EXPLORING THE GROWING ROLE OF CYANOBACTERIA IN INDUSTRIAL BIOTECHNOLOGY AND SUSTAINABILITY

EDITED BY: David John Lea-Smith, Tina Summerfield, Daniel Ducat, Xuefeng Lu, Alistair McCormick and Saul Purton

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

EXPLORING THE GROWING ROLE OF CYANOBACTERIA IN INDUSTRIAL BIOTECHNOLOGY AND SUSTAINABILITY

Topic Editors:
David John Lea-Smith, University of East Anglia, United Kingdom
Tina Summerfield, University of Otago, New Zealand
Daniel Ducat, Michigan State University, United States
Xuefeng Lu, Chinese Academy of Sciences (CAS), China
Alistair McCormick, University of Edinburgh, United Kingdom
Saul Purton, University College London, United Kingdom

Table of Contents

frontiers
in Microbiology

Editorial: Exploring the Growing Role of Cyanobacteria in Industrial Biotechnology and Sustainability

David J. Lea-Smith[1]*, Tina C. Summerfield[2], Daniel C. Ducat[3], Xuefeng Lu[4], Alistair J. McCormick[5] and Saul Purton[6]

[1] School of Biological Sciences, University of East Anglia, Norwich, United Kingdom, [2] Department of Botany, University of Otago, Dunedin, New Zealand, [3] Michigan State University-Department of Energy (MSU-DOE) Plant Research Laboratory and Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI, United States, [4] Shandong Provincial Key Laboratory of Synthetic Biology, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, China, [5] School of Biological Sciences, SynthSys and Institute of Molecular Plant Sciences, University of Edinburgh, Edinburgh, United Kingdom, [6] Department of Structural and Molecular Biology, University College London, London, United Kingdom

Editorial on the Research Topic

Exploring the Growing Role of Cyanobacteria in Industrial Biotechnology and Sustainability

INTRODUCTION

A major challenge of the 21st Century is the development of innovative, sustainable biotechnologies able to replace fossil fuel derived synthetic routes for production of bulk chemicals and high-value materials. Potentially, this challenge could be met in part through the use of cyanobacteria, the only prokaryotes capable of oxygenic photosynthesis, as microbial hosts for development of next generation industrial biotechnologies. The production of cyanobacterial biomass and synthesis of bioproducts does not require arable land, thereby avoiding competition with food production. Minimal nutrients are needed for cyanobacterial growth, and many species can be cultured in seawater, avoiding use of limited freshwater supplies. In addition, cyanobacteria can synthesize a range of high-value commercial compounds e.g., healthcare relevant pharmaceuticals, nutraceuticals, and industrial materials (Ducat et al., 2011), and their derived photo-production technology has been industrialized in several countries, including the USA, China, and Japan (Zahra et al., 2020).

However, expanding the use of cyanobacteria for production of a wider range of compounds is still restricted by multiple factors. These include the naturally slow growth rate and low biomass accumulation of most cyanobacterial species, although the recent discovery of several species, such as *Synechococcus* sp. PCC 11901 (Wlodarczyk et al., 2020), with doubling times as short as 2 h, is a promising development. The challenges of low cost culturing in outdoor photobioreactors, development of strains capable of synthesizing commercial quantities of the desired compound, and a lack of low-cost and sustainable methods for compound extraction also impede commercialization of cyanobacterial biotechnology. Strain development is hampered by genetic tools that are less developed than those available for established heterotrophic platforms such as *Escherichia coli* and *Saccharomyces cerevisiae*. Finally, our understanding of many core physiological and biochemical processes in cyanobacteria, even in the most widely studied model cyanobacterium, *Synechocystis* sp. PCC 6803 (6803), is incomplete (Mills et al., 2020).

The publications in this special issue address many of these considerations. In terms of strain selection for biotechnology applications, many factors need to be considered including growth and biomass accumulation, long term storage of strains at −80°C, growth in seawater and genetic tractability. Jeong et al. sequenced a new *Synechocystis* species, PCC 7338 (7338), which can be cultured in seawater, and compared it to multiple freshwater species, including 6803. Although the majority of genes were conserved in all the species examined, a number were unique to 7338 and likely involved in salt tolerance. Shaw et al. examined 26 photosynthetic co-cultures (consisting of a cyanobacterium and associated heterotrophic microbes) collected from a range of geographical locations and different ecosystems, focusing upon extreme environments such as hot springs and Antarctic ponds. Sequence analysis of these strains highlighted the diversity of species surviving within these extreme conditions, and provides potential opportunities for discovery of new proteins and biotechnologically-valuable compounds.

A greater understanding of cyanobacterial biology will aid commercialization of cyanobacterial biotechnology. Karlsen et al. examined why protein levels are relatively constant in 6803 subjected to artificial day-night cycles, despite the transcriptional profile of many genes altering under these conditions. Their data demonstrates that slow protein turnover and not translational regulation is the main factor in controlling protein amounts. These results have ramifications for synthetic biology and that controlling the abundance of heterologous proteins not only has to take into account expression but also degradation of the target protein. Thirumurthy et al. investigated the role of type IV pili in extracellular electron transport in 6803. This process may play a role in photoprotection by exporting excess electrons but can also be exploited for electricity production using biophotovoltaic devices, a type of microbial fuel cell (McCormick et al., 2015). Using pili-deficient mutants, their data shows that deleting these appendages has no effect on electron export, suggesting that other compounds, likely soluble electron carriers, may perform this role.

