

**GLOBAL PROTEOMICS REVEAL AN ATYPICAL STRATEGY FOR C/N ASSIMILATION BY
A CYANOBACTERIUM UNDER DIVERSE ENVIRONMENTAL PERTURBATIONS**

Kimberly M. Wegener^{1,*}, Abhay K. Singh^{1,*}, Jon M. Jacobs², Thanura Elvitigala¹, Eric A. Welsh^{1, 3}, Nir Keren^{1, 4}, Marina A. Gritsenko², Bijoy K. Ghosh⁵, David G. Camp II², Richard D. Smith², and Himadri B. Pakrasi¹

¹Department of Biology, Washington University, St. Louis, MO 63130.

²Environmental Molecular Sciences Biological Sciences Division and Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99352.

³Pfizer, Inc., St. Louis, MO 63124.

⁴Department of Plant and Environmental Sciences, Silberman Institute of Life Science, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel.

⁵Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409.

*- Equal Contribution

Correspondence and requests for materials should be addressed to H.B.P. (E-mail: Pakrasi@wustl.edu)

Manuscript Information: 16 text pages, 3 figures, 2 tables.

All proteomics datasets have been deposited with NCBI Peptidome under the accession number PSE117.

ABSTRACT

Cyanobacteria, the only prokaryotes capable of oxygenic photosynthesis, are present in diverse ecological niches and play crucial roles in global carbon and nitrogen cycles. To proliferate in the nature, cyanobacteria utilize a host of stress responses to accommodate periodic changes in environmental conditions. A detailed knowledge of the composition as well as the dynamic changes in the proteome is necessary to gain fundamental insights into such stress responses. Toward this goal, we have performed a large-scale proteomic analysis of the model cyanobacterium *Synechocystis* sp. PCC 6803 under 33 different environmental conditions. The resulting high-quality dataset consists of 22318 unique peptides corresponding to 2369 proteins, a coverage of 65% of the predicted proteome. Quantitative determination of protein abundances have led to the identification of 1221 differentially regulated proteins. Notably, our analysis revealed that a common stress response in cyanobacteria under various environmental perturbations, irrespective of their amplitudes and durations, is the activation of atypical pathways for the acquisition of carbon and nitrogen from urea and arginine. In particular, arginine is catabolized via putrescine to produce succinate and glutamate, sources of carbon and nitrogen, respectively. This study provides the most comprehensive functional and quantitative analysis of the proteome of a phototroph to date, and shows that a significant stress response of cyanobacteria is associated with an unexpected mode of acquisition of carbon and nitrogen.

\body

INTRODUCTION

Most organisms experience daily changes in environmental conditions in their natural habitats. Typically, they tightly coordinate growth with cellular energy levels to survive unfavorable conditions. Photosynthetic organisms, whose energy requirements for cellular metabolism are derived from sunlight, offer attractive model systems to understand the impact of environmental perturbations on organismal physiology. This is particularly true for cyanobacteria, oxygenic photosynthetic prokaryotes of ancient lineage. During their evolution, cyanobacteria have survived large changes in environmental conditions [1]. Additionally, they can readily adapt their cellular metabolism to daily changes in light quality and quantity. Integration of nutrient specific pathways with photosynthetic processes is a key survival mechanism employed by cyanobacteria under changing environmental conditions [2-4]. Such adaptation strategies allow cyanobacteria to balance the supply of electrons from photosynthetic processes with the demands of cellular metabolism and prevent the generation of damaging reactive oxygen species by the excess reducing power.

Assimilation of carbon (C) and nitrogen (N) in photosynthetic organisms is one of the main sinks for the reducing power produced by photosynthetic light reactions. Accordingly, cyanobacteria have developed intricate mechanisms to control several pathways involved in the acquisition of C and N. For example, PII, a regulatory protein, has been suggested to balance the acquisition of the two nutrients by sensing the C/N ratio [5]. Thioredoxins have also been shown to link the activity of photosynthetic electron transport chain with the C and N assimilation [2]. Despite the active participation of these proteins, cyanobacteria can assimilate C and N at disparate levels in excess of cellular demands. Excess C and N are stored in the forms of glycogen and cyanophycin granules, respectively, and are subsequently utilized under limiting conditions. *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*), a model cyanobacterium, utilizes oxidative pentose phosphate and glycolytic pathways to obtain C from glycogen granules [6]. However, the pathway for the utilization of cyanophycin is not well understood. Cyanophycin is a polymer of Asp and Arg which must be further catabolized to meet the N-requirement for cellular metabolism. There are at least five known pathways for the catabolism of Arg in prokaryotes. Among them, a pathway utilizing arginase and the urea cycle has been shown to be active in *Synechocystis* [7].

Our understanding of gene regulation linked to the assimilation of C and N, as well as broader cellular adaptation mechanisms under different environmental conditions, have significantly benefited from global transcript profiling in *Synechocystis* [8]. Generally, it has been observed that such perturbations lead to downregulation of genes involved in light absorption and photosystems, as well as

genes involved in C fixation and N assimilation. However, many of the studies have reported a complex regulation of genes involved in the C and N assimilation. For example, genes involved in N assimilation using nitrate (NO_3^-) as a substrate respond negatively whereas those involved in utilization of either ammonia (NH_3), urea or Arg as substrate respond positively to preferential illumination of PSII [9]. Importantly, the impact of transcript regulation on protein levels remains poorly understood, in part due to several previous studies showing poor correlation between transcriptomic and proteomic datasets [10, 11]. Because proteins are directly responsible for cellular functions, measurements of protein abundances could provide significant clues to the modulation of cellular functions during different environmental perturbations. Several proteomic studies under diverse environmental conditions have been undertaken in *Synechocystis* [12-30]. However, such studies have not yielded a comprehensive understanding of cellular adaptations, either due to low proteome coverage or due to the limited information on the changes in proteins abundance. In the current study, we have analyzed the proteome of *Synechocystis* across 33 different environmental conditions. Our efforts have led to a 65% proteome coverage, resulting in the most complete functional and qualitative proteome of any photosynthetic organism to date. Our analysis of differentially regulated proteins showed that *Synechocystis* activates alternate pathways for the acquisition of C and N under diverse environmental conditions.

