Proteome mapping of a cyanobacterium reveals distinct compartment organisation and cell-dispersed metabolism

Laura L. Baers[1], Lisa M. Breckels[1,2], Lauren A. Mills[3], Laurent Gatto[1,2], Michael J. Deery[1], Tim J. Stevens[4], Christopher J. Howe[1*], Kathryn S. Lilley[1*], David J. Lea-Smith[1,3*]

[1] Department of Biochemistry, University of Cambridge, CB2 1QW, United Kingdom
[2] Computational Proteomics Unit, Cambridge Centre for Proteomics, University of Cambridge, CB2 1QW, United Kingdom
[3] School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom
[4] MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, United Kingdom

[*]Corresponding author emails: ch26@cam.ac.uk, k.s.lilley@bioc.cam.ac.uk, D.Lea-Smith@uea.ac.uk

Short title: Mapping the proteome of a cyanobacterium

One sentence summary: The most extensive proteome map of an entire cyanobacterial cell demonstrates that thylakoid and plasma membrane proteins have distinct functions and that metabolic pathways are dispersed throughout the cell.

L.L.B., C.J.H., K.S.L., D.J.L-S. conceived the original screening and research plans; C.J.H., K.S.L., D.J.L-S. supervised the experiments; L.L.B. performed most of the experiments; L.M.B, L.G., M.J.D., T.J.S., L.A.M. performed bioinformatics analysis; L.L.B., C.J.H., K.S.L., D.J.L-S. designed the experiments and analyzed the data; L.L.B., C.J.H., K.S.L., D.J.L-S. conceived the project and wrote the article with contributions of all the authors; D.J.L-S. supervised and completed the writing. D.J.L-S. agrees to serve as the author responsible for contact and ensures communication.

Abstract

Cyanobacteria are complex prokaryotes, incorporating a Gram-negative cell wall and internal thylakoid membranes (TMs). However, localisation of proteins within cyanobacterial cells is poorly understood. Using subcellular fractionation and quantitative proteomics we produced an extensive subcellular proteome map of an entire cyanobacterial cell, identifying ~67% of proteins in *Synechocystis* sp. PCC 6803, ~1000 more than previous studies. 1,712 proteins were assigned to six specific subcellular regions. Proteins involved in energy conversion localised to TMs. The majority of transporters, with the exception of a TM-localised copper importer, resided in the plasma membrane (PM). Most metabolic enzymes were soluble although numerous pathways terminated in the TM (notably those involved in peptidoglycan monomer, $NADP^+$, heme, lipid and carotenoid biosynthesis), or PM (specifically, those catalysing lipopolysaccharide, molybdopterin, FAD and phylloquinol biosynthesis). We also identified the proteins involved in the TM and PM electron transport chains. The majority of ribosomal proteins and enzymes synthesising the storage compound polyhydroxybuyrate formed distinct clusters within the data, suggesting similar subcellular distributions to one another, as expected for proteins operating within multi-component structures. Moreover, heterogeneity within membrane regions was observed, indicating further cellular complexity. Cyanobacterial TM protein localisation was conserved in *Arabidopsis thaliana* chloroplasts, suggesting similar proteome organisation in more developed photosynthetic organisms. Successful application of this technique in *Synechocystis* suggests it could be applied to mapping the proteomes of other cyanobacteria and single-celled organisms. The organisation of the cyanobacterial cell revealed here substantially aids our understanding of these environmentally and biotechnologically important organisms.

Introduction

Cyanobacteria (oxygenic photosynthetic bacteria) are a widespread and abundant phylum of environmental and biotechnological importance (Zwirglmaier et al., 2008; Ducat et al., 2011). Amongst prokaryotes they are distinguished by the presence of a highly differentiated series of internal thylakoid membranes (TM), parts of which are in close contact, but do not fuse with the plasma membrane (PM) (Rast et al., 2019). The cell envelope is similar to other Gram-negative bacteria, consisting of the PM, peptidoglycan layer and outer membrane (OM) (Stanier and Cohen-Bazire, 1977) (Fig. 1).

Cytoplasmic compartments such as the carboxysome, a proteinaceous structure in which carbon fixation occurs, and various storage bodies containing glycogen, cyanophycin, polyhydroxybutyrate (PHB), lipids and polyphosphate, add further complexity to the cell (Liberton et al., 2006; van de Meene et al., 2006). Many species also contain multiple chromosomal copies (Griese et al., 2011), and in the case of the model cyanobacterium, *Synechocystis* sp. PCC 6803 (*Synechocystis*), approximately 70% of ribosomes are localised in the central cytoplasm with the remainder in the cytoplasmic periphery between the PM and TM (20%) or within the TM stacks (10%) (van de Meene et al., 2006).

Given this intricate organisation, characterising the distribution of the subcellular proteome is critical in understanding the biochemical and physiological processes within the cell and the role of individual cellular components, as their spatial organisation will reflect protein function (Dreger, 2003). Moreover, the chloroplasts of algal and plant cells are descended from an internalised cyanobacterium (Howe et al., 2008), with many cyanobacterial genes (De Las Rivas et al., 2002; Martin et al., 2002) and structural features (Hinterstoisser et al., 1993) conserved in photosynthetic eukaryotes (Fig. 2). Therefore, knowledge of cyanobacterial protein localisation will help in understanding the evolution of chloroplast ultrastructure from its cyanobacterial ancestors.

Multiple studies have attempted to verify the distribution of proteins in cyanobacteria, via analysis of isolated cellular fractions. This approach has been used to elucidate the proteomes of the membranous (Wang et al., 2000; Huang et al., 2002; Herranen

et al., 2004; Huang et al., 2004; Srivastava et al., 2005; Huang et al., 2006; Pisareva et al., 2007; Wang et al., 2009; Zhang et al., 2009; Agarwal et al., 2010; Rowland et al., 2010; Wegener et al., 2010; Pisareva et al., 2011; Li et al., 2012; Plohnke et al., 2015; Liberton et al., 2016) and soluble (Simon et al., 2002; Huang et al., 2006; Kurian et al., 2006a; Kurian et al., 2006b; Slabas et al., 2006; Suzuki et al., 2006; Zhang et al., 2009; Rowland et al., 2010; Wegener et al., 2010; Plohnke et al., 2015) compartments that constitute *Synechocystis* (Supplemental Table S1). In these studies membranes were typically isolated using two-phase aqueous polymer partitioning and/or sucrose density ultracentrifugation, followed by gel based or shotgun proteomic analysis.

This approach has been applied to investigate PM (Huang et al., 2002; Pisareva et al., 2007; Pisareva et al., 2011; Liberton et al., 2016), TM (Wang et al., 2000; Srivastava et al., 2005; Agarwal et al., 2010; Pisareva et al., 2011; Liberton et al., 2016), OM (Huang et al., 2004) and soluble fractions (Simon et al., 2002). However, there are numerous inconsistencies in the assignment of protein localisation to subcellular fractions between these studies (Srivastava et al., 2005; Pisareva et al., 2007; Pisareva et al., 2011; Liberton et al., 2016), suggesting that this approach of membrane fractionation could have limitations due to technical difficulties in separating cellular compartments and/or the complicated organisation of cyanobacterial cells (Pisareva et al., 2011). For example, these methods have been shown to give ‘purified’ PM fractions that actually contain detectable amounts of TM e.g. (Zhang et al., 2015; Lea-Smith et al., 2016b). In addition, isolating membranes via two-phase aqueous polymer partitioning results in considerable losses of cellular material and under-sampling of the proteome. Furthermore, both the PM and TM may be heterogeneous (Srivastava et al., 2006; Agarwal et al., 2010; Pisareva et al., 2011) and previous work has suggested that only a hydrocarbon-rich fraction of the TM, and not the whole membrane, is purified via two-phase partitioning (Lea-Smith et al., 2016b). For example, a highly curved ‘convergence membrane’ substructure in the TM was recently observed, which was in close contact with the PM, and may play a role in biogenesis of thylakoid proteins (Rast et al., 2019).

Recently, a study was published by Liberton *et al* on the distribution of proteins between the PM and TM in *Synechocystis* (Liberton et al., 2016). Two-phase separation was used to separate the cellular membranes into two partitions

representative of the PM and TM. Proteins within these two fractions were then labelled using isobaric tags and analysed via mass spectrometry (MS), resulting in the quantification of 1,496 proteins. Looking at the distribution of proteins across the two phases, the authors were able to assign 459 and 176 proteins to the PM or TM, respectively. This study eliminated the need to obtain complete purification of either membrane. However, much of the cellular material was discarded during the purification stages, and the simplified approach of partitioning into two phases meant that other subcellular compartments, such as the OM, the soluble proteins from the cytosol, thylakoid lumen and periplasmic space, the carboxysome and storage bodies, were not taken into account. Additionally, the method was insensitive to proteins residing in multiple compartments. Furthermore, quantitative variation within the biological replicates, noted by the authors, rendered the dataset limited in its utility to assign membrane proteins to specific subcellular structures.

In this study we adapted the hyperLOPIT approach to map the proteins of the entire *Synechocystis* cell using spatial proteomics applied to cellular fractions enriched with various subcellular membranes (Mulvey et al., 2017; Thul et al., 2017). This method relies on the correlation of proteins within these subcellular fractions using stable isotope tagging coupled with machine learning approaches to assign similar fractionation behaviour. The output of this method is the steady state location of a protein within a cell. This approach resulted in the identification of 2,445 proteins. This study provides the most complete description of the *Synechocystis* proteome to date, covering ~67% of the predicted proteome, and assigns 1,712 proteins to specific regions of the cell, which can be interrogated via an interactive database. These regions include the PM, TM, small and large ribosomal subunits, PHB storage body and soluble fraction, adding a further layer of complexity compared to previous studies. This work uses a simplified strategy to separate the contents of the cell, overcoming problems in the purification of membrane systems and loss of cellular components, leading to a more thorough understanding of the spatial distribution of proteins within a cyanobacterial cell.

For interactive data mining and data visualisation we have deployed a dedicated online data app for the community at https://lgatto.shinyapps.io/synechocystis/. The app contains a searchable and clickable data table, visualisation of the quantitative protein profiles across both replicates, and a fully interactive PCA plot.

Results

Fractionation of *Synechocystis* cell extracts by sucrose density ultracentrifugation

In order to fractionate cellular components, *Synechocystis* cells were cultured to late-logarithmic phase (Supplemental Fig. S1) under continuous moderate light (60 µmol photons m^{-2} s^{-1}) with air-bubbling at 30°C. Growth conditions and cell harvesting are similar as those performed in studies where membranes were isolated using two-phase aqueous polymer partitioning (e.g. (Norling et al., 1998; Pisareva et al., 2007)), allowing a comparison of protein localisation between these datasets. Cells were subsequently lysed and the extract fractionated via sucrose density centrifugation (Schottkowski et al., 2009). Separation on a step gradient resulted in cellular material accumulating in the heaviest fraction (Supplemental Fig. S2A).

Further separation of this fraction on a continuous sucrose gradient was therefore required. This resulted in 12 fractions with varying protein-pigment composition (Fig. 3A), as determined by absorption spectra measurements (Supplemental Fig. S2B), diverse protein profiles, as evaluated by SDS-PAGE (Supplemental Fig. S2C), and different distributions of TM and PM, as indicated by immunoblot analysis using antibodies against TM (photosystem II core light harvesting protein; PsbB (CP47)) and PM (Sodium-dependent bicarbonate transporter; SbtA) specific marker proteins (Fig. 3B). These results demonstrate the validity of this approach in effectively separating and enriching cellular components, a necessary prerequisite for labelling and subsequent analysis.

Extensive coverage of the *Synechocystis* proteome by mass spectrometry reveals sub-clustering of different compartments

Of the twelve fractions obtained from the continuous sucrose gradient, both the lightest two and the heaviest two were deemed to be most similar to one another compared with other fractions by SDS-PAGE and were thus combined in pairs to yield ten fractions, reflecting the number of Tandem Mass Tags (TMT) tags in a 10-plex reagent set. These ten fractions were then labelled with the TMT reagents (Fig. 3C). RP-HPLC was used to separate the proteins according to their hydrophobicity (Fig. 3D) and provide better resolution before subsequent MS/MS analysis (Fig. 3E).

In total, the MS analysis resulted in the identification of 2,445 proteins (Supplemental Table S2; Supplemental Table S3) across both biological replicates, out of a potential 3,672 listed in the CyanoBase database (http://genome.annotation.jp/cyanobase). This included 397 predicted integral membrane, 768 hypothetical and 400 unknown proteins.

