therefore, we isolated all isoforms of these two proteins through SEC and conducted affinity purification of them in variant oligomer forms with the cell lysate of *Synechococcus* CC9311 strain. The 1D SDS-PAGE gels showing purified Sync_1857 and Sync_2828 phosphatase proteins and the pull down assay of the phosphatases is depicted in Figure 4.2a-f.
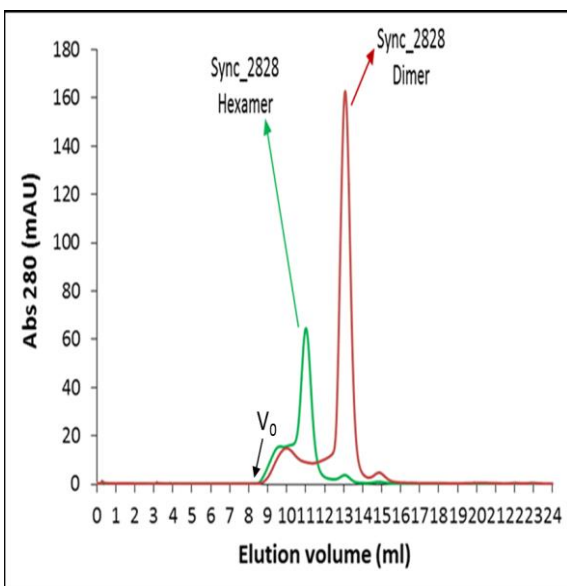
a)
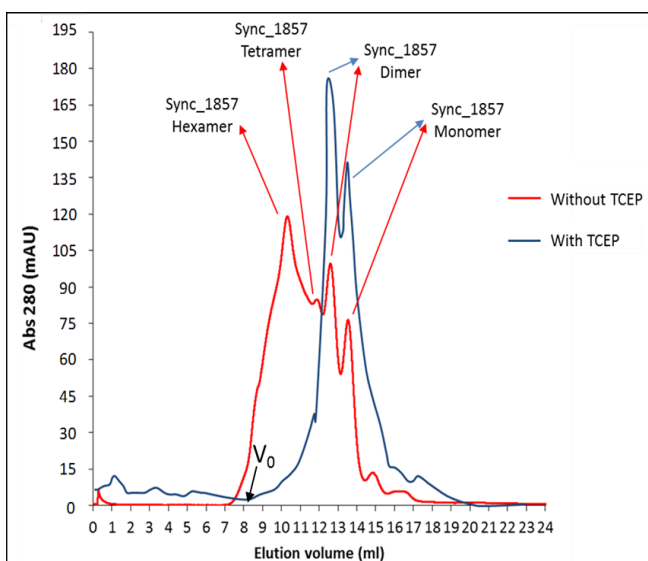
b)

Figure 4.1. Size exclusion chromatography (SEC) of the proteins Sync_1857 and Sync_2828. V_0 represents void volume. a) SEC of Sync_1857 in TCEP running buffer and without TCEP. Four oligomerization states (hexamer, tetramer, dimer and monomer) of protein Sync_ 1857 are observed when using running buffer without TCEP, while two oligomers (dimer and monomer) are seen when using running buffer with TCEP. b) SEC of Sync_2828 in TCEP running buffer. SEC trace of protein Sync_2828 remains the same without TCEP (not shown).
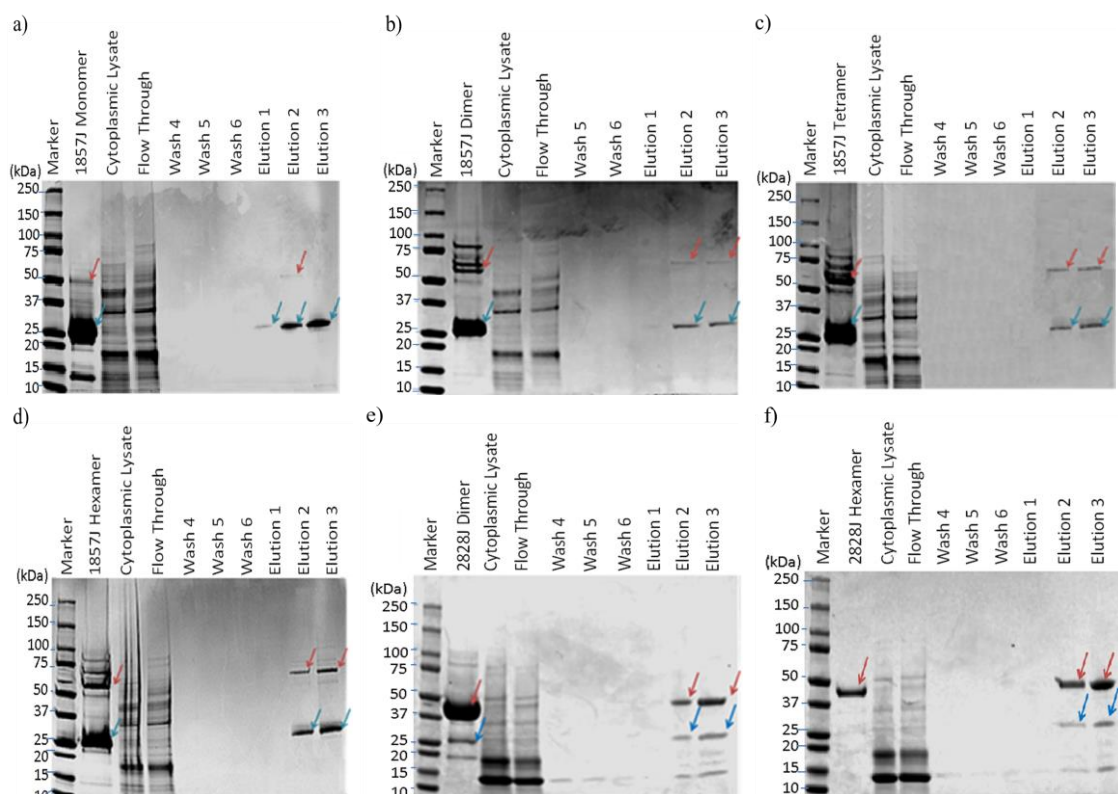
Figure 4.2. 1D SDS-PAGE analysis of the affinity purified phosphatases with cytoplasmic lysate of marine *Synechococcus* sp. CC9311. Affinity purified cytoplasmic lysate (prey proteins) using baits of a) monomeric Sync_1857, b) dimeric Sync_1857, c) tetrameric Sync_1857, d) hexameric Sync_1857, e) dimeric Sync_2828 f) hexameric Sync_2828. Red arrows indicate the bait proteins band size while blue arrows indicate the GST proteins (GST-tagged recombinant proteins are prone to cleavage when boiling in SDS). Elution 2 and 3 was selected for AP-MS analysis and results analysed from elution 2 was used to generate phosphatase interactome network).

The Protein Phosphatase Interactome

In total, 84 and 26 *Synechococcus* sp. CC9311 prey proteins were affinity purified from the cytoplasmic lysate of CC9311 using Sync_1857 and Sync_2828 as baits, respectively (Figure 4.3), with a confidence limit of 99% by Q-Exactive mass spectrometry. Bioinformatic analyses of subcellular localization using Signal P and psortB show that majority of the prey proteins identified belongs to cytoplasmic proteins (Supplementary Info 4.3).

For quality control, two negative controls were employed for the study. The first negative control consisted of affinity purification of the cytoplasmic lysate of *Synechococcus* sp. CC9311 without the phosphatase proteins as bait to exclude other highly expressed *Synechococcus* proteins. The second negative control consisted of affinity purification of the

cytoplasmic lysate of *Synechococcus* sp. CC9311 expressing the GST protein tag alone to exclude proteins pull-downed by the GST tag. In total, 18 proteins were identified as either highly expressed mainly photosynthesis-related proteins that eluted out or proteins that bind to GST and were excluded from the analysis (Supplementary Info 4.4).
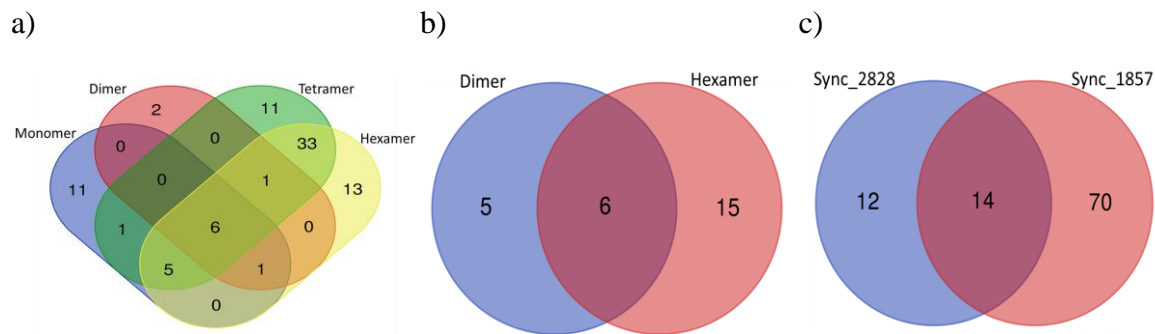

Figure 4.3. Venn diagram of the number of affinity purified proteins using Sync_1857 and Sync_2828 as baits with cytoplasmic lysate of marine *Synechococcus* sp. CC9311. a) Venn diagram of the number of proteins found in bait-prey interactions between Sync_1857 monomer, dimer, tetramer and hexamer forms with the cytoplasmic lysate of marine *Synechococcus* sp. CC9311. b) Venn diagram of the number of proteins found in bait-prey interactions between Sync_2828 dimer and hexamer forms with the cytoplasmic lysate of marine *Synechococcus* sp. CC9311. c) Venn diagram of all proteins found in the AP-MS analysis of Sync_1857 and Sync_2828. Complete lists of all of the affinity purified proteins are provided in Supplementary Info 4.3.