RESULTS

Determination of the composition of *Synechocystis* proteome

To obtain a comprehensive proteomic description of *Synechocystis*, we collected samples from cells grown under 33 different environmental conditions. These included time series studies of *Synechocystis* growth under nutrient-limiting conditions followed by recovery under nutrient-sufficient conditions. The kinetics of pigment loss, a typical observable phenotype associated with nutrient starvation in cyanobacteria, were quite variable between nutritional conditions, possibly due to the ability of cyanobacteria to store several nutrients in the form of inclusion bodies that can be utilized during starvation (Fig. S1). Therefore, we used a strategy that involved prolonged starvation for 6 days followed by recovery with the addition of the limiting nutrient. This “starve and recovery strategy” resulted in a uniform recovery response. We also exposed cells to excess sodium chloride (2 M), carbon dioxide (3%) with a recovery under ambient air, Glc (5 mM), and low (20°C) and high (38°C) temperatures.

We utilized a sensitive LC-MS peptide-based “bottom-up” approach to maximize the proteome coverage of *Synechocystis* [31]. To increase the coverage of membrane proteins, total cellular extracts from 33 environmental conditions were separated into membrane and soluble fractions by centrifugation prior to tryptic digestion. We identified a total of 22318 unique tryptic peptides out of the predicted 48534 with a confidence criterion of 95% (Dataset S1). These observed peptides correspond to 2369 proteins of the predicted 3663 for the *Synechocystis* genome, a coverage of 65% (Dataset S2). Bias analyses show that the observed proteins were uniformly distributed among different functional categories (Fig. 1A), including photosynthesis, which contains a large number of membrane proteins. The proteome data observed in this study consisted of 67% and 55% of predicted soluble and membrane proteins, respectively (inset, Fig. 1B). Importantly, increasing number of trans-membrane helices had little impact on the identification of membrane proteins (Fig. 1B). However, we noticed that number of identified peptides for membrane proteins decreased with increasing negative or positive hydrophobicity values (Fig. S2A). Analysis of a subgroup of membrane proteins showed that most of the peptides identified were from the cytosolic loop regions of the proteins. Other bias analyses determined that detection of the observed peptide was not dependent on peptide mass (Fig. S2B), length (Fig. S2C), or isoelectric point (Fig. S2D).

Previous proteomic studies in *Synechocystis* have resulted in the identification of 1099 proteins (Dataset S3). A comparative analysis of these proteins with those identified in the present study shows that 1010 proteins were commonly identified (Fig. 1C). 89 previously identified proteins were not observed in our study whereas 1359 proteins were uniquely identified in the present study. A large number of these proteins (758) are currently annotated as hypothetical in Cyanobase [32]. Thus, our

results have provided direct proof of the presence of over one half of hypothetical proteins in *Synechocystis*.

Quantitative analysis of protein response to various perturbations

We re-examined 12 environmental conditions for quantitative determination of protein abundance (Table 1). A total of 1221 differentially regulated proteins were identified (Dataset S4). The number of differentially regulated proteins in each condition varied from a low of 267 (cold shock) to a high of 553 (N depletion) proteins (Table 1). A majority of differentially regulated proteins (56%-76%) in most conditions, with the exception of cold shock and N depletion, were upregulated. Under cold stress and N depletion, approximately 62 and 86% of the differentially regulated proteins, respectively, were downregulated. However, addition of either NO_3^- or NH_3 to the N depleted cells had significant effect on the expression patterns of proteins. Under these conditions, the number of upregulated proteins increased from 77 to 231 (with NO_3^-) and 257 (with NH_3). At the same time, the number of downregulated proteins decreased from 476 to 150 (with NO_3^-) and 141 (with NH_3). The number of up- and down-regulated proteins during the recovery in the presence of either NO_3^- or NH_3 was similar to those of other nutrients. A small number of proteins were differentially regulated in a stress-specific manner (Table 1). Most of such proteins have no known functions.

Examination of cellular processes based on differential regulation of their associated proteins showed that response varies widely between conditions and, surprisingly, is not correlated with physiological responses (Fig. 2). For example, depletion of Fe, P, S and N is accompanied by significant chlorosis and slow growth. However, we observed that most processes were downregulated significantly only under N depletion and to some extent, under Fe depletion. In fact, in most other conditions, we noticed that majority of proteins in cellular processes were upregulated. This is even true for ribosomal proteins whose expression has been typically linked with the growth of an organism. Though, the number of differentially regulated ribosomal proteins varied depending on conditions, we found that ribosomal proteins were downregulated under N, P, and Fe depletion, and upregulated under S depletion and heat shock. As expected, majority of ribosomal proteins were upregulated during all recovery stages. A large number of proteins with unknown functions (~ 40%) showed significant differential regulation. Several of these proteins showed stress specific regulation, providing evidence to the crucial roles in the cellular adaptation (Fig. 2).