Similar scale proteome coverage (2,461 proteins) was recently reported by Spat *et al* (Spat et al., 2018). In their study MS analysis was performed on cells cultured under similar environmental conditions (40 µmol photons m^{-2} s^{-1} with air-bubbling at 26°C) to those used here, but which were nitrogen deprived and then harvested 2, 8, 24 and 55 hours after resuscitation via addition of nitrate. A comparison of protein coverage between our data and Spat *et al* showed that 2,127 proteins (~58%) were detected in both studies (Supplemental Table S4), suggesting that this may be the core proteome expressed under these laboratory conditions. 318 proteins were only detected in our study (Supplemental Table S5), while 334 were unique to Spat *et al* (Supplemental Table S6). These differences are likely due to the physiological response induced during resuscitation from nitrogen deprived to replete media or variation in cell preparation and proteome detection methods. Moreover, 109 proteins were only detected in some of the five Spat *et al* samples and 82 were detected at very low quantities. 856 (~25%) were not detected in either study (Supplemental Table S7), which included 112 with transposon related functions, 290 hypothetical and 275 unknown proteins. This portion of the proteome may be dormant under these laboratory conditions.

In order to localise proteins to specific regions of the cell, the abundance profile of each protein along the sucrose gradient was first quantified using the distribution of TMT reporter ions generated by tandem MS. Assuming that proteins which reside together in the cell would co-fractionate in the sucrose gradient, we therefore used this data to interpret the distribution of proteins within the cell. Resulting abundance profiles of proteins were subjected to principal component analysis (PCA) for visualisation purposes. The PCA plot represents a map of all 2,445 proteins identified in both biological replicates, in which proteins with similar distribution profiles along the gradient are clustered together (Fig. 4A). Marker proteins for subcellular compartments, including the PM and TM, small and large ribosomal subunits, and soluble proteins (including cytosolic, thylakoid lumen and periplasmic

proteins) (Fig. 4B; Supplemental Table S8) were used to identify which clusters on the plot correspond to which subcellular regions. This resulted in identification of distinct clusters corresponding to certain subcellular regions, including the PM, TM, small ribosomal subunit, large ribosomal subunit and soluble proteins, without the need to obtain pure membrane fractions.

The localisation of previously unclassified proteins was achieved by matching their profiles along the sucrose gradient to the marker protein profiles. This was carried out using supervised classification with a support vector machine (SVM) (Gatto et al., 2014) to assign unclassified proteins, defining the boundaries of the subcellular regions (Fig. 4C), and producing an SVM score for each protein and a predicted localisation. The SVM score is a measure of the confidence with which the protein was classified. The majority of assigned proteins (1,054) were found to be soluble, followed by those that were localised to the PM (436) or TM (147), with only a small number associated with the small (29) and large (45) ribosomal subunits, including the protein markers themselves (Supplemental Table S3). No integral membrane proteins localised to the soluble fraction (Fig. 4D), although a large number of proteins lacking transmembrane helical domains (TMHs) (Supplemental Table S3) localised to the PM and TM. The remaining 734 proteins were not classified into any of these subcellular locations, and were thus given an ‘unclassified’ allocation. Of the 1,168 unknown and hypothetical proteins, 56 were TM localised, 233 PM localised and 467 were found to be soluble. Seven and five proteins were associated with the small and large ribosomal subunit fractions, respectively. Further description of the localisation of sets of proteins including those with a previously assigned function is given in detail in the supplemental information, along with comparisons with published localisation information.

Further subcellular regions and compartmentalisation within the cell were observed. For example, the PM proteome grouped into two distinct regions (Fig. 4C, 5A). A small proportion of transport and binding proteins were sub-localised within the PM cluster, in close association with the cell division protein FtsZ, which forms the septal ring, and the MinCDE proteins, which control the position and shape of the septal ring. Large ribosomal subunits also grouped into two distinct regions with five proteins (L16, L28, L27, L19 and L35) forming a distinct cluster close to the PM region (Fig. 5B). This region also contains the high molecular weight Class A

penicillin binding proteins (PBPs) PBP1-3, thought to operate in cell elongation and cytokinesis (Marbouty et al., 2009b). While little is known about the OM proteome, four proteins designated as ‘probable porin; major OM proteins’ by CyanoBase, and PilQ, the OM subunit of the pili, were grouped together in a distinct cluster between the PM and TM regions (Fig. 5C). Moreover, the subunits of certain complexes clustered together. These included RNA polymerase, RuBisCO, and hydrogenase, as well as complexes involved in chlorophyll (light-independent protochlorophyllide reductase subunits ChlN/ChlB) and tryptophan/folate biosynthesis (anthranilate synthase component I/II (TrpE/TrpG)) (Fig. 5D). This indicates that some complexes are not disassociated by cell rupture and sucrose gradient separation of cellular contents.

Comparison with previous subcellular localisation data for the *Synechocystis* proteome.

Of the previous studies on subcellular distributions of *Synechocystis* proteins, the most comprehensive list was achieved by Liberton and co-workers who used quantitative proteomics coupled with two-phase separation of cellular membranes to determine the protein content of the PM and TM (Liberton et al., 2016). Supplemental figure S3A shows the comparison of the Liberton data with those presented here. Of note, where both studies assign a protein to either the PM or TM, there is a high degree of overlap between the assignment and very few proteins assigned to the PM by Liberton *et al* are assigned to the TM in this study and *vice versa*. There is only limited overall overlap between TM assignments and PM assignments, however, between the two studies (Supplemental Fig. S3B). This is in part due to the facts that different proteins were identified in both studies and that the study presented here represents the whole cell, whereas the Liberton study analysed only a subset of proteins. Many proteins thought to be TM or PM localised by the Liberton study are not assigned to either membrane here. It is not clear whether the additional PM and TM proteins presented in the Liberton study represent contamination of their TM and PM enriched fractions with proteins from other parts of the cell, or that the lack of overlap is a result of the fact that the study presented here returns the steady state location of proteins. Hence, if a TM and PM protein were also elsewhere in the cell, our study would flag it up as ‘mixed location’. It is interesting to note that many of the results for the TM and particularly the PM in

Liberton's study are assigned to the soluble protein set in the data presented here, demonstrating the importance of mapping the whole cell and not just isolated fractions. Analysis of these proteins shows that only 7% have a predicted single transmembrane domain and the remainder have no predicted membrane spanning regions, so a location in the TM or PM seems less likely.

Metabolic pathways are distributed throughout the cell

Enzymes involved in metabolism predominantly localised to the soluble region, including those synthesising amino acids, cofactors, prosthetic groups and carriers, glycolysis, tricarboxylic acid cycle and pentose phosphate pathway intermediates, cell wall components, purines and pyrimidines, fatty acids, phospholipids, sterols and hydrocarbons (Fig. 6; Supplemental Table S3). However, some enzymes, predominantly those involved in the final catalytic steps of certain metabolites, localised to membranes. These included enzymes synthesising membrane lipids (acyltransferase PlsC, fatty acid/phospholipid synthesis protein PlsX, monogalactosyl diacylglycerol synthase MgdA and phosphatidate cytidylyltransferase CdsA), all of which localised to the TM. This is likely due to the thylakoids constituting the bulk of the membranes in the cell and it is possible that a minor percentage of these proteins are PM localised.

Other TM localised enzymes include those synthesising heme (ferrochelatase HemH, protoporphyrinogen IX oxidase HemJ) and transhydrogenation of $NADP^+$ (PntA, PntB). HemJ converts protoporphyrinogen IX to protoporphyrin IX, the precursor of heme and chlorophyll (Skotnicova et al., 2018). A recent study in *Chlamydomonas reinhardtii* indicates that HemJ likely requires plastoquinone as an electron acceptor (Brzezowski et al., 2019). Localisation of HemJ to the TM in *Synechocystis* suggests a similar enzymatic reaction is possible. TM localisation of PntA/B is consistent with the majority of $NADP^+$ undergoing reduction to NADPH via ferredoxin-NADP reductase in the TM photosynthetic electron transport chain, and heme acting as a precursor for phycobilins, subsequently incorporated into phycobilisomes.

Enzymes synthesising phylloquinol (2-phytyl-1,4-benzoquinone methyltransferase MenG, MenH), flavin adenine dinucleotide (RibF) and molybdopterin cofactors (MoeA), were associated with the PM. It is unclear why RibF is PM localised. MenG

is closely associated with the type two NAD(P)H dehydrogenase, NdbB, on the PCA plot. Both proteins are required for the final biosynthetic step of phylloquinol biosynthesis and their close association suggests they may form a complex (Fatihi et al., 2015). PM localisation of MoeA may aid incorporation of imported molybdate into the molybdopterin cofactor.

In addition, several enzymes catalysing carotenoid biosynthesis localised to the membranes. Carotenoids play a key role in assembly of photosynthetic complexes (Toth et al., 2015), membrane integrity and thylakoid organisation (Mohamed et al., 2004), and as light harvesting and photoprotective pigments. Seven carotenoids have been detected in *Synechocystis*: synechoxanthin, myxol-2'-dimethylfucoside (myxoxanthophyll), zeaxanthin, 3'-hydroxy-echinenone, *cis*-zeaxanthin, echinenone and β-carotene (Graham and Bryant, 2008). Carotenoids have been localised to both membrane fractions (Zhang et al., 2015) but the enzymes involved in biosynthesis of these compounds have not been completely elucidated or their intracellular location determined (the pathway is detailed in supplemental figure S4). Enzymes involved in γ-carotene (CruF) and β-carotene (CrtL and CruA) biosynthesis (Maresca et al., 2007) were TM localised, as were the only enzymes identified in synechoxanthin (CruE, CruH) and myxoxanthophyll (CruG) biosynthesis (Graham and Bryant, 2009). The only carotenoid biosynthetic enzyme localised to the PM was the carotene isomerase CrtH, involved in cis-to-trans conversion of carotenes (Masamoto et al., 2001). However, carotenoid biosynthesis in a ΔCrtH mutant is only affected under dark conditions, not light, and its exact role in the cell has not been determined (Masamoto et al., 2001).

A few proteins involved in intermediate enzymatic steps localised to membranes. For example, the long-chain-fatty-acid CoA ligase Aas, involved in the cycling of free fatty acids via activation by acyl carrier protein (ACP), localised to the PM, which is in agreement with Liberton *et al* (Liberton et al., 2016). This supports the proposed role of Aas in mediating fatty acid import (von Berlepsch et al., 2012). Dihydroorotate dehydrogenase (PyrD), the only membrane associated enzyme involved in nucleotide metabolism, also localised to the PM. In *E. coli*, PyrD requires a respiratory quinone as an electron acceptor (Nørager et al., 2002). Our data suggest that *Synechocystis* PyrD may utilise plastoquinone (PQ) as an electron acceptor, which could be one of the roles of the PM electron transport chain.

Assembly of cell wall components occurs in both membranes

A similar pattern was observed with enzymes involved in biosynthesis of cell wall components (Fig. 7). The enzymes catalysing the initial steps of the core region of lipopolysaccharides (LpxACD) were soluble, while the one catalysing the final step of lipid A disaccharide bisynthesis (LpxB), localised to the PM. MsbA, the flippase that translocates lipid A disaccharide across the PM (Ruiz et al., 2009), has not been identified in cyanobacteria. However, four genes with high sequence similarity to *E. coli msbA* (*slr2019, sll1276, sll1725, slr1149*; 70.5, 69.6, 64.9, 66.3% similarity, respectively) were identified in our study. All localised to the PM, so further genetic and biochemical studies will be required to identify cyanobacterial MsbA. Several putative glycosyltransferases (RfbU, 2 x RfbW, RfbJ, RffM), postulated to add sugar groups to the outer core of this molecule (Fisher et al., 2013), also localised to the PM. Homologs of the proteins in the Lpt transport complex, responsible for transporting lipopolysaccharides from the PM to the outer leaflet of the OM in *E. coli* (Ruiz et al., 2009), are not present in *Synechocystis*, suggesting an alternate system must perform this role.