Development of new molecular tools, such as CyanoGate, a modular Golden Gate cloning kit (Vasudevan et al., 2019), aids research into understanding cyanobacterial biology and the engineering of strains for biotechnology applications. Gale et al. described the development of a CyanoGate-compatible dual reporter system to quantify and compare the efficiency of transcriptional terminators within and between species. A suite of 34 terminators were characterized and five were identified with high efficiencies in 6803, *Synechococcus elongatus* UTEX 2973 and *E. coli*. Zhang et al. utilized knock out/down and overexpression strategies to characterize twelve genes potentially

involved in cell division and/or elongation in *Synechococcus elongatus* PCC 7942. Their work has advanced our understanding of these processes and identified several new targets for engineering cell morphology in cyanobacteria.

The remaining publications focus on engineering cyanobacteria for development of efficient strains more suitable for biotechnology or for higher production of a range of compounds. Wang et al. detail in their comprehensive review the recent progress in manipulating cyanobacteria for production of a range of compounds, including fatty acids, alcohols, hydrocarbons, and fatty acid esters. Song et al. analyzed the effect of minimizing photorespiratory carbon losses by expression of the formate-tetrahydrofolate ligase in 6803. Strains accumulated higher amounts of photorespiratory intermediates but had altered regulation of the carbon/nitrogen ratio. This paper highlighted both the robustness of cyanobacteria as chassis for heterologous protein expression and the value of empirical experimentation in evaluating its impact. Wang et al. engineered 6803 for production of *myo*-inositol, a compound of interest to the food and pharmaceutical industries. Introduction of the *S. cerevisiae myo*-inositol-1-phosphate synthase gene and overexpression of native genes encoding putative *myo*-inositol-1-monophosphates, together with the re-direction of carbon flux into production of the precursor compound, glucose-6-phosphate and the optimization of cultivation conditions, resulted in production of 12.72 mg L^{-1}.

Overall, the work published in these studies will contribute to the development of cyanobacteria for biotechnology applications, resulting in more sustainable industries for production of a range of chemicals currently derived from agricultural or fossil fuel sources.

AUTHOR CONTRIBUTIONS

DL-S wrote the first draft. All authors developed the overall concept, provided input, and comments to the draft. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

REFERENCES

Ducat, D. C., Way, J. C., and Silver, P. A. (2011). Engineering cyanobacteria to generate high-value products. *Trends Biotechnol.* 29, 95–103. doi: 10.1016/j.tibtech.2010.12.003

McCormick, A. J., Bombelli, P., Bradley, R. W., Thorne, R., Wenzel, T., Howe, C. J., et al. (2015). Biophotovoltaics: oxygenic photosynthetic organisms in the world of bioelectrochemical systems. *Energy Environ.Sci.* 8, 1092–1109. doi: 10.1039/C4EE03875D

Mills, L. A., McCormick, A. J., and Lea-Smith, D. J. (2020). Current knowledge and recent advances in understanding metabolism of the model

cyanobacterium Synechocystis sp. PCC 6803. *Biosci. Rep.* 40:BSR20193325. doi: 10.1042/BSR20193325

Vasudevan, R., Gale, G. A. R., Schiavon, A. A., Puzorjov, A., Malin, J., Gillespie, M. D., et al. (2019). CyanoGate: a modular cloning suite for engineering cyanobacteria based on the plant MoClo syntax. *Plant Physiol.* 180, 39–55. doi: 10.1104/pp.18.01401

Wlodarczyk, A., Selao, T. T., Norling, B., and Nixon, P. J. (2020). Newly discovered Synechococcus sp. PCC 11901 is a robust cyanobacterial strain for high biomass production. *Commun. Biol.* 3:215. doi: 10.1038/s42003-020-0910-8

Zahra, Z., Choo, D. H., Lee, H., and Parveen, A. (2020). Cyanobacteria: review of current potentials and applications. *Environments* 7:13. doi: 10.3390/environments7020013

Check for
updates

Photosynthetic Conversion of Carbon Dioxide to Oleochemicals by Cyanobacteria: Recent Advances and Future Perspectives

Li Wang, Liyuan Chen, Shihui Yang* and Xiaoming Tan*

State Key Laboratory of Biocatalysis and Enzyme Engineering, Environmental Microbial Technology Center of Hubei Province, School of Life Sciences, Hubei University, Wuhan, China

Sustainable production of biofuels and biochemicals has been broadly accepted as a solution to lower carbon dioxide emissions. Besides being used as lubricants or detergents, oleochemicals are also attractive biofuels as they are compatible with existing transport infrastructures. Cyanobacteria are autotrophic prokaryotes possessing photosynthetic abilities with mature genetic manipulation systems. Through the introduction of exogenous or the modification of intrinsic metabolic pathways, cyanobacteria have been engineered to produce various bio-chemicals and biofuels over the past decade. In this review, we specifically summarize recent progress on photosynthetic production of fatty acids, fatty alcohols, fatty alk(a/e)nes, and fatty acid esters by genetically engineered cyanobacteria. We also summarize recent reports on fatty acid and lipid metabolisms of cyanobacteria and provide perspectives for economic cyanobacterial oleochemical production in the future.