Proteins involved in amino acid (AA) biosynthesis, Glc metabolism, TCA cycle, and cytochrome b_6f complex showed strong upregulation in majority of environmental conditions. In general, enzymes known to catalyze key reactions in any given pathway were differentially regulated. For example, glycogen phosphorylase, which catalyzes the release of Glc from glycogen, was strongly upregulated in

all studied conditions. We also found that fructose-bisphosphate aldolase, which catalyzes the formation of glyceraldehyde 3-phosphate and dihydroxyacetone phosphate from fructose 1,6-bisphosphate, and pyruvate dehydrogenase, which converts pyruvate into acetyl-CoA, are strongly upregulated in all conditions. Similarly, key proteins involved in the biosynthesis of AAs belonging to aromatic, Asp, branched chain, Ser and Glu families were upregulated. In contrast, most photosynthetic proteins, including phycobilin proteins, did not show significant changes in protein abundance. Similarly, very few proteins involved in pigment biosynthesis showed differential regulation. However, proteins with critical functions were differentially regulated. For example, heme oxygenase, involved in the multi-step monooxygenase reaction to produce biliverdin IX α and CO from protoheme, was downregulated during nutrient depletion conditions.

Concordance between transcriptomic and proteomic datasets

The global scale of proteome coverage and protein expression profiles obtained in this study enabled the first large-scale comparison of gene expression at RNA and protein levels in *Synechocystis*. The differentially regulated proteins identified under cold shock, P, S, N, and Fe depletions were compared with the differentially regulated genes from the respective DNA microarray studies [33-37]. To allow for a uniform comparison, all five transcriptomic datasets were reanalyzed. We used a fold change of 1.5 to identify differentially regulated genes from these datasets. Concordance analysis showed that the expression changes between these two studies were quite low. Agreement between the two studies was lowest for S depletion and highest for N depletion (Table 2). Analysis of correlated and anti-correlated genes revealed interesting results. Stress-specific genes showed similar expression patterns in both transcriptomic and proteomic studies. For example, expression of nutrient specific transporters showed strong concordance. The relatively higher concordance seen under N depletion was due to similar expression patterns of ribosomal and photosynthesis genes. On the other hand, expression patterns of photosynthetic genes were anti-correlated under Fe depletion with downregulation at transcript levels and upregulation at protein levels.

Alternate pathway for assimilation of N and C under various perturbations

Several proteins involved in N assimilation showed significant differential regulation. Generally, proteins involved in transport of NO₃⁻ were downregulated in cells grown under nutrients depletion, cold and heat shock. Repletion of nutrients to starved cells led to the upregulation of these transporter proteins. Interestingly, recovery of N depleted cells in the presence of NH₃ did not lead to the upregulation of NO₃⁻ transport proteins. To compensate for reduced NO₃⁻ uptake, our data shows that cells upregulate proteins involved in transport and utilization of urea and Arg. UrtE and UrtD involved in transport of urea were strongly upregulated under most environmental conditions. Our data also shows that urease which

converts urea into CO₂ and NH₃ is upregulated by 4-fold under most conditions. Similarly, BgtB and to some extent BgtA, the periplasmic and ATP-binding components of an Arg transporter, are upregulated under most conditions. We also find that cyanophycinase, involved in the breakdown of cyanophycin into Arg and Asp, was somewhat upregulated under certain conditions. These results suggest that a common response in *Synechocystis* under perturbation is to reduce the uptake of NO₃⁻ and increase the uptake of alternate N sources.

Arg and Asp must be further catabolized to acquire N for cellular metabolism. An Arg catabolic pathway has been described in *Synechocystis* which combines the arginase pathway and urea cycle [7]. All proteins in this pathway were detected in our study, however, none showed significant regulation suggested that this pathway is not the preferred route for Arg catabolism under perturbations. Further analysis shows that Arg is preferentially broken down into agmatine via Arg decarboxylase. Arg decarboxylase (Slr1312 and Slr0662), the first committed enzymes in this pathway, are upregulated by 1.5 to 3 fold as compared to control conditions (Fig. 3, Dataset S6). Agmatine can subsequently be catabolized into putrescine via the putative agmatinase (Sll1077) or the arginase (Sll0228). These two proteins were not observed in the control cells; however, peptides corresponding to these proteins were detected under various perturbations (Dataset S2). Putrescine is known to play a critical role in DNA, RNA and protein synthesis as well as in cell proliferation and differentiation. Furthermore, putrescine also serves as a source for C and N in *E. coli* and *Pseudomonas* [38, 39] where it is converted to succinate. Examination of proteins involved in putrescine degradation suggested that this pathway is also active in *Synechocystis* (Fig. 3, Dataset S6). Slr1022 shows strong similarity to the proteins (YgjG and SpuC) involved in degradation of putrescine into aminobutyrate in *E. coli* and *Pseudomonas* [38, 39]. Further evidence that succinate is produced from Arg comes from the strong upregulation of succinate dehydrogenase and malate dehydrogenase (Fig. 3, Dataset S6). Thus conversion of Arg into succinate not only allows generation of the TCA cycle intermediate but in the process releases Glu, NH₃ and CO₂. NH₃ is assimilated into Glu whereas CO₂ is fixed by the Rubisco. Indeed, we find that transporters of Ci were not differentially regulated. However, several carbon concentrating mechanism proteins, required for concentrating intracellular Ci, showed significant upregulation.

DISCUSSION

Here, we report the most comprehensive functional and quantitative analysis of *Synechocystis* proteome to date. The resulting proteome consists of 2369 unique proteins (65% of the predicted proteins), 1221 of which have been identified as differentially regulated under 12 different environmental conditions. Several bias analyses show that proteins identified in this study are representative of the entire genome. Importantly, functional categories based analysis shows that the observed proteins were

uniformly distributed. Identification of 758 proteins with unknown functions, of which 326 were differentially regulated, provides direct evidence of their roles in *Synechocystis* physiology. Taken together, this study has revealed the global proteome makeup of *Synechocystis* and facilitated a systems-level analysis of cellular response under different perturbations.