The enzymes catalysing the initial steps of peptidoglycan monomer biosynthesis (MurABCDEF) were soluble. Somewhat suprisingly, the final two steps of peptidoglycan monomer biosynthesis (MraY, MurG) localised to the TM, not the PM as would be expected. MurG has been identified as TM specific in a previous study (Pisareva et al., 2011). This would suggest that monomers are assembled at the TM, and subsequently transported to the PM. A single homolog of MurJ (slr0488), the flippase which translocates peptidoglycan monomers across the PM (Sham et al., 2014), is present in *Synechocystis* but was not detected in our study or in Spat *et al* or Liberton *et al* (Liberton et al., 2016; Spat et al., 2018). Neither was FtsW, responsible for peptidoglycan polymerisation in association with PBPs (Taguchi et al., 2019). Our knowledge of the role of cyanobacterial PBPs is limited, although all eight putative PBPs, separated into class A (PBP 1-3), B (PBP4/FtsI) and C (PBP 5-8), were detected. While PBP4 is essential in *Synechocystis*, single mutants deficient in one class A or C PBP have been generated, although not mutants lacking two of each class (Marbouty et al., 2009b). PBP1-3 co-localised in a unique cluster on the PCA plot, PBP4 and PBP6/8 localised to different PM regions, while PBP5/7 was

soluble (Fig. 5B). Both class A and B PBPs are believed to be involved in peptidoglycan polymerization, with class A primarily involved in synthesis of the cell wall linked to cell elongation, while class B interacts with other proteins of the divisome, with a primary role in cell division (Sauvage et al., 2008). Other components of the divisome including Cdv3, ZipN and ZipS (Marbouty et al., 2009a), also localised to the PM in our study. In *Synechocystis*, the Type C PBPs are divided into two classes, type 4 (PBP 5/8) and AmpH (PBP 6/7) (Marbouty et al., 2009b). PBP5/7 are soluble, presumably in the periplasm, while PBP6/8 are PM associated. Their primary role is likely in disassembling the peptidoglycan heteropolymer with other proteins such as the N-acetylmuramoyl-L-alanine amidases, which were also PM localised (Slr1744) or soluble (Slr0891) (van Heijenoort, 2011).

The thylakoid membrane proteome is predominantly involved in energy conversion

As expected, the majority of subunits in photosynthetic complexes, including Photosystem I and II (PSI and PSII), and cytochrome b_6f (cyt b_6f), were TM localised (Fig. 8A; Supplemental Table S3). Other proteins associated with photosystems including the PSII assembly protein RubA, Ycf48 and Ycf39 (Garcia-Cerdan et al., 2019; Kiss et al., 2019), the putative PSI assembly proteins Ycf4 and Ycf37, and IsiA, which is required for PSI formation and state transitions under iron starvation, were also TM localised. In addition, CpcG2, an integral protein of the phycobilisome, the light harvesting complex of cyanobacteria, localised to this compartment although other phycobilisome subunits were predominantly soluble. Respiration has previously been established to occur in the TM (Lea-Smith et al., 2016a), although the location of electron transport complexes has not been fully established. Of the respiratory electron donors, only NADH dehydrogenase type 1 subunits were TM localised (Fig. 8B). The membrane subunits of succinate dehydrogenase have not been identified (Lea-Smith et al., 2016a), although it has been suggested as the main TM localised respiratory donor (Cooley and Vermaas, 2001). Subunits of two terminal oxidases, cytochrome-*c* oxidase and cytochrome *bd*-quinol oxidase, localised to the TM. Interestingly, ATP synthase subunits localised to the TM, in agreement with Liberton *et al* (Liberton et al., 2016). Overall, this suggests that energy conversion is predominantly localised to the TM. Other proteins of note that

localised to the TM include three FtsH proteins involved in PSII repair (FtsH2, FtsH3, FtsH4), the thiol:disulphide interchange protein TrxA and the detoxification protein Slr0236. Only six proteins involved in transport localised to the TM, including three Na^+/H^+ antiporters (NhaS1, NhaS3, NhaS6), the copper importer CtaA, the H^+/Ca^{2+} exchanger SynCAX and an ABC transporter (Sll0759). Of the 83 characterised proteins localised to the TM, 63 are involved in energy conversion, photosystem repair/assembly or synthesis of lipids required for membrane assembly or photosystem function.

The plasma membrane proteome is predominantly involved in transport and regulatory functions

The majority of proteins involved in transport localised to the PM (Fig. 6; Supplemental Table S3). These included the transporters of ammonium, basic and neutral amino acids, glutamate, bicarbonate, inorganic iron and iron dicitrate, glucosylglycerol, manganese, molybdate, nitrate/nitrite, phosphate, potassium, sulfate, urea and zinc. Copper is required in both the cytoplasm and thylakoid lumen. Previously it has been thought that copper is transported into the cytosol and thylakoid lumen via PM localised CtaA and TM localised PacS, respectively, based on studies performed in *Synechococcus elongatus* (Kanamaru et al., 1994; Tottey et al., 2012). In contrast, our results placed CtaA in the TM and PacS in the PM.

A second, poorly characterised, electron transport chain localises to the PM (Lea-Smith et al., 2016a). Two NAD(P)H dehydrogenase type 2 electron donor proteins (NdbB, NdbC) and subunits of the alternative respiratory terminal oxidase localised to the PM, suggesting the presence of a simpler electron transport chain in this compartment (Fig. 8C). NdbB is required for phylloquinol biosynthesis (Fatihi et al., 2015). Deletion of NdbB resulted in almost a complete loss of phylloquinol and accumulation of the precursor molecule, 2-phytyl-1,4-naphthoquinone. NdbB was shown to reduce 2-phytyl-1,4-naphthoquinone to 2-phytyl-1,4-naphthoquinol using electrons derived from NADPH (Fatihi et al., 2015), which is subsequently methylated to phylloquinol by MenG (Sakuragi et al., 2002). Other proteins of note that localised to the PM included the cell division proteins MinD and FtsH1, the chaperone DnaK3, chemotaxis proteins PixJ1 and TaxD2, the competence protein ComE involved in DNA uptake, the detoxification protein Gst1 and the sigma factor

SigF. Pili proteins localised to the PM, including 8/11 PilA designated subunits (another, PilA6, is unclassified but is in the PM region of the PCA plot), with the exception of PilQ, the OM subunit, and PilH, which was soluble. PilA1 is required for formation of thick pili (Yoshihara et al., 2001), but expression of the other 8 PilA proteins suggests they have a functional role in the cell under these growth conditions. Two proteins involved in DNA replication, DnaG, the DNA primase, which synthesises oligonucleotides, and DnaX, a DNA polymerase II subunit, were both PM localised. The PM may therefore play an active role in DNA replication or regulation, which has been suggested to occur in *E. coli* (Saxena et al., 2013; Magnan et al., 2015).

Protein translocation pathways localise to the thylakoid membrane

The mechanism by which cyanobacteria target proteins to different membranes is poorly characterized. Single copy homologues encoding proteins involved in the Secretory (Sec), Twin-Arginine Translocation (Tat) and Signal Recognition Particle (SRP) protein translocation pathways are present in the *Synechocystis* genome (Kaneko et al., 1996). Components of each pathway were either soluble or TM localised.

Two leader peptidases (LepB1, LepB2), which are involved in generation of mature proteins and may also have a role in releasing proteins into the correct compartment, have been identified in *Synechocystis*. Only LepB2 is essential for cell viability, and the two are not functionally redundant (Zhbanko et al., 2005). Both leader peptidases were identified in the study; LepB1 localised to the PM, whilst LepB2 was unclassified. In contrast to this work, previous proteomic studies and investigations into the leader peptidases have identified LepB1 as a TM specific protein, with a suggested function in maturation of the photosynthetic machinery (Srivastava et al., 2005; Zhbanko et al., 2005; Pisareva et al., 2011; Liberton et al., 2016).

Various intracellular organelles localise to distinct regions of the cytosol

Transmission electron microscopy indicates that carboxysomes in *Synechocystis* are located in the central cytoplasm (van de Meene et al., 2006). Most carboxysome subunits were found to be soluble, with the exception of CcmM, which was PM localised, and CcmN and CcaA, which were localised to an unclassified fraction.

CcmM and CcmN are core shell proteins and CcaA is the carbonic anhydrase, converting HCO_3^- to CO_2 (Gonzalez-Esquer et al., 2015). This suggests that certain subunits may interact with the PM or that cell disruption and subsequent separation caused the carboxysome to break apart due to its large size (between 80 and 150 nm in diameter), resulting in distribution of various subunits across the sucrose gradient and in the PCA plot (Supplemental Fig. S5). Interestingly, the enzyme catalysing the initial step of photorespiration (Pgp), the conversion of phosphoglycolate to glycolate, was also PM localised. The two subunits of RuBisCO, RbcS and RbcL, which are assembled into the carboxysome (Wang et al., 2019), were found in a different area and grouped in a distinct unclassified fraction.

Of the enzymes involved in forming compounds which aggregate into storage bodies, only heterodimeric PHB synthase (PhaE/PhaC), catalysing the final step of PHB biosynthesis, was found. PhaE/PhaC, along with PhaP (ssl2501) which is the surface coding protein of PHB granules, mapped to a unique unclassified region separate from any other proteins on the PCA plot (Fig. 5C). This suggests PHB synthesis may occur in a specific, distinct part of the cytosol (Hauf et al., 2015). GFP labelling of PhaC, PhaE, and PHB granules indicate that these biosynthetic steps are localised to the cell periphery (Hauf et al., 2013).

Profiles of ribosomal subunits show clustering in a specific region of the PCA plot

The majority of the large ribosomal subunit proteins localised to a specific fraction separate from the TM, PM and soluble regions (Fig. 4C). Likewise, the majority of the small ribosomal proteins clustered in a specific region of the plot, distinct from the large ribosomal subunit protein area (Fig. 4C). However, three small ribosomal proteins were found in other locations on the plot. Two poorly characterised Rps1 homologues (Rps1A, Rps1B) localised to the soluble fraction, whilst Rps3 localised to the TM. Rps1 subunits are not present in all bacteria, and participate in recruiting mRNA to the 30S subunit where it is localised on the solvent side (Yusupova and Yusupov, 2014). All sequenced cyanobacteria with the exception of *Gloeobacter kilaueensis* JS1 and *Gloeobacter violaceus* PCC 7421, which lack TMs, encode two Rps1 subunits (Supplemental Fig. S6 and S7). Therefore, it is possible Rps1 subunits may play a role in determining protein localisation to different subcellular

505 locations. Rps3 is thought to form the mRNA entry tunnel along with Rps4 and Rps5
506 in bacteria (Ito and Chiba, 2014) and it is possible that it may play an ancillary role in
507 anchoring a particular fraction of ribosomes to the TM. A few other proteins localised
508 to this fraction. For example, HemA, a transfer RNA-Glutamyl reductase which
509 catalyses the first step in the heme biosynthesis pathway and uses charged tRNA-
510 Glutamyl as a substrate, localised to the large ribosomal subunit protein fraction. In
511 addition, Vipp1, a protein implicated in thylakoid membrane biogenesis, localised to
512 the small ribosomal subunit protein fraction. The subcellular location and exact
513 function of this protein in *Synechocystis* has been a matter of some controversy
514 (Westphal et al., 2001; Hennig et al., 2015). However, localisation to the ribosomal
515 fractions is consistent with a proposed role in organising localised protein assembly
516 centres, as suggested by Bryan *et al* (Bryan et al., 2014).

517 **Homologues of *Synechocystis* thylakoid membrane proteins localise to the**
518 **same compartment in *Arabidopsis***

519 In order to determine whether localisation of *Arabidopsis* homologues of
520 *Synechocystis* proteins have been conserved in the corresponding region of the
521 chloroplast, proteins that have been assigned to either the TM or envelope from
522 *Arabidopsis* (Ferro et al., 2010) were compared with the results obtained in this study
523 (Supplemental Table S9). Of the TM-specific *Arabidopsis* homologues, six PSI, eight
524 PSII, four cyt b_6f and four ATP synthase membrane bound components were
525 identified here, in addition to nine homologues of the chloroplast NADH
526 dehydrogenase like complex (NDH), which is known to localise to the chloroplast
527 thylakoid membrane (Shikanai, 2016). Out of three TM-specific *Arabidopsis*
528 homologues not found in these complexes, all localised to the TM in *Synechocystis*,
529 including two hypothetical proteins (sll1390, slr1470). Therefore, 34 out of 34 TM-
530 specific *Arabidopsis* homologues localised to the same membrane in *Synechocystis*.
531 Of the 31 homologous *Arabidopsis* chloroplast envelope proteins, 22 were identified
532 in *Synechocystis*, with ten in the PM and seven in the TM, while the remainder were
533 unclassified. Of these seven, two are involved in lipid biosynthesis. In *Arabidopsis*,
534 the essential pathway for thylakoid lipid biosynthesis requires export of fatty acids
535 from the chloroplast to the endoplasmic reticulum (Xu et al., 2005). This suggests
536 that a number of TM localised processes have been transferred to the envelope in

chloroplasts during evolutionary remodelling, presumably to accommodate the requirements of organelle function in a eukaryotic cell. One protein, Sll0269, associated with the small ribosomal subunit region. Proteins homologous to TM specific proteins in *Arabidopsis* are nearly all exclusive to the TM in *Synechocystis*. Of the remaining 62 uncharacterised TM localised proteins in *Synechocystis*, 10 (slr1747, sll0862, sll0875, sll1071, sll1399, sll1925, slr0575, slr1591, slr1821, slr1919) have homologues in *Chlamydomonas reinhardtii* and *Arabidopsis*, suggesting a conserved role throughout the photosynthetic lineage (Highlighted in red in Supplemental Table S10). In contrast, the *Arabidopsis* envelope proteins are distributed in both the PM and TM of *Synechocystis*.