Keywords: cyanobacterium, oleochemicals, metabolic engineering, carbon dioxide conversion, lipid metabolism

INTRODUCTION

Since the Industrial Revolution, the level of global carbon dioxide together with other greenhouse gases (GHGs) significantly increased due to human activities (Ainsworth et al., 2020). Cumulative anthropogenic emissions of CO_2 have been considered as the main driver of global warming (Ainsworth et al., 2020). Transportation is a major contributor to the global CO_2 emission, representing 65% of the world oil consumption and 24% of global CO_2 emissions due to the direct combustion of fuels (Li et al., 2019; Solaymani, 2019). With the worldwide concerns about global

warming, biofuels have been embraced as promising alternatives to fossil fuels, because they are renewable and generally can lower carbon emissions (Demirbas, 2009; Gaurav et al., 2017).

Oleochemicals are a large group of fatty acid derivatives, including fatty acids, fatty alcohols, fatty alk (a/e)nes, and fatty acid methyl/ethyl esters and waxes (Pfleger et al., 2015). They can be used as biodiesels, lubricants, and surfactants, and others (Yu et al., 2014; Pfleger et al., 2015; Marella et al., 2018). Compared with ethanol, which is another popular biofuel molecule, lipid-derived biodiesels have been considered to be better biofuel molecules due to their high energy density and compatibility with the existing liquid fuel infrastructure (i.e., fuel engines, refinery equipment, and transportation pipelines) (Lu, 2010).

Traditionally, crop oils and animal fats (**Figure 1**) were used as feedstocks for the production of oleochemicals by chemical or enzymatic processes (Pfleger et al., 2015). However, this traditional route for oleochemical production will compete with crops for arable land, decrease food production, and raise serious concerns about food security (Graham-Rowe, 2011). Microalgae have been considered as promising feedstocks for oleochemicals (**Figure 1**) because of their higher lipid productivities per ground area than oleaginous agricultural crops, as well as the lack of competition they would provide for agricultural land (Mata et al., 2010; Wijffels and Barbosa, 2010). Besides, abundant lignocellulosic biomass has become another ideal feedstock for the production of oleochemicals (**Figure 1**), in the context of large-scale metabolic engineering efforts in microbial systems (Lee et al., 2008; Alper and Stephanopoulos, 2009; Peralta-Yahya and Keasling, 2010; Keasling, 2012). Some heterotrophic model microorganisms, such as *Escherichia coli* and *Saccharomyces cerevisiae*, have been genetically modified to produce many kinds of biofuels and bio-chemicals including oleochemicals from lignocellulosic sugars (Atsumi et al., 2008; Steen et al., 2010; Buijs et al., 2013). Some recent review articles have summarized biosynthesis pathways, metabolic engineering strategies, and challenges for the production of oleochemicals by heterotrophic microbes (Janssen and Steinbuchel, 2014; Pfleger et al., 2015; Marella et al., 2018).

Cyanobacteria are the only prokaryotes capable of performing oxygen-evolving photosynthesis (Hagemann and Hess, 2018), and have been the genetic models for photosynthesis research for decades (Wang et al., 2018). They were initially not considered to be useful to the Aquatic Species Program for biofuel production, because most of them do not naturally accumulate storage lipids in the form of triacylglycerol (TAG) as some oleaginous eukaryotic microalgae do (Sheehan et al., 1998). However, cyanobacteria have emerged as novel chassis strains for the production of biofuels and bio-chemicals since 2009, owing to their photosynthetic abilities and reliable genetic systems (Angermayr et al., 2009; Atsumi et al., 2009; Dexter and Fu, 2009; Lindberg et al., 2009). Engineered cyanobacteria are able to produce various compounds directly from CO_2, bypassing the need for fermentable sugars and arable land (Lai and Lan, 2015). In the past decade, photosynthetic production of various compounds, including

oleochemicals (**Figure 1**), has been achieved in several model cyanobacteria through metabolic engineering (Zhou and Li, 2010; Angermayr et al., 2015; Savakis and Hellingwerf, 2015; Oliver et al., 2016; Xiong et al., 2017). This review summarizes current knowledge on the metabolism of fatty acids and membrane lipids in cyanobacteria, provides the current status of metabolic engineering strategies for producing oleochemicals, and discusses key challenges and possible solutions in the field.