Analysis of 1221 differentially regulated proteins shows that *Synechocystis* utilizes few stress-specific proteins to optimize of cellular functions under perturbations (Table 1). Many of these proteins have no known function. In contrast, a large number of proteins associated with housekeeping functions were commonly differentially regulated. For example, key proteins involved in biosynthesis of all AA families were strongly upregulated. These results suggest that despite the prolonged starvation for essential nutrients, cells continue to maintain a metabolically active state by seeking either the limiting nutrients or alternate nutrients for growth. Typically, transporters involved in acquisition of Fe, S and P were upregulated. While S transporters were specifically upregulated under S depletion, expression of Fe transporters was also upregulated under P depletion and vice versa. In contrast, NO_3^- transporters were downregulated under N depletion. In fact, they were also downregulated under other environmental conditions. Previous studies using DNA microarrays have also shown strong downregulation of NO_3^- transporters under environmental conditions [3, 9]. It has been suggested that changes in environmental conditions leads to reduced transport of NO_3^- while simultaneously activating the pathway involved in the transport of alternate N substrates including NH_3 , urea and Arg. A similar mechanism under changing nutrient conditions is apparent from the analysis of proteomic datasets. Our data shows that the preferred substrate for N acquisition is urea and Arg.

The most striking cellular strategy from the analysis of differentially regulated proteins is revealed in the way cells acquire C and N under different perturbations. Our results suggest that any change in conditions, irrespective of the amplitudes and/or duration, immediately leads to the activation of alternate pathways towards the acquisition of C and N. Although majority of perturbations had little impact on levels of proteins involved in photosynthesis, most perturbations affected the efficiency of photosynthetic light reactions resulting in a lower production of energy. It should be mentioned that most photosynthetic proteins were identified as strongly downregulated under N depletion and therefore, we do not believe that the lack of differential regulation is related to an inability in the detection of peptides. The assimilation of C and N is an energy intensive process requiring significant amount of ATP, NADPH and reduced ferredoxin. Therefore, decreased energy production leads to the activation of alternate C and N assimilation pathways. Our data strongly suggest that urea and Arg are the preferred substrates for both C and N. Cells also actively seek for internal/external carbon sources as apparent from upregulation of proteins involved in glycolysis and Glc transport. Arg is preferentially catabolized via putrescine using

the pathway recently characterized in *E. coli* and *Pseudomonas* [38, 39]. Activation of this pathway allows cells to obtain both C in the form of succinate, and CO₂ and N in the form of Glu and NH₃. Asp, which can be generated from cyanophycin, can serve two purposes. It can be utilized for the synthesis of Met, Lys and Thr. Additionally, it can be combined with 2-oxoglutarate to produce Gln, which can then be combined with NH₃ to produce Glu. This result also indirectly suggests that Arg is not directly converted to Glu using the pathway identified [7]. The utility of this arrangement enables that 2-oxoglutarate generated by the glycolysis is channeled towards the production of Glu.

In conclusion, the proteome analysis presented in this study has provided a unique catalogue of the proteome makeup of *Synechocystis* under various environmental conditions. We believe that the knowledge of the functional information of when and how proteins with known and unknown functions are expressed is going to be a foothold for future experimental studies. Analysis of dynamic changes in the proteome has provided insights into cellular adaptations under various environmental perturbations. Our results showed that a key cellular adaptation leads to the activation of alternate pathways for the acquisition of C and N, which are the two major sinks for reducing powers generated by the photosynthetic light reactions in cyanobacteria.

METHODS

Culture conditions. *Synechocystis* cultures were grown to a density of 2×10^8 cells/mL as described [40]. Cells were washed twice with 100 mM TES, pH 8.0. The washed cells were inoculated into complete BG11 and sampled at 0, 4, 6, 8, and 16 days. For growth under nutrient depletion conditions, washed cells were grown in BG11 depleted of either NO_3^- (N depletion), sulfate (S depletion) or phosphate (P depletion) and sampled after 6 days. For Fe depletion, cells were incubated twice in 20 mM MES, 10 mM EDTA, pH 5.0 for 10 min each, inoculated in BG11 depleted of Fe and sampled at 6 days. After 6 days, starved cultures were fed with NO_3^- , NH_3 , S, P, or Fe as appropriate and sampled at 4 and 24 h. For salt stress, 0.5 M NaCl was added to cells grown in complete BG11 and sampled at 0.5, 6, and 24 h. Cells were also collected after 3 and 6 days growth in the presence of 5 mM Glc and 10 μM DCMU. Cells grown in 3% CO_2 were sampled at 1 and 25 h and then moved back to 0.3% CO_2 and sampled at 2 h. Lastly, cells were subjected to heat (38°C) or cold (22°C) shock and sampled at 1, 4, and 24 h.

Sample preparation. Cells were harvested by centrifugation at 6000xg for 5 min. Membrane and soluble fractions from total cell extracts were prepared as described [40] with following modifications: 1 min break/rest, 6 cycles without detergent. Membrane fractions were washed with 100 mM ammonium bicarbonate buffer, pH 8.0. Protein concentrations were determined by BCA Assay (Pierce, Rockford, IL). Fractions were denatured using 7 M urea, 2 M thiourea, 1% CHAPS, and DTT (5 mM for soluble samples, 10 mM for membrane samples) at 60°C for 45 min. All samples were diluted 1:10 in ammonium bicarbonate (100 mM for soluble samples, 50 mM for membrane samples) prior to digestion by sequencing-grade modified trypsin (Promega, Madison, WI) at a 1:50 (w/w) trypsin:protein ratio for 5 h at 37°C. Digestion was interrupted by rapid freezing in liquid nitrogen. Samples were fractionated via SCX using a PolySulfoethyl A 200 mm x 2.1mm, 5 μM , 300-Å column with 10 mm x 2.1 mm guard column (PolyLC, Inc., Columbia, MD) as described [40]. The peptides were resuspended in mobile phase A, and separated on an Agilent 1100 system (Agilent, Palo Alto, CA) equipped with a quaternary pump, degasser, diode array detector, Peltier-cooled autosampler and fraction collector (set at 4 °C).

