ORIGINAL ARTICLE

Microarray analysis of phosphate regulation in the marine cyanobacterium *Synechococcus* sp. WH8102

Sasha G Tetu[1], Bianca Brahamsha[2], Daniel A Johnson[3], Vera Tai[2], Katherine Phillippy[3,4], Brian Palenik[2] and Ian T Paulsen[1,3]

[1]*Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, New South Wales, Australia;* [2]*Marine Biology Research Division, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, CA, USA and* [3]*J Craig Venter Institute, Rockville, MD, USA*

Primary productivity of open ocean environments, such as those inhabited by marine picocyanobacteria, is often limited by low inorganic phosphate (P). To observe how these organisms cope with P starvation, we constructed a full genome microarray for *Synechococcus* sp. WH8102 and compared differences in gene expression under P-replete and P-limited growth conditions, including both early P stress, during extracellular alkaline phosphatase induction, and late P stress. A total of 36 genes showed significant upregulation ($>\log_2$ fold) whereas 23 genes were highly downregulated at the early time point; however, these changes in expression were maintained during late P stress for only 5 of the upregulated genes. Knockout mutants were constructed for genes SYNW0947 and SYNW0948, comprising a two-component regulator hypothesized to have a key function in regulating P metabolism. A high degree of overlap in the sets of genes affected by P stress conditions and in the knockout mutants supports this hypothesis; however, there is some indication that other regulators may be involved in this response in *Synechococcus* sp. WH8102. Consistent with what has been observed in many other cyanobacteria, the Pho regulon of this strain is comprised largely of genes for alkaline phosphatases, P transport or P metabolism. Interestingly, however, the exact composition and arrangement of the Pho regulon appears highly variable in marine cyanobacteria.

Introduction

Ocean environments comprise a vast component of the Earth's biosphere and have a key function in global biogeochemical processes. Unicellular marine cyanobacteria, including *Synechococcus* species, are estimated to constitute 20–40% of total marine chlorophyll biomass and carbon fixation, and hence significantly impact the carbon cycle and global climate processes (Partensky *et al.*, 1999). Cyanobacterial cell growth is frequently limited by deprivation of essential nutrients, leading to strong selective pressure toward adaptations that

aid survival during periods of nutrient limitation (Schwarz and Forchhammer, 2005). Genome sequences are now available for a number of marine cyanobacteria, including *Synechococcus* sp. WH8102, isolated from tropical Atlantic Ocean waters (Palenik *et al.*, 2003) and others (Palenik *et al.*, 2006; Dufresne *et al.*, 2008). The availability of whole-genome sequences allows comparative genomic analyses and techniques such as microarray analysis to be used to elucidate mechanisms by which cyanobacteria cope with challenges, such as nutrient deprivation, posed by the marine environment.

Phosphorus is an essential nutrient, necessary for the synthesis of a number of biomolecules, such as DNA, RNA, ATP and nucleotides, and in regulatory processes involving phosphorylation. Bacteria, both in aquatic and terrestrial environments, have been shown to encode regulatory mechanisms, controlling the uptake, storage and metabolism of phosphate (P). In *Escherichia coli*, where the response to P limitation has been studied in detail, a two-component system is known to regulate P

Phosphate regulation in *Synechococcus* sp. WH8102
SG Tetu *et al*

836

assimilation genes. Under P stress conditions, autophosphorylation of a specific histidine kinase (*phoR*) is triggered, followed by transfer of the P group to the cognate response regulator (*phoB*). Thus activated, *phoB* alters the expression of genes that constitute the Pho regulon, including alkaline phosphatase and P transporters, among others (Wanner, 1996; Baek and Lee, 2007).

In many oceans, especially oligotrophic areas, the dissolved P concentrations are generally in the nanomolar range and evidence for P stress and P limitation has been found (Palenik and Dyhrman, 1998; Wu *et al.*, 2000; Thingstad *et al.*, 2005; Zubkov *et al.*, 2007). Phosphate uptake regulation has been the focus of investigation in a number of cyanobacterial species with the goal of understanding the effects of P availability on distribution of these important marine photosynthetic bacteria.

Under P-starvation conditions, freshwater *Synechococcus* cells have been shown to significantly increase the uptake of P (Grillo and Gibson, 1979). Work in *Synechococcus* sp. PCC7942 identified a pair of genes encoding a sensory kinase and response regulator involved in P metabolism, named *sphS* (homologous to *E. coli phoR*) and *sphR* (homologous to *phoB*), respectively (Aiba *et al.*, 1993). Subsequent work showed that, as with *phoB*, *sphR* affected the expression of multiple genes by binding upstream of their promoter regions in response to P limitation, with genes *sphX* (encoding a binding protein of the P-specific ABC transporter) and *phoA* (alkaline phosphatase) constituting members of the Pho regulon in this organism (Aiba and Mizuno, 1994; Nagaya *et al.*, 1994). A P-regulated two-component sensory system has likewise been characterized in the marine cyanobacterium *Synechococcus* sp. WH7803 (Watson *et al.*, 1996) and in the freshwater cyanobacterium *Synechocystis* sp. PCC 6803 (Hirani *et al.*, 2001).

More recently, microarray analysis has been used to further define the Pho regulon in cyanobacterial species. In *Synechocystis* sp. PCC 6803 such experiments showed the expression of 12 genes to be strongly induced (whereas 1 gene was clearly repressed) under P-limited conditions (Suzuki *et al.*, 2004). The 12 upregulated genes were included in three clusters: *pst1* and *pst2*, which encode P transporters, and a third containing *phoA* and *nucH* (extracellular nuclease). Inactivation of either of *sphS* or *sphR* was shown to completely eliminate the P-limitation inducible expression (and repression), suggesting that this system is the dominant (or perhaps only) controller of gene expression in response to P limitation in this organism (Scanlan *et al.*, 1993; Hirani *et al.*, 2001).

Subsequent work on marine species of *Prochlorococcus* indicated that the Pho regulon of cyanobacteria, while containing an overlapping set of genes, can be quite variable. Comparison of the P stress response of two strains of *Prochlorococcus* MED4, which can grow on organic phosphates as a sole phosphorus source, and MIT9313, which cannot, revealed significant differences within this single species, perhaps due to variations in P availability in their marine niches (Martiny *et al.*, 2006). When 11 different *Prochlorococcus* strains were compared, the presence and arrangement of different genes comprising the Pho regulon appeared to be highly variable and not related to phylogeny, as defined by 16S rRNA. Instead, gene distribution suggested that P-acquisition genes were more often found in strains inhabiting P-limited environments and may be subject to horizontal gene transfer (Martiny *et al.*, 2006). Only a subset of these genomes were found to contain functional *phoB* and *phoR* homologues (Moore *et al.*, 2005; Martiny *et al.*, 2006). Likewise, *Synechococcus* sp. CC9311 obtained from a coastal, typically P-replete, environment was noted to lack PhoB/PhoR homologues (Palenik *et al.*, 2006). These findings suggest that some cyanobacteria regulate P uptake using other means, or that in some environments P levels may be high enough that selective pressure to maintain this system is reduced.

In this study we identify changes in global gene expression levels during the growth of *Synechococcus* sp. WH8102 under P stress conditions, using microarray and qRT-PCR analyses. Transcriptome profiling of SYNW0947 and SYNW0948 knockout mutants, putative homologues of *phoB* and *phoR*, indicates that these genes comprise a key response regulator of P metabolism in *Synechococcus* sp. WH8102 and provides further information regarding the Pho regulon in this organism.

Materials and methods

Strains and growth conditions
Axenic *Synechococcus* sp. WH8102 cultures were maintained in either SN medium (Waterbury and Willey, 1988) made with seawater from the Scripps Pier (Scripps Institution of Oceanography, La Jolla, CA, USA), or in an artificial seawater medium (synthetic ocean water, SOW) prepared as described previously (Su *et al.*, 2006) with 9.0 mM NaNO$_3$ standard. Phosphate concentrations were varied from 5 to 87 μM as noted. Batch cultures were grown in glass flasks, gently stirred at 25 °C under 30 μmol photons (μEinstein) m^{-2} s^{-1} continuous white light. Kanamycin was added to a final concentration of 15 μg ml^{-1} for liquid cultures, where appropriate, to maintain selection for insertions in the SYNW0947 and SYNW0948 mutants.

For P-replete cells (wild type or mutants), a small inoculation of P-replete cells was added to P-replete (87 μM P) SOW media and the cells were grown to mid-log phase (approximately 1.5×10^8 cells per ml). Knockout mutants of SYNW0947 and SYNW0948 were grown similarly for comparison with wild type. For P-stressed cells, *Synechococcus* sp. WH8102 was grown with 5 μM P SOW media

Phosphate regulation in *Synechococcus* sp. WH8102
SG Tetu *et al*

837

until alkaline phosphatase activity was first detected (early P stress) or later until cell growth had largely ceased (late P stress, or P-limited cells). Thus phenotypes (induction of alkaline phosphatase activity or growth cessation) were used for determining cell harvesting rather than arbitrary time points in a growth curve. Cultures grew for several days between these two time points as they used up intracellular phosphorus storage pools.