METABOLISMS OF FATTY ACIDS AND MEMBRANE LIPIDS IN CYANOBACTERIA

The biosynthesis of membrane lipids in cyanobacteria has been investigated since the 1980s (Naoki and Norio, 1982) and was followed by systematical works by Murata and co-workers in the 1990s (Wada and Murata, 1990; Wada and Murata, 1998) and 2000s (Sato and Wada, 2009). Unlike heterotrophic prokaryotes, the vast majority of cyanobacteria have thylakoid membranes in their cytoplasm where photosynthesis takes place (Rexroth et al., 2011). Both cytoplasmic (plasma) and thylakoid membranes of cyanobacteria include four major polar glycerolipids: monogalactosyl diacylglycerol (MGDG), digalactosyl diacylglycerol (DGDG), sulfoquinovosyldiacylglycerol (SQDG), and phosphatidylglycerol (PG) (Los and Mironov, 2015). Despite a report indicating the occurrence of neutral lipid droplets including triacylglycerol (TAG) in the cyanobacterium *Nostoc punctiforme* PCC73102 (hereafter Npu73102) (Peramuna and Summers, 2014), it is noteworthy that the above four polar lipids still serve as the dominant sink for fatty acids in cyanobacteria.

Fatty Acid Biosynthesis Pathway

Same as the widely studied fatty acid biosynthesis pathway in *E. coli*, cyanobacterial fatty acid biosynthesis pathways are composed of reactions catalyzed by two protein complexes, namely, acetyl-CoA carboxylase (ACCase) and type II fatty acid synthase (FAS) encoded by *fab* genes (Sato and Wada, 2009). In brief, acetyl-CoA is firstly converted to malonyl-CoA by acetyl-CoA carboxylase, and then to malonyl-ACP by malonyl-CoA:ACP transacylase (FabD) (**Figure 2**). Subsequently, butyryl-ACP is generated by sequential reactions catalyzed by β-ketoacyl-ACP synthase III (FabH), β-ketoacyl-ACP reductase (FabG), β-hydroxyacyl-ACP dehydrase (FabZ), and enoyl-ACP reductase (FabI). The fatty acid chain is then elongated with an acetyl unit from malonyl-ACP for each condensation-reduction-dehydration-reduction cycle (**Figure 2**) (Sato and Wada, 2009). For most cyanobacteria, palmitoyl-ACP (C16) and stearoyl-ACP (C18) are used as precursors for the biosynthesis of membrane lipids. Contrary to previous findings in *E. coli* (Yu et al., 2011), it was proved that FabH, which condenses malonyl-ACP with acetyl-CoA to form acetoacetyl-ACP, is the sole rate-limiting enzyme of FAS in *Synechococcus* sp. PCC7002 (hereafter Syn7002) (Kuo and Khosla, 2014).

FIGURE 1 | Overview of both traditional and emerging technologies for the bio-production of oleochemicals. The green lines show enzymatic or biological conversion, whereas physical or chemical processes are in black. Vegetable oil seeds are traditionally utilized to produce oleochemicals. Oleaginous eukaryotic microalgae are attractive alternatives to plant oil. Abundant cellulosic biomass is first hydrolyzed to sugars, and the latter is then fermented to produce oleochemicals by the engineered heterotrophic microbes (indicated by the asterisk). Harboring photosynthesis ability, cyanobacteria can be genetically modified to direct convert CO_2 to oleochemicals.

Metabolisms of Membrane Lipids

For the biosynthesis of the four polar glycerolipids mentioned above, phosphatidic acid (PA) is synthesized as the common precursor by the acylation of both sn-1 and 2 positions of glycerol-3-phosphate (G3P) with the long-chain fatty acyl-A (C16 or C18) by different acyltransferases (Sato and Wada, 2009). Specifically, the fatty acyl-ACPs are first activated by an inorganic phosphate group by phosphate acyltransferase (PlsX), and subsequently transferred to the sn-1 position of G3P by acylglycerol-phosphate acyltransferase (AGPAT or PlsY), resulting in lysophosphatidic acid (LPA) (**Figure 2**). Secondly, lysophosphatidic acid acyltransferase (LPAAT or PlsC) catalyzes the transfer of fatty acid chains to the sn-2 position of LPA in the PA biosynthesis, resulting in PA (**Figure 2**). Although the over-expression of the putative PlsX enhanced lipid production in *Synechocystis* sp. PCC 6803 (hereafter Syn6803) (Towijit et al., 2018), the detailed enzymatic characteristics of both PlsX and PlsY are still unknown. Sll1848 was identified as the primary LPAAT with a high specificity for 16:0-ACP (Weier et al., 2005), whereas Sll1752 was characterized as the secondary LPAAT that prefers stearoyl and oleoyl substrates (C18) in Syn6803 (Okazaki et al., 2006) (**Figure 2**).

As summarized by previous reviews (Wada and Murata, 1998; Sato and Wada, 2009), different polar head groups are further transferred to the sn-3 position of PA to synthesize four major polar glycerolipids in cyanobacteria. Finally, MGDG, DGDG, SQDG, and PG

have a head group of 1β-galactose, digalactose, 6-deoxy-6-sulfo-α1-glucose, and sn-glycerol 1-phosphate at their sn-3 position of the glycerol moiety, respectively, besides two acyl groups esterified at the sn-1 and sn -2 positions (Sato and Wada, 2009).