Discussion

Here we detail a method for separating and analysing the cellular components of *Synechocystis*, resulting in the most extensive proteome mapping of a cyanobacterium to date. The importance of examining the whole cell compared to fractions enriched in individual compartments is highlighted by the assignment of a large number of proteins, most lacking membrane spanning domains, to the soluble fraction in our study which had previously been assigned to membranes in the Liberton study or earlier reports using ‘purified’ fractions e.g. (Pisareva et al., 2007). In the cells examined in this study, which were cultured under continuous moderate light and carbon replete conditions, approximately two-thirds of the proteome was detected, demonstrating the advantages of this proteomics technique compared to those previously applied to map proteins in cyanobacteria. In certain cases the technique described here allowed identification of the isoenzyme catalysing specific biosynthetic steps under these conditions. For example only one of the two possible aspartate aminotransferases (Sll0402) was detected. The remaining proteome may not have been detected for a variety of reasons. Only proteins which were identified in both replicates were included, and, whilst MS is a sensitive method, some proteins may be expressed at levels too low to be detected via this approach. Other proteins may simply not be expressed under these conditions. Examples of this include proteins expressed only under microoxic conditions such as Ho2, involved in phycobiliprotein biosynthesis, and PsbA1, a subunit of PSII (Summerfield et al., 2008), and conditions of low carbon dioxide availability, such as the flavodiiron

proteins Flv2 and Flv4 (Zhang et al., 2012). Of the 1227 potential proteins not detected, 444 were hypothetical proteins and 360 were unknown. It is possible that the genes encoding these proteins may not produce functional products or be transcriptionally inactive. Regardless, the development of a robust technique for separating cellular components will facilitate proteomics of *Synechocystis* cultured under a range of environmental conditions. This technique may also be useful for analysing the proteome of other cyanobacteria and possibly microalgae, especially since membrane separation techniques are poorly developed in unicellular photosynthetic species apart from *Synechocystis* and are not ideal due to large amounts of cellular material being lost. Other prokaryotes which have complicated internal structures, such as purple photosynthetic bacteria, or complex multi-layered cell walls, for example Corynebacterineae, may also benefit from analysis via these methods.

The higher proportion of proteins detected and localised to specific regions of the cell in this study compared to published data using purified membranes further emphasises the advantages of this method. Purification of only a sub-fraction of cellular components in past studies may explain this difference. The heterogeneous nature of the membranes and cytoplasm of *Synechocystis* is illustrated by the existence of sub-regions within the PCA plot (Fig. 5A). Particularly intriguing was the presence of possible sub fractions in the PM and a region that may correspond to the OM. While it is not possible for us to define these regions currently, due to our lack of knowledge of their composition, previous studies have suggested a heterogeneous distribution of proteins within the PM and TM (Srivastava et al., 2006; Agarwal et al., 2010; Straskova et al., 2019). As our understanding of processes within the cells increases, other regions, or sub-regions, may be identified. For example, as the proteins embedded within the OM become better identified and characterised we can integrate this into our model to carry out further predictions of the proteome of this region.

The complexity of cyanobacteria compared to other prokaryotes is likely to be due to the requirement to separate photosynthesis into a separate compartment, which is supported by our results. The majority of metabolic enzymes are soluble, whereas the TM and PM have specialised roles focusing primarily on energy conversion and

transport, respectively (Fig. 6). While this is obviously a successful evolutionary strategy, the presence of multiple compartments, further complicated by the presence of sub-regions within the membranes and possibly the cytosol, means that these organisms require a complex targeting system capable of directing proteins to the correct location. How this occurs is still poorly understood (Frain et al., 2016). Subunits of the protein translocation systems localised only to the TM, although it is possible that a small proportion are present in the PM. Intriguingly, the leader peptidase LepB1, localised specifically to the PM. Therefore, it is possible that this protein has a role in targeting proteins specifically to this membrane. Another possibility is that mRNAs migrate to specific subcellular locations (Nevo-Dinur et al., 2011; Moffitt et al., 2016) and that following translation proteins are inserted into the membrane or region in closest proximity. This is a distinct possibility given the spatial distribution of ribosomes throughout the cell. Furthermore, ribosomes on membrane-like structures connected to the TM have been observed in *Synechocystis* (van de Meene et al., 2006). Certain ribosomal subunits, such as TM localised Rps3 and cytosolic Rps1A and Rps1B, may have a role in anchoring ribosomes to different cellular regions. Our study has also provided insights into the proteomic remodelling associated with the evolution of a chloroplast from a cyanobacterium.

Although the method developed as part of this study has achieved the most extensive subcellular map of *Synechocystis* to date, the approach is not without some limitations. While subunits of some protein complexes co-localised on the PCA plot, others may have dissociated from one another during sample preparation, and in future it would be interesting to compare these data with those obtained using a workflow that employs protein crosslinking reagents (Liu et al., 2015; Leitner et al., 2016). Furthermore, the data visualisation methods employed use a dimension reduction approach and it cannot be ruled out that the apparent resolution of some un-related cellular substructures is lost as a result of this or by the physical subcellular separation methods employed. In future it would be interesting to see how the map presented here compares with similar data achieved using different cell fractionation methods such as differential centrifugation and free flow electrophoresis, or other spatial approaches involving proximity tagging (Lam et al., 2015; Kim et al., 2016; Loh et al., 2016). Ultimately, our knowledge of many aspects of cyanobacterial biology is poor, with function assigned to only about 50% of genes

in *Synechocystis* (http://genome.annotation.jp/cyanobase), the most highly characterised species within the phylum. Since the majority of the proteins identified in this study have no assigned function, understanding their location in the cell will aid future studies characterising their exact role. For example, Slr0060, currently classified as an unknown protein, may be associated with PHB granules due to its proximity to PhaE, PhaC and PhaP in our data. Of particular interest are the 10 TM localised, uncharacterised proteins that have homologues in *C. reinhardtii* and *Arabidopsis*, which are likely to have a conserved role in photosynthesis.

This database is the largest and most extensive list of the *Synechocystis* TM and PM proteome and is an invaluable tool to identify how proteins are targeted to each compartment and how these mechanisms could be utilised to insert recombinant proteins into different membrane compartments for biotechnology applications, i.e. insertion of transporters into the PM for export of biofuels and industrial compounds.

Methods

Bacterial strains, media, and growth conditions

Synechocystis sp. PCC 6803 was routinely cultured in liquid BG11 medium with 10 mM sodium bicarbonate (Castenholz, 1988) at 30°C and grown under continuous moderate white light (50 µmol photons m^{-2} s^{-1}) with vigorous air bubbling and shaking at 160 rpm. For growth of larger cultures, two 50 ml starter cultures were grown for 3-4 days in BG11 medium with 10 mM sodium bicarbonate to OD_{750nm} = ~1 and used to inoculate 2 x 2 L flasks containing 1 L of BG11 medium with 10 mM sodium bicarbonate. Cultures were air bubbled and harvested at OD_{750nm} = ~2.

Cell lysis and subcellular fractionation

Whole-cell lysate was fractionated by sucrose density ultracentrifugation, as previously described (Schottkowski et al., 2009), with modifications. All steps were carried out at 4°C. Cells were harvested from 2 l cultures, by centrifugation at 5,000*g* for 10 min. The cell pellet was washed in 50 ml Buffer I (5 mM Tris-HCl, pH 6.8) and centrifuged at 5,000*g* for 10 min. The resulting cell pellet was re-suspended in 75 ml Buffer II (10 mM Tris-HCl, 1 mM PMSF, 600 mM sucrose, 5 mM EDTA, 0.2% (w/v) lysozyme, pH 6.8), and shaken at 160 rpm for 2 h at 30°C before centrifugation at 5,000*g* for 10 min. The cell pellet was washed twice with Buffer III (20 mM Tris-HCl, 1 mM PMSF, 600 mM sucrose, pH 6.8) and re-suspended in 17.5 ml of the same buffer, to which half the volume of 425-600 µm acid-washed glass beads was added. Cells were disrupted in a Mini Bead Beater-16 (BioSpec Products) for 10 min at 3,450 oscillations/min, with 1 min intervals on ice. The cell suspension was centrifuged at 3,000*g* for 10 min to pellet unbroken cells. The supernatant was concentrated to 50% sucrose by the addition of 80% sucrose (w/w) in Buffer II to a final volume of 10 ml. The refractive index of sucrose solutions was measured to ensure correct concentrations by using a hand-held refractometer (Reichert). A discontinuous sucrose gradient containing Buffer II was made, consisting of 10 ml 50% (w/w) including cell lysate, 8 ml 39% (w/w), 6 ml 30% (w/w), and 6 ml 10% (w/w), and centrifuged at 125,000*g* for 17 h (SW 32 Ti Swinging Bucket Rotor, Beckman Coulter Optima L-100 XP Ultracentrifuge). Fractions 10% (I), 30% (II), 39% (III), and 50% (IV) were collected using a fraction collector (LabConco). Fraction V was diluted with 5 mM Tris-HCl buffer (pH 6.8) to a concentration of 20% (w/w) and

added onto a continuous sucrose gradient from 30% (w/w) to 60% (w/w) and centrifuged at 125,000*g* for 17 h. 2.5 ml fractions were collected (1-12) using a fraction collector.

Protein precipitation was performed using a methanol-chloroform system (chilled methanol/chloroform/water, 4:1:3 (v/v/v)) (Wessel and Flügge, 1984). Protein was recovered at the interphase after vigorous vortexing for 30 s and centrifugation at 13,000*g* for 90 s at 4°C. The upper phase was discarded and the protein disc washed in 3 volumes of methanol before further centrifugation (13,000*g*, 90 s, 4°C) to pellet the protein, which was air-dried after removal of the supernatant. Protein pellets were solubilised by re-suspension in 150 µl 50 mM HEPES-NaOH, 0.2% SDS (w/v) (pH 7.4), and incubated at 42°C for 15 min. Protein concentration was determined using the DC Protein Assay kit (Bio-Rad).

SDS-PAGE and immunoblotting

Samples from each of the fractions collected were boiled in 4 x Laemmli sample buffer for 10 min. Proteins were resolved on a 4-20% SDS-PAGE gel (Bio-Rad), transferred to PVDF membrane (Amersham Hybond-P, 0.45 µm; GE Healthcare), and detected with antibodies against PM (SbtA, 1/2,000; Agrisera) and TM (CP47, 1/2,000; Agrisera) specific proteins (Norling et al., 1998; Zhang et al., 2004) by chemiluminescence using WesternBright Quantum Blotting Detection Reagent (Advansta). Visualisation was carried out using a G:Box imaging system (Syngene).

Protein digestion and TMT 10-plex labelling

Sucrose gradient fractions 1 and 2, as well as 11 and 12, were combined, leaving 10 samples for TMT 10-plex labelling. Each sample was normalised to 100 µg of protein in 25 mM TEAB, before being reduced, alkylated and digested with trypsin. Each sample was made up to a total volume of 50 µl with 25 mM TEAB. Disulphide bonds were reduced with 5 µl 200 mM tris(2-carboxyethyl)phosphine for 1h at 55°C, followed by alkylation of cysteine residues with 5 µl of 375 mM iodoacetamide for 20 min at room temperature in the dark. Protein was precipitated from the samples by addition of 6 volumes of ice-cold acetone, vortexing and incubation at -20°C overnight. The protein pellet was recovered by centrifugation at 16,000*g* for 10 min, air-dried, and solubilised in 100 µl 100 mM HEPES (pH 8.5). Samples were digested with 2.5 µg sequencing grade trypsin (Promega) for 1h at 37°C. A second aliquot of

2.5 µg trypsin was added to the samples, and incubated at 37°C overnight. Trypsin digests were centrifuged for 10 min at 13,000*g* to remove any insoluble material.