Among the 84 prey proteins affinity purified with phosphatase Sync_1857, we found that 6 proteins were bound to all 4 oligomeric states of phosphatase Sync_1857. These include the periplasmic binding protein urea ABC transporter (Sync_2872), hypothetical proteins (Sync_1862, Sync_1071, Sync_0950), translation elongation factor Tu (Tuf) and DNA-directed RNA polymerase, gamma subunit (RpoC1) (represented in PPIs and chemical-protein interactions network in Figure 4.4a). Meanwhile, the number of proteins found solely in monomer, dimer, tetramer and hexamer forms were 11, 2, 11, 13 respectively.

Of the 26 prey proteins affinity purified with phosphatase Sync_2828, 6 proteins are present in both oligomeric forms (dimer and hexamer) of this protein. These include cysteine synthase (CysK-1), O-acetylhomoserine sulfhydrylase (Thiol), polyribonucleotide nucelotidyltransferase (PnpA), glycerol kinase (GlpK), GMP synthase (GuaA) and Ni^{2+} superoxide dismutase (SodN) (represented in PPIs and chemical-protein interactions network
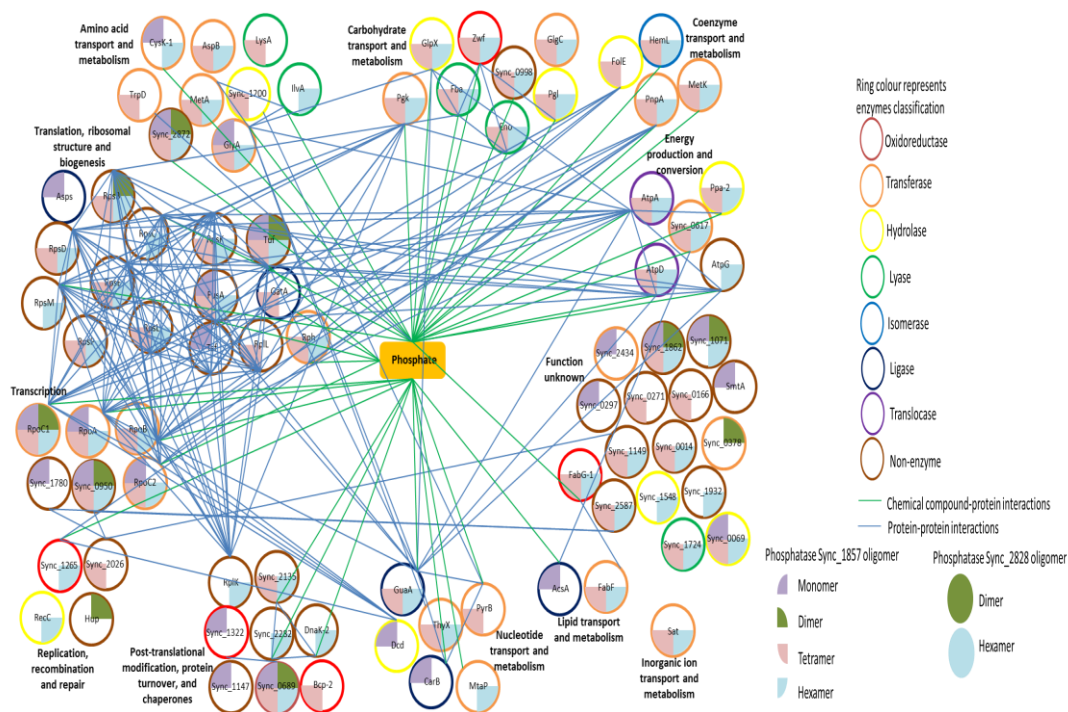
200

in Figure 4.4b). A total of 11 and 21 proteins have been isolated from the dimeric and hexameric forms of the phosphatase respectively (Figure 4.3b).

As depicted in Figure 4.3c, 14 common interacting proteins were seen in both the phosphatases studied. The PPIs and chemical-protein interaction network of Sync_1857 and Sync_2828 were generated using the STITCH software package (Kuhn et al., 2008), using all of the prey proteins found in the study and adding phosphate as a chemical compound for chemical-protein interaction analysis (Figure 4.4a, b). As the role of typical protein phosphatase involves removing a phosphate group from a phosphorylated protein, therefore the proteins should be phosphorylated to interact with phosphatase. Owing to this, phosphate was added as a chemical compound in the STITCH analysis to check which proteins that have a known interaction with phosphate from the published literature. Protein-protein and protein chemical nodes scoring from STITCH network is provided in Supplementary Info 4.5.

Prey proteins isolated with phosphatase Sync_1857 affinity purification include all seven classes of enzymes namely oxidoreductase, transferase, hydrolase, lyase, isomerase, ligase, translocase (Supplementary Info 4.6a), meanwhile phosphatase Sync_2828 prey proteins didn't include lyase and translocase enzymes (Supplementary Info 4.6b). The GO category of biological process from the STITCH network analysis suggested that the majority of the proteins are involved in nitrogen related metabolic process and primary metabolic process for both of the phosphatases used in this study (Supplementary Info 4.2).

The functional categories of the respective phosphatase prey proteins were assigned using EggNOG v4.5 (Huerta-Cepas et al., 2016) (Figure 4.4 and Figure 4.5). This functional category analysis showed that cyanobacterial phosphatases are potentially involved in a broad range of cellular activities in *Synechococcus* cells. It has been suggested that Ser/Thr/Tyr kinases and phosphatases may be involved in the regulation of diverse cellular activities in cyanobacteria, such as stress adaptation, cell differentiation, cell motility, carbon and nitrogen metabolism and photosynthesis (Zhang et al., 2005). Proteins identified from both phosphatases PPIs network in this study show a similar pattern of functional category assignment. The top three most abundant protein functional categories for prey proteins found in this study were translation, ribosomal structure and biogenesis; function unknown; and amino acid transport and metabolism (Figure 4.5a, b).
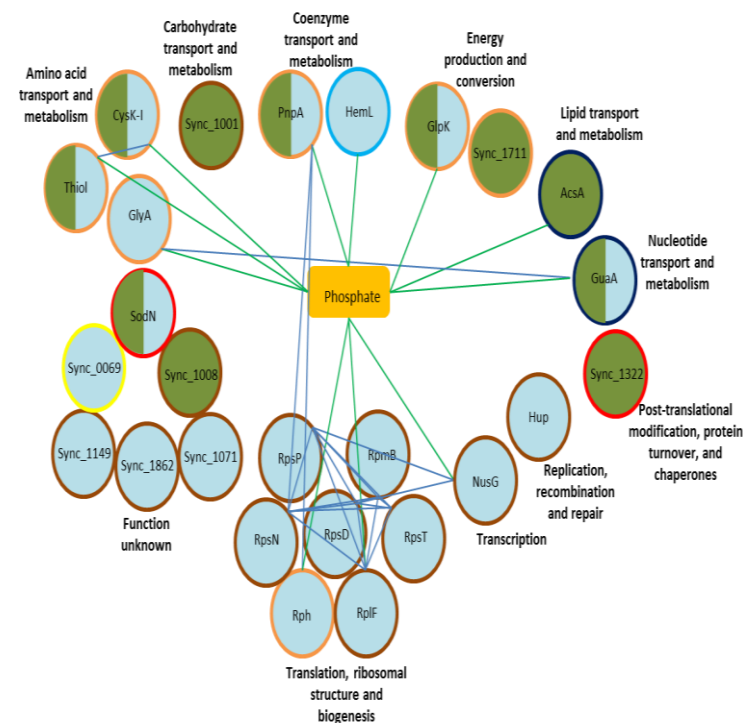
Figure 4.4. Phosphatase protein-protein and chemical-protein interactions network. PPIs and chemical-protein interactions network was generated from the identified proteins found in this AP-MS study. PPIs and chemical-protein interactions networks were extracted from STITCH network (confidence score 0.75). Phosphatase proteins were not included in the network as no PPIs of phosphatases Sync_1857 and Sync_2828 have been identified before this study.

a) Sync_1857 PPIs and chemical compound-protein interactions network.

b) Sync_2828 PPIs and chemical compound-protein interactions network.