Inactivation of SYNW0947 and SYNW0948
Insertional inactivation was accomplished by cloning a completely internal fragment of either SYNW0947 or SYNW0948 into the suicide vector pMUT100 as previously described (Brahamsha, 1996a, b; McCarren and Brahamsha, 2005, 2007). In this method, the coding sequence of the gene is interrupted by homologous recombination with a suicide vector carrying an internal fragment of the gene. For SYNW0947, this fragment was amplified from *Synechococcus* sp. WH8102 chromosomal DNA using primers 1782F (5′-GAGGGATATGACG TTCTCAC-3′) and 1783R (5′-CGGGCCACCAGCTCT CTGAG-3′). For SYNW0948, the primers were 1783F (5′-ACTCAGGGCTGGTTGATCCT-3′) and 1783R (5′-ATCCCGGCAGCACGAAAGC-3′). The amplified fragments were cloned into pCR2.1-TOPO (Invitrogen, Carlsbad, CA, USA), excised using *Eco*RI and ligated into the *Eco*RI site of pMUT100. The pMUT100 constructions were introduced into *Synechococcus* sp. WH8102 by conjugation with *E. coli* followed by selection of exconjugants on solidified SN medium containing kanamycin (25 µg ml^{-1}) (Brahamsha, 1996a). Verification of the segregation of mutant chromosomes was carried out as described in McCarren and Brahamsha (2005) using primer pairs 947CKF2 (5′-ACACCAACAGGC GTGATC-3′) and 947CKR (5′-AGAGATTGAGGTCG TTCCCA-3′) for SYNW0947, and 948CKF (5′-ATGT CGGTCATCCAGCGTCG-3′) and 948CKR (5′-GAGCT CATGTGCCACATCG-3′) for SYNW0948. Each primer pair flanks the fragment used for inactivation and is used to confirm the absence of a fragment of the wild-type size in the mutant's chromosomal DNA, confirming the absence of an intact gene. Controls include using this primer pair to amplify a fragment of the expected size from wild-type DNA and another set of primers directed to another open reading frame as a positive control to confirm that the mutant's chromosomal DNA was of sufficient quality for PCR amplification. In addition, primers EcoRICC (5′-GCTTATCGATGATAAGCTGTCAAAC-3′), which anneal near the *Eco*RI site of pMUT100 and 947CKF2 (for SYNW0947) or 948CKR (for SYNW0948), were used with chromosomal DNA from each mutant to amplify the region where the insertion of the suicide plasmid had taken place, hence confirming integration of the construct in the mutant's chromosome. Wild-type chromosomal DNA was used as a negative control.

Cell growth, P analysis and alkaline phosphatase assays
Cell counts were typically determined using a FACSort flow cytometer (Becton Dickinson, Franklin Lakes, NJ, USA) using the highest flow rate, as described earlier (Collier and Palenik, 2003), but without applying a compensation matrix. As shown in Supplementary Figure 1, cell counts were determined with a Petroff-Hausser counting chamber (Hausser Scientific, Horsham, PA, USA).

Phosphate concentrations were determined by a standard ammonium molybdate assay (Strickland and Parsons, 1972) with the volumes adjusted for measurement with a microplate reader (the detection limit of the method was 0.5 µM). Alkaline phosphatase activity was measured by incubating 180 µl of the sample (cell culture) with 20 µl of 135 µM *p*-nitrophenyl phosphate (Sigma, St Louis, MO, USA) in 1 M Tris (pH 8.0). Absorbance at 405 nm was read every 30 min for 90 min with a microplate reader. Known concentrations of *p*-nitrophenol (Sigma) were used to generate a standard curve. Phosphate concentrations and alkaline phosphatase activity were measured with unfiltered and 0.2 µm filtered samples to determine if any background alkaline phosphatase activity was observed in the medium, but this was not found. For an example experiment, shown in Supplementary Figure 1, 1 liter of SOW was inoculated with 1×10^6 WH8102 cells per ml and measured daily for cell counts, P concentrations and alkaline phosphatase activity. Subsequent experiments used 5 µM P.

RNA isolation
From 1 liter of SOW media, *Synechococcus* sp. WH8102 cells were harvested by centrifugation at 9000 *g* for 10 min at room temperature. Total RNA was extracted from the cell pellet using TRIzol Reagent (Invitrogen) according to the manufacturer's instructions. The RNA was resuspended in 100 µl of diethylpyrocarbonate-treated water. Using the RNeasy Mini kit (Qiagen, Valencia, CA, USA) and following the manufacturer's protocol, the RNA was twice digested with DNase I (Qiagen) then eluted from the columns. The purity and yield of RNA was determined spectrophotometrically by measuring optical density at wavelengths of 260 and 280 nm. Samples were stored at −80 °C.

DNA microarray design and construction
A full genome microarray for *Synechococcus* sp. WH8102 was constructed consisting of a mixed population of PCR amplicons (2142 genes) and 70-mer oligonucleotides (389 genes). Unique PCR amplicons representing each gene were approximately 800 bp in size, or smaller if the gene size was smaller. Unique 70-mer oligonucleotides were used for genes under 200 bp in size and for two genes that we were not able to PCR amplify successfully. PCR

838

products and oligonucleotides were suspended in 50% dimethyl sulfoxide (DMSO) and deposited onto aminosilane-coated UltraGAPS glass slides (Corning Inc., Corning, NY, USA) using an Intelligent Automation Systems array spotter. Each slide included six replicates of each PCR amplicon or oligonucleotide, negative controls such as 50% DMSO/50% deionized water and positive controls including a total mix of *Synechococcus* sp. WH8102 PCR amplicons, *Arabidopsis* PCR amplicons and 70-mer oligonucleotides as per Hegde *et al.* (2000). After printing, slides were air-dried for 30 min, and DNA was irreversibly cross-linked to the surface by using a Stratalinker (Stratagene, La Jolla, CA, USA) to deliver $250\,mJ\,cm^{-2}$ of short-wavelength UV energy and stored in a desiccator at room temperature.

DNA microarray transcriptional profiling
An indirect labeling method was used to label cDNA as previously described (Peterson *et al.*, 2004), where cDNA is synthesized in the presence of a nucleoside triphosphate analogue containing a reactive aminoallyl group to which a fluorescent dye molecule, either cyanine 3 or cyanine 5 (Cy3/Cy5), is coupled. Approximately 4 μg of total RNA was used for indirect labeling, leading to the production of approximately 4 μg of cDNA with approximately 200 pmol of dye molecule incorporated per microgram of cDNA synthesized. Before hybridization, labeled cDNA was scanned spectrophotometrically to ensure optimal dye incorporation per sample for adequate signal intensity. Each experiment typically included a minimum of two biological replicates and a minimum of three technical replicates for each biological sample, and at least one flip-dye experiment per biological replicate.

All hybridizations were performed as previously described (Peterson *et al.*, 2004) and slides were promptly scanned at a 10-μm resolution using an Axon 4000B scanner with GenePix 4.0 software (Molecular Devices, Sunnyvale, CA, USA). Processing of the TIFF images from hybridized arrays was performed using TIGR Spotfinder (www.tigr.org/software), and the data sets were normalized by applying the LOWESS algorithm (after removing positive controls), using block mode and a smooth parameter of 0.33, available in the TIGR MIDAS package (www.tigr.org/software). Statistical analysis was performed on the mean of $\log_2$-transformed signal ratios of the replicate spots using the statistical analysis of microarrays (SAM) algorithms (Tusher *et al.*, 2001) with a false discovery rate of less than 5%.

To identify sets of genes that displayed similar patterns of expression when exposed to different test conditions, *k*-means clustering analysis (Soukas *et al.*, 2000) was performed using the TIGR Multi-Experiment Viewer software (Saeed *et al.*, 2003). A total of 20 clusters were generated, requiring 33

iterations. Gene trees were subsequently generated by average linkage hierarchical clustering with Euclidean distance as the distance metric.

All microarray data presented here are in accordance with the Microarray Gene Expression Data (MGED) Society's minimum information about a microarray experiment recommendations (Brazma *et al.*, 2001). Descriptions of the microarray experiments, quantitation data and array design have been deposited into the gene expression omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) and have been assigned accession number GSE13170.

Array validation by qRT-PCR
RNA underwent DNase digestion using the DNA-free Turbo DNase Digestion kit (Ambion, Austin, TX, USA). qRT-PCR reactions were carried out using the Superscript III Platinum SYBR Green One-Step qRT-PCR kit (Invitrogen) in an ABI PRISM 7900 (HT) (PE Applied Biosystems, Foster City, CA, USA). Primers were designed with Primer3 (Rozen and Skaletsky, 2000) for products of 100–250 bp in length. As an internal control, reverse transcriptase–PCR was carried out using primers for the amplification of SYNW0547 (0547F: 5′-TGTGGGTCCAGGCCAACCG CTATT-3′; 0547R: 5′-CCGCCCACGGTGATGGATTC AA-3′) and SYNW1090 (1090F: 5′-AACTGGACCAC GATGAAAGC-3′; 1090R: 5′-CACCAGAGGGATGTC CAGTT-3′), which encode 50S ribosomal protein L21 and 30S ribosomal protein S2, respectively. Reactions were performed at least three times for each gene. Cycle threshold (C_T) values for each reaction were determined with ABI PRISM SDS 2.1 software (PE Applied Biosystems). Analysis of relative RNA abundance was performed using the $\Delta\Delta C_T$ method (PE Applied Biosystems).