The 10 TMT tags were equilibrated to room temperature and re-suspended in 41 µl acetonitrile before being added to each of the 10 peptide samples. Samples were placed onto a shaker for 2 h at room temperature. TMT labelling efficiency was between 93-95%. Un-reacted TMT tags were quenched with 8 µl 5% (w/v) hydroxylamine in 100 mM HEPES (pH 8.5) for 1 h at room temperature. 100 µl of ultrapure water was added and the samples incubated at 4°C overnight. The samples were then combined and reduced to dryness by vacuum centrifugation.

The solid-phase extraction of TMT-labelled peptides was performed according to the method previously described (Villén and Gygi, 2008), with modifications. The samples were re-suspended in 1 ml of 0.4% (v/v) formic acid, and placed onto 100 mg Sep Pak tC28 solid phase extraction cartridges (Waters Corporation). Cartridges were conditioned using 1.8 ml 100% (v/v) acetonitrile, followed by 50% (v/v) acetonitrile and 0.5% (v/v) acetic acid, and equilibrated with 1.8 ml 0.1% (v/v) formic acid. The peptides were de-salted after loading in 1.8 ml 0.1% (v/v) formic acid, re-equilibrated with 500 µl 0.5% (v/v) acetic acid. Samples were eluted with 0.5 ml 75% (v/v) methanol with 0.5% (v/v) acetic acid, followed by 75% (v/v) acetonitrile with 0.5% (v/v) acetic acid, and reduced to dryness by vacuum centrifugation before re-suspension in 0.1 ml 20 mM ammonium formate (pH 10), 4% (v/v) acetonitrile, for high pH reversed-phase liquid chromatography.

Sample fractionation

Peptides were loaded onto an Acquity bridged ethyl hybrid C18 UPLC column (Waters; 2.1 mm inner diameter x 150 mm, 1.7 µm particle size), and profiled with a linear gradient of 5-75% acetonitrile + 20 mM ammonium formate (pH 10) over 50 min, at a flow rate of 50 µl/min. Chromatographic performance was monitored by sampling eluate with a diode array detector (Acquity UPLC, Waters) scanning between wavelengths of 200 and 400 nm. 44 fractions were collected from 11 min onwards in 1 min intervals. Fractions 1-8 were pooled together, and the rest were pooled pair-wise, with fraction 9 pooled with fraction 26, 10 with 27 and so on to yield 19 samples for mass spectrometry analysis.

Mass spectrometry

All LC-MS/MS experiments were performed using a Dionex Ultimate 3000 RSLC nanoUPLC (Thermo Fisher Scientific) system and a Lumos Fusion Orbitrap mass spectrometer (Thermo Fisher Scientific) using synchronous precursor selection (SPS)-MS. Each of the fractionated samples was resuspended in 35 µL 0.1% (v/v) formic acid and between 1-5 µL of these was applied to LC-MS/MS analysis using an Orbitrap Fusion Lumos coupled with a Proxeon EASY-nLC 1000 (Thermo Fisher Scientific). Separation of peptides was performed by reverse-phase chromatography at a flow rate of 300 nl/minute and a Thermo Scientific reverse-phase nano Easy-spray column (Thermo Scientific PepMap C18, 2 µm particle size, 100A pore size, 75 µm i.d. x 50 cm length). Peptides were loaded onto a pre-column (Thermo Scientific PepMap 100 C18, 5 µm particle size, 100A pore size, 300 µm i.d. x 5 mm length) from the Ultimate 3000 autosampler with 0.1% formic acid for 3 minutes at a flow rate of 10 µl/minute. After this period, the column valve was switched to allow elution of peptides from the pre-column onto the analytical column. Solvent A was water + 0.1% formic acid and solvent B was 80% acetonitrile, 20% water + 0.1% formic acid. The linear gradient employed was 4-140 B in 100 minutes (the total run time including column washing and re-equilibration was 120 minutes).

An electrospray voltage of 2.1 kV was applied to the eluent via the EASY-Spray column electrode. The following workflow in the Method Editor was used: MS OT (Detector type: Orbitrap, Resolution: 120000, Mass range: Normal, Use Quadrupole Isolation (Yes), Scan Range: 380-1500, RF Lens (%): 30, AGC Target: 4e5, Max Inject Time: 50 ms, Microscans: 1, Data Type: Profile, Polarity: Positive) - Monoisotopic Precursor Selection (MIPS) (Monoisotopic Peak Determination: Peptide, Relax restrictions when too few precursors are found: Yes) - Charge State (Include charge state(s): 2-7) - Dynamic Exclusion (Exclude after n times: 1, Exclusion duration (s): 70, Mass Tolerance; ppm, Low: 10, High: 10, Exclude Isotopes: Yes, Perform dependent scan on single charge state per precursor only: Yes) - Intensity Threshold (5.0e3) - Decisions (Data dependent mode: Top Speed, Number of Scan Event Types: 1, Scan Event Type 1: No Condition) - ddMS2 IT CID (MSn Level: 2, Isolation Mode: Quadrupole, Isolation Window (m/z): 0.7, Activation Type: CID), CID Collision Energy (%): 35, Activation Q: 0.25, Detector Type: Ion Trap, Scan Range Mode: Auto, m/z: Normal, Ion Trap Scan Rate: Turbo, AGC Target; 1.0e4, Max Inject Time (ms): 50, Microscans: 1, Data Type: Centroid) -

Precursor Selection Range (Mass Range: 400-1200) - Precursor Ion Exclusion (Exclusion mass width: m/z, Low: 18, High: 5) - Isobaric Tag Loss Exclusion (Reagent: TMT) - Decisions (Precursor Priority: Most Intense, Scan event type 1: No Condition) - ddMS3 OT HCD (Synchronous Precursor Selection: Yes, Number of Precursors: 10, MS Isolation Window: 0.7, Activation Type: HCD, HCD Collision Energy (%): 65, Detector Type: Orbitrap, Scan Range Mode: Define m/z range, Orbitrap Resolution: 60000, Scan Range (m/z): 100-500, AGC Target: 1.0e5, Max Inject Time (ms): 120, Microscans: 1, Data Type: Profile). Total run time was 120 minutes.

Data processing

Raw data files were processed using Proteome Discoverer (v1.4.1.14, Thermo Fisher Scientific), interfaced with Mascot server (v.2.3.02, Matrix Science). Mascot searches were performed against the CyanoBase database, with carbamidomethylation of cysteine, and TMT 10-plex modification of lysine and peptide N termini set as modifications. Precursor and fragment ion tolerances of ±20 p.p.m and ±0.1 Da were applied. Up to 2 missed tryptic cleavages were permitted. Proteins were reported with a FDR of 0.5%.

TMT 10-plex quantification was also performed via Proteome Discoverer by calculating the sum of centroided ions within a ±2 mmu window around the expected m/z for each of the 10 TMT reporter ions. For protein-level reporting, protein grouping was enabled, and values were calculated from the median of all quantifiable peptide spectral matches (PSMs) for each group. TMT values were then reported as a ratio to the sum of reporters in each spectrum.

Machine learning, multivariate analysis, and visualisation of data

The Bioconductor (Gentleman et al., 2004) packages MSnbase (Gatto and Lilley, 2012) and pRoloc (Gatto et al., 2014) for the R statistical programming language (R Core Team, 2013) were used for handling of the quantitative proteomics data and the protein-localisation prediction. pRolocGUI (Gatto et al., 2014) was employed for interactive visualisation of the data. Protein markers for the plasma membrane, thylakoid membrane, cytosol, and small and large ribosomal subunits were curated from a literature review (Supplemental Table S8). A Support Vector Machine (SVM) classifier was employed on the combined dataset, with a radial basis function kernel,

using class specific weights for classification of unassigned proteins to one of the five defined sub-cellular niches, TM, PM, soluble, small ribosomal subunit, large ribosomal subunit. The weights used in classification were set to be inversely proportional to the subcellular class frequencies to account for class imbalance. Algorithmic performance of the SVM on the dataset was estimated (as described in Trotter *et al* (Trotter et al., 2010)). Scoring thresholds were calculated per subcellular niche and were set based on concordance with existing subcellular knowledge annotation to attain a 7.5% false discovery rate (FDR). Unassigned proteins were then classified to 1 of the 5 compartments according to the SVM prediction if greater than the calculated class threshold.

All protein level datasets are available in the R Bioconductor pRolocdata package (https://bioconductor.org/packages/pRolocdata version 1.19.2) and can be interactively explored using the pRolocGUI package (https://bioconductor.org/packages/pRolocGUI) or using the standalone online interactive app (https://lgatto.shinyapps.io/synechocystis/).

The mass spectrometry data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD014662.

Accession Numbers

Gene/protein names, products and accession numbers of all genes/proteins identified in this study are listed in Supplemental Table S3

Supplemental Data

Supplemental Figures

Supplemental Figure S1: Growth of *Synechocystis* under continuous moderate light (60 µmol photons m^{-2} s^{-1}) with air-bubbling at 30°C.

Supplemental Figure S2: Partial fractionation of *Synechocystis* by sucrose density ultracentrifugation. Lysed cells were fractionated based on the method by Schottkowski *et al* (Schottkowski et al., 2009) with modifications. The first biological

replicate is used as a representative example. **A.** Initial step sucrose gradient (left) producing fractions I-IV and their corresponding absorption spectra (right). Fractions I-III demonstrated similar absorption spectra. Only the heaviest fraction (IV) showed any detectable absorbance or protein content. Asterisk indicates the fraction (IV) carried forward; **B.** Continuous sucrose gradient (left) resulting in fractions 1-12 and their corresponding absorption spectra (right). The lightest fractions (1-5) showed peaks of varying intensity at approximately 620 nm, corresponding to an enrichment of phycocyanin, whilst the highest density fractions (8-12) showed peaks of differing intensities at approximately 430 and 680 nm, corresponding to chlorophyll *a*. Fractions 6-7 exhibited substantially less absorption across the spectrum; **c.** Continuous sucrose gradient (left) resulting in fractions 1-12 and the separation of proteins by SDS-PAGE (right), visualised by Instant Blue staining.

Supplemental Figure S3: Comparison of assignment of proteins from this study with the Liberton *et al* (2016) data set between: A. Those found in the membranes in both studies and B. Those found in the soluble fraction in this study. Analysis of the data from the current study with that published by Liberton *et al* (Liberton et al., 2016) reveals some interesting observations about assignments to the plasma and thylakoid membranes in both studies. Liberton and co-workers presented their TM and PM in two different ways. Firstly in the 'TM_PM Sig. Protein 635.' tab of supplementary table 1, they listed all TM or PM proteins assigned by virtue of their quantitative log2 iTRAQ ratios and an arbitrary cut off +/- log2 0.5 was chosen. These data we denote as Liberton_Full. Secondly, the authors provided additional reduced lists, Top_TM and Top_PM, where a more stringent but equally arbitrary cut off of log2 +/- 2.0 was employed resulting in a list of 83 TM and 89 PM proteins. When comparing the full list with the data presented here, it is interesting to note that very few of Liberton's PM proteins were assigned as TM in this study and even fewer TM proteins assigned as PM, showing consistency between the membranes to which they have been assignments and the results presented in this study. There is only limited overlap between TM assignments and PM assignments, however, between the two studies. This is in part due to the fact that different proteins were identified in both studies. It is most likely due to the fact that the study presented here represents the whole cell, whereas the Liberton study analysed only a subset of proteins. It is not clear whether the additional PM and TM proteins presented in the Liberton study represent

contamination of their TM and PM enriched fractions with proteins from other parts of the cell, or the fact that this study returns the steady state location of proteins and hence if a TM and PM protein is also elsewhere in the cell, this study would flag it up as 'mixed location'. It is interesting to note that there is some overlap with Liberton's TM and particularly PM data with the soluble assignments in the data presented here. Analysis of these proteins shows that only 7% have a predicted single TMD and the remainder have no predicted membrane spanning regions.