Desaturation of Membrane Lipids

As reviewed previously, cyanobacterial desaturases were classified as acyl-lipid desaturases rather than acyl-CoA or acyl-ACP desaturases, which means the fatty acid chain would be desaturated only when fatty acids are bound to membrane lipids (Murata and Wada, 1995; Sato and Wada, 2009; Los and Mironov, 2015). In response to the cold stress, the fatty acid chains can be stepwise desaturated at the Δ9, Δ12, ω3, and Δ6 positions by four specific desaturases, namely, DesC, DesA, DesB, and DesD, respectively. The fatty acid chain length of cyanobacteria varies from C14 to C18, whereas the number of double bonds in the fatty acid chains may vary from 0 to 4, which is controlled by the activities of the above desaturases (Los and Mironov, 2015). The fatty acid composition determined by the chain length and the numbers of double bonds can be used for the classification of cyanobacterial strains (Wada and Murata, 1998).

Regulation of Fatty Acid and Lipid Metabolisms in Cyanobacteria

As the physical barrier of cells and sites of photosynthesis and respiration, the cytoplasmic and thylakoid membranes of

FIGURE 2 | Schematic overview of native and synthetic pathways for biosynthesis of membrane lipids and oleochemicals in cyanobacteria. The illustration shows the native biosynthesis pathways for membrane lipids, together with synthetic pathways for cyanobacterial oleochemical production. For synthetic pathways, enzymes and reaction direction are shown in green. For native cyanobacterial pathways, enzymes and reaction direction are in black. Oleochemicals discussed in this work are shown in green texts and in gray boxes.

cyanobacteria are sensitive to various environmental stimuli. On the one hand, glycerolipid and fatty acid compositions were observed to change with alterations of growth temperature, light illumination intensity, carbon dioxide, and pH (Wada and Murata, 1990; Sakamoto and Bryant, 2002; Cuellar-Bermudez et al., 2015). On the other hand, membrane lipids play an active role in the acclimation of cyanobacteria to different environmental conditions, including high temperature (Nanjo et al., 2010) and low temperature (Murata and Wada, 1995; Sakamoto and Bryant, 2002). Despite these observations on the physiological roles of lipids, little is known about the regulation of fatty acid and lipid metabolisms in cyanobacteria.

Hik33 has been identified as the sensory histidine kinase of a two-component system which perceives the low-temperature signal and controls the expression of *desB* gene (Suzuki et al., 2000; Mikami et al., 2002; Murata and Los, 2006). Additionally, the global regulator LexA was found to repress the expression of some *fab* genes (Kizawa et al., 2017), while a transcriptional regulator, CyAbrB2, was found to inhibit the FFA production in Syn6803 (Kawahara et al., 2016). The global nitrogen regulator PII protein was shown to negatively regulate cyanobacterial fatty acid biosynthesis by transcriptional control (Verma et al., 2018) or by interacting with biotin carboxyl carrier protein (BCCP) which is a subunit of ACCase (Hauf et al., 2016).

ENGINEERING CYANOBACTERIA TO PRODUCE FREE FATTY ACIDS

A decade ago, it was shown that the elimination of fatty acid β-oxidation by disrupting the *fadD* or *fadE* gene (**Figure 2**), over-expression of a thioesterase (TE) gene to release FFAs, and over-expression of ACCase have been demonstrated to be effective approaches for the overproduction of free fatty acids (FFAs) in some heterotrophic microbes like *E. coli* (Lu et al., 2008). Similar strategies were also adopted soon afterward for cyanobacterial FFA production (**Table 1**).

Disruption of Fatty Acid Re-activation

There is no complete fatty acid β-oxidation pathway for FFA degradation in cyanobacteria, based on bioinformatics analysis. However, acyl–acyl carrier protein synthetases (Aas) were proven to be able to re-activate FFAs to acyl-ACPs (**Figure 2**), and the latter can be incorporated into the membrane lipids through the above-mentioned acyltransferases in cyanobacteria (Kaczmarzyk and Fulda, 2010). The disruption of Aas led to the FFA accumulation and secretion in both Syn6803 and *Synechococcus elongatus* PCC7942 (hereafter Syn7942) (Kaczmarzyk and Fulda, 2010). This strategy has been adopted by most research efforts on cyanobacterial FFA production (Liu et al., 2011a; Ruffing and Jones, 2012; Ruffing, 2014) (**Table 1**).

TABLE 1 | A summary of oleochemical production by engineered cyanobacteria.