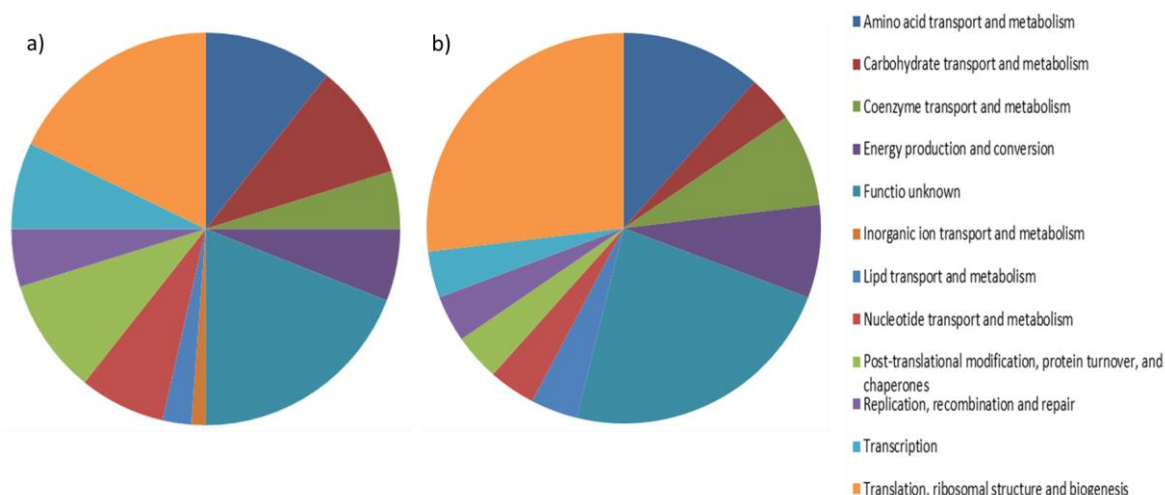**Figure 4.5. Functional categorization of the proteins identified in this AP-MS study.**

a) Functional categories of the prey proteins identified for phosphatase Sync_1857.
b) Functional categories of the prey proteins identified for phosphatase Sync_2828.

Possible biological roles of Sync_1857

Protein Sync_1857 is a Ser/Thr phosphatase that is predicted to contain a metallophos domain (PF00149). The protein families which contain this domain include a broad range of phosphodiesterases such as protein phosphoserine/threonine phosphatases, nucleotidases, sphingomyelin phosphodiesterases and 2'-3'cAMP phosphodiesterase and nucleases (Chen et al. 2004). Metallo-phosphodiesterase enzymes are characterized by metal-dependent phosphoesterase activity against a diverse range of substrates (Connelly and Leach, 2002, McLoughlin et al., 2004, Khalid et al., 2005, Shenoy et al., 2005, Keppetipola and Shuman, 2008, Wang et al., 2013). In our PPIs study, we identified a broad set of proteins that bind with phosphatase Sync_1857, suggesting this enzyme may be involved in a broad range of protein dephosphorylation and cell signaling events.

Metallo-phosphodiesterase enzymes play a vital role in DNA repair, cyclic, nucleotide metabolism and RNA processing (Aravind and Koonin, 1998, Matange et al., 2014). Our affinity purification assay has found a number of interacting proteins of phosphatase Sync_1857 which belong to the functional groups of transcription, translation, ribosomal structure and biogenesis, nucleotide transport and metabolism and replication, recombination and repair (Figure 4.4a). In comparison, the number of proteins found in

these functional categories is relatively lower for Sync_2828. Phosphorylation events involving the ribosome have been found in various prokaryotic and eukaryotic phosphoproteomic studies (Chen et al., 2014, Tian et al., 2014, Nakedi et al., 2015, Verma et al., 2017). However, the current roles of phosphorylation in proteins involved in translation, ribosomal structure and biogenesis are not well understood.

We identified CysK-1 and GlyA proteins as binding proteins of proteins Sync_1857 and also Sync_2828. Phosphorylation of serine hydroxymethyltransferase (GlyA) proteins has been identified in a cyanobacterial phosphoproteomic study (Spat et al., 2015). Serine hydroxymethyltransferase, a pyridoxal phosphate-dependent enzyme, catalyzes the interconversion of serine and glycine, both of which are major sources of one-carbon units vital for the synthesis of purine, thymidylate, methionine and so on (Herbig et al. 2002).

ABC transporters are the primary transporters in marine picocyanobacteria (Scanlan et al., 2009). They play an important role in the transport of ions, nutrients and substrate in/out of cells. A previous phosphoproteomic study on cyanobacteria has detected phosphorylation of the periplasmic binding protein urea ABC transporter, UrtA (Spat et al., 2015). In our AP-MS study, we found that all 4 different oligomer forms of protein Sync_1857 bound to UrtA. This suggests that phosphorylation modifications on UrtA may be vital for activation of the transport of urea. However, the kinase that involved in the phosphorylation of UrtA is still unknown, hence, future investigation should be carried out to investigation the effect of phosphorylation and dephosphorylation protein modifications on UrtA, given that urea ABC transporter appears to be an important nitrogen acquisition transporter for marine *Synechococcus* (Christie-Oleza et al., 2015) (Chapter 2 in this thesis).

Possible biological roles of Sync_2828

Sync_2828 is a predicted phosphatase 2A regulatory B protein. In yeast, regulatory B subunit protein associates with a regulatory subunit A protein and a catalytic C subunit forming protein phosphatase 2A (PP2A) that regulates the enzyme localization and activities (Herzog et al., 2012). Previous studies showed that PP2A is involved in reactive oxidative species (ROS) signaling in plants (Trotta et al., 2011, Rahikainen et al., 2016).

Mutagenesis of the gene encoding the PP2A regulatory B protein in plants showed the upregulation of stress response proteins such as superoxide dismutase and carbonic anhydrase (Trotta et al., 2011). Therefore, it was suspected that PP2A regulatory B is required in monitoring the stress response in cells (Trotta et al., 2011).

Our affinity purification assay of Sync_2828 identified interactions with thioredoxin peroxidase (Sync_1322) and SodN which are important antioxidants involved in maintaining ROS balance in cells. Additional prey proteins that showed an interaction with Sync_2828 included cysteine synthase (CysK-1) and O-acetylhomoserine sulfhydrylase (Thiol) enzymes which play an important role in methionine-cysteine metabolism. Methionine metabolism is vital for the regulation of oxidative stress in bacteria (Hondorp and Matthews, 2004). Therefore, we hypothesize that Sync_2828 could be vital in regulating oxidative stress signaling in marine cyanobacteria CC9311. Cyanobacteria perform oxygenic photosynthesis, hence, regulation of intracellular ROS homeostasis is pivotal, so that ROS generated from photosynthetic activities does not damage the cells (Rahikainen et al., 2016).

Glycolysis, gluconeogenesis and oxidative pentose phosphate pathway

A number of proteins identified in our phosphatase Sync_1857 PPIs study were involved in glycolysis (metabolic pathway that converts glucose into pyruvate) and gluconeogenesis (metabolic process which produce glucose for catabolic reactions from non-carbohydrate precursors) metabolism. These proteins including fructose-biphosphate aldolase (Fba), fructose-1,6-biphosphatase class II (GlpX), phosphoglycerate kinase (Pgk) and phosphopyruvate hydratase (Eno) (as illustrated in Figure 4.6). All of these proteins interact with the tetrameric and hexameric forms of phosphatase Sync_1857 suggesting the possibility that an allosteric switch regulates the protein-protein binding (Smith et al., 2016). Fba is a core carbon metabolic enzyme, responsible for the aldol cleavage of fructose-biphosphate into dihydroxyacetonephosphate (DHAP) and glyceraldehyde 3-phosphate (GAP) sugars in glycolytic reactions, and the reverse aldol condensation of these triose sugars to fructose-biphosphate in gluconeogenic reactions and Calvin cycle (Gefflaut et al., 1995, Haake et al., 1998). The enzyme family can be divided into two groups: one is the class-I Fba (CI-Fba), in which the enzyme forms a Schiff base with

substrates and does not depend on divalent ions for its activity. The other is the class-II Fba (CII-Fba), whose activity requires divalent ions. CI-Fbas, are present in most eukaryotes especially in animals (Rutter, 1964, Marsh and Lebherz, 1992) and higher plants (Rutter, 1964, Schnarrenberger et al., 1990), whereas CII-Fbas occur mainly in prokaryotes (Henze et al., 1998) and to a lesser degree, in some eukaryotes such as fungi and yeast (Rutter, 1964, Schwelberger et al., 1989). Both classes of Fba have been detected in the genome of freshwater cyanobacteria *Synechocystis* (Kaneko et al., 1996, Nakahara et al., 2003,) and coastal marine strains clade I and IV including *Synechococcus* sp. CC9311 used in this study, however the class CI-Fba is absent in open ocean oligotrophic marine clade III strains genome (Rogers et al., 2007). CI- and CII-Fbas share no homology between their gene sequences (Kelly and Tolan, 1986, Marsh and Lebherz, 1992, Pelzer-Reith et al., 1993), and it is proposed that the genes for the two Fba classes evolved independently from each other (Rogers et al., 2007). We found that only class II Fba protein interacted with Sync_1857, suggesting that phosphorylation could be important to switch on/off the glycolysis, Calvin cycle and gluconeogenesis pathways in cyanobacteria.

GlpX catalyzes the formation of fructose 6-phosphate by hydrolysis of fructose 1,6-biphosphate. It is highly conserved among prokaryotes and eukaryotes and in cyanobacteria, it takes part in the photosynthetic reduction of carbon (PRC) cycle, and gluconeogenesis (Marcus et al., 1987). Pgk is an enzyme that catalyzes the reversible transfer of a phosphate group from 1,3-biphosphoglycerate to ADP producing 3-phosphoglycerate and ATP. The phosphorylation site(s) of Pgk have been identified in *E. coli* (Macek et al., 2008) and the cyanobacteria *Synechocystis* PCC 6803 (Spat et al., 2015). Eno is a metalloenzyme which catalyzes the conversion of 2-phosphoglycerate (2-PG) to phosphoenolpyruvate (PEP), the penultimate step of glycolysis. Eno has been known to be phosphorylated by the conserved Ser/Thr kinase PrkC and dephosphorylated by phosphatase PrpC in *Bacillus subtilis* (Gaidenko et al., 2002). In marine cyanobacteria, kinases involved in phosphorylation of Eno remains unknown, but we have detected the interaction of phosphatase Sync_1857 with Eno.