Supplemental Figure S4: Carotenoid biosynthesis in *Synechocystis*. Carotenoids that accumulate in the cell are highlighted in red. Uncharacterised biosynthetic steps are indicated by broken arrows. Cellular location of proteins is indicated by the colour of the box surrounding the protein name: Yellow- TM; Blue- PM; Orange- soluble; Black- Unclassified.

Supplemental Figure S5: Distribution of carboxysome subunits and internal proteins in the PCA plot. Shell proteins of the carboxysome are localised predominantly in the soluble fraction (CcmAK1234LO) with the exception of CcmM (PM) and CcmN (unclassified). The carbonic anhydrase (CcaA) and RuBisCo subunits (RbcS, RbcL) are also in unclassified regions of the PCA plot.

Supplemental Figure S6: Alignment of Rps1A subunits from sequenced cyanobacterial species.

Supplemental Figure S7: Alignment of Rps1B subunits from sequenced cyanobacterial species. This protein is not conserved in *Gloeobacter kilaueensis* JS1 and *Gloeobacter violaceus* PCC 7421.

Supplemental Figure S8: Comparison of the TM and PM proteomes in terms of their functional categories. Proteins are classified into functional categories according to CyanoBase.

Supplementary Tables

Supplemental Table S1. Large-scale proteomic studies of *Synechocystis*. Comparative analysis was used to investigate responses to environmental changes, whilst a targeted approach focuses on a specific cellular sub-region without changing

environmental parameters. **Gel-based:** proteins separated by PAGE; **Shotgun:** proteins digested in solution, with peptides separated by fractionation; **iTRAQ:** peptides labelled with isobaric tags for relative and absolute quantification.

Supplemental Table S2: TMT quantitation data for two LOPIT replicate experiments and length, weight and pI of proteins identified.

Supplemental Table S3: Proteins identified in both replicates, the predicted localisations of proteins in *Synechocystis* by machine learning, using marker proteins as a training set. Protein size and the number of transmembrane helical domains (TMHs) present is also listed.

Supplemental Table S4: Proteins identified in this study and the one performed by Spat *et al* (2018)

Supplemental Table S5: Proteins identified in this study but not the one performed by Spat *et al* (2018)

Supplemental Table S6: Proteins not identified in this study but identified in the one performed by Spat *et al* (2018)

Supplemental Table S7: Proteins not identified in this study or in the one performed by Spat *et al* (2018)

Supplemental Table S8: Marker proteins used to identify subcellular regions.

Supplemental Table S9: Comparison of the localisation of *Arabidopsis* chloroplast envelope and thylakoid membrane proteins with homologs in *Synechocystis*. Excluded are proteins from the PSI, PSII, cyt b_6f, ATP synthase and NDH complexes, all of which are TM localised in both species.

Supplemental Table S10: BLAST analysis of uncharacterized *Synechocystis* TM localised proteins. Sequence similarity with proteins in *Chlamydomonas reinhardtii* and *Arabidopsis thaliana* are shown. Proteins highlighted in red are highly conserved in all three species.

Acknowledgements

D.J.L-S. was supported by the Environmental Services Association Education Trust. L.L.B was supported by a BBSRC Doctoral Training Grant (BB/F017464/1). L.M.B was supported by BBSRC grant BB/N023129/1. L.G was supported by Wellcome Trust Grant 110071/Z/15/Z. L.A.M was supported by a BBSRC Norwich Research Park Doctoral Studentship (BB/S507404/1). We thank Conrad Mullineaux (QMUL) for useful discussions and manuscript feedback, and Ravendran Vasudevan for help with culturing *Synechocystis*.

Figure Legends

Figure 1: The ultrastructure of *Synechocystis* showing various subcellular components. L: Lipid body; **C:** Carboxysome; **PHB:** Polyhydroxybutyrate granule; **PP:** Polyphosphate body; **Glyc:** Glycogen granule; **Cyano:** Cyanophycin granule. SEMs taken from Van de Meene *et al*. **Membrane-like structure** in close association with ribosomes (black arrow head) and seemingly continuous with TM (white arrow head). **Convergence site** of the PM and TM (white arrow). Bar = 50 nm.

Figure 2: Structural similarities between cyanobacteria and chloroplasts. Schematic depictions of the similar membrane organisation within a cyanobacterial cell and chloroplast.

Figure 3: Outline of the proteomic workflow. A. Total protein was extracted from each of the gradient fractions and quantified. **B.** The different distributions of TM and PM, as indicated by immunoblot analysis using antibodies against TM (CP47) and PM (SbtA) specific marker proteins **C.** Fractions 1-2 and 11-12 were merged to yield 10 gradient fractions and each labelled with a different tag using a 10-plex TMT kit. These fractions were merged as they exhibited similar protein profiles according to SDS-PAGE and immunoblot analysis. **D.** RP-HPLC was used to separate the proteins according to their hydrophobicity. **E.** This provided better resolution before subsequent MS/MS analysis. Proteins were identified by comparison to the database held by CyanoBase, and quantified using Proteome Discoverer Software 1.4.1.14 (Thermo Fisher Scientific).

Figure 4: Principal component analysis plots. A. Principal component analysis of the combined biological replicates. **B.** PCA plot showing the location of protein markers. **C.** PCA plot showing the assignment of proteins to subcellular regions. A cut-off of 0.75 (corresponding to 75%) was used for the boundaries of the TM, PM, small and large ribosomal subunits, and 0.65 for the soluble proteins. Grey circles indicate proteins with an unclassified localisation. **D.** Integral membrane proteins highlighted on the PCA plot of combined datasets.

Figure 5: Clustering of proteins with similar functions indicates potential further subcellular regions and compartmentalization. A. Two distinct sub-clusters of transport and binding proteins can be seen within the PM region. The smaller of these two groups is in close proximity to FtsZ, which forms the septal ring, and the MinCDE proteins which control the position and shape of the spectral ring; **B.** Sub-clustering of certain large ribosomal subunit proteins was observed, in close association with PBP1-3 to the PM region. The location of PBP4-8 are shown; **C.** Proteins thought to reside in the OM were found to localise to a distinct and unclassified region in between the PM and TM regions. Proteins involved in PHB biosynthesis are highlighted in purple; **D.** Numerous proteins which form complexes were found in very close proximity to each other on the PCA plot.

Figure 6: Predicted localization of proteins and biosynthetic pathways in *Synechocystis*. Enzymatic steps within a pathway which are localized to different regions of the cell are separated into appropriate colours/styles. **Green:** TM; **Brown:** PM; **Solid line:** Soluble; **Broken line:** Unclassified. **TCA cycle:** Tricarboxylic cycle; **PPP:** Pentose phosphate pathway; **Flv 1/3:** Flavodiiron protein 1/3. Refer to Supplemental Table S3 for protein abbreviations.

Figure 7: Schematic diagram detailing biosynthesis of lipopolysaccharides (LPSs) and assembly and polymerization of peptidoglycan (PG) monomers. LpxACDB enzymes synthesize the LPS disaccharide precursor. In *E. coli*, the flippase MsbA transfers the disaccharide to the periplasmic side of the PM, although the cyanobacterial MsbA has not been identified. RfbJUW are hypothesised to glycosylate the disaccharide. The LPS is transported to the OM by an uncharacterized protein complex. PG monomers are synthesised by MurABCDEFG and MraY enzymes. Localisation of MraY and murG in the TM suggests that the

monomers are subsequently transported to the PM, where the flippase, MurJ, transfers the monomers to the periplasmic side. Penicillin binding proteins Pbp1-4 and FtsW are involved in PG polymerization, while Pbp5-8 are likely involved in PG depolymerisation. A question mark indicates uncharacterized processes.

Figure 8: Schematic diagram detailing localisation of the electron transport complexes in cyanobacteria. Shown are the thylakoid membrane (A) photosynthetic and (B) respiratory electron transport chains, and the (C) plasma membrane electron transport chain. PSII- Photosystem II, PQ- plastoquinone, HemJ- protoporphyrinogen IX oxidase, cyt b_6f- cytochrome b_6f, Pc- plastocyanin, PSI- Photosystem I, Fd- ferredoxin, FNR- ferredoxin-NADP+-reductase, NDH-1- NDH dehydrogenase 1, SDH- Succinate dehydrogenase, Cyd- bd-quinol oxidase, COX- cytochrome-c oxidase, NdhB- NAD(P)H dehydrogenase 2 B, NdbC- NAD(P)H dehydrogenase 2 C, MenG- Demethylphyloquinone methyltransferase, PyrD- Dihydroorotate dehydrogenase, ARTO- Alternative respiratory terminal oxidase. Also shown are the PSII assembly proteins RubA (Rubredoxin A), Ycf48 and Ycf39 and the putative PSI assembly proteins Ycf4 and Ycf37. Localisation of SDH in the PM has not been confirmed. Dotted lines indicate possible electron transport routes.

References

Agarwal, R., Matros, A., Melzer, M., Mock, H.-P., and Sainis, J.K. (2010). Heterogeneity in thylakoid membrane proteome of Synechocystis sp. PCC 6803. Journal of proteomics **73,** 976-991.

Bryan, S.J., Burroughs, N.J., Shevela, D., Yu, J., Rupprecht, E., Liu, L.-N., Mastroianni, G., Xue, Q., Llorente-Garcia, I., Leake, M.C., Eichacker, L.A., Schneider, D., Nixon, P.J., and Mullineaux, C.W. (2014). Localisation and interactions of the Vipp1 protein in cyanobacteria. Molecular microbiology **94,** 1179-1195.

Brzezowski, P., Ksas, B., Havaux, M., Grimm, B., Chazaux, M., Peltier, G., Johnson, X., and Alric, J. (2019). The function of PROTOPORPHYRINOGEN IX OXIDASE in chlorophyll biosynthesis requires oxidised plastoquinone in Chlamydomonas reinhardtii. Commun Biol **2**.

Castenholz, R.W. (1988). Culturing methods for Cyanobacteria. Method Enzymol **167,** 68-93.

Cooley, J.W., and Vermaas, W.F. (2001). Succinate dehydrogenase and other respiratory pathways in thylakoid membranes of Synechocystis sp. strain PCC 6803: capacity comparisons and physiological function. J Bacteriol **183,** 4251-4258.

De Las Rivas, J., Lozano, J.J., and Ortiz, A.R. (2002). Comparative analysis of chloroplast genomes: functional annotation, genome-based phylogeny, and deduced evolutionary patterns. Genome research **12,** 567-583.

Dreger, M. (2003). Proteome analysis at the level of subcellular structures. European journal of biochemistry **270,** 589-599.

Ducat, D.C., Way, J.C., and Silver, P.A. (2011). Engineering cyanobacteria to generate high-value products. Trends Biotechnol **29,** 95-103.

Fatihi, A., Latimer, S., Schmollinger, S., Block, A., Dussault, P.H., Vermaas, W.F., Merchant, S.S., and Basset, G.J. (2015). A dedicated Type II NADPH Dehydrogenase performs the penultimate step in the biosynthesis of Vitamin K1 in Synechocystis and Arabidopsis. Plant Cell.

Ferro, M., Brugière, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., Ramus, C., Miras, S., Mellal, M., Le Gall, S., Kieffer-Jaquinod, S., Bruley, C., Garin, J., Joyard, J., Masselon, C., and Rolland, N. (2010). AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. Molecular & cellular proteomics : MCP **9,** 1063-1084.

Fisher, M.L., Allen, R., Luo, Y., and Curtiss 3rd, R. (2013). Export of extracellular polysaccharides modulates adherence of the Cyanobacterium synechocystis. PLoS One **8,** e74514-e74514.

Frain, K.M., Gangl, D., Jones, A., Zedler, J.A.Z.Z., and Robinson, C. (2016). Protein translocation and thylakoid biogenesis in cyanobacteria. Biochimica et biophysica acta **1857,** 266-273.

Garcia-Cerdan, J.G., Furst, A.L., McDonald, K.L., Schunemann, D., Francis, M.B., and Niyogi, K.K. (2019). A thylakoid membrane-bound and redox-active rubredoxin (RBD1) functions in de novo assembly and repair of photosystem II. Proceedings of the National Academy of Sciences of the United States of America **116,** 16631-16640.

Gatto, L., and Lilley, K.S. (2012). MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. Bioinformatics **28,** 288-289.

Gatto, L., Breckels, L.M., Wieczorek, S., Burger, T., and Lilley, K.S. (2014). Mass-spectrometry-based spatial proteomics data analysis using pRoloc and pRolocdata. Bioinformatics **30,** 1322-1324.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A.J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J.Y.H., and Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome biology **5,** R80.