No.	Hosts	Products	Genetic modifications		Culture optimization	Titer (mg/L)	Yield[c] (mg/g DCW)	Productivity (mg/L/h)	References
			Over-expression	**Deletion**					
1	Syn6803	FFA	'TesA, UcfatB1 from *U. californica*, ChfatB2 from *C. hookeriana*, CcfatB1 from *C. camphorum* and TesA137, AccBCDA	Aas, PhaAB, CphAB, Pta, S-layer protein and PBP2	1% CO_2, 140 $\mu E/m^2/s$	211.2	167.2[a]	ND	Liu et al., 2011a
2	Syn6803	FFA	FatB from *A. thaliana*	Aas	Air	95.1	24.5	ND	Hu et al., 2013
3	Syn6803	FFA	Membrane-located expression of 'AcTesA from *A. baylyi*		1% CO_2, 50 $\mu E/m^2/s$	331.0[s]	199.2[s]	1.97[s]	Afrin et al., 2018
4	Syn6803	FFA	AhFatA, AhFatB from *A. hypogaea*		Air, 40 $\mu E/m^2/s$	ND	ND	ND	Chen et al., 2017
5	Syn6803	FFA	'TesA	Aas	1% CO_2, 60 $\mu E/m^2/s$	ND	209.0	ND	Yunus et al., 2018
6	Syn6803	FFA	Tes3 from *Anaerococcus tetradius*	Aas	1% CO_2, 60 $\mu E/m^2/s$	97.1	ND	ND	Yunus and Jones, 2018
7	Syn7942	FFA	Fat1 from *C. reinhardtii*, AccBCDA from *C. reinhardtii*, RbcLS	Aas	1% CO_2, 60 $\mu E/m^2/s$	23.4	155.3	0.05	Ruffing, 2013a
8	Syn7942	FFA	'TesA, RndA1B1	Aas	2% CO_2, 180 $\mu E/m^2/s$, overlaid with isopropyl myristate	640[s]	360.0[s]	1.48[s]	Kato et al., 2017
9	Syn7002	FFA	'TesA, RbcLS from Syn7942	Aas	1% CO_2	131.5[s]	70.0[s]	0.27	Ruffing, 2014
10	Syn7002	FFA	UcfatB1	Aas, GlgC	1% CO_2, 160 $\mu E/m^2/s$	ND	ND	ND	Work et al., 2015
11	Syn7002	FFA	UcfatB1, FabH from *Chaetoceros* GSL56	Aas, FabH	1% CO_2	ND	ND	ND	Gu et al., 2016
12	Syn6803	Alk(a/e)nes	Two copies of Ado-Aar		5% CO_2, 100 $\mu E/m^2/s$	26.0	11.0	0.11	Wang et al., 2013
13	Syn6803	Alk(a/e)nes	Ado-Aar		Air	ND	1.9	ND	Hu et al., 2013
14	Syn6803	Alk(a/e)nes	'TesA, 'FAP from *C. variabilis*	Aas	1% CO_2, 300 $\mu E/m^2/s$	111.2	77.1	0.46	Yunus et al., 2018
15	Syn7002	Alk(a/e)nes	Ado and Aar from Syn7942		5% CO_2, 300 $\mu E/m^2/s$	ND	7.5	ND	Knoot and Pakrasi, 2019
16	Ana7120	Alk(a/e)nes	Ado, Aar from *Aphanothece halophytica*		40 $\mu E/m^2/s$, 140 mM NaCl	ND	1.3	ND	Kageyama et al., 2015
17	Npu73102	Alk(a/e)nes	Ado, Aar, Npun_F5141 (Lipase)		25°C, Air, 135~160 $\mu E/m^2/s$, MA medium	ND	129	ND	Peramuna et al., 2015
18	Syn6803	Fatty alcohols	Far from jojoba		5% CO_2, 100 $\mu E/m^2/s$	0.2	0.1[a]	0.00046	Tan et al., 2011
19	Syn6803	Fatty alcohols	Two copies of Far from jojoba, At3g11980 from *A. thaliana*		Air, 30~50 $\mu E/m^2/s$	ND	0.8	ND	Qi et al., 2013
20	Syn6803	Fatty alcohols	Maqu_2220 from *Marinobacter aquaeolei* VT8	Aar, Ado	Air, 50 $\mu E/m^2/s$	1.3	2.9	0.0032	Yao et al., 2014

(Continued)

TABLE 1 | Continued

No.	Hosts	Products	Genetic modifications		Culture optimization	Titer (mg/L)	Yield[c] (mg/g DCW)	Productivity (mg/L/h)	References
			Over-expression	Deletion					
21	Syn6803	Fatty alcohols	Maqu_2220 from *M. aquaeolei* VT8	PlsX (Slr1510) transcriptionally inhibited by CRISPRi	1% CO_2, 60 µE/m²/s	ND	10.4	ND	Kaczmarzyk et al., 2018
22	Syn6803	Fatty alcohols	'TesA, Sfp, CAR from *M. marinum*	Aas	1% CO_2, 60 µE/m²/s	ND	68.0	ND	Yunus et al., 2018
23	Syn6803	Fatty alcohols	Tes3 from *A. tetradius*, Sfp, CAR from *M. marinum*	Aas	1% CO_2, 60 µE/m²/s, overlaid with isopropyl myristate	100.0	80.0	0.42	Yunus and Jones, 2018
24	Syn6803	Fatty alcohols	Sfp, CAR from *M. marinum*	Aas	1% CO_2, 60 µE/m²/s, overlaid with isopropyl myristate, with octanoic acid feeding	905.7[b]	ND	4.71	Yunus and Jones, 2018
25	Syn7942	FAEEs	AtfA from *A. baylyi* ADP1, Pdc and Adh from *Zymomonas mobilis*, XpkA from *A. nidulans*, Pta from *B. subtilis*		5% CO_2, 100 µE/m²/s, 20% hexadecane overlay	15.11	50.0	0.06	Lee et al., 2017
26	Syn7942	Wax	AtfA, Aar, Slr1192 from Syn6803 or ACIAD3612 from *A. baylyi* ADP1		1% CO_2, 50 µE/m²/s	ND	ND	ND	Kaiser et al., 2013