Three oxidative pentose phosphate (OPP) pathway proteins, namely glucose 6-phosphate dehydrogenase (Zwf), 6-phosphogluconolactonase (Pgl) and phosphoribulokinase (Prk) (Sync_1711) were detected in our AP-MS phosphatase study (Figure 4.6). The former two

proteins have been identified in Sync_1857 PPIs network while the latter one protein in Sync_2828 PPIs network. Generally, cyanobacteria oxidize carbohydrates via two parallel pathways: glycolysis and the oxidative pentose phosphate pathway (OPP) pathway. The OPP pathway is a process of glucose turnover that produces NADPH as reducing equivalents and pentoses as essential parts of nucleotides. There are two different phases in the pathway: one is the irreversible oxidative phase in which glucose-6-phosphate is converted to ribulose-5-phosphate by oxidative decarboxylation with the production of NADPH. The second phase is the reversible non-oxidative phase in which phosphorylated sugars are interconverted to generate xylulose-5-phosphate, ribulose-5-Phosphate, and ribose-5-phosphate. Regulation of the activity of Zwf, the enzyme that controls the entry of carbon into the OPP cycle, is complex and not well understood (Hagen and Meeks, 2001). We have identified a potential binding protein of Sync_2828, named Sync_1001, it's gene was predicted to encode OpcA protein (Figure 4.4b). The *sync_1001* gene is located adjacent to the *zwf* gene and is proposed to play a role in the functional assembly of Zwf protein (Hagen and Meeks, 2001). After catalysis reaction of glucose-6-phosphate by enzyme Zwf, Pgl catalyzes the second enzymatic reaction in OPP pathway that hydrolyzes 6-phosphogluconolactonase to 6-phosphogluconic acid.

Prk is an essential photosynthetic enzyme that catalyzes the ATP-dependent phosphorylation of ribulose 5-phosphate (RuP) into ribulose 1,5-biphosphate (RuBP), both intermediates in the Calvin cycle. It plays a role to regenerate RuBP, which is the initial substrate and CO_2-acceptor molecule of the Calvin Cycle. Therefore, Prk activity is an important indicator of metabolic rate in organisms like cyanobacteria for which carbon fixation is the vital key to their survival (Miziorko, 2000). The identification of these key enzymes in the glycolytic and OPP pathways in our phosphatase PPIs study suggests a potential key role of protein kinase/phosphatase activity in regulating carbon metabolism of cyanobacteria.
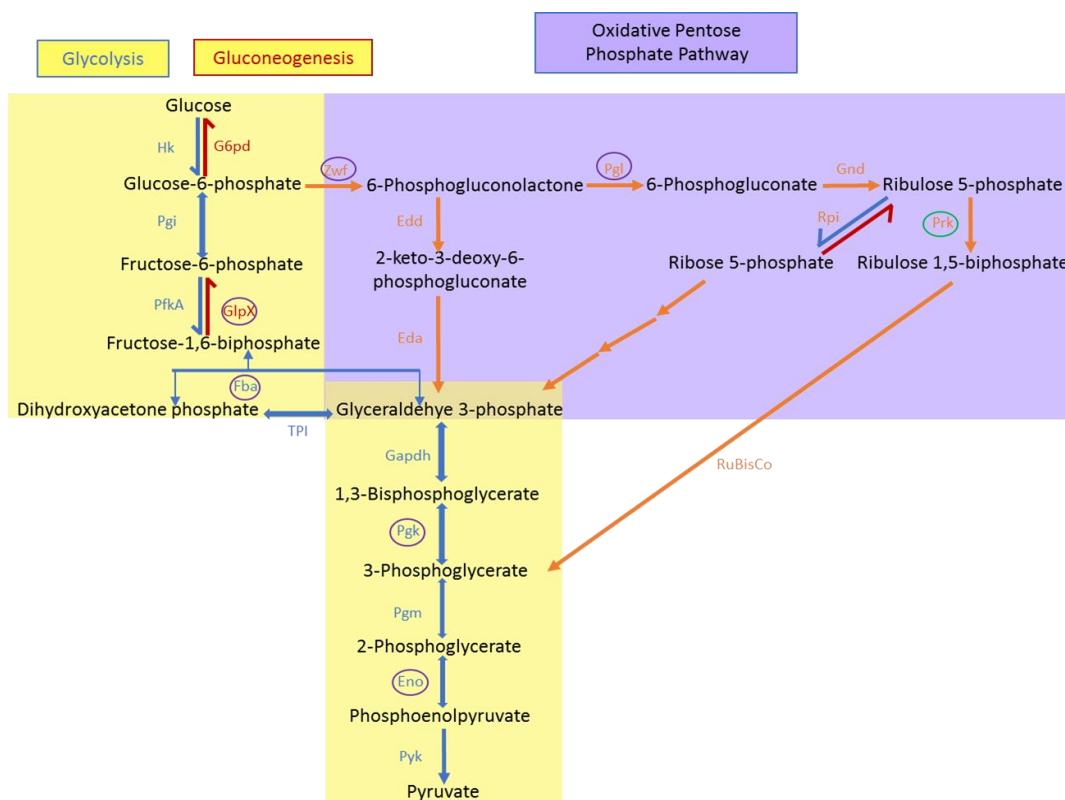
Figure 4.6. Glycolysis, gluconeogenesis and oxidative pentose phosphate (OPP) pathway in cyanobacteria. The background yellow colour represents the glycolysis and gluconeogenesis pathway, while purple represents the oxidative pentose phosphate (OPP) pathway. The arrow direction represents enzyme catalysis reaction. The arrow colour and enzyme word font colour represents different metabolic pathways: blue represents glycolysis pathway; red, represent gluconeogenesis pathway; orange represents OPP pathway. Purple circles represent proteins identified in this study to bind to Sync_1857 and green circles represent proteins identified in this study to bind to Sync_2828. Abbreviations: Hk, hexokinase; G6pd, glucose-6-phosphate dehydrogenase; Pgi, phosphoglucoisomerase; PfkA, 6-phosphofrucktokinase; GlpX, fructose-1,6-biphosphatase, class II; Fba, fructose-biphosphate aldolase; Tpi, triosephosphate isomerase; Gapdh, glyceraldehyde 3-phosphate dehydrogenase; Pgk, phosphoglycerate kinase; Pgm, phosphoglycerate mutase; Eno, phosphopyruvate hydratase; Pyk, pyruvate kinase; Zwf, glucose-6-phosphate dehydrogenase; Edd, phosphogluconate dehydratase; Eda, Entner-Douderoff aldolase; Pgl, 6-phosphogluconolactonase; Gnd, 6-phosphogluconate dehydrogenase; Rpi, ribose-5-phosphate isomerase; Prk, phosphoribulokinase; RuBisCO, ribulose biphosphate carboxylase oxygenase. Notes: only the oxidative phase is showed for OPP pathway.

CONCLUSIONS

Few phosphoproteins found in other species can be directly used as a reference for the phosphorylated sites in marine cyanobacteria. However, every species has varied phosphorylation events that are induced by external and/or internal stimuli in cells. Organisms from the same genus can also have different forms of phosphorylation, for example glycerol kinase proteins were phosphorylated in the slow growing *Mycobacterium bovis* BCG, while this is not the case in the fast growing *Mycobacterium smegmatis* (Nakedi et al., 2015). Therefore, phosphorylation sites of proteins in marine *Synechococcus* remains to be explored to generate a clearer picture of phosphorylation events in the cellular metabolism of marine cyanobacteria.

Our study highlights the PPIs of phosphatases in marine *Synechococcus*. Both Sync_1857 and Sync_2828 show potential active involvement in cellular metabolic activities in cyanobacteria cells. Their involvement in glycolysis, gluconeogenesis and oxidative phosphate pathway, suggest their vital roles in carbon metabolism. Phosphatase Sync_1857, a metallophosphodiesterase enzyme has abroad range of interacted proteins suggesting its pivotal role in cellular metabolism in cells. Urea periplasmic ABC transporter, UrtA was shown to bind with phosphatase Sync_1857, suggesting phosphorylation may be an important PTM event monitoring the transport of urea into marine cyanobacterial cells. Meanwhile, Sync_2828 bound with important antioxidants thioredoxin peroxidase and superoxide dismutase proteins suggesting its role in stress signaling regulation in marine cyanobacteria. Overall, phosphatases play a pivotal role in cellular metabolism and possibly have a major role in carbon metabolism in marine cyanobacteria.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Supplementary Info 4.1. PCR amplification primers.

Gene	Primer (5'→3') *	Length	Tm (^{0}C)
sync_1857_F	AAGTTCTGTTTCAGGGCCCGatgcaaacaagtcg	34	66
sync_1857_R	CACAAACTGGTCTAGAAAGCTTTAttaagagccataggac	40	64
sync_2828_F	AAGTTCTGTTTCAGGGCCCGttgacctccggcagg	35	70
sync_2828_R	CTGGTCTAGAAAGCTTTAtcaggattcagcacgtttccaaggg	43	68

(Notes: F: forward primer; R: reverse primer; * Uppercase letter represents extension primer, while lowercase letter represents template DNA primer)

Supplementary Info 4.2. GO term of molecular function and biological process of prey proteins identified in this AP-MS study.

Available in the attached file named Suppl 4.2_GO term of molecular function and biological process of prey proteins identified in this AP-MS study.