Gonzalez-Esquer, C.R., Shubitowski, T.B., and Kerfeld, C.A. (2015). Streamlined Construction of the Cyanobacterial CO2-Fixing Organelle via Protein Domain Fusions for Use in Plant Synthetic Biology. Plant Cell **27,** 2637-2644.

Graham, J.E., and Bryant, D.A. (2008). The Biosynthetic Pathway for Synechoxanthin, an Aromatic Carotenoid Synthesized by the Euryhaline, Unicellular Cyanobacterium Synechococcus sp Strain PCC 7002. Journal of Bacteriology **190,** 7966-7974.

Graham, J.E., and Bryant, D.A. (2009). The biosynthetic pathway for myxol-2' fucoside (myxoxanthophyll) in the cyanobacterium Synechococcus sp. strain PCC 7002. J Bacteriol **191,** 3292-3300.

Griese, M., Lange, C., and Soppa, J. (2011). Ploidy in cyanobacteria. FEMS Microbiology Letters **323,** 124-131.

Hauf, W., Watzer, B., Roos, N., Klotz, A., and Forchhammer, K. (2015). Photoautotrophic Polyhydroxybutyrate Granule Formation Is Regulated by Cyanobacterial Phasin PhaP in Synechocystis sp. Strain PCC 6803. Applied and environmental microbiology **81,** 4411-4422.

Hauf, W., Schlebusch, M., Hüge, J., Kopka, J., Hagemann, M., and Forchhammer, K. (2013). Metabolic Changes in Synechocystis PCC6803 upon Nitrogen-Starvation: Excess NADPH Sustains Polyhydroxybutyrate Accumulation. Metabolites **3,** 101-118.

Hennig, R., Heidrich, J., Saur, M., Schmuser, L., Roeters, S.J., Hellmann, N., Woutersen, S., Bonn, M., Weidner, T., Markl, J., and Schneider, D. (2015). IM30 triggers membrane fusion in cyanobacteria and chloroplasts. Nat Commun **6,** 7018.

Herranen, M., Battchikova, N., Zhang, P.P., Graf, A., Sirpiö, S., Paakkarinen, V., Aro, E.-M.M., Sirpio, S., Paakkarinen, V., and Aro, E.-M.M. (2004). Towards functional proteomics of membrane protein complexes in Synechocystis sp PCC 6803. Plant Physiology **134,** 470-481.

Hinterstoisser, B., Cichna, M., Kuntner, O., and Peschek, G.A. (1993). Cooperation of plasma and thylakoid membranes for the biosynthesis of chlorophyll in cyanobacteria: the role of the thylakoid centers. Journal of Plant Physiology **142,** 407-413.

Howe, C.J., Barbrook, A.C., Nisbet, R.E.R., Lockhart, P.J., and Larkum, A.W.D. (2008). The origin of plastids. Philos Trans R Soc Lond B Biol Sci **363,** 2675-2685.

Huang, F., Fulda, S., Hagemann, M., and Norling, B. (2006). Proteomic screening of salt-stress-induced changes in plasma membranes of Synechocystis sp. strain PCC 6803. Proteomics **6,** 910-920.

Huang, F., Parmryd, I., Nilsson, F., Persson, A.L., Pakrasi, H.B., Andersson, B., and Norling, B. (2002). Proteomics of Synechocystis sp. strain PCC 6803: identification of plasma membrane proteins. Mol Cell Proteomics **1,** 956-966.

Huang, F., Hedman, E., Funk, C., Kieselbach, T., Schroder, W.P., Norling, B., Schröder, W.P., and Norling, B. (2004). Isolation of outer membrane of Synechocystis sp PCC 6803 and its proteomic characterization. Molecular & Cellular Proteomics **3,** 586-595.

Ito, K., and Chiba, S. (2014). Regulatory Nascent Polypeptides. (Tokyo: Springer Japan).

Kanamaru, K., Kashiwagi, S., and Mizuno, T. (1994). A Copper-Transporting P-Type Atpase Found in the Thylakoid Membrane of the Cyanobacterium Synechococcus Species Pcc7942. Molecular Microbiology **13,** 369-377.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions (supplement). DNA Res **3,** 185-209.

Kim, D.I., Jensen, S.C., Noble, K.A., Kc, B., Roux, K.H., Motamedchaboki, K., and Roux, K.J. (2016). An improved smaller biotin ligase for BioID proximity labeling. Mol Biol Cell **27,** 1188-1196.

Kiss, E., Knoppova, J., Aznar, G.P., Pilny, J., Yu, J.F., Halada, P., Nixon, P.J., Sobotka, R., and Komenda, J. (2019). A Photosynthesis-Specific Rubredoxin-Like Protein Is Required for Efficient Association of the D1 and D2 Proteins during the Initial Steps of Photosystem II Assembly. Plant Cell **31,** 2241-2258.

Kurian, D., Jansèn, T., and Mäenpää, P. (2006a). Proteomic analysis of heterotrophy in Synechocystis sp. PCC 6803. Proteomics **6,** 1483-1494.

Kurian, D., Phadwal, K., and Mäenpää, P. (2006b). Proteomic characterization of acid stress response in Synechocystis sp. PCC 6803. Proteomics **6,** 3614-3624.

Lam, S.S., Martell, J.D., Kamer, K.J., Deerinck, T.J., Ellisman, M.H., Mootha, V.K., and Ting, A.Y. (2015). Directed evolution of APEX2 for electron microscopy and proximity labeling. Nat Methods **12,** 51-54.

Lea-Smith, D.J., Bombelli, P., Vasudevan, R., and Howe, C.J. (2016a). Photosynthetic, respiratory and extracellular electron transport pathways in cyanobacteria. Biochimica et Biophysica Acta - Bioenergetics **1857,** 247-255.

Lea-Smith, D.J., Ortiz-Suarez, M.L., Lenn, T., Nurnberg, D.J., Baers, L.L., Davey, M.P., Parolini, L., Huber, R.G., Cotton, C.A.R., Mastroianni, G., Bombelli, P., Ungerer, P., Stevens, T.J., Smith, A.G., Bond, P.J., Mullineaux, C.W., and Howe, C.J. (2016b). Hydrocarbons are essential for optimal cell size, division and growth of cyanobacteria. Plant Physiology **172,** 1928-1940.

Leitner, A., Faini, M., Stengel, F., and Aebersold, R. (2016). Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. Trends Biochem Sci **41,** 20-32.

Li, T., Yang, H.-m.M., Cui, S.-x.X., Suzuki, I., Zhang, L.-F.F., Li, L., Bo, T.-t.T., Wang, J., Murata, N., and Huang, F. (2012). Proteomic Study of the Impact of Hik33 Mutation in Synechocystis sp PCC 6803 under Normal and Salt Stress Conditions. Journal of Proteome Research **11,** 502-514.

Liberton, M., Howard Berg, R., Heuser, J., Roth, R., and Pakrasi, H.B. (2006). Ultrastructure of the membrane systems in the unicellular cyanobacterium Synechocystis sp. strain PCC 6803. Protoplasma **227,** 129-138.

Liberton, M., Saha, R., Jacobs, J.M., Nguyen, A.Y., Gritsenko, M.A., Smith, R.D., Koppenaal, D.W., and Pakrasi, H.B. (2016). Global proteomic analysis reveals an exclusive role of thylakoid membranes in bioenergetics of a model cyanobacterium. Molecular & Cellular Proteomics **15,** 2021-2032.

Liu, F., Rijkers, D.T.S., Post, H., and Heck, A.J.R. (2015). Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. Nature methods **12,** 1179-1184.

Loh, K.H., Stawski, P.S., Draycott, A.S., Udeshi, N.D., Lehrman, E.K., Wilton, D.K., Svinkina, T., Deerinck, T.J., Ellisman, M.H., Stevens, B., Carr, S.A., and Ting, A.Y. (2016). Proteomic Analysis of Unbounded Cellular Compartments: Synaptic Clefts. Cell **166,** 1295-1307

Magnan, D., Joshi, Mohan C., Barker, Anna K., Visser, Bryan J., and Bates, D. (2015). DNA Replication Initiation Is Blocked by a Distant Chromosome–Membrane Attachment. Current Biology **25,** 2143-2149.

Marbouty, M., Saguez, C., Cassier-Chauvat, C., and Chauvat, F. (2009a). ZipN, an FtsA-like orchestrator of divisome assembly in the model cyanobacterium Synechocystis PCC6803. Molecular Microbiology **74,** 409-420.

Marbouty, M., Mazouni, K., Saguez, C., Cassier-Chauvat, C., and Chauvat, F. (2009b). Characterization of the Synechocystis strain PCC 6803 penicillin-binding proteins and cytokinetic proteins FtsQ and FtsW and their network of interactions with ZipN. J Bacteriol **191,** 5123-5133.

Maresca, J.A., Graham, J.E., Wu, M., Eisen, J.a., and Bryant, D.a. (2007). Identification of a fourth family of lycopene cyclases in photosynthetic bacteria. Proceedings of the National Academy of Sciences of the United States of America **104,** 11784-11789.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M., and Penny, D. (2002). Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proceedings of the National Academy of Sciences of the United States of America **99,** 12246-12251.

Masamoto, K., Wada, H., Kaneko, T., and Takaichi, S. (2001). Identification of a gene required for cis-to-trans carotene isomerization in carotenogenesis of the cyanobacterium Synechocystis sp. PCC 6803. Plant Cell Physiol **42,** 1398-1402.

Moffitt, J.R., Pandey, S., Boettiger, A.N., Wang, S.Y., and Zhuang, X.W. (2016). Spatial organization shapes the turnover of a bacterial transcriptome. Elife **5**.

Mohamed, H.E., Vermaas, W., and Myxoxanthophyll, R. (2004). Slr1293 in Synechocystis sp. strain PCC 6803 Is the C-3',4' desaturase (CrtD) involved in myxoxanthophyll biosynthesis. J Bacteriol **186,** 5621-5628.

Mulvey, C.M., Breckels, L.M., Geladaki, A., Britovšek, N.K., Nightingale, D.J.H., Christoforou, A., Elzek, M., Deery, M.J., Gatto, L., and Lilley, K.S. (2017). Using hyperLOPIT to perform high-resolution mapping of the spatial proteome. Nature Protocols **12,** 1110-1135.

Nevo-Dinur, K., Nussbaum-Shochat, A., Ben-Yehuda, S., and Amster-Choder, O. (2011). Translation-independent localization of mRNA in E. coli. Science **331,** 1081-1084.

Nørager, S., Jensen, K.F., Björnberg, O., and Larsen, S. (2002). E. coli dihydroorotate dehydrogenase reveals structural and functional distinctions between different classes of dihydroorotate dehydrogenases. Structure (London, England : 1993) **10,** 1211-1223.

Norling, B., Zak, E., Andersson, B., and Pakrasi, H. (1998). 2D-isolation of pure plasma and thylakoid membranes from the cyanobacterium Synechocystis sp. PCC 6803. FEBS Lett **436,** 189-192.

Perez-Riverol, Y., Csordas, A., Bai, J.W., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., Perez, E., Uszkoreit, J., Pfeuffer, J., Sachsenberg, T., Yilmaz, S., Tiwary, S., Cox, J., Audain, E., Walzer, M., Jarnuczak, A.F., Ternent, T., Brazma, A., and Vizcaino, J.A. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Research **47,** D442-D450.

Pisareva, T., Shumskaya, M., Maddalo, G., Ilag, L., and Norling, B. (2007). Proteomics of Synechocystis sp PCC 6803 - Identification of novel integral plasma membrane proteins. Febs Journal **274,** 791-804.

Pisareva, T., Kwon, J., Oh, J., Kim, S., Ge, C., Wieslander, Å., Choi, J.S., and Norling, B. (2011). Model for membrane organization and protein sorting in the cyanobacterium synechocystis sp. PCC 6803 inferred from proteomics and multivariate sequence analyses. Journal of Proteome Research **10,** 3617-3631.

Plohnke, N., Seidel, T., Kahmann, U., Rögner, M., Schneider, D., and Rexroth, S. (2015). The proteome and lipidome of Synechocystis sp. PCC 6803 cells grown under light-activated heterotrophic conditions. Molecular & cellular proteomics : MCP **14,** 572-584.

R Core Team, R. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Rast, A., Schaffer, M., Albert, S., Wan, W., Pfeffer, S., Beck, F., Plitzko, J.M., Nickelsen, J., and Engel, B.D. (2019). Biogenic regions of cyanobacterial thylakoids form contact sites with the plasma membrane. Nat Plants **5,** 436-446.