Reversed phase LC separation and MS/MS analysis of peptides. Analyses were performed with Finnigan LCQ and LTQ ion trap mass spectrometers (ThermoFinnigan, San Jose, CA) using an in-house manufactured ESI source as described [40]. Each intact sample and SCX fraction was analyzed via capillary LC-MS/MS. SEQUEST analysis software was used to match MS/MS fragmentation spectra to sequences from the 2004 annotation of *Synechocystis* (3663 total entries, no enzyme search, ± 3 Da tolerance for parent MS peak) [32]. The criteria selected for filtering of LCQ and LTQ data followed methods based upon a reverse database false positive model that provides a target of 95% confidence in peptide identifications [41]. An additional 8 proteins and 180 peptides were included from wild type spectra MSMS_01 - MSMS_06 (NCBI GEO accession GSE9577) [40].

Data Processing and Analysis. Peptides matching multiple proteins were assigned to each of the matching proteins. Protein spectral counts were calculated by summing numbers of observed peptides for each protein in all fractions. For sample replicates, all combinations of soluble and membrane replicate pairings were summed, and the average and standard deviation of these combinations were used for the final values. Differentially expressed proteins were identified using three criteria: [$mean1/mean2 \geq 1.5$, $mean1 - mean2 > 1$, $(mean1 - 2stddev1) - (mean2 + 2stddev2) > 0$] where mean1 and stddev1 are the values of largest mean of the treatment or control. Proteins were categorized as up- or down-regulated, based on whether abundance was higher or lower in the treatment compared to the control experiment. Transmembrane helices (TMHs) were predicted with TMHMM [42]. SignalP [43] was used to predict cleavage sites for signal peptides, using the Gram-negative bacteria setting. TMHMM predicted helices shorter than 15 AAs and those overlapping with signal peptides were discarded. Proteins were considered to be membrane proteins if at least one TMH was predicted. Peptide hydrophobicities were calculated by summing the hydrophobicities of the AA sequences using the Kyte and Doolittle scale. Peptide hydrophobicity, length, and mass histograms were generated for the subset of observed fully tryptic peptides of ≥ 5 AAs in length and ≥ 500 Da and compared to ideal tryptic digests of the observed proteins, using the same constraints.

Acknowledgements We thank the members of the Pakrasi and Smith groups for collegial discussions. We thank Dr. N. Murata for providing a transcriptomic dataset for *Synechocystis* under cold stress. This work was supported by the National Science Foundation grants- FIBR EF0425749 and MCB 0745611 programs to H. B. P This work was additionally supported as part of the Membrane Biology Scientific Grand Challenge project at the W. R. Wiley Environmental Molecular Science Laboratory, a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research program (Pacific Northwest National Laboratory).

REFERENCES

- 1 Kasting JF (2004) When methane made climate. *Sci Am* 291: 78-85.
- 2 Lindahl M, Florencio FJ (2003) Thioredoxin-linked processes in cyanobacteria are as numerous as in chloroplasts, but targets are different. *Proc Natl Acad Sci USA* 100: 16107-16112.
- 3 Singh AK et al. (2008) Integration of carbon and nitrogen metabolism with energy production is crucial to light acclimation in the cyanobacterium *Synechocystis*. *Plant Physiol* 148: 467-478.

- 4 Tsinoremas NF, Castets AM, Harrison MA, Allen JF, Tandeau de Marsac N (1991) Photosynthetic electron transport controls nitrogen assimilation in cyanobacteria by means of posttranslational modification of the *glnB* gene product. *Proc Natl Acad Sci USA* 88: 4565-4569.
- 5 Forchhammer K (2004) Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signals to targets. *FEMS Microbiol Rev* 28: 319-333.
- 6 Yang C, Hua Q, Shimizu K (2002) Quantitative analysis of intracellular metabolic fluxes using GC-MS and two-dimensional NMR spectroscopy. *J Biosci Bioeng* 93: 78-87.
- 7 Quintero MJ, Muro-Pastor AM, Herrero A, Flores E (2000) Arginine catabolism in the cyanobacterium *Synechocystis* sp. Strain PCC 6803 involves the urea cycle and arginase pathway. *J Bacteriol* 182: 1008-1015.
- 8 Singh AK et al. (2010) Integrative Analysis of Large Scale Expression Profiles Reveals Core Transcriptional Response and Coordination between Multiple Cellular Processes in a Cyanobacterium. *BMC Systems Biology*.
- 9 Singh AK et al. (2009) A systems-level analysis of the effects of light quality on the metabolism of a cyanobacterium. *Plant Physiol* 151: 1596-1608.
- 10 Corbin RW et al. (2003) Toward a protein profile of *Escherichia coli*: comparison to its transcription profile. *Proc Natl Acad Sci USA* 100: 9232-9237.
- 11 Ideker T et al. (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292: 929-934.
- 12 Fulda S, Huang F, Nilsson F, Hagemann M, Norling B (2000) 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.
- 13 Fulda S et al. (2006) Proteome analysis of salt stress response in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Proteomics* 6: 2733-2745.
- 14 Gan CS, Chong PK, Pham TK, Wright PC (2007) Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *J Proteome Res* 6: 821-827.
- 15 Gan CS, Reardon KF, Wright PC (2005) Comparison of protein and peptide prefractionation methods for the shotgun proteomic analysis of *Synechocystis* sp. PCC 6803. *Proteomics* 5: 2468-2478.
- 16 Herranen M et al. (2004) Towards functional proteomics of membrane protein complexes in *Synechocystis* sp. PCC 6803. *Plant Physiol* 134: 470-481.
- 17 Huang F, Fulda S, Hagemann M, Norling B (2006) Proteomic screening of salt-stress-induced changes in plasma membranes of *Synechocystis* sp. strain PCC 6803. *Proteomics* 6: 910-920.
- 18 Huang F et al. (2004) Isolation of outer membrane of *Synechocystis* sp. PCC 6803 and its proteomic characterization. *Mol Cell Proteomics* 3: 586-595.