Bioinformatic analyses
All available *Synechococcus* and selected *Prochlorococcus* genome sequences were subjected to searches to detect homologues of genes observed to be strongly upregulated in *Synechococcus* sp. WH8102. This was carried out using the genomic BLASTP tool on the NCBI microbial genomes resources webpage (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi). Genes were considered homologous where expect values of $\leqslant E-50$ were recorded (provided coverage of query sequence was >70%). For genes SYNW1453 and SYNW1454 (<100 amino acids in length), matches sharing >50% identity over entire length of putative protein were considered homologous.

Results and discussion

Bioinformatic identification of likely Pho regulon components in the WH8102 genome
Examination of the genome of *Synechococcus* sp. WH8102 provides some insight into P utilization

Phosphate regulation in *Synechococcus* sp. WH8102
SG Tetu *et al*

839

in this organism. Encoded within this genome are multiple putative solute-binding proteins for P (SYNW1018, SYNW1286, SYNW1815 and SYNW2507), an ABC transporter consisting of *pstC*, *pstA*, *pstB* homologues (SYNW1270, SYNW1271 and SYNW1272) and putative genes for the transport of phosphonates (SYNW1169, SYNW1168 and SYNW1170), which may allow these organisms to derive P from organic sources or incorporate the phosphonate directly. Multiple putative alkaline phosphatase genes (SYNW0120, SYNW0196, SYNW2391 and SYNW2390) were also observed, which may be used to obtain P from other organic sources (Palenik *et al.*, 2003). Also observed in this genome are homologues of the two-component regulator genes, *phoB* and *phoR* (SYNW0947 and SYNW0948) that are hypothesized to have a function in regulating the cellular response to P limitation (Palenik *et al.*, 2006).

Gene expression changes in response to P stress
As an initial step in characterizing P regulation in *Synechococcus* sp. WH8102 we assayed the temporal profile of alkaline phosphatase activity across a time course. Alkaline phosphatase activity has been shown to be an indicator of upregulation of the Pho regulon in cyanobacteria (Ray *et al.*, 1991; Aiba *et al.*, 1993; Hirani *et al.*, 2001). Under the P stress conditions used here alkaline phosphatase activity was induced at 7 days and peaked at 11 days (Supplementary Figure 1). In between, cells presumably grew on residual P and intracellular storage pools.

To investigate transcriptional profiles related to P regulation in *Synechococcus* sp. WH8102, we used a full genome microarray, consisting of a mix of PCR amplicons and 70-mer oligonucleotides, as described in the Materials and methods. Gene expression in *Synechococcus* sp. WH8102 was compared between cells grown in P-replete and -depleted conditions. For P-deplete cultures, cells were sampled at early P stress, when alkaline phosphatase was first detectable, or at late P stress, corresponding to late log phase and greatly reduced growth (Supplementary Figure 1). This approach allowed us to identify genes that were affected by the onset of P stress or by extended P stress.

In response to early P stress 617 genes were found to be significantly upregulated, whereas the expression of 649 genes was negatively affected (based on SAM analysis; Supplementary Table 1). Genes whose early expression was induced or repressed more than twofold under P stress conditions are listed in Table 1. Using this cutoff a set of 36 upregulated genes and 23 downregulated genes was compiled. Most of the 36 genes observed to be strongly induced under early P stress conditions have not been directly studied in this organism (with the exception of the motility genes *swmA* and *swmB*; Brahamsha, 1996b; McCarren and Brahamsha, 2007). However, bioinformatic analyses provide

clues to the function of 24 of these, whereas 12 encode hypothetical (3) or conserved hypothetical (9) proteins. A high proportion of the strongly repressed genes, 14 of 23, are annotated as hypothetical (10) or conserved hypothetical (4) proteins. BLASTP searches and analyses of genomic context were conducted for genes annotated as either hypothetical or conserved hypothetical proteins, to determine if any comment could be made regarding the biological significance of changes in their expression.

Genes induced under conditions of P stress
Among the 24 strongly induced genes for which a function was predicted, one-third are hypothesized to be involved in transport. Two of these, SYNW2223 and SYNW2224, encode putative outer membrane porins, which we hypothesize to be involved in P transport. The putative proteins encoded by these genes show similarity to each other and to outer membrane porins *somA* and *somB*, which have been characterized in *Synechococcus* sp. PCC 6301 (Hansel and Tadros, 1998) and to porins observed to be upregulated under reduced nitrogen conditions in *Synechococcus* sp. PCC7942 (Sauer *et al.*, 2001). The remaining six are predicted to encode components of P-specific ABC transporters. All three genes of the ABC transporter consisting of *pstC*, *pstA* and *pstB* homologues (SYNW1270, SYNW1271 and SYNW1272) were significantly upregulated (Table 1). Of the four putative P-specific solute-binding proteins encoded within this genome, three were upregulated by more than twofold (SYNW1018, SYNW1286, SYNW1815) whereas the fourth, SYNW2507, was also upregulated significantly, but fell just below the more than twofold cutoff (1.9-fold). Under the conditions used in this study, expression of genes comprising a putative phosphonate ABC transporter (SYNW1169, SYNW1168 and SYNW1170) was not significantly affected.

Also prevalent among the strongly induced genes were putative alkaline phosphatases. Contiguous genes SYNW2390 and SYNW2391 show homology to domains within an atypical alkaline phosphatase, PhoA (Moore *et al.*, 2005), previously characterized in *Synechococcus* sp. PCC7942 (Ray *et al.*, 1991) and in *Synechocystis* sp. PCC6803 (sll0654) (Hirani *et al.*, 2001), where they were also found to be highly expressed under P stress conditions. The other induced putative phosphatase, SYNW0196, has no homologue in any other sequenced *Synechococcus* genome. However, BLASTP searches indicate that related phosphatases are present in other cyanobacterial genera, such as *Gloeobacter*, *Nostoc* and *Anabaena*. Conserved hypothetical protein SYNW0762 also contained a region with similarity to putative alkaline phosphatase genes. However, other matches are to phytase genes that would specifically remove P from phosphorylated inositols. The presence of a phytase-like insertion in a homologous atypical alkaline phosphatase,

Table 1 *Synechococcus* sp. WH8102 genes whose expression was strongly (more than twofold) upregulated or downregulated under conditions of early P stress