Rowland, J.G., Simon, W.J., Nishiyama, Y., and Slabas, A.R. (2010). Differential proteomic analysis using iTRAQ reveals changes in thylakoids associated with Photosystem II-acquired thermotolerance in Synechocystis sp. PCC 6803. Proteomics **10,** 1917-1929.

Ruiz, N., Kahne, D., and Silhavy, T.J. (2009). Transport of lipopolysaccharide across the cell envelope: the long road of discovery. Nat Rev Microbiol **7,** 677-683.

Sakuragi, Y., Zybailov, B., Shen, G., Jones, A.D., Chitnis, P.R., van der Est, A., Bittl, R., Zech, S., Stehlik, D., Golbeck, J.H., and Bryant, D.A. (2002). Insertional inactivation of the menG gene, encoding 2-phytyl-1,4-naphthoquinone methyltransferase of Synechocystis sp. PCC 6803, results in the incorporation of 2-phytyl-1,4-naphthoquinone into the A(1) site and alteration of the equilibrium constant betw. Biochemistry **41,** 394-405.

Sauvage, E., Kerff, F., Terrak, M., Ayala, J.A., and Charlier, P. (2008). The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. FEMS Microbiol Rev **32,** 234-258.

Saxena, R., Fingland, N., Patil, D., Sharma, A.K., and Crooke, E. (2013). Crosstalk between DnaA protein, the initiator of Escherichia coli chromosomal replication, and acidic phospholipids present in bacterial membranes. International journal of molecular sciences **14,** 8517-8537.

Schottkowski, M., Gkalympoudis, S., Tzekova, N., Stelljes, C., Schünemann, D., Ankele, E., Nickelsen, J., Schunemann, D., Ankele, E., and Nickelsen, J. (2009). Interaction of the Periplasmic PratA Factor and the PsbA (D1) Protein during Biogenesis of Photosystem II in Synechocystis sp PCC 6803. Journal of Biological Chemistry **284,** 1813-1819.

Sham, L.T., Butler, E.K., Lebar, M.D., Kahne, D., Bernhardt, T.G., and Ruiz, N. (2014). MurJ is the flippase of lipid-linked precursors for peptidoglycan biogenesis. Science **345,** 220-222.

Shikanai, T. (2016). Chloroplast NDH: A different enzyme with a structure similar to that of respiratory NADH dehydrogenase. Biochimica Et Biophysica Acta-Bioenergetics **1857,** 1015-1022.

Simon, W.J., Hall, J.J., Suzuki, I., Murata, N., and Slabas, A.R. (2002). Proteomic study of the soluble proteins from the unicellular cyanobacterium Synechocystis sp. PCC6803 using automated matrix-assisted laser desorption/ionization-time of flight peptide mass fingerprinting. Proteomics **2,** 1735-1742.

Skotnicova, P., Sobotka, R., Shepherd, M., Hajek, J., Hrouzek, P., and Tichy, M. (2018). The cyanobacterial protoporphyrinogen oxidase HemJ is a new b-type heme protein functionally coupled with coproporphyrinogen III oxidase. Journal of Biological Chemistry **293,** 12394-12404.

Slabas, A.R., Suzuki, I., Murata, N., Simon, W.J., and Hall, J.J. (2006). Proteomic analysis of the heat shock response in Synechocystis PCC6803 and a thermally tolerant knockout strain lacking the histidine kinase 34 gene. Proteomics **6,** 845-864.

Spat, P., Klotz, A., Rexroth, S., Macek, B., and Forchhammer, K. (2018). Chlorosis as a Developmental Program in Cyanobacteria: The Proteomic Fundament for Survival and Awakening. Molecular & Cellular Proteomics **17,** 1650-1669.

Srivastava, R., Pisareva, T., and Norling, B. (2005). Proteomic studies of the thylakoid membrane of Synechocystis sp. PCC 6803. Proteomics **5,** 4905-4916.

Srivastava, R., Battchikova, N., Norling, B., and Aro, E.-M.M. (2006). Plasma membrane of Synechocystis PCC 6803: a heterogeneous distribution of membrane proteins. Archives of Microbiology **185,** 238-243.

Stanier, R.Y., and Cohen-Bazire, G. (1977). Phototrophic prokaryotes: the cyanobacteria. Annual review of microbiology **31,** 225-274.

Straskova, A., Steinbach, G., Konert, G., Kotabova, E., Komenda, J., Tichy, M., and Kana, R. (2019). Pigment-protein complexes are organized into stable microdomains in cyanobacterial thylakoids. Biochim Biophys Acta Bioenerg.

Summerfield, T.C., Toepel, J., and Sherman, L.A. (2008). Low-Oxygen Induction of Normally Cryptic psbA Genes in Cyanobacteria. Biochemistry **47,** 12939-12941.

Suzuki, I., Simon, W.J., and Slabas, A.R. (2006). The heat shock response of Synechocystis sp. PCC 6803 analysed by transcriptomics and proteomics. Journal of experimental botany **57,** 1573-1578.

Taguchi, A., Welsh, M.A., Marmont, L.S., Lee, W., Sjodt, M., Kruse, A.C., Kahne, D., Bernhardt, T.G., and Walker, S. (2019). FtsW is a peptidoglycan polymerase that is functional only in complex with its cognate penicillin-binding protein. Nature Microbiology **4,** 587-594.

Thul, P.J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., Alm, T., Asplund, A., Björk, L., Breckels, L.M., Bäckström, A., Danielsson, F., Fagerberg, L., Fall, J., Gatto, L., Gnann, C., Hober, S., Hjelmare, M., Johansson, F., Lee, S., Lindskog, C., Mulder, J., Mulvey, C.M., Nilsson, P., Oksvold, P., Rockberg, J., Schutten, R., Schwenk, J.M., Sivertsson, Å., Sjöstedt, E., Skogs, M., Stadler, C., Sullivan, D.P., Tegel, H., Winsnes, C., Zhang, C., Zwahlen, M., Mardinoglu, A., Pontén, F., von Feilitzen, K., Lilley, K.S., Uhlén, M., and Lundberg, E. (2017). A subcellular map of the human proteome. Science **356,** 6340.

Toth, T.N., Chukhutsina, V., Domonkos, I., Knoppova, J., Komenda, J., Kis, M., Lenart, Z., Garab, G., Kovacs, L., Gombos, Z., and van Amerongen, H. (2015). Carotenoids are essential for the assembly of cyanobacterial photosynthetic complexes. Biochim Biophys Acta **1847,** 1153-1165.

Tottey, S., Patterson, C.J., Banci, L., Bertini, I., Felli, I.C., Pavelkova, A., Dainty, S.J., Pernil, R., Waldron, K.J., Foster, A.W., and Robinson, N.J. (2012). Cyanobacterial metallochaperone inhibits deleterious side reactions of copper. Proceedings of the National Academy of Sciences of the United States of America **109,** 95-100.

Trotter, M.W., Sadowski, P.G., Dunkley, T.P., Groen, A.J., and Lilley, K.S. (2010). Improved sub-cellular resolution via simultaneous analysis of organelle proteomics data across varied experimental conditions. Proteomics **10,** 4213-4219.

van de Meene, A.M.L., Hohmann-Marriott, M.F., Vermaas, W.F.J., and Roberson, R.W. (2006). The three-dimensional structure of the cyanobacterium Synechocystis sp. PCC 6803. Archives of Microbiology **184,** 259-270.

van Heijenoort, J. (2011). Peptidoglycan hydrolases of Escherichia coli. Microbiol Mol Biol Rev **75,** 636-663.

Villén, J., and Gygi, S.P. (2008). The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nature protocols **3,** 1630-1638.

von Berlepsch, S., Kunz, H.-H., Brodesser, S., Fink, P., Marin, K., Flügge, U.-i., and Gierth, M. (2012). The acyl-acyl carrier protein synthetase from Synechocystis sp. PCC 6803 mediates fatty acid import. Plant physiology **159,** 606-617.

Wang, H., Yan, X., Aigner, H., Bracher, A., Nguyen, N.D., Hee, W.Y., Long, B.M., Price, G.D., Hartl, F.U., and Hayer-Hartl, M. (2019). Rubisco condensate formation by CcmM in beta-carboxysome biogenesis. Nature.

Wang, Y., Sun, J., and Chitnis, P.R. (2000). Proteomic study of the peripheral proteins from thylakoid membranes of the cyanobacterium Synechocystis sp. PCC 6803. Electrophoresis **21,** 1746-1754.

Wang, Y., Xu, W., and Chitnis, P.R. (2009). Identification and bioinformatic analysis of the membrane proteins of synechocystis sp. PCC 6803. Proteome science **7,** 11-11.

Wegener, K.M., Singh, A.K., Jacobs, J.M., Elvitigala, T., Welsh, E.a., Keren, N., Gritsenko, M.a., Ghosh, B.K., Camp, D.G., Smith, R.D., and Pakrasi, H.B. (2010). Global proteomics reveal an atypical strategy for carbon/nitrogen assimilation by a cyanobacterium under diverse environmental perturbations. Molecular & cellular proteomics : MCP **9,** 2678-2689.

Wessel, D., and Flügge, U.I. (1984). A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Analytical biochemistry **138,** 141-143.

Westphal, S., Heins, L., Soll, J., and Vothknecht, U.C. (2001). Vipp1 deletion mutant of Synechocystis: A connection between bacterial phage shock and thylakoid biogenesis? Proc Natl Acad Sci U S A **98,** 4243-4248.

Xu, C.C., Fan, J., Froehlich, J.E., Awai, K., and Benning, C. (2005). Mutation of the TGD1 chloroplast envelope protein affects phosphatidate metabolism in Arabidopsis. Plant Cell **17,** 3094-3110.

Yoshihara, S., Geng, X., Okamoto, S., Yura, K., Murata, T., Go, M., Ohmori, M., and Ikeuchi, M. (2001). Mutational analysis of genes involved in pilus structure, motility and transformation competency in the unicellular motile cyanobacterium Synechocystis sp. PCC 6803. Plant Cell Physiol **42,** 63-73.

Yusupova, G., and Yusupov, M. (2014). High-resolution structure of the eukaryotic 80S ribosome. Annual review of biochemistry **83,** 467-486.

Zhang, L.-f.F., Yang, H.-m.M., Cui, S.-x.X., Hu, J., Wang, J., Kuang, T.-y.Y., Norling, B., and Huang, F. (2009). Proteomic analysis of plasma membranes of cyanobacterium Synechocystis sp. Strain PCC 6803 in response to high pH stress. J Proteome Res **8,** 2892-2902.

Zhang, L.F., Selao, T.T., Selstam, E., Norling, B., Selão, T.T., Selstam, E., and Norling, B. (2015). Subcellular Localization of Carotenoid Biosynthesis in Synechocystis sp PCC 6803. PLoS One **10,** e0130904-e0130904.

Zhang, P., Battchikova, N., Jansen, T., Appel, J., Ogawa, T., and Aro, E.-M.M. (2004). Expression and functional roles of the two distinct NDH-1 complexes and the carbon acquisition complex NdhD3/NdhF3/CupA/Sll1735 in Synechocystis sp PCC 6803. Plant Cell **16,** 3326-3340.

Zhang, P., Eisenhut, M., Brandt, A.-M.M., Carmel, D., Silén, H.M., Vass, I., Allahverdiyeva, Y., Salminen, T.A., Aro, E.-M.M., Silen, H.M., Vass, I., Allahverdiyeva, Y., Salminen, T.A., and Aro, E.-M.M. (2012). Operon flv4-flv2 provides cyanobacterial photosystem II with flexibility of electron transfer. Plant Cell **24,** 1952-1971.

Zhbanko, M., Zinchenko, V., Gutensohn, M., Schierhorn, A., Klosgen, R.B., and Klösgen, R.B. (2005). Inactivation of a predicted leader peptidase prevents photoautotrophic growth of Synechocystis sp. strain PCC 6803. J Bacteriol **187,** 3071-3078.

Zwirglmaier, K., Jardillier, L., Ostrowski, M., Mazard, S., Garczarek, L., Vaulot, D., Not, F., Massana, R., Ulloa, O., and Scanlan, D.J. (2008). Global phylogeography of marine Synechococcus and Prochlorococcus reveals a distinct partitioning of lineages among oceanic biomes. Environ Microbiol **10,** 147-161.

Competing Financial Interests

The authors declare no competing financial interest**.**