Supplementary Info 4.3. Complete lists of all prey proteins identified in this AP-MS study.

Available in the attached file named Suppl 4.3_Complete lists of all prey proteins identified in this AP-MS study.

Supplementary Info 4.4. Proteins identified in the negative control of this AP-MS study.

Available in the attached file named Suppl 4.4_Proteins identified in the negative control of this AP-MS study.

Supplementary Info 4.5. Protein-protein and protein-chemical interaction nodes scoring of prey proteins of phosphatases Sync_1857 and Sync_2828 from STITCH network.

Available in the attached file named Suppl 4.5_Protein-protein and protein-chemical interaction nodes scoring.

Supplementary Info 4.6. Enzyme classification of protein identified in this AP-MS.
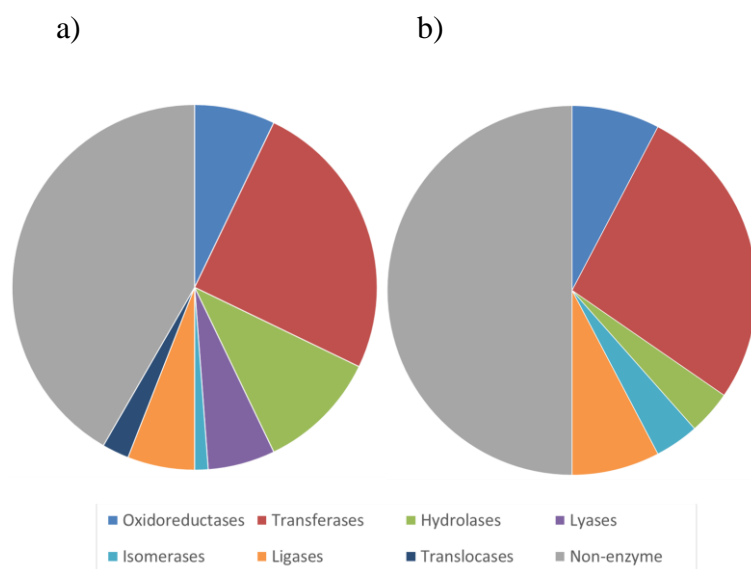
a) Enzyme classification of protein identified in this AP-MS study of protein Sync_1857.

b) Enzyme classification of protein identified in this AP-MS study of protein Sync_2828.

REFERENCES

ALMAGRO ARMENTEROS, J. J., TSIRIGOS, K. D., SONDERBY, C. K., PETERSEN, T. N., WINTHER, O., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol,* 37**,** 420-423.

AMSTER-CHODER, O. & WRIGHT, A. 1992. Modulation of the dimerization of a transcriptional antiterminator protein by phosphorylation. *Science,* 257**,** 1395-1398.

ARAVIND, L. & KOONIN, E. V. 1998. Phosphoesterase domains associated with DNA polymerases of diverse origins. *Nucleic Acids Res,* 26**,** 3746-3752.

BRADFORD, M. M. 1976. Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.,* 72**,** 248-254.

CHEN, S., YAKUNIN, A.F., KUZNETSOVA, E., BUSSO, D., PUFAN, R., PROUDFOOT, M., KIM, R. & KIM, S. 2004. *J Biol Chem,* 279, 31854-31862.

CHEN, Z., YANG, M. K., LI, C. Y., WANG, Y., ZHANG, J., WANG, D. B., ZHANG, X. E. & GE, F. 2014. Phosphoproteomic analysis provides novel insights into stress responses in *Phaeodactylum tricornutum*, a model diatom. *J Proteome Res,* 13**,** 2511-23.

CHRISTIE-OLEZA, J. A., ARMENGAUD, J., GUERIN, P. & SCANLAN, D. J. 2015. Functional distinctness in the exoproteomes of marine *Synechococcus*. *Environ Microbiol,* 17**,** 3781-3794.

COHEN, P. 1989. The structure and regulation of protein phosphatases. *Annu Rev Biochem,* 58**,** 453-508.

CONNELLY, J. C. & LEACH, D. R. F. 2002. Tethering on the brink- the evolutionarily conserved Mre11–Rad50 complex. *Trends in Biochemical Sciences,* 27**,** 410-418.

EMANUELSSON, O., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc,* 2**,** 953-71.

GAIDENKO, T. A., KIM, T. J. & PRICE, C. W. 2002. The PrpC Serine/Threonine Phosphatase and PrkC Kinase Have Opposing Physiological Roles in Stationary-Phase *Bacillus subtilis* Cells. *J Bacteriol ,* 184**,** 6109-6114.

GEFFLAUT, T., BLONSKI, C., PERIE, J. & WILLSON, M. 1995. Class I aldolases-Substrate specificity, mechanism, inhibitors and structural aspects. *Prog. Biophys. Molec. Biol.,* 63**,** 301-340.

HAAKE, V., ZRENNER, R., SONNEWALD, U. & STITT, M. 1998. A moderate decrease of plastid aldolase activity inhibits photosynthesis, alters the levels of sugars and starch, and inhibits growth of potato plants. *Plant J,* 14**,** 147-157.

HAGEN, K. D. & MEEKS, J. C. 2001. The unique cyanobacterial protein OpcA is an allosteric effector of glucose-6-phosphate dehydrogenase in *Nostoc punctiforme* ATCC 29133. *J Biol Chem,* 276**,** 11477-86.

HENZE, K., MORRISON, H. G., SOGIN, M. L. & MULLER, M. 1998. Sequence and phylogenetic position of a class II aldolase gene in the amitochondriate protist, *Giardia lamblia. Gene,* 222**,** 163-168.

HERBIG, K. CHIANG, E. P., LEE, L. R., HILLS. J., SHANE, B., STOVER, P. J. 2002. *J Biol Chem,* 277, 38381-38389.

HERZOG, F., KAHRAMAN, A., BOEHRINGER, D., MAK, R., BRACHER, A., WALZTHOENI, T., LEITNER, A., BECK, M., HARTL, F.-U., BAN, N., MALMSTROM, L. & AEBERSOLD, R. 2012. Structural Probing of a Protein Phosphatase 2A Network by Chemical Cross-Linking and Mass Spectrometry. *Science,* 337**,** 1348-1352.

HONDORP, E. R. & MATTHEWS, R. G. 2004. Oxidative stress inactivates cobalamin-independent methionine synthase (MetE) in *Escherichia coli. . PLoS Biol,* 2**,** e336.

HUERTA-CEPAS, J., SZKLARCZYK, D., FORSLUND, K., COOK, H., HELLER, D., WALTER, M. C., RATTEI, T., MENDE, D. R., SUNAGAWA, S., KUHN, M., JENSEN, L. J., VON MERING, C. & BORK, P. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res,* 44**,** D286-93.

KANEKO, T., SATO, S., KOTANI, H., TANAKA, A., ASAMIRU, E., NAKAMURA, Y., MIYAJIMA, N., HIROSAWA, M., SUGIURA, M., SASAMOTO, S., KIMURA, T., HOSOUCHI, T., MATSUNO, A., MURAKI, A., NAKAZAKI, N., NARUO, K., OKUMURA, S., SHIMPO, S., TAKEUCHI, C., WADA, T., WATANABE, A., YAMADA, M., YASUDA, M. & TABATA, S. 1996. Sequence Analysis of the Genome of the Unicellular Cyanobacterium *Synechocystis* sp. Strain PCC6803. II. Sequence Determination of the Entire Genome and Assignment of Potential Protein-coding Regions. *DNA Res.,* 3**,** 185-209.

KELLY, P. M. & TOLAN, D. R. 1986. The Complete Amino Acid Sequence for the Anaerobically Induced Aldolase from Maize Derived from cDNA Clones. *Plant Physiol,* 82 1076-1080.

KEPPETIPOLA, N. & SHUMAN, S. 2008. A Phosphate-binding Histidine of Binuclear Metallophosphodiesterase Enzymes Is a Determinant of 2′,3′-Cyclic Nucleotide Phosphodiesterase Activity. *J Biol Chem,* 283**,** 30942-30949.

KHALID, M. F., DAMHA, M. J., SHUMAN, S. & SCHWER, B. 2005. Structure-function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. *Nucleic Acids Res,* 33**,** 6349-60.

KINTER, M. & SHERMAN, N. E. 2000. *Protein sequencing and identification using tandem mass spectrometry.*, John Wiley and Sons Inc.

KLUMPP, S. & KRIEGLSTEIN, J. 2002. Phosphorylation and dephosphorylation of histidine residues in proteins. *Eur. J. Biochem.,* 269**,** 1067-1071.

KUHN, M., VON MERING, C., CAMPILLOS, M., JENSEN, L. J. & BORK, P. 2008. STITCH: interaction networks of chemicals and proteins. *Nucleic Acids Res,* 36**,** D684-8.

LAZAZZERA, B. A., BATES, D. M. & KILEY, P. 1993. The activity of the *Escherichia coli* transcription factor FNR is regulated by a change in oligomeric state. *Genes Dev,* 7**,** 1993-2005.

LOS, D. A., ZORINA, A., SINETOVA, M., KRYAZHOV, S., MIRONOV, K. & ZINCHENKO, V. V. 2010. Stress sensors and signal transducers in cyanobacteria. *Sensors (Basel),* 10**,** 2386-415.