Gene[a]	Function	SAM score Early	qRT-PCR	Early	Late	0947 mutant	0948 mutant	Cluster[b]
				Log₂ fold change				
(A) Upregulated								
SYNW1018	ABC transporter, substrate-binding protein, phosphate	18.0	5.7	**4.08**	0.24	**−2.63**	**−2.56**	A
SYNW0165	Conserved hypothetical protein	20.4	8.6	**3.98**	0.45	**−1.06**	**−1.44**	A
SYNW2390	Putative alkaline phosphatase/5 nucleotidase	20.9	8.8	**3.98**	−0.01	**−2.45**	**−2.58**	A
SYNW2391	Putative alkaline phosphatase	6.3	7.7	**3.95**	−0.04	**−2.12**	−0.76	A
SYNW2224	Possible porin	8.5	5.7	**3.30**	−1.06	−0.63	**−1.53**	A
SYNW0196	Putative alkaline phosphatase	19.5	5.6	**3.07**	0.62	**−1.13**	**−1.31**	A
SYNW1019	Possible transcriptional regulator	19.4	7.0	**2.71**	−0.26	**−1.22**	**−1.49**	A
SYNW0762	Conserved hypothetical protein	16.3	7.0	**2.47**	0.45	**−1.13**	**−1.41**	A
SYNW1333	Conserved hypothetical protein	4.4	4.1	**1.99**	**1.60**	**−1.78**	**−1.44**	A
SYNW1270	Putative phosphate ABC transporter	14.0	4.1	**1.84**	0.10	−0.23	−0.38	B
SYNW1562	Conserved hypothetical protein	8.3	ND[c]	**1.84**	**1.18**	**1.56**	0.95	
SYNW2028	Possible rare lipoprotein A	14.0	ND	**1.61**	**−1.49**	−0.30	0.01	
SYNW1454	Conserved hypothetical protein	6.9	ND	**1.57**	**2.69**	**1.51**	−0.32	
SYNW1815	ABC transporter, substrate-binding protein, phosphate	10.4	1.7	**1.54**	0.05	−0.13	−0.55	B
SYNW0406	Hypothetical protein	5.5	ND	**1.37**	−0.12	0.40	0.37	B
SYNW1334	NUDIX family protein	9.5	3.6	**1.37**	0.79	−0.68	**−1.32**	B
SYNW1272	Putative phosphate ABC transporter, ATP-binding subunit	9.0	2.7	**1.37**	−0.22	−0.47	−0.68	B
SYNW1529	Cytochrome *c* oxidase subunit I	13.6	ND	**1.36**	0.71	−0.46	−0.82	B
SYNW1453	Conserved hypothetical protein	6.9	ND	**1.31**	**1.34**	0.77	−0.68	
SYNW1271	Putative phosphate ABC transporter	13.9	4.0	**1.31**	0.02	−0.04	−0.08	B
SYNW1120	Putative 6-phosphogluconolactonase	14.2	ND	**1.30**	0.81	−0.56	**−1.07**	B
SYNW1119	6-Phosphogluconate dehydrogenase	11.4	2.4	**1.30**	0.17	−0.41	−0.70	B
SYNW1213	Thioredoxin peroxidase	6.7	ND	**1.27**	−0.44	**−1.49**	**−1.03**	B
SYNW0953	SwmB: cell-surface protein required for swimming motility	6.7	ND	**1.24**	**1.47**	−0.05	−0.58	
SYNW1527	Conserved hypothetical protein	21.2	ND	**1.24**	0.32	−0.27	−0.63	B
SYNW2415	Hypothetica–l protein	26.0	ND	**1.23**	−0.19	−0.26	−0.38	B
SYNW2508	Molecular chaperone DnaK2, heat-shock protein hsp70-2	16.3	ND	**1.23**	**−2.77**	**−3.75**	**−1.33**	
SYNW0799	Glyceraldehyde-3-phosphate dehydrogenase	13.9	ND	**1.20**	−0.74	−0.52	−0.87	B
SYNW2223	Possible porin	10.1	ND	**1.16**	0.03	−0.61	−0.91	B
SYNW0353	Hypothetical protein	8.8	ND	**1.13**	0.71	0.56	0.40	
SYNW0160	Conserved hypothetical protein	14.5	ND	**1.12**	−0.32	−0.37	−0.43	B
SYNW0085	SwmA: cell-surface protein required for swimming motility	3.2	ND	**1.09**	−1.02	0.21	−0.17	
SYNW0958	Similar to leukotoxin secretion protein	7.1	ND	**1.07**	0.85	**1.15**	0.27	
SYNW0806	Conserved hypothetical protein	6.3	ND	**1.05**	0.22	0.23	−0.15	B
SYNW1286	ABC transporter, substrate-binding protein, possibly phosphate	6.8	6.5	**1.02**	0.04	−0.31	−0.20	B
SYNW0156	Phosphorylase	5.3	ND	**1.01**	0.62	**−1.10**	**−1.00**	B

Table 1 Continued

Gene	Function	SAM score Early	Log$_2$ fold change			
			Early	Late	0947 mutant	0948 mutant
(B) Downregulated						
SYNW1149	Peptide chain release factor RF-3	−7.7	**−2.12**	0.67	**1.11**	−0.02
SYNW1532	Hypothetical protein	−12.7	**−1.85**	0.60	**1.17**	0.77
SYNW1559	Hypothetical protein	−5.9	**−1.82**	0.31	**1.47**	**1.26**
SYNW1682	Conserved hypothetical protein	−7.3	**−1.76**	0.10	**1.19**	0.64
SYNW2340	50S ribosomal protein L7/L12	−14.7	**−1.73**	**1.27**	**1.35**	−0.63
SYNW2030	Conserved hypothetical protein	−10.6	**−1.73**	−0.48	**−1.13**	−0.16
SYNW0542	Hypothetical protein	−7.3	**−1.56**	−0.48	**2.33**	**1.36**
SYNW0541	Hypothetical protein	−5.6	**−1.54**	−0.29	**1.16**	0.44
SYNW0888	Hypothetical protein	−9.3	**−1.47**	**−1.83**	**1.45**	**1.91**
SYNW2341	50S ribosomal protein L10	−7.1	**−1.42**	**1.34**	**1.08**	−0.37
SYNW1946	PhoH family protein	−9.4	**−1.36**	0.26	0.40	0.39
SYNW1061	Hypothetical protein	−2.1	**−1.35**	−0.31	**−1.19**	0.15
SYNW2520	Putative sucrose phosphate synthase	−6.3	**−1.25**	−0.87	−0.39	−0.24
SYNW2454	Putative aspartate carbamoyltransferase	−6.1	**−1.18**	0.47	0.37	−0.35
SYNW0561	Hypothetical protein	−9.5	**−1.18**	**−1.28**	0.05	0.40
SYNW1366	Hypothetical protein	−8.7	**−1.14**	−0.94	0.61	−0.25
SYNW2453	Putative asparagine synthetase protein	−25.2	**−1.12**	−0.01	0.33	−0.55
SYNW1893	Hypothetical protein	−6.2	**−1.06**	0.90	**1.39**	0.45
SYNW0432	Putative short-chain dehydrogenase family protein	−8.6	**−1.06**	0.44	**1.07**	0.61
SYNW0264	Conserved hypothetical protein	−4.5	**−1.04**	0.11	0.68	0.25
SYNW2082	50S ribosomal protein L18	−12.0	**−1.03**	0.04	**−1.83**	**−1.76**
SYNW1698	Hypothetical protein	−5.8	**−1.03**	−0.39	**1.05**	0.66
SYNW1104	Conserved hypothetical protein	−7.1	**−1.02**	0.08	**1.27**	**1.34**

[a]Genes are ordered according to log$_2$ fold change under early P stress. Bold values indicate log$_2$ fold changes greater than ±1.
[b]Indicates gene is present in one of the two clusters of interest (A or B) described in Figure 1.
[c]'ND' indicates fold change was not determined by qRT-PCR.

Phosphate regulation in *Synechococcus* sp. WH8102
SG Tetu *et al*

842

SynWH7803_0111, in *Synechococcus* sp. WH7803 was previously noted by Moore *et al.* (2005).

The remaining 16 significantly upregulated genes for which a functional prediction is available are involved in a range of cellular processes. SYNW1019, which lies directly downstream of the P-binding protein SYNW1018 (*pstS*), was previously designated *ptrA* (putative transcriptional regulator) and speculated to be involved in global P regulation, based on its proximity to *pstS* in *Synechococcus* sp. WH7803 and identity to the CRP family of regulators (Scanlan *et al.*, 1996). The regulation of *ptrA* in response to P limitation supports the possibility this regulator is involved in a P-specific response.

SYNW1334 (*mutT*) belongs to the Nudix family of hydrolyses whose substrates span a wide range of organic pyrophosphates, including nucleoside di- and triphosphates, dinucleoside and diphosphoinositol polyphosphates, nucleotide sugars and RNA caps, with varying degrees of specificity (McLennan, 2006). The upregulation of SYNW1334 under P stress conditions may function to allow the cell to better utilize existing supplies of P by releasing it from certain molecules. One of the upregulated conserved hypothetical proteins, SYNW0806, contains a conserved domain for polyphosphate kinase 2 and therefore may also be involved in altering use of available P within the cell under conditions of P stress.

A number of genes that are involved in aspects of carbohydrate metabolism were upregulated under low P conditions. These include SYNW1119 (*gnd*) and SYNW1120 (*pgl*), which catalyze steps in the pentose P pathway, and SYNW0799 (*gap3*), which catalyzes the conversion of D-glycerate 1,3-bisphosphate to glyceraldehyde 3-phosphate (and the reverse reaction) during glycolysis/glyconeogenesis. SYNW0156, which encodes a putative starch phosphorylase, was also among the set of significantly upregulated genes. In addition, one of the upregulated conserved hypothetical genes, SYNW0160, contains a conserved domain, GalM, indicating possible galactose mutarotase activity. Previous work in other microorganisms has likewise indicated that aspects of carbohydrate metabolism are affected by P starvation. In *E. coli* the *gnd* homologue of SYNW1119 was similarly observed to be significantly upregulated under conditions of P limitation (VanBogelen *et al.*, 1996) whereas *gap1* was observed to be one of only two genes upregulated in both MED4 and MIT9313 strains of *Prochlorococcus* following P stress (Martiny *et al.*, 2006). This *gap1* appears to be the likely orthologue of SYNW0799 *gap3*, also upregulated by P stress.

Other aspects of cellular metabolism may also be affected by P stress. Significant upregulation of SYNW1529, cytochrome *c* oxidase subunit I, was observed in this work, indicating that aspects of energy metabolism may alter under this stress condition. Interestingly, SYNW1213, encoding a thioredoxin peroxidase was also significantly upregulated. In *Synechocystis* sp. PCC 6803, this enzyme has been shown to be coupled to photosynthetic electron transport (Yamamoto *et al.*, 1999) and it is unclear why this enzyme appears responsive to P stress.

The upregulation of SYNW2028 (possible rare lipoprotein A, *rlpA*) and SYNW2508 (*dnaK2*) may constitute part of a general stress response in cells grown under conditions of P stress, rather than a specific response to low P. Previous work has shown rare lipoprotein A to be highly induced by hyperosmotic stress in *Synechocystis* sp. PCC 6803 (Kanesaki *et al.*, 2002). The expression of DnaK2, also known as heat-shock protein hsp70-2, is also known to be induced by stress conditions including salt, UV light and heat stress (Varvasovszki *et al.*, 2003; Mary *et al.*, 2004; Fulda *et al.*, 2006).