- 19 Huang F et al. (2002) Proteomics of *Synechocystis* sp. strain PCC 6803: identification of plasma membrane proteins. *Mol Cell Proteomics* 1: 956-966.
- 20 Kurian D, Jansen T, Maenpaa P (2006) Proteomic analysis of heterotrophy in *Synechocystis* sp. PCC 6803. *Proteomics* 6: 1483-1494.
- 21 Kurian D, Phadwal K, Maenpaa P (2006) Proteomic characterization of acid stress response in *Synechocystis* sp. PCC 6803. *Proteomics* 6: 3614-3624.
- 22 Mata-Cabana A, Florencio FJ, Lindahl M (2007) Membrane proteins from the cyanobacterium *Synechocystis* sp. PCC 6803 interacting with thioredoxin. *Proteomics* 7: 3953-3963.
- 23 Perez-Perez ME, Florencio FJ, Lindahl M (2006) Selecting thioredoxins for disulphide proteomics: target proteomes of three thioredoxins from the cyanobacterium *Synechocystis* sp. PCC 6803. *Proteomics* 6 Suppl 1: S186-195.
- 24 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 J* 274: 791-804.
- 25 Sazuka T, Ohara O (1997) Towards a proteome project of cyanobacterium *Synechocystis* sp. strain PCC6803: linking 130 protein spots with their respective genes. *Electrophoresis* 18: 1252-1258.
- 26 Sergeyenko TV, Los DA (2000) Identification of secreted proteins of the cyanobacterium *Synechocystis* sp. strain PCC 6803. *FEMS Microbiol Lett* 193: 213-216.
- 27 Simon WJ, Hall JJ, Suzuki I, Murata N, Slabas AR (2002) Proteomic study of the soluble proteins from the unicellular cyanobacterium *Synechocystis* sp. PCC 6803 using automated matrix-assisted laser desorption/ionization-time of flight peptide mass fingerprinting. *Proteomics* 2: 1735-1742.
- 28 Srivastava R, Pisareva T, Norling B (2005) Proteomic studies of the thylakoid membrane of *Synechocystis* sp. PCC 6803. *Proteomics* 5: 4905-4916.
- 29 Wang Y, Sun J, Chitnis PR (2000) Proteomic study of the peripheral proteins from thylakoid membranes of the cyanobacterium *Synechocystis* sp. PCC 6803. *Electrophoresis* 21: 1746-1754.
- 30 Sazuka T, Yamaguchi M, Ohara O (1999) Cyano2Dbase updated: linkage of 234 protein spots to corresponding genes through N-terminal microsequencing. *Electrophoresis* 20: 2160-2171.
- 31 Page JS, Masselon CD, Smith RD (2004) FTICR mass spectrometry for qualitative and quantitative bioanalyses. *Curr Opin Biotechnol* 15: 3-11.
- 32 Nakamura Y, Kaneko T, Hirosawa M, Miyajima N, Tabata S (1998) CyanoBase, a www database containing the complete nucleotide sequence of the genome of *Synechocystis* sp. strain PCC6803. *Nucleic Acids Res* 26: 63-67.
- 33 Osanai T et al. (2006) Nitrogen induction of sugar catabolic gene expression in *Synechocystis* sp. PCC 6803. *DNA Res* 13: 185-195.

- 34 Singh AK, McIntyre LM, Sherman LA (2003) Microarray analysis of the genome-wide response to iron deficiency and iron reconstitution in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol* 132: 1825-1839.
- 35 Suzuki I, Kanesaki Y, Mikami K, Kanehisa M, Murata N (2001) Cold-regulated genes under control of the cold sensor Hik33 in *Synechocystis*. *Mol Microbiol* 40: 235-244.
- 36 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.
- 37 Zhang Z, Pendse ND, Phillips KN, Cotner JB, Khodursky A (2008) Gene expression patterns of sulfur starvation in *Synechocystis* sp. PCC 6803. *BMC Genomics* 9: 344.
- 38 Chou HT, Kwon DH, Hegazy M, Lu CD (2008) Transcriptome analysis of agmatine and putrescine catabolism in *Pseudomonas aeruginosa* PAO1. *J Bacteriol* 190: 1966-1975.
- 39 Kurihara S et al. (2005) A novel putrescine utilization pathway involves gamma-glutamylated intermediates of *Escherichia coli* K-12. *J Biol Chem* 280: 4602-4608.
- 40 Wegener KM et al. (2008) High sensitivity proteomics assisted discovery of a novel operon involved in the assembly of photosystem II, a membrane protein complex. *J Biol Chem* 283: 27829-27837.
- 41 Qian WJ et al. (2005) Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: the human proteome. *J Proteome Res* 4: 53-62.
- 42 Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305: 567-580.
- 43 Emanuelsson O, Brunak S, von Heijne G, Nielsen H (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2: 953-971.

FIGURE LEGENDS

Fig. 1. Proteome Coverage and Bias Analysis. (A) Distribution of the observed proteins in various functional categories. The numbers in parentheses indicate the observed proteins (bold font) out of the predicted proteins. (B) Detection of membrane proteins as a function of transmembrane helices. Black and white bars indicate the numbers of predicted and observed proteins, respectively. The inset shows the numbers of soluble and membrane proteins detected in the present study. (C) Comparative analysis of proteins (gray circle) identified in this study with previously published proteomic studies (black circle) as a fraction of the predicted proteome (white circle).