MACEK, B., GNAD, F., SOUFI, B., KUMAR, C., OLSEN, J. V., MIJAKOVIC, I. & MANN, M. 2008. Phosphoproteome Analysis of *E. coli* Reveals Evolutionary Conservation of Bacterial Ser/Thr/Tyr Phosphorylation. *Mol Cell Proteomics,* 7**,** 299-307.

MANN, M. & JENSEN, O. N. 2003. Proteomic analysis of post-translational modifications. *Nat Biotech,* 21**,** 255-261.

MARCUS, F., RITTENHOUSE, J., GONTERO, B. & HARRSCH, P. B. 1987. Function, structure and evolution of fructose-1,6-bisphosphatase. *Arch Biol Med Exp,* 20**,** 371-8.

MARSH, J. J. & LEBHERZ, H. G. 1992. Fructose-bisphosphate aldolases- an evolutionary history. *Trends Biochem Sci,* 17**,** 110-113.

MATANGE, N., PODOBNIK, M. & VISWESWARIAH, S. S. 2014. The non-catalytic "cap domain" of a mycobacterial metallophosphoesterase regulates its expression and localization in the cell. *J Biol Chem,* 289**,** 22470-81.

MCLOUGHLIN, S. Y., JACKSON, C., LIU, J. W. & OLLIS, D. L. 2004. Growth of *Escherichia coli* Coexpressing Phosphotriesterase and Glycerophosphodiester Phosphodiesterase, Using Paraoxon as the Sole Phosphorus Source. *Appl Environ Microbiol,* 70**,** 404-412.

MIZIORKO, H. M. 2000. Phosphoribulokinase: Current perspectives on the structure/function basis for regulation and catalysis. *In:* PURICH, D. L. (ed.) *Advances n Enzymology and Related Areas of Molecular Biology.* John Wiley and Sons.

MOREL, F. M. M., RUETER, J. G., ANDERSON, D. M. & GUILLARD, R. R. L. 1979. Aquil: a chemically defined phytoplankton culture medium for trace metal studies. *J Phycol,* 15**,** 135-141.

NAKAHARA, K., YAMAMOTO, H., MIYAKE, C. & YOKOTA, A. 2003. Purification and Characterization of Class-I and Class-II Fructose-1,6-bisphosphate Aldolases from the Cyanobacterium *Synechocystis sp*. PCC 6803. *Plant Cell Physiol. ,* 44**,** 326-333.

NAKEDI, K. C., NEL, A. J., GARNETT, S., BLACKBURN, J. M. & SOARES, N. C. 2015. Comparative Ser/Thr/Tyr phosphoproteomics between two mycobacterial species: the fast growing *Mycobacterium smegmatis* and the slow growing *Mycobacterium bovis* BCG. *Front Microbiol,* 6**,** 237.

PAITHOONRANGSARID, K., SHOUMSKAYA, M. A., KANESAKI, Y., SATOH, S., TABATA, S., LOS, D. A., ZINCHENKO, V. V., HAYASHI, H., TANTICHAROEN, M., SUZUKI, I. & MURATA, N. 2004. Five histidine kinases perceive osmotic stress and regulate distinct sets of genes in *Synechocystis*. *J Biol Chem,* 279**,** 53078-86.

PELZER-REITH, B., PENGER, A. & SCHNARRENBERGER, C. 1993. Plant aldolase-cDNA and deduced amino-acid sequences of the chloroplast and cytosol enzyme from spinach. *Plant Mol Biol,* 21**,** 331-340.

PEREIRA, S. F., GOSS, L. & DWORKIN, J. 2011. Eukaryote-like serine/threonine kinases and phosphatases in bacteria. *Microbiol Mol Biol Rev,* 75**,** 192-212.

PITTERA, J., HUMILY, F., THOREL, M., GRULOIS, D., GARCZAREK, L. & SIX, C. 2014. Connecting thermal physiology and latitudinal niche partitioning in marine *Synechococcus*. *ISME J,* 8**,** 1221-36.

RAHIKAINEN, M., PASCUAL, J., ALEGRE, S., DURIAN, G. & KANGASJARVI, S. 2016. PP2A Phosphatase as a Regulator of ROS Signaling in Plants. *Antioxidants (Basel),* 5.

RICHARDSON, T. L. & JACKSON, G. A. 2007. Small phytoplankton and carbon export from the surface ocean. *Science,* 315**,** 838-40.

ROGERS, M. B., PATRON, N. J. & KEELING, P. J. 2007. Horizontal transfer of a eukaryotic plastid-targeted protein gene to cyanobacteria. *BMC Biology,* 5.

RUBIOLO, J. A., LÓPEZ-ALONSO, H., ALFONSO, A., VEGA, F. V., VIEYTES, M. R. & BOTANA, L. M. 2014. Bioengineered protein phosphatase 2A. *Bioengineered,* 4**,** 72-77.

RUTTER, W. J. 1964. Evolution of aldolase. *Fed Proc Fed Amer Soc Exp Biol. ,* 23**,** 1248-1257.

SCANLAN, D. J., OSTROWSKI, M., MAZARD, S., DUFRESNE, A., GARCZAREK, L., HESS, W. R., POST, A. F., HAGEMANN, M., PAULSEN, I. & PARTENSKY, F. 2009. Ecological genomics of marine picocyanobacteria. *Microbiol Mol Biol Rev,* 73**,** 249-99.

SCHNARRENBERGER, C., GROSS, W., PLEZER-REITH, B. & WIEGAND, S. 1990. Evolution of isozymes of sugar phosphate metabolism in green algae. *Prog Clin Biol Res,* 344**,** 743-764.

SCHWELBERGER, H. G., KOHLWEIN, S. D. & PALTAUF, F. 1989. Molecular cloning, primary structure and disruption of the structural gene of aldolase from *Saccharomyces cerevisiae*. *Eur J Biochem,* 180**,** 301-308.

SHENOLIKAR, S. & NAIM, A. C. 1991. Protein phosphatases: recent progress. *Adv Second Messenger Phosphoprotein Res,* 23**,** 1-121.

SHENOY, A. R., SREENATH, N., PODOBNIK, M., KOVACEVIC, M. & VISWESWARIAH, S. S. 2005. The Rv0805 Gene from *Mycobacterium tuberculosis* Encodes a 3',5'-Cyclic Nucleotide Phosphodiesterase- Biochemical and Mutational Analysis. *Biochemistry,* 44**,** 15695-15704.

SHI, L., POTTS, M. & KENNELLY, P. J. 1998. The serine, threonine, and/or tyrosine-specific protein kinases and protein phosphatases of prokaryotic organisms- a family portrait. *FEMS Microbiol Rev,* 22**,** 229-253.

SMITH, C. A., BAN, D., PRATIHAR, S., GILLER, K., PAULAT, M., BECKER, S., GRIESINGER, C., LEE, D. & DE GROOT, B. L. 2016. Allosteric switch regulates protein-protein binding through collective motion. *PNAS,* 113**,** 3269-74.

SPAT, P., MACEK, B. & FORCHHAMMER, K. 2015. Phosphoproteome of the cyanobacterium *Synechocystis* sp. PCC 6803 and its dynamics during nitrogen starvation. *Front Microbiol,* 6**,** 248.

STANDISH, A. J. & MORONA, R. 2014. The role of bacterial protein tyrosine phosphatases in the regulation of the biosynthesis of secreted polysaccharides. *Antioxid Redox Signal,* 20**,** 2274-89.

SU, Z., MAO, F., DAM, P., WU, H., OLMAN, V., PAULSEN, I. T., PALENIK, B. & XU, Y. 2006. Computational inference and experimental validation of the nitrogen assimilation regulatory network in cyanobacterium *Synechococcus* sp. WH 8102. *Nucleic Acids Res,* 34**,** 1050-65.

TIAN, M., CHEN, X., XIONG, Q., XIONG, J., XIAO, C., GE, F., YANG, F. & MIAO, W. 2014. Phosphoproteomic Analysis of Protein Phosphorylation Networks in *Tetrahymena thermophila*, a Model Single-celled Organism. *Mol Cell Proteomics,* 13**,** 503-519.

TROTTA, A., WRZACZEK, M., SCHARTE, J., TIKKANEN, M., KONERT, G., RAHIKAINEN, M., HOLMSTROM, M., HILTUNEN, H. M., RIPS, S., SIPARI, N., MULO, P., WEIS, E., VON SCHAEWEN, A., ARO, E. M. & KANGASJARVI, S. 2011. Regulatory subunit B'gamma of protein phosphatase 2A prevents unnecessary defense reactions under low light in *Arabidopsis*. *Plant Physiol,* 156**,** 1464-80.

VERMA, R., PINTO, S. M., PATIL, A. H., ADVANI, J., SUBBA, P., KUMAR, M., SHARMA, J., DEY, G., RAVIKUMAR, R., BUGGI, S., SATISHCHANDRA, P., SHARMA, K., SUAR, M., TRIPATHY, S. P., CHAUHAN, D. S., GOWDA, H., PANDEY, A., GANDOTRA, S. & PRASAD, T. S. 2017. Quantitative Proteomic and Phosphoproteomic Analysis of H37Ra and H37Rv Strains of *Mycobacterium tuberculosis. J Proteome Res,* 16**,** 1632-1645.

WANG, L. K., SMITH, P. & SHUMAN, S. 2013. Structure and mechanism of the 2',3' phosphatase component of the bacterial Pnkp-Hen1 RNA repair system. *Nucleic Acids Res,* 41**,** 5864-73.

WESSEL, D. & FLUGGE, U. I. 1984. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem,* 138**,** 141-143.