Among the most interesting observations to emerge from this examination of the Pho regulon in *Synechococcus* is that the genes for three known cell-surface proteins are among this set of significantly upregulated genes. SYNW0953 (*swmB*) and SYNW0085 (*swmA*) encode polypeptides involved in the unique swimming motility observed in this genus (Brahamsha, 1996b; McCarren and Brahamsha, 2007), and SYNW0406 encodes the heavily glycosylated 70 kDa integral outer membrane polypeptide found in *Synechococcus* sp. WH8102 (Brahamsha, 1996b and B Brahamsha, unpublished). One explanation for this may be that P stress results in a reorganization of the cell envelope, perhaps to accommodate P-specific porins, such as SYNW2224, and that this requires increased expression of other cell-surface proteins. Another possibility is that increased expression of the motility genes facilitates movement of cells toward regions of higher P concentration. Although Willey and Waterbury (1989) showed chemotaxis toward nitrogenous compounds by *Synechococcus* sp. WH8102, it is possible that chemotaxis toward organophosphate compounds also occurs. However, although SwmA and SwmB are required for swimming, it is not yet clear whether they constitute part of the motility apparatus. Because they are such abundant cell-surface proteins, it is possible that their presence is required for an intact cell envelope in which the true components of the motility apparatus can assemble and interact and hence that they have only an indirect function in motility.

Only five of the genes strongly (more than twofold) upregulated under early P stress were also strongly upregulated at the second time point, indicating that the transcriptional response to early P stress is, in most cases, transitory (Table 1). This may indicate that transitory increases in gene expression are sufficient to generate a protein response that can support P utilization over more extended time periods. As there is information regarding function for only one of these five genes, *swmB* (SYNW1333,

Phosphate regulation in *Synechococcus* sp. WH8102
SG Tetu *et al*

843

SYNW1453, SYNW1454, SYNW1562, all encode conserved hypothetical proteins), it is difficult to comment on the difference between transiently and permanently affected genes.

Very few of the genes strongly upregulated only after extended P stress are involved in P metabolism. Many of the most affected are conserved hypotheticals or of unknown function (Supplementary Table 1). One exception is phosphoglycerate mutase (SYNW0519), involved in transfer of P groups within glycerate molecules during glycolysis, which was strongly upregulated after prolonged P limitation. Several photosystem I genes were upregulated whereas various photosystem II genes were downregulated under late P stress (Supplementary Table 1). This might be consistent with an enhanced use of photosystem I-dependent cyclic photophosphorylation to generate ATP. Interestingly, two genes with putative regulatory functions were among the strongly upregulated at this time point, including SYNW2037, encoding a possible *pex* (period extender) gene, thought to be involved in regulation of the circadian clock in these cyanobacteria (Kutsuna *et al.*, 2007) and SYNW1660, also a putative transcriptional regulator (Supplementary Table 1).

To validate the results of the microarray experiments, qRT-PCR was performed on a subset of the genes observed to be highly induced under conditions of early P stress. All of these also showed significant upregulation in qRT-PCR analysis (Table 1). A comparison of the data derived from both experiments shows that qRT-PCR detected a higher fold change in all examined upregulated genes, compared to microarray analyses (Table 1). The tendency for microarray experiments to underestimate the relative change in gene expression, compared to qRT-PCR, has been observed previously in studies comparing data derived from these two methodologies (Yuen *et al.*, 2002; Dallas *et al.*, 2005).

Genes repressed under conditions of P stress

Among the nine functionally characterized genes whose expression was significantly repressed under early P stress conditions, four are involved in translation. SYNW1149, peptide chain release factor RF-3, involved in termination of protein synthesis, showed the greatest degree of downregulation. Although SYNW2340, SYNW2341 and SYNW2082 encoding 50S ribosomal proteins were also significantly repressed under early P stress, the suppression of genes involved in protein synthesis has been observed in other microorganisms in response to both P limitation (Ishige *et al.*, 2003; Martiny *et al.*, 2006) and N limitation (Silberbach *et al.*, 2005) and has been suggested to be a result of a general reduction in the cells metabolic rate.

Other genes negatively affected by low P include a putative sucrose P synthase (SYNW2520), putative asparagine synthetase (SYNW2453, putative aspartate carbamoyltransferase (SYNW2454) and a putative dehydrogenase (SYNW0432). After extended P stress, genes SYNW1717 and SYNW1718, encoding ribulose bisphosphate carboxylase (both large and small chain) were among the most strongly repressed (Supplementary Table 1), indicating that such nutrient limitation, if experienced for extended periods, may be extremely detrimental to fundamental carbon fixation processes in these cells.

Interestingly, SYNW1946, annotated as a *phoH* (phosphate starvation inducible) family protein, was observed to be downregulated in this organism, whereas previously observed to be upregulated under P-limited conditions in other organisms, including *E. coli* and *Corynebacterium glutamicum* (Wanner, 1993; Ishige *et al.*, 2003) and unaffected in *Prochlorococcus* strains MED4 and MIT9313 (Martiny *et al.*, 2006). This suggests that the encoded protein, whose function is at present unknown, may have a different function regarding the response to P stress in different microorganisms. Interestingly phoH orthologues are also found in marine phages (Rohwer *et al.*, 2000).

Gene knockouts of the putative phosphate two-component system

In other well-studied bacteria, including cyanobacteria, a two-component regulator, composed of a sensory kinase (*phoB/sphR*) and response regulator gene (*phoR/sphS*), are commonly observed to be involved the cellular response to P limitation (Aiba *et al.*, 1993; Wanner, 1996; Suzuki *et al.*, 2004). On the basis of sequence similarity, the most apparent candidate genes for controlling P regulation in *Synechococcus* sp. WH8102 are SYNW0947 and SYNW0948, encoding a probable P sensor kinase and response regulator (Scanlan and West, 2002; Palenik *et al.*, 2003). To investigate the function of these genes in this organism, we generated knockouts in each of these genes by insertional inactivation. Comparison of the transcript levels in the knockout mutants of SYNW0947 and SYNW0948 with their isogenic parent (all grown under P-replete conditions) supported the hypothesis that the two-component regulatory system encoded by these genes is important to the regulation of P metabolism of *Synechococcus* sp. WH8102. Intergenic distances (52 bp) as well as operon prediction methods (Price *et al.*, 2005) suggest that SYNW0947 and SYNW0948 may be co-transcribed. Hence, insertional inactivation of SYNW0947 may have polar effects on SYNW0948. On the basis of the microarray data, there was a minor decrease in SYNW0948 expression (log_2 fold change of -0.09) in the SYNW0947 mutant (Supplementary Table 1). In addition, downstream of SYNW0948 is a gene, *rsuA* (SYNW0949), whose predicted start site overlaps the stop codon of SYNW0948, although the operon prediction model does not predict it to be co-transcribed, nor was it observed to be significantly affected in either of the
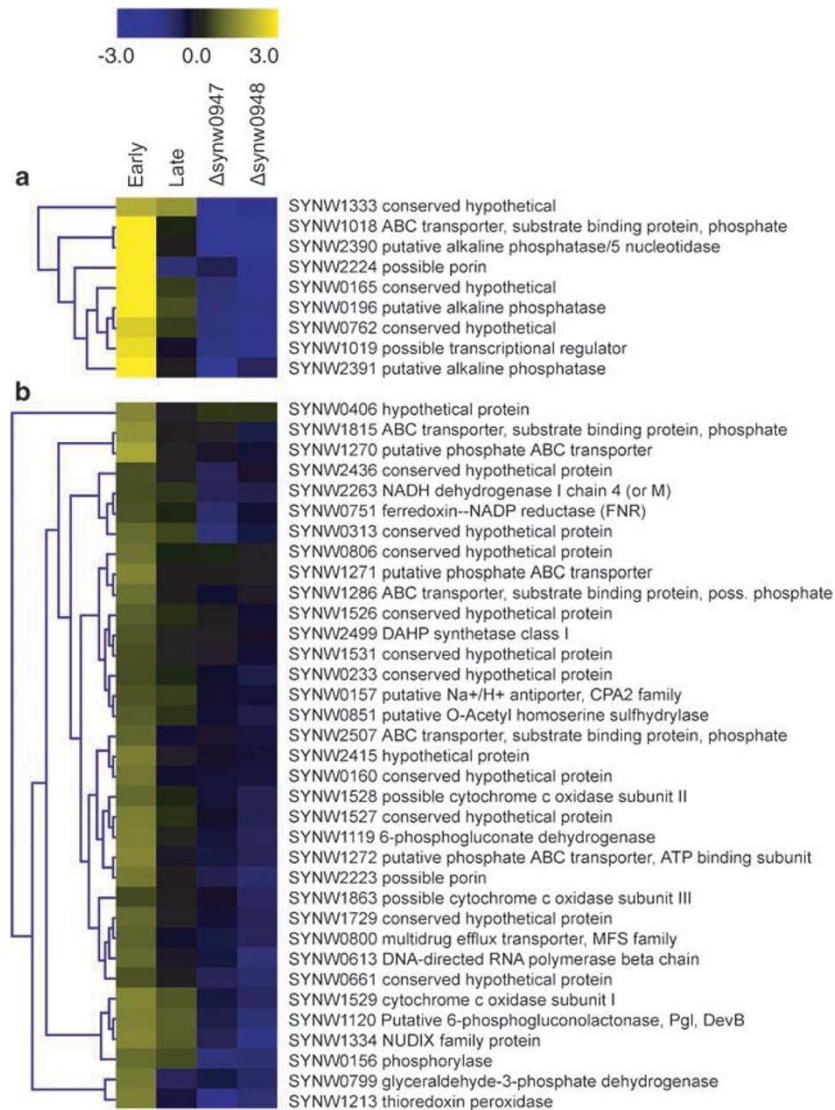
Phosphate regulation in *Synechococcus* sp. WH8102
SG Tetu *et al*