aThe number was converted using a calibration between DCW and optical density (1.0 OD_{750} equals approximately 200 mg DCW/L) (Yao et al., 2014); boctanoic acid titer; cyield relative to biomass; sfor oleochemicals secreted out of cells, otherwise, for the total oleochemicals; ND, not determined.

Heterologous Metabolic Pathway Engineering Toward FFAs and Chain Length Control

Hydrolysis of acyl-ACP to FFA by thioesterases can release the feedback inhibition of acyl-ACP to some enzymes of FAS-II and has in turn been confirmed to be an effective strategy to enhance FFA production in *E. coli* (Heath and Rock, 1996a,b; Lu et al., 2008). However, there is no gene encoding a thioesterase in cyanobacteria. Engineering efforts for improving FFA production in cyanobacteria began nearly a decade ago (Liu et al., 2011a). In this work, a truncated *E. coli* TE gene *'tesA* and three plant TE genes were heterologously expressed in the *aas* mutant of Syn6803 to achieve the production and secretion of FFAs (Liu et al., 2011a). Other thioesterases from *Arabidopsis thaliana* (FatB) (Hu et al., 2013), *Chlamydomonas reinhardtii* (Fat1) (Ruffing, 2013a), *Acinetobacter baylyi* ('AcTesA) (Afrin et al., 2018), and *Arachis hypogaea* L. (AhFatA, AhFatB) (Chen et al., 2017) were also functionally expressed in cyanobacteria for FFA production (**Table 1**), yielding long-chain (C16–C18) FFAs in most cases.

Medium-chain fatty acids (MCFAs, C4–C12) are valuable precursors to gasoline, but are not typical products of microbial fatty acid synthesis (Torella et al., 2013). Different from the above-mentioned thioesterases, thioesterases from *Cinnamomum camphorum* (CcFatB1), *Umbellularia californica* (UcFatB1), *Cuphea hookeriana* (ChFatB1), and *Anaerococcus tetradius* (Tes3) prefer medium chain length acyl-ACP substrates. When producing MCFAs, they were expressed in cyanobacteria to control the chain lengths of the FFA products (Murata and Wada, 1995; Work et al., 2015; Yunus and Jones, 2018). It is noteworthy that these short or medium chain length specific thioesterases should be expressed in the *aas* mutant to avoid the reactivation and the elongation of FFAs (**Table 1**). In addition, the replacement of the native FabH with a *Chaetoceros* ketoacyl-ACP synthase III in the lauric acid-secreting strain of Syn7002 increased MCFA synthesis up to five-fold (Gu et al., 2016).

Translocation of FFA Out of the Cells

Besides the activity of acyl–acyl carrier protein synthetase, cyanobacterial Aas was also identified as a FFA importer (von Berlepsch et al., 2012) (**Figure 2**). And the inactivation of Aas resulted in the FFA secretion in some cyanobacterial strains, such as Syn6803, Syn7942 (Kaczmarzyk and Fulda, 2010), and Syn7002 (Ruffing, 2014), indicating that FFAs can be exported out of cyanobacterial cells by active or passive transport. It was proved that weakening cell walls by the deletion of the possible surface protein (Sll1951) and the peptidoglycan assembly protein (PBP2) as well as by ampicillin treatment led to the decrease of intracellular FFA amounts and the increase of overall FFA production in Syn6803 (Liu et al., 2011a).

Moreover, a RND-type FFA efflux system (RndA1B1) was identified by genomic analysis of a spontaneous mutant of the FFA-producing strain of Syn6803 (Kato et al., 2015) (**Figure 2**). Furthermore, the highest FFA yield (0.36 g/g dry cell weight) up to now has been achieved in the RndA1B over-expressing strain of Syn7942 through *in situ* removal of the FFA product

from the culture medium by an isopropyl myristate (IM) overlay (Kato et al., 2017) (**Figure 3**). Recently, Sll0180 and Slr2131, homologs to AcrA and AcrB of *E. coli* respectively, were identified to be another FFA efflux system (Bellefleur et al., 2019) (**Figure 2**). Replacing the native *slr2131* with the *E. coli acrB* gene significantly increased the extracellular FFA concentration of Syn6803 (Bellefleur et al., 2019).