Fig. 2. Differential Regulation of Proteins under Different Perturbations. 1221 differentially regulated proteins were grouped in functional categories and a heat map was generated using Spotfire 7.0. Color bar indicates protein fold change in experimental conditions as compared to levels in complete BG11.

Fig 3. Global Stress Response under Different Perturbations. Numbers denote the enzymes required for various reactions whereas colors indicate either increased (red) or decreased (green) protein abundances under stress conditions compared to those in the BG11 control sample. Black color indicates that proteins are not observed in BG11 control sample. Numbers correspond to the following enzymes: 1=cyanophycinase, 2 & 3=Arg decarboxylase, 4=agatinase, 5=arginase, 6 & 7=4-Aminobutryaldehyde dehydrogenase, 8-10=4-Aminobutyrate transaminase, 11=aldehyde dehydrogenase, 12=adenylosuccinate lyase, 13=L-argininosuccinate lyase, 14=L-Asp oxidase, 15=Asp aminotransferase, 16=succinyl-CoA synthetase, 17=succinate dehydrogenase, 18=malate dehydrogenase, 19=Glu synthase, 20=urease, 21=Glu-NH₃ ligase, 22=RubisCo, 23=CCM proteins. Abundance values for these proteins are provided in Dataset S6.

Table 1. Numbers of differentially regulated proteins under 12 environmental conditions.

Treatment	Number of Proteins			
	Differentially Abundant	Upregulated	Downregulated	Stress Specific
CO ₂	312	192 (62%)	120 (38%)	16
Cold Shock	267	102 (38%)	165 (62%)	12
Heat Shock	382	214 (56%)	168 (44%)	12
Fe Depletion	401	235 (59%)	166 (41%)	12
Fe Repletion	375	244 (65%)	131 (35%)	21
N Depletion	553	77 (14%)	476 (86%)	26
N Repletion	381	231 (61%)	150 (39%)	29
NH ₄ Repletion	398	257 (65%)	141 (35%)	18
P Depletion	356	268 (75%)	88 (25%)	4
P Repletion	415	316 (76%)	99 (24%)	5
S Depletion	395	247 (63%)	148 (37%)	18
S Repletion	376	275 (73%)	101 (27%)	19

Table 2. Concordance between microarray and proteomic studies in *Synechocystis*.

	Cold Stress	Fe depletion	P depletion	S depletion	N depletion
Number of Genes Correlated	8	56	57	14	110
Number of Genes Anti-Correlated	13	53	16	5	26
Total genes differentially expressed in proteomic and transcriptomic datasets	21	109	73	19	136

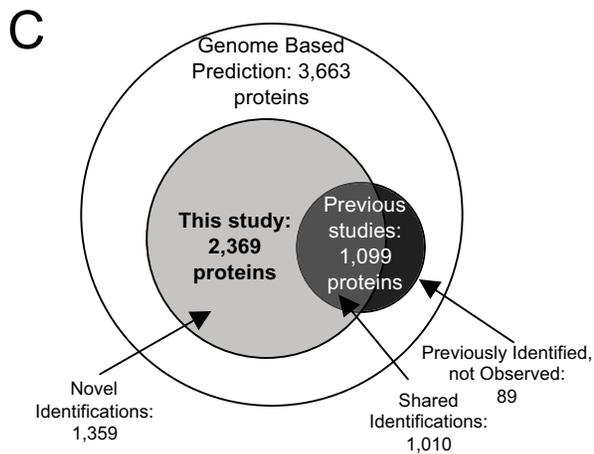
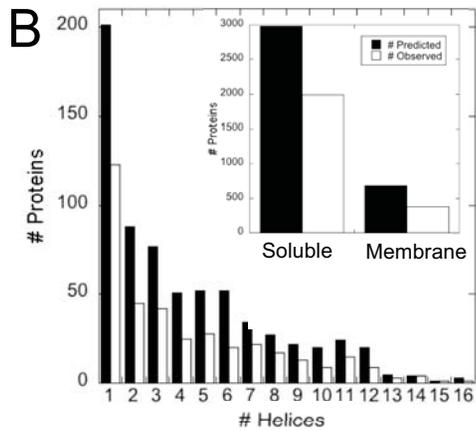
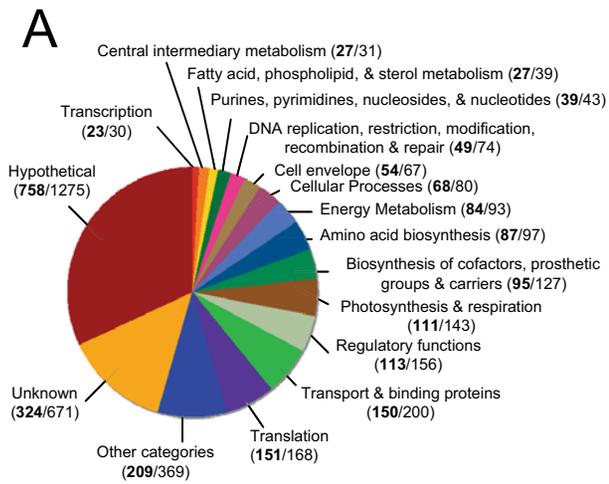


Figure 1. Wegener et al.