844

Figure 1 Hierarchical trees showing clusters of *Synechococcus* sp. WH8102 genes whose expression was affected similarly by growth under P stress conditions and in SYNW0947 and SYNW0948 knockout mutants. The expression levels under early and late phosphate stress (5 μM initial P) compared with phosphate-replete (87 μM initial P) conditions and in the SYNW0948 and SYNW0947 knockout mutants compared to the isogenic parental strains (all under P-replete conditions) was used in the analysis. For each set of experiments data from two independent biological replicates was used. Two clusters of interest (**a** and **b**) were observed. The color bar indicates the log₂ ratio.

knockout mutants (Supplementary Table 1). *rsuA* codes for a putative pseudouridylate synthase, responsible for the synthesis of pseudouridine from uracil in rRNA. This gene arrangement is found in two other cyanobacteria whose genomes have been sequenced, the freshwater unicellular *Synechococcus elongatus* strains PCC6301 and PCC7942 (Aiba *et al.*, 1993). Its involvement, if any, in stress responses, has yet to be investigated.

Regulation of the P stress response
There is considerable overlap in the set of genes whose expression was significantly affected by inactivation of SYNW0947 and SYNW0948 (even

under P-replete conditions), and those observed to be responsive to P stress (Table 1). A cluster analysis of the microarray data from each set of experiments provides statistical support for the observation that similar sets of genes were affected by low (compared to replete) P and loss of SYNW0947 or SYNW0948 (Figure 1). Two clusters of interest emerge. The first includes genes encoding alkaline phosphatases and two genes involved in aspects of P transport (Figure 1a; Table 1). Interestingly, cluster 1 also includes the possible transcriptional regulator SYNW1019, which may also be involved in the P stress response in this organism. All members of this first cluster were strongly affected by P stress and in the sensor kinase/response regulator knockout

Phosphate regulation in *Synechococcus* sp. WH8102
SG Tetu *et al*

npg

845

Table 2 Distribution of *Synechococcus* sp. WH8102 Pho regulon gene homologues across other sequenced cyanobacterial genomes

WH8102 genes[a]	Function	Sequenced genomes[b]															
		CC9605	CC9311	BL107	CC9902	RCC307	RS9916	RS9917	WH 5701	WH 7803	WH 7805	PCC 6301	PCC 7942	JA-2-3B'a(2-13)	JA-3-3Ab	MIT 9313*	CCMP1986*
SYNW0947	putative response regulator (PhoB)																
SYNW0948	putative sensor histidine kinase (PhoR)																
SYNW0085	SwmA - cell surface protein																
SYNW0156	phosphorylase																
SYNW0160	conserved hypothetical protein																
SYNW0165	conserved hypothetical protein																
SYNW0196	putative alkaline phosphatase																
SYNW0353	hypothetical protein																
SYNW0406	hypothetical protein																
SYNW0762	conserved hypothetical protein																
SYNW0799	glyceraldehyde-3-phosphate dehydrogenase																
SYNW0806	conserved hypothetical protein																
SYNW0953	SwmB - cell surface protein																
SYNW0958	leukotoxin secretion protein																
SYNW1018	ABC transporter component																
SYNW1019	possible transcriptional regulator																
SYNW1119	6-phosphogluconate dehydrogenase																
SYNW1120	putative 6-phosphoglucono-lactonase																
SYNW1213	thioredoxin peroxidase																
SYNW1270	ABC transporter component																
SYNW1271	ABC transporter component																
SYNW1272	ABC transporter component																
SYNW1286	ABC transporter component																
SYNW1333	conserved hypothetical protein																
SYNW1334	NUDIX family protein																
SYNW1453	conserved hypothetical protein																
SYNW1454	conserved hypothetical protein																
SYNW1527	conserved hypothetical protein																
SYNW1529	cytochrome c oxidase subunit I																
SYNW1562	conserved hypothetical protein																
SYNW1815	ABC transporter component																
SYNW2028	possible rare lipoprotein A																
SYNW2223	possible porin																
SYNW2224	possible porin																
SYNW2390	putative alkaline phosphatase																
SYNW2391	putative alkaline phosphatase																
SYNW2415	hypothetical protein																
SYNW2508	DnaK2 heat shock protein hsp70-2																

[a]BLASTP searches were carried out using putative PhoB and PhoR homologues (listed first) and *Synechococcus* sp. WH8102 genes whose expression was strongly (more than twofold) upregulated under early P stress (listed in the order they are observed in the genome). Black squares indicate a homologue was detected by BLASTP, according to the criteria described in the Materials and methods section. Gray squares indicate that a *phoR* homologue was observed that was positioned next to a *phoB* homologue, but fell below the *E*-value cutoff used.
[b]*Synechococcus* and *Prochlorococcus* (marked with asterisk) genome protein sequences subjected to BLASTP searches for *Synechococcus* sp. WH8102 Pho regulon homologues. The names listed under sequenced genomes indicate strains. Two *Prochlorococcus* strains are included (*P. marinus* str. MIT 9313 and *P. marinus* subsp. pastoris str. CCMP1986, also referred to as MED4), PCC6301 and PCC7942 are strains of *S. elongatus* whereas the remaining strains are of *Synechococcus* sp.

Phosphate regulation in *Synechococcus* sp. WH8102
SG Tetu *et al*

846

mutants, with observed $\log_2$ fold changes of 2.0–4.1 under P stress conditions at the early time point.

A second cluster (Figure 1b) was observed, which was less strongly affected, with members showing $\log_2$ fold changes of 0.6–1.8 under P stress conditions at the early time point. This set of genes included members of the pentose phosphate pathway and others linked to P metabolism such as the Nudix protein, SYNW1334, as well as many expected to be involved in the transport of P. Interestingly, genes found in this cluster also encode key enzymes in oxidative phosphorylation and amino-acid metabolism pathways. The upregulation of both cytochrome *c* oxidase subunits I and II may function to ensure that sufficient ATP is generated from the available P in the cell.

The Pho regulon of WH8102

The response regulator, *phoB*, affects transcription by binding to a conserved sequence motif, known as a Pho box, located upstream of genes in the Pho regulon. The Pho box motif has been well characterized in enteric bacteria (Wanner, 1993) and *Bacillus subtilis* (Qi *et al.*, 1997). Work on cyanobacteria has likewise shown that the response regulator, *sphR*, functions by binding to conserved Pho box sequences (Nagaya *et al.*, 1994; Suzuki *et al.*, 2004). Comparison of these regions indicated that the consensus sequence of the Pho box in cyanobacteria differed from that of *E. coli*, in this case being composed of repetitive PyTTAAPyPy(T/A)-like sequences (Suzuki *et al.*, 2004). More recently, computational prediction was used to generate a profile for Pho boxes in 16 sequenced cyanobacterial genomes, to explore the Pho regulons of these organisms (Su *et al.*, 2007). This work predicted Pho boxes for 38 independent transcriptional units in *Synechococcus* sp. WH8102, only 5 of which correlate to genes observed in this study to be upregulated by more than twofold under early P stress (SYNW2390, SYNW2391, SYNW1019, SYNW0085 and SYNW0165). None of the ⩾2-fold downregulated genes was observed to be part of these Pho-box-associated transcriptional units.

To explore the distribution of genes involved in the P stress response, the genome sequences of all available *Synechococcus* strains, plus two well-studied *Prochlorococcus marinus* strains (Rocap *et al.*, 2003), were examined for homologues of *phoB*, *phoR* and the 36 early P stress-responsive *Synechococcus* sp. WH8102 genes. Homologues of *phoB* and *phoR* were observed in some of the marine *Synechococcus* strains (Table 2). However, microbial mat representatives, *Synechococcus* sp. JA-2-3Ab and JA-2-3B'a(2–13), contained *phoB* homologues only, whereas two coastal isolates *Synechococcus* sp. CC9311 and CC9902, and *Synechococcus* sp. BL107, isolated from the Mediterranean Sea, lacked both of these genes (Table 2). In freshwater *S. elongatus* strains and the *P. marinus* strain MIT 9313, adjacent *phoB* and *phoR* candidate genes were

observed, however the *phoR* homologues were more divergent.