YANG, M. K., QIAO, Z. X., ZHANG, W. Y., XIONG, Q., ZHANG, J., LI, T., GE, F. & ZHAO, J. D. 2013. Global phosphoproteomic analysis reveals diverse functions of serine/threonine/tyrosine phosphorylation in the model cyanobacterium *Synechococcus* sp. strain PCC 7002. *J Proteome Res,* 12**,** 1909-23.

YU, N. Y., WAGNER, J. R., LAIRD, M. R., MELLI, G., REY, S., LO, R., DAO, P., SAHINALP, S. C., ESTER, M., FOSTER, L. J. & BRINKMAN, F. S. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics,* 26**,** 1608-15.

ZHANG, C.-C., JANG, J., SAKR, S. & WANG, L. 2005. Protein Phosphorylation on Ser, Thr and Tyr Residues in Cyanobacteria. *JMol Microb Biotech,* 9**,** 154-166.

ZORINA, A., STEPANCHENKO, N., NOVIKOVA, G. V., SINETOVA, M., PANICHKIN, V. B., MOSHKOV, I. E., ZINCHENKO, V. V., SHESTAKOV, S. V., SUZUKI, I., MURATA, N. & LOS, D. A. 2011. Eukaryotic-like Ser/Thr protein kinases SpkC/F/K are involved in phosphorylation of GroES in the Cyanobacterium *Synechocystis. DNA Res,* 18**,** 137-51.

ZWIRGLMAIER, K., JARDILLIER, L., OSTROWSKI, M., MAZARD, S., GARCZAREK, L., VAULOT, D., NOT, F., MASSANA, R., ULLOA, O. & SCANLAN, D. J. 2008. Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. *Environ Microbiol,* 10**,** 147-161.

Chapter 5: Conclusions and future perspectives

5.1 Conclusions

Chapter 2 investigates the membrane proteomes of marine *Synechococcus* spp. This is the first membrane proteomic work conducted on marine *Synechococcus*. Four marine *Synechococcus* strains, namely, CC9311 (Clade I), CC9605 (Clade II), WH8102 (Clade III) and CC9902 (Clade IV) showed different adaptation strategies for nutrients/elements acquisition in the oceanic water. Marine *Synechococcus* WH8102 shows high expression of transporters involved in iron, phosphorus and nitrogen uptake. This is consistent with an adaptation strategy of scavenging nutrients or substrates found in low concentrations in their oligotrophic living environment. Coastal mesotrophic strains CC9311 and CC9902 display different adaptation strategies though they share similar ecological niches: CC9311 showed high expression of transporters involved in nitrogen acquisition, and mechanosensitive ion channels for osmotic regulation, while CC9902 strain demonstrates high expression of transporters for phosphorus and iron acquisition. CC9902 appears to have the highest acquisition for phosphorus in nutrients replete conditions among the four marine *Synechococcus* strains, although the reasons for this remain unclear. *Synechococcus* CC9605, a representative of Clade II shows high expression of transporters involved in nitrogen acquisition, while low or moderate expression of other transporters. This may suggest why clade II strains were the most ubiquitous clade found around the global ocean as they have a generalist strategy that has minimized energy investment in nutrient acquisition. One protein annotated as a protein phosphatase 2C (Sync_2505 and Syncc9902_0387) was very highly expressed in the coastal mesotrophic strains CC9311 and CC9902, suggesting it serves as a pivotal function, for strains living in temperate mesotrophic environments. However, the exact role of protein phosphatase 2C in marine cyanobacteria remains to be investigated.

Although the putative protein phosphatase 2C appears to play a key role in coastal mesotrophic *Synechococcus* strains, there is very little experimental data available on the roles of protein phosphatases in marine cyanobacteria. *Synechococcus* sp. CC9311 has a large predicted gene of phosphatases and response regulators compared to other marine picocyanobacteria, hence it was chosen as a model strain for a global investigation of the function of protein phosphatases and response regulators.

We found nine protein phosphatase and seventeen response regulator genes in marine *Synechococcus* sp. CC9311. The protein phosphatase and response regulator genes were amplified, cloned and transformed in an *E. coli* expression system (Chapter 3). We were able to successfully purify soluble protein for two putative protein phosphatases, Sync_1857 and Sync_2828, and 9 response regulators, Sync_0574, Sync_0586, Sync_0707, Sync_0807, Sync_1116, Sync_1219, Sync_1232, Sync_2591 and Sync_2642. SEC analyses show that the protein phosphatase Sync_2828 exists as dimeric and hexameric forms in reducing and non-reducing conditions while Sync_1857 has four oligomeric forms: monomer, dimer, tetramer and hexamer in non-reducing conditions and exists as two oligomeric forms (monomer and dimer) in reducing conditions. In contrast, all 9 purified response regulator proteins exist as monomeric form, while Sync_2591 also present in dimeric forms. All of the purified proteins were used for pull-down assays to study PPIs between the protein phosphatases and response regulators. Mass spectrometry analyses from pull-down assays suggested that the protein phosphatases Sync_1857 and Sync_2828 interact with 7 out of the 9 response regulators. In short, pull-down assays coupled with mass spectrometry can be used as an approach for *in vitro* phosphatase PPIs study, however, a bi-directional pull-down assay is recommended to enhance validation of the results. Although this study was inspired by the high expression of the putative protein phosphatase 2C (Chapter 2), as this protein was not expressed in soluble form, it was not possible to further investigate its role in this thesis.

The putative protein phosphatases Sync_1857 and Sync_2828 were used to investigate PPIs of protein phosphatases at a global level in marine cyanobacteria *Synechococcus* sp. CC9311 cells (Chapter 4). The AP-MS approach was used for the PPIs investigation of the putative protein phosphatases: phosphatase was used as baits to affinity purify prey proteins from cytoplasmic lysate that interact with it. Based on this data, a PPI network of proteins Sync_1857 and Sync_2828 was generated and the AP-MS results showed that both protein phosphatases displayed PPIs with proteins involved in glycolysis, gluconeogenesis and oxidative pentose pathway suggesting their possible important roles in controlling carbon metabolism in cells. Sync_1857 interacted with a periplasmic urea ABC transporter, UrtA suggesting phosphorylation might be an important PTM event for transportation of urea into marine cyanobacterial cells. As a metallophosphodiestearase family member, Sync_1857 displayed a broad diversity of PPIs. Meanwhile, Sync_2828 displayed interactions with stress related proteins suggesting a role in the regulation of stress responses in cyanobacteria. Response regulators are present in low abundance in

marine cyanobacterial cells from other studies (Mackey et al. 2013; Varkey et al. 2016). This might be the main reasons why response regulators were not present in the prey proteins list in this PPIs study. Therefore, the AP-MS experimental results of phosphatases with cytoplasmic lysates of *Synechococcus* sp. CC9311 (Chapter 4) cannot be used to cross-check with the *in vitro* PPIs results of phosphatases with response regulators (Chapter 3). Nonetheless, the experimental results of Chapter 4 have revealed the unknown diverse potential roles of phosphatases in cyanobacterial cells.

5.2 Future perspectives

Future work should be carried out on further investigation of the marine cyanobacteria cell signalling. This section suggests a number of possible approaches to enhance the experimental work in Chapter 3 and Chapter 4 on the study of cyanobacterial protein phosphatases.

Phosphorylation is one of the major PTMs in bacteria, involved in a wide variety of metabolic, developmental and stress responses in cells (Bourett et al., 1991, McCartney et al., 1997, Pei et al., 2017). However, little is known about the phosphorylation sites in marine cyanobacteria and most of the phosphorylation events we known are derived from the freshwater cyanobacteria such as *Synechocystis* sp. PCC6803 (Spat et al., 2015), and euryhaline species *Synechococcus* sp. PCC7002 (Yang et al., 2013). Therefore, phosphoproteomics study on marine *Synechococcus* should be conducted to explore the phosphorylation sites and physiological conditions of marine cyanobacteria undergo this phosphorylation modification. A few different phosphoproteomic approaches to analyse phosphorylation sites of proteomes have been developed (McLachlin and Chait, 2001, Thingholm et al., 2009, Larsen et al., 2005, Thingholm et al., 2006, Villen and Gygi, 2008, Schroeder et al., 2004, Chi et al., 2007). These approaches include enhance analysis of phosphoproteins/phosphopeptides by utilizing various types of chemistries for enrichment of phosphorylated peptides and/or proteins, and different technologies for ionization, fragmentation and mass-analysis of phosphorylated peptides. For example, enrichment of phosphopeptides can be enhanced by immobilized metal affinity chromatography, metal oxide surfaces using titanium oxide (TiO_2) and/or antibody-based enrichment (Silva-Sanchez et al. 2014).

The identification of phosphorylation sites in proteins/peptides and conditions when proteins undergo phosphorylation/dephosphorylation modifications would definitely

enhance the phosphatase PPI studies. Once protein phosphorylation sites are known, we can design an experiment to target specific phosphorylation site(s) and answer a more specific biological question addressed on interested protein(s). For instance, this enables us to determine the precise functional roles of the proteins and their physiological states in cells. For Chapter 3, *in vitro* PPIs of marine cyanobacteria of phosphatases and response regulators can be enhanced by bi-directional pull-down assay which use phosphatase as baits and response regulators as preys (as in Chapter 3) and vice versa. Besides that, the remaining response regulators which have not been tested should be conducted with pull-down assay with phosphatase, to generate a clearer picture of phosphatase functional roles. Moreover, pull-down assay should be conducted on all the oligomer forms of putative phosphatase protein Sync_1857 to check any difference among the PPIs between different phosphatase oligomers with response regulators.