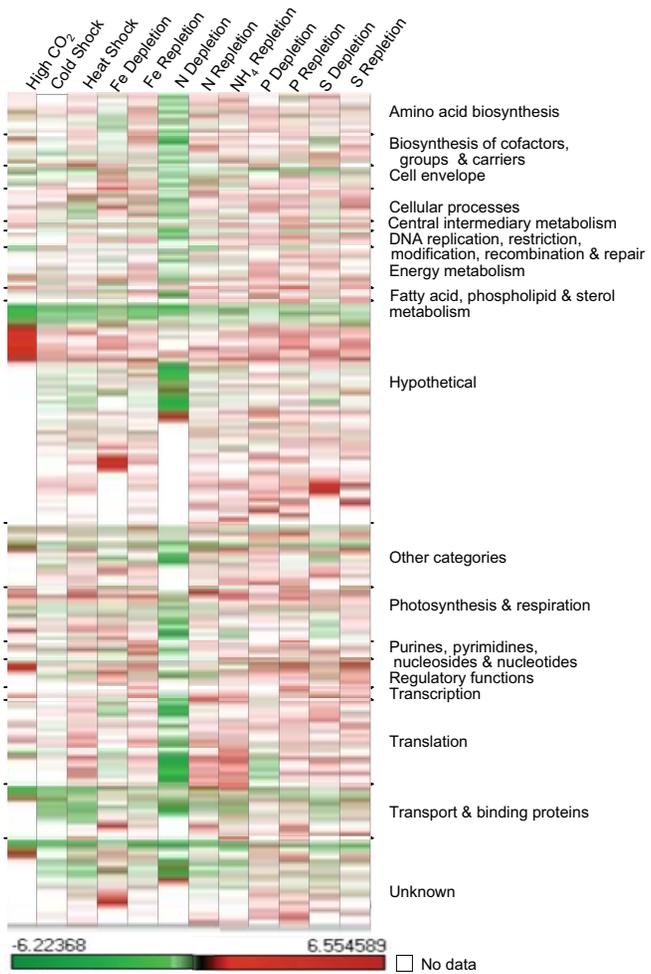


Figure 2. Wegener et al.

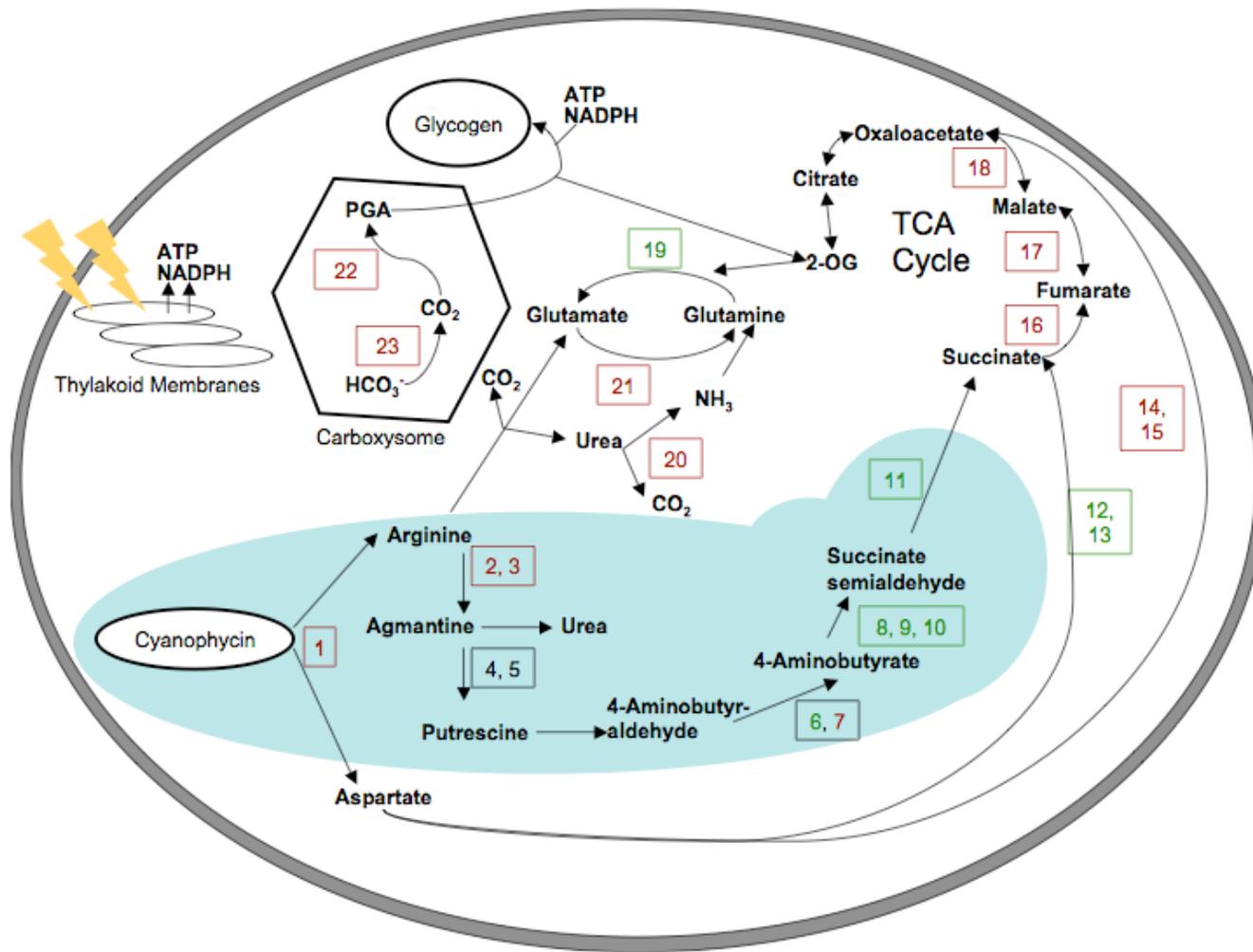


Figure 3. Wegener et al.