Many of the Pho regulon genes likewise show a patchy distribution (Table 2). The porin genes (SYNW2223 and SYNW2224), a putative ABC transporter (SYNW1286) and alkaline phosphatase (SYNW2390) gene and the possible transcriptional regulator (SYNW1019) were found to have homologues in some, but not all examined cyanobacteria (Table 2). Other members, including the remaining ABC transporter components, were ubiquitous in marine species. A small number of genes, including those that encode key metabolic enzymes, were present in all examined cyanobacterial genomes. In contrast, two putative alkaline phosphatase genes (SYNW2391 and SYNW0196), motility associated genes *swmA* and *swmB*, and a number of hypothetical genes were not observed in any *Synechococcus* genome other than WH8102.

The distribution of SYNW2390, one of most strongly induced genes under early P stress conditions, was particularly interesting. Homologues were observed in both *S. elongatus* PCC 6301 and PCC 7942 and *Synechococcus* sp. JA-2-3Ab and JA-2-3B'a, isolated from hot spring microbial mats but in only one marine isolate, *Synechococcus* sp. WH7803 (Table 2). As marine *Synechococcus* are phylogenetically very distant from freshwater and hot spring *Synechococcus*, the distribution of this gene is clearly inconsistent with the phylogeny of this group. Such phylogeny-independent distribution of Pho regulon genes has previously been observed in *Prochlorococcus* sp. (Moore *et al.*, 2005; Martiny *et al.*, 2006) and may be common in cyanobacteria. Such findings highlight the need to evaluate the effects of P stress in multiple strains to gain a reliable understanding of the nature and variability of this response in these organisms in the environment.

Conclusions

Many open ocean environments are oligotrophic, containing low concentrations of nutrients, such as N and P. Inhabitants of these environments show adaptations to these conditions in the form of a range of traits designed to best utilize the available resources. In areas of the Atlantic Ocean, P levels are sufficiently low to result in phosphorus control of primary production (Wu *et al.*, 2000), which may be an issue for *Synechococcus* strains present in these environments, such as *Synechococcus* sp. WH8102.

In response to P stress, *Synechococcus* sp. WH8102 cells exhibit significant alterations in gene expression. As in other studied bacteria, the Pho regulon in this strain includes a number of genes involved in the transport and metabolism of P. Upregulation of these genes under conditions of low P presumably functions to increase P import from outside the cell and alters metabolism in ways that result in utilization of P from organic sources. It is

Phosphate regulation in *Synechococcus* sp. WH8102
SG Tetu *et al*

847

worth noting that although the Pho regulon in this strain appears sizable, constituting more than 20 genes, these genes do not appear to be clustered around *phoB*, as was observed in *Prochlorococcus* strains originating from low P waters (Martiny *et al.*, 2006). This suggests that both the nature and the organization of the Pho regulon may be highly variable within and between cyanobacterial species.

In this work we have gained significant experimental support for the hypothesis that the two-component regulator encoded by SYNW0947 and SYNW0948 is important in the response to P stress. Inactivation of each of these genes impacted expression of a substantial proportion of the genes that constitute the Pho regulon in *Synechococcus* sp. WH8102. Further work is now needed to elucidate the mechanism by which these genes impact expression of the Pho regulon in this organism and determine the extent to which this adheres or deviates from what has been seen in other studied bacteria. In particular, examining the effect of early P stress on the mutants would be useful but was beyond the scope of this paper. The lack of an observable Pho box (*phoB*-binding site) upstream of the majority of Pho regulon members suggests that *phoB* does not directly influence the expression of the entire regulon. However, one gene that does encode a Pho box is the putative transcriptional regulator, *ptrA* (SYNW1019), which was upregulated by P stress and affected in both SYNW0947 and SYNW0948 knockout mutants. One possibility is that *phoB* activates PtrA by binding to the Pho box region, which then acts to upregulate expression of other members of the Pho regulon that lack a *phoB*-binding site. PtrA was named in recognition of its similarity to regulators of the Crp/Fnr family (Scanlan *et al.*, 1996). This family comprises DNA-binding proteins that predominantly function as positive transcription factors, members of which have been shown to respond to a broad spectrum of intracellular and exogenous signals (Korner *et al.*, 2003). Two members of this family have previously been characterized in cyanobacteria: *cysR* involved in regulation of sulfur assimilation (Laudenbach and Grossman, 1991) and *ntcA* involved in regulating genes subjected to nitrogen control by ammonium, achieved through binding to a palindromic sequence motif in the promoter region of nitrogen-related genes (Luque *et al.*, 1994). Although no obvious conserved palindromes were observed in the remaining Pho regulon genes, the function of this gene in the P stress response could be tested in the future by observing the effect of a *ptrA* knockout on members of the Pho regulon.

Considering the importance of P to marine microorganisms, it is not surprising that complex regulatory systems are encoded in the genomes of these organisms, facilitating a rapid response to conditions of P stress. What is perhaps surprising is the degree of variation in the nature of this response between different studied cyanobacteria, with mounting evidence that the genes involved in this response are highly variable, as is their organization, and perhaps also, their mode of activation. The results of this and previous studies (Suzuki *et al.*, 2004; Moore *et al.*, 2005; Martiny *et al.*, 2006) indicate that different species and even different strains of the same species are likely to react quite differently to P stress. As the impact of human activities on nutrient cycling in marine environments continues to increase, it is important that we improve our understanding of how key photosynthetic organisms are likely to be affected.

Acknowledgements

References

Aiba H, Mizuno T. (1994). A novel gene whose expression is regulated by the response-regulator, *SphR*, in response to phosphate limitation in *Synechococcus* species PCC7942. *Mol Microbiol* **13**: 25–34.

Aiba H, Nagaya M, Mizuno T. (1993). Sensor and regulator proteins from the cyanobacterium *Synechococcus* species PCC7942 that belong to the bacterial signal-transduction protein families: implication in the adaptive response to phosphate limitation. *Mol Microbiol* **8**: 81–91.

Baek JH, Lee SY. (2007). Transcriptome analysis of phosphate starvation response in *Escherichia coli*. *J Microbiol Biotechnol* **17**: 244–252.

Brahamsha B. (1996a). A genetic manipulation system for oceanic cyanobacteria of the genus *Synechococcus*. *Appl Environ Microbiol* **62**: 1747–1751.

Brahamsha B. (1996b). An abundant cell-surface polypeptide is required for swimming by the nonflagellated marine cyanobacterium *Synechococcus*. *Proc Natl Acad Sci USA* **93**: 6504–6509.

Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C *et al.* (2001). Minimum information about a microarray experiment (MIAME) toward standards for microarray data. *Nat Genet* **29**: 365–371.

Collier JL, Palenik B. (2003). Phycoerythrin-containing picoplankton in the Southern California Bight. *Deep Sea Res II* **50**: 2405–2422.

Dallas P, Gottardo N, Firth M, Beesley A, Hoffmann K, Terry P *et al.* (2005). Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR—how well do they correlate? *BMC Genomics* **6**: 59.

Dufresne A, Ostrowski M, Scanlan DJ, Garczarek L, Mazard S, Palenik BP *et al.* (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol* **9**: R90.

Fulda S, Mikkat S, Huang F, Huckauf J, Marin K, Norling B *et al.* (2006). Proteome analysis of salt stress response in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Proteomics* **6**: 2733–2745.

Grillo JF, Gibson J. (1979). Regulation of phosphate accumulation in the unicellular cyanobacterium *Synechococcus*. *J Bacteriol* **140**: 508–517.

Hansel A, Tadros MH. (1998). Characterization of two pore-forming proteins isolated from the outer membrane of *Synechococcus* PCC 6301. *Curr Microbiol* **36**: 321–326.

Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R *et al.* (2000). A concise guide to cDNA microarray analysis. *Biotechniques* **29**: 548–550, 552–544, 556 passim.

Hirani TA, Suzuki I, Murata N, Hayashi H, Eaton-Rye JJ. (2001). Characterization of a two-component signal transduction system involved in the induction of alkaline phosphatase under phosphate-limiting conditions in *Synechocystis* sp. PCC 6803. *Plant Mol Biol* **45**: 133–144.

Ishige T, Krause M, Bott M, Wendisch VF, Sahm H. (2003). The phosphate starvation stimulon of *Corynebacterium glutamicum* determined by DNA microarray analyses. *J Bacteriol* **185**: 4519–4529.

Kanesaki Y, Suzuki I, Allakhverdiev SI, Mikami K, Murata N. (2002). Salt stress and hyperosmotic stress regulate the expression of different sets of genes in *Synechocystis* sp. PCC 6803. *Biochem Biophys Res Commun* **290**: 339–348.

Korner H, Sofia HJ, Zumft WG. (2003). Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol Rev* **27**: 559–592.

Kutsuna S, Kondo T, Ikegami H, Uzumaki T, Katayama M, Ishiura M. (2007). The circadian clock-related gene pex regulates a negative *cis* element in the *kaiA* promoter region. *J Bacteriol* **189**: 7690–7696.

Laudenbach DE, Grossman AR. (1991). Characterization and mutagenesis of sulfur-regulated genes in a cyanobacterium: evidence for function in sulfate transport. *J Bacteriol* **173**: 2739–2750.