Protein phosphatase interactions with response regulators could be transient, hence further analysis with surface plasmon resonance is recommended to investigate the binding affinity of phosphatases with response regulators. Surface plasmon resonance is an optical approach that uses the evanescent wave phenomenon to measure changes in refractive index close to a sensor surface (McDonnell, 2001). It allows the analysis of receptor ligand interactions with a broad range of different molecular weights, affinities and binding rate (Davis and Wilson, 2000, Quinn et al., 2000). By using a surface plasmon resonance approach, transient interaction between proteins can be easily detected and the conditions required for such reactions to be occurred can be optimized by different buffers used and/or abiotic factors change. Previous study has utilized surface plasmon resonance as an approach for the PPI study of protein phosphatase in eukaryotes (Toth et al., 2000) and hence it should be tested for the applicability in cyanobacterial study as well.

Cell signalling regulation in cyanobacteria can be enhanced by combined studies of kinases and phosphatases. Reversible phosphorylation of proteins is monitored by concerted reactions of kinases and phosphatases; Kinases catalyses the transfer of γ-phosphate from ATP to the protein substrate while phosphatases remove the phosphate group from a phosphorylated protein substrate. In the Chapter 4, we identified different proteins that bound to phosphatases; however, the kinases that phosphorylate the proteins remain unknown. Therefore, the kinases PPIs analysis should be conducted in the future, combining with the phosphatase PPIs analysis to generate a comprehensive cell signalling regulation network in cyanobacteria.

In addition, PPIs study approaches can be enhanced by chemical cross-linking mass spectrometry approach. This approach is a powerful approach to study protein complex structure and also PPIs (Kluger and Alagic, 2004, Herzog et al., 2012, Navare et al., 2015). Cross-linking mass spectrometry provides distance constraints to study the PPIs which are pivotal for the understanding of protein function. It is based on the usage of a bifunctional reagent that covalently links two reactive residues (depends on cross-linkers used) under near-physiological conditions, to generate an informative linkage in a protein or between proteins. The linkage has a defined length, for example, the bissulfosuccinimidyl suberate cross-linker introduces a maximal distance constraint between the linked residues of 30Å (sum of the length of the lysine side chains and the bissulfosuccinimidyl suberate cross-linker) (Barysz and Malmstrom, 2018) and provides information about which residues in the proteins are within the distant constraint defined by the length of the selected cross-linker. The stable covalent linkage irreversibly locks the interacting proteins and tolerates stringent processing, potentially enabling the identification of weak and transient PPIs with high specificity, accuracy and reliability (Sinz, 2010). Cross-linkers with different specificity on amino acid residues have been developed (reactive toward Lys, Asp, Glu, Cys, Arg, nonspecific), although the lysine reactive are most commonly used (Barysz and Malmstrom, 2018).

PPI study is a big challenge in biological science research owing to the dynamic and complex structure of proteins, limitation of experimental techniques, difficulty of bioinformatics analysis etc. Nonetheless, the comprehensive mapping of PPIs is highly desired for one to gain deep insight into bacterial cells' biological pathways and processes. Therefore, extensive research on cell signalling network of marine cyanobacteria should be conducted to generate a comprehensive protein interactome network of marine cyanobacterium and using it as a model to deepen our understanding of the ocean ecosystem.

References

BARYSZ, H. M. & MALMSTROM, J. 2018. Development of Large-scale Cross-linking Mass Spectrometry. *Mol Cell Proteomics,* 17**,** 1055-1066.

BOURETT, R. B., BORKOVICH, K. A. & SIMON, M. I. 1991. Signal transduction pathways involving protein phosphorylation in prokaryotes. *Annu Rev Biochem,* 60**,** 401-441.

CHI, A., HUTTENHOWER, C., GEER, L. Y., COON, J. J., SYKA, J. E. P., BAI, D. L., SHABANOWITZ, J., BURKE, D. J., TROYANSKAYA, O. G. & HUNT, D. F. 2007. Analysis of phosphorylation sites on proteins from *Saccharomyces cerevisiae* by electron transfer dissociation (ETD) mass spectrometry. *PNAS,* 104**,** 2193-2198.

DAVIS, T. M. & WILSON, W. D. 2000. Determination of the refractive index increments of small molecules for correction of surface plasmon resonance data. *Anal Biochem,* 284**,** 348-353.

HERZOG, F., KAHRAMAN, A., BOEHRINGER, D., MAK, R., BRACHER, A., WALZTHOENI, T., LEITNER, A., BECK, M., HARTL, F.-U., BAN, N., MALMSTROM, L. & AEBERSOLD, R. 2012. Structural Probing of a Protein Phosphatase 2A Network by Chemical Cross-Linking and Mass Spectrometry. *Science,* 337**,** 1348-1352.

KLUGER, R. & ALAGIC, A. 2004. Chemical cross-linking and protein–protein interactions—a review with illustrative protocols. *Bioinorg Chem,* 32**,** 451-472.

LARSEN, M. R., THINGHOLM, T. E., JENSEN, O. N., ROEPSTORFF, P. & JORGENSEN, T. J. D. 2005. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics,* 4**,** 873-886.

MACKEY, K. R., PAYTAN, A., CALDEIRA, K., GROSSMAN, A. R., MORAN, D., MCILVIN, M. & SAITO, M. A. 2013. Effect of temperature on photosynthesis and growth in marine *Synechococcus* spp. *Plant Physiol,* 163**,** 815-29.

MCCARTNEY, B., HOWELL, L. D., KENNELLY, P. J. & POTTS, M. 1997. Protein Tyrosine Phosphorylation in the Cyanobacterium *Anabaena* sp. Strain PCC 7120. *J Bacteriol ,* 179**,** 2314-2318.

MCDONNELL, J. M. 2001. Surface plasmon resonance- towards an understanding of the mechanisms of biological molecular recognition. *Curr Opin Chem Biol,* 5**,** 572-577.

MCLACHLIN, D. T. & CHAIT, B. T. 2001. Analysis of phosphorylated proteins and peptides by mass spectrometry. *Curr Opin Chem Biol,* 5**,** 591-602.

NAVARE, A. T., CHAVEZ, J. D., ZHENG, C., WEISBROD, C. R., ENG, J. K., SIEHNEL, R., SINGH, P. K., MANOIL, C. & BRUCE, J. E. 2015. Probing the protein interaction network of *Pseudomonas aeruginosa* cells by chemical cross-linking mass spectrometry. *Structure,* 23**,** 762-73.

PEI, H., HAN, S., YANG, S., LEI, Z., ZHENG, J. & JIA, Z. 2017. Phosphorylation of bacterial L9 and its functional implication in response to starvation stress. *FEBS Lett,* 591**,** 3421-3430.

QUINN, J. G., O''NEILL, S., DOYLE, A., MCATAMNEY, C., DIAMOND, D., MACCRAITH, B. D. & O'KENNEDY, R. 2000. Development and application of surface plasmon resonance-based biosensors for the detection of cell-ligand interactions. *Anal Biochem,* 281**,** 135-143.

SCHROEDER, M. J., SHABANOWITZ, J., SCHWARTZ, J. C., HUNT, D. F. & COON, J. J. 2004. A Neutral Loss Activation Method for Improved Phosphopeptide Sequence Analysis by Quadrupole Ion Trap Mass Spectrometry. *Anal Chem,* 76**,** 3590-3598.

SILVA-SANCHEZ, C., LI, H. & CHEN, S. 2014. Recent advances and challenges in plant phosphoproteomics. *Proteomics*, 15, 1127-1141.

SINZ, A. 2010. Investigation of protein-protein interactions in living cells by chemical crosslinking and mass spectrometry. *Anal Bioanal Chem,* 397**,** 3433-40.

SPAT, P., MACEK, B. & FORCHHAMMER, K. 2015. Phosphoproteome of the cyanobacterium *Synechocystis* sp. PCC 6803 and its dynamics during nitrogen starvation. *Front Microbiol,* 6**,** 248.

THINGHOLM, T. E., JENSEN, O. N. & LARSEN, M. R. 2009. Analytical strategies for phosphoproteomics. *Proteomics,* 9**,** 1451-1468.

THINGHOLM, T. E., JORGENSEN, T. J., JENSEN, O. N. & LARSEN, M. R. 2006. Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nat Protoc,* 1**,** 1929-35.

TOTH, A., KISS, E., HERBERG, F. W., GERGELY, P., HARTSHORNE, D. J. & ERDODI, F. 2000. Study of the subunit interactions in myosin phosphatase by surface plasmon resonance. *Eur J Biochem,* 267**,** 1687-1697.

VARKEY, D., MAZARD, S., OSTROWSKI, M., TETU, S. G., HAYNES, P. & PAULSEN, I. T. 2016. Effects of low temperature on tropical and temperate isolates of marine *Synechococcus*. *ISME J,* 10**,** 1252-63.

VILLEN, J. & GYGI, S. P. 2008. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat Protoc,* 3**,** 1630-8.

YANG, M. K., QIAO, Z. X., ZHANG, W. Y., XIONG, Q., ZHANG, J., LI, T., GE, F. & ZHAO, J. D. 2013. Global phosphoproteomic analysis reveals diverse functions of serine/threonine/tyrosine phosphorylation in the model cyanobacterium *Synechococcus* sp. strain PCC 7002. *J Proteome Res,* 12**,** 1909-23.