Luque I, Flores E, Herrero A. (1994). Molecular mechanism for the operation of nitrogen control in cyanobacteria. *EMBO J* **13**: 2862–2869.

Martiny AC, Coleman ML, Chisholm SW. (2006). Phosphate acquisition genes in *Prochlorococcus* ecotypes: evidence for genome-wide adaptation. *Proc Natl Acad Sci USA* **103**: 12552–12557.

Mary I, Tu CJ, Grossman A, Vaulot D. (2004). Effects of high light on transcripts of stress-associated genes for the cyanobacteria *Synechocystis* sp. PCC 6803 and *Prochlorococcus* MED4 and MIT9313. *Microbiology* **150**: 1271–1281.

McCarren J, Brahamsha B. (2005). Transposon mutagenesis in a marine *Synechococcus* strain: isolation of swimming motility mutants. *J Bacteriol* **187**: 4457–4462.

McCarren J, Brahamsha B. (2007). SwmB, a 1.12-megadalton protein that is required for nonflagellar swimming motility in *Synechococcus*. *J Bacteriol* **189**: 1158–1162.

McLennan A. (2006). The Nudix hydrolase superfamily. *Cell Mol Life Sci* **63**: 123–143.

Moore LR, Ostrowski M, Scanlan DJ, Feren K, Sweetsir T. (2005). Ecotypic variation in phosphorus acquisition mechanisms within marine picocyanobacteria. *Aquat Microb Ecol* **39**: 257–269.

Nagaya M, Aiba H, Mizuno T. (1994). The *sphR* product, a two-component system response regulator protein, regulates phosphate assimilation in *Synechococcus* sp. strain PCC 7942 by binding to two sites upstream from the *phoA* promoter. *J Bacteriol* **176**: 2210–2215.

Palenik B, Brahamsha B, Larimer FW, Land M, Hauser L, Chain P *et al.* (2003). The genome of a motile marine *Synechococcus*. *Nature* **424**: 1037–1042.

Palenik B, Dyhrman ST. (1998). Recent progress in understanding the regulation of marine primary productivity by phosphorus. In: Lynch JP, Deikman J (eds). *Phosphorus in Plant Biology: Regulatory Roles in Molecular, Cellular, Organismic, and Ecosystem Processes*. American Society of Plant Physiologists: Rockville, MD, pp 26–38.

Palenik B, Ren Q, Dupont CL, Myers GS, Heidelberg JF, Badger JH *et al.* (2006). Genome sequence of *Synechococcus* CC9311: insights into adaptation to a coastal environment. *Proc Natl Acad Sci USA* **103**: 13555–13559.

Partensky F, Hess WR, Vaulot D. (1999). *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Rev* **63**: 106–127.

Peterson SN, Sung CK, Cline R, Desai BV, Snesrud EC, Luo P *et al.* (2004). Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. *Mol Microbiol* **51**: 1051–1070.

Price MN, Huang KH, Alm EJ, Arkin AP. (2005). A novel method for accurate operon predictions in all sequenced prokaryotes. *Nucleic Acids Res* **33**: 880–892.

Qi Y, Kobayashi Y, Hulett FM. (1997). The *pst* operon of *Bacillus subtilis* has a phosphate-regulated promoter and is involved in phosphate transport but not in regulation of the Pho regulon. *J Bacteriol* **179**: 2534–2539.

Ray JM, Bhaya D, Block MA, Grossman AR. (1991). Isolation, transcription, and inactivation of the gene for an atypical alkaline phosphatase of *Synechococcus* sp. strain PCC 7942. *J Bacteriol* **173**: 4297–4309.

Rocap G, Larimer FW, Lamerdin J, Malfatti S, Chain P, Ahlgren NA *et al.* (2003). Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* **424**: 1042–1047.

Rohwer F, Segall A, Steward G, Seguritan V, Breitbart M, Wolven F *et al.* (2000). The complete genomic sequence of the marine phage Roseophage SIO1 shares homology with nonmarine phages. *Limnol Oceanogr* **45**: 408–418.

Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N *et al.* (2003). TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34**: 374–378.

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, Bourne J, Mann N. (1996). A putative transcriptional activator of the Crp/Fnr family from

Phosphate regulation in *Synechococcus* sp. WH8102
SG Tetu *et al*

849

the marine cyanobacterium *Synechococcus* sp.
WH7803. *J Appl Phycol* **8**: 565–567.

Scanlan DJ, Mann NH, Carr NG. (1993). The response of
the picoplanktonic marine cyanobacterium *Synechococcus* species WH7803 to phosphate starvation
involves a protein homologous to the periplasmic
phosphate-binding protein of *Escherichia coli*. *Mol
Microbiol* **10**: 181–191.

Scanlan DJ, West NJ. (2002). Molecular ecology of the
marine cyanobacterial genera *Prochlorococcus* and
Synechococcus. *FEMS Microbiol Ecol* **40**: 1–12.

Schwarz R, Forchhammer K. (2005). Acclimation of
unicellular cyanobacteria to macronutrient deficiency:
emergence of a complex network of cellular responses.
Microbiology **151**: 2503–2514.

Silberbach M, Huser A, Kalinowski J, Puhler A, Walter B,
Kramer R *et al.* (2005). DNA microarray analysis of
the nitrogen starvation response of *Corynebacterium
glutamicum*. *J Biotechnol* **119**: 357–367.

Soukas A, Cohen P, Socci ND, Friedman JM. (2000).
Leptin-specific patterns of gene expression in white
adipose tissue. *Genes Dev* **14**: 963–980.

Strickland JDH, Parsons TR. (1972). *A Practical Handbook
of Seawater Analysis*, 2nd edn. Canada Fisheries
Research Board: Canada.

Su Z, Mao F, Dam P, Wu H, Olman V, Paulsen IT *et al.*
(2006). Computational inference and experimental
validation of the nitrogen assimilation regulatory
network in cyanobacterium *Synechococcus* sp. WH
8102. *Nucleic Acids Res* **34**: 1050–1065.

Su Z, Olman V, Xu Y. (2007). Computational prediction
of Pho regulons in cyanobacteria. *BMC Genomics*
8: 156.

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.

Thingstad TF, Krom MD, Mantoura RF, Flaten GA, Groom
S, Herut B *et al.* (2005). Nature of phosphorus
limitation in the ultraoligotrophic eastern Mediterranean. *Science* **309**: 1068–1071.

Tusher VG, Tibshirani R, Chu G. (2001). Significance
analysis of microarrays applied to the ionizing
radiation response. *Proc Natl Acad Sci USA* **98**:
5116–5121.

VanBogelen RA, Olson ER, Wanner BL, Neidhardt FC.
(1996). Global analysis of proteins synthesized during
phosphorus restriction in *Escherichia coli*. *J Bacteriol*
178: 4344–4366.

Varvasovszki V, Glatz A, Shigapova N, Josvay K, Vigh L,
Horvath I. (2003). Only one *dnaK* homolog, *dnaK2*, is
active transcriptionally and is essential in *Synechocystis*. *Biochem Biophys Res Commun* **305**: 641–648.

Wanner BL. (1993). Gene regulation by phosphate in
enteric bacteria. *J Cell Biochem* **51**: 47–54.

Wanner BL. (1996). Phosphorus assimilation and control
of the phosphate regulon. In: Neidhardt FC, Curtiss III
R, Ingraham JL and other authors (eds). *Escherichia
coli and Salmonella: Cellular and Molecular Biology*.
ASM Press: Washington, DC, pp 1357–1381.

Waterbury JB, Willey JM. (1988). Isolation and growth of
marine planktonic cyanobacteria. *Methods Enzymol*
167: 100–105.

Watson GM, Scanlan DJ, Mann NH. (1996). Characterization of the genes encoding a phosphate-regulated two
component sensory system in the marine cyanobacterium *Synechococcus* sp. WH7803. *FEMS Microbiol Lett*
142: 105–109.

Willey JM, Waterbury JB. (1989). Chemotaxis toward
nitrogenous compounds by swimming strains of
marine *Synechococcus* spp. *Appl Environ Microbiol*
55: 1888–1894.

Wu J, Sunda W, Boyle EA, Karl DM. (2000). Phosphate
depletion in the western North Atlantic Ocean.
Science **289**: 759–762.

Yamamoto H, Miyake C, Dietz K-J, Tomizawa K-I, Murata
N, Yokota A. (1999). Thioredoxin peroxidase in the
cyanobacterium *Synechocystis* sp. PCC 6803. *FEBS
Lett* **447**: 269–273.

Yuen T, Wurmbach E, Pfeffer RL, Ebersole BJ, Sealfon SC.
(2002). Accuracy and calibration of commercial
oligonucleotide and custom cDNA microarrays. *Nucl
Acids Res* **30**: e48.

Zubkov MV, Mary I, Woodward EM, Warwick PE, Fuchs
BM, Scanlan DJ *et al.* (2007). Microbial control of
phosphate in the nutrient-depleted North Atlantic
subtropical gyre. *Environ Microbiol* **9**: 2079–2089.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)