Native Biosynthesis Pathway for Free Fatty Acids in Cyanobacteria

In fact, little FFAs was found in the cells of cyanobacteria grown under the normal culture condition (Kaczmarzyk and Fulda, 2010). Isotope labeling experiments indicated that they are released from membrane lipids (Kaczmarzyk and Fulda, 2010). Lipases were considered to be responsible for the releasing of FFAs from membrane lipids, and heterologous expression of the foreign lipase resulted in the increase of FFAs in Syn6803 (Liu et al., 2011b). *sll1969* is the only candidate lipase gene in the genome of Syn6803. The deletion of this gene decreased the FFA production, but did not completely block the FFA production (Gao et al., 2012a), suggesting it is not the only pathway for endogenous FFAs biosynthesis in Syn6803.

Besides the lipase, two cyanobacterial aldehyde dehydrogenases (AldE), namely Synpcc7942_0489 from Syn7942 (Kaiser et al., 2013) and Slr0091 from Syn6803 (Trautmann et al., 2013), were proven sufficient to oxidize fatty aldehyde precursors into fatty acids (**Figure 2**). It is noteworthy that these two aldehyde dehydrogenases are also able to utilize aldehyde substrates with shorter chain lengths (C8 to 12) (Kaiser et al., 2013) or apocarotenals (Trautmann et al., 2013), besides long-chain fatty aldehydes. For the purpose of FFA over-production in cyanobacteria, overexpression of acyl-ACP reductase (Aar) in the presence of AldE was proven to be a successful strategy (Kaiser et al., 2013) (**Table 1**).

Over-Production of Polyunsaturated Fatty Acids by Introduction of Heterologous Desaturases

Syn7942 encodes only one Δ9 dedaturase gene (*desC*) in its genome and has only saturated and monounsaturated (Δ9) fatty acid chains in its membranes. Heterologous expression of the Δ12 desaturase gene (*desA*) from Syn6803 led to the conversion of endogenous monounsaturated fatty acids into dienoic fatty acids (Δ9, 12) and in turn changed the fatty acid compositions of Syn7942 (Wada et al., 1990). This modification further enhanced host tolerance to chilling (Wada et al., 1990) and strong light illumination (Gombos et al., 1997). In a recent study, the heterogeneous expression of two desaturases (DesA and DesB) from Syn7002 conferred an ability of producing alpha-linolenic acid (ALA; Δ9, 12, 15) to Syn7942 (Santos-Merino et al., 2018). Further, the ALA content of the desaturases-expressing mutant was improved to levels as high as 22.6% of the total lipids, by two metabolic engineering approaches designated as the *fabF* overexpression and the *fadD* disruption (Santos-Merino et al., 2018).

METABOLISM OF FATTY ALK(A/E)NES IN CYANOBACTERIA

Since 1960s, cyanobacteria were known to be able to naturally produce linear and branched fatty alkanes and alkenes with carbon chain lengths ranging from 15 to 19, besides membrane lipids (Han et al., 1968; Winters et al., 1969). However, the cyanobacterial biosynthesis pathways of fatty alk(a/e)nes were not identified until 2010 (Schirmer et al., 2010).

Fatty Alk(a/e)nes Biosynthesis Pathways in Cyanobacteria

A two-enzyme pathway, consisting of acyl-ACP reductase (Aar) and aldehyde-deformylating oxygenase (Ado), was first identified by both comparative genomic and enzymatic analysis (Schirmer et al., 2010). In the Aar-Ado pathway, Aar catalyzes the conversion of acyl-ACP to fatty aldehydes, and Ado oxidizes and deformylates aldehydes to alk(a/e)nes (**Figure 2**), including pentadecane, heptadecane, 8-heptadecene, or 7-methylheptadecane.

A year later, the second native cyanobacterial alkene biosynthesis pathway (olefin synthase, Ols) was characterized in Syn7002 (Mendez-Perez et al., 2011). Harnessing a modular type I polyketide synthase (Ols), fatty acyl-ACP precursors are elongated and decarboxylated to synthesize terminal 1-alkenes (**Figure 2**), including 1-heptadecene, 1-non-adecene, or 1,14-nanadecadiene. It is worth noting that all alk(a/e)nes-producing cyanobacteria harbor only one of the two above-mentioned pathways but never both in nature (Coates et al., 2014; Klahn et al., 2014). However, it was proved that both Ols and Aar-Ado pathways can co-exist in one engineered marine cyanobacterium (Yoshino et al., 2015; Knoot and Pakrasi, 2019). And the Aar-Ado pathway, as well as two non-cyanobacterial alkane biosynthesis genes, can complement an Ols knockout strain of Syn7002 (Knoot and Pakrasi, 2019).

Physiological Effects of Alk(a/e)nes in Cyanobacteria

Despite the fact that alkane biosynthesis pathways were characterized, little is known about the physiological roles of alk(a/e)nes in cyanobacteria. Alkanes were shown to accumulate in thylakoid and cytoplasmic membranes of Syn6803 (Lea-Smith et al., 2016) as well as in lipid droplets of Npu73102 (Peramuna and Summers, 2014; Peramuna et al., 2015). Through reverse genetic approaches, it was found that cyanobacterial alkanes might play roles in regulating redox balance and reductant partitioning in photosynthesis (Berla et al., 2015), and modulating membrane flexibility, which is required for optimal cell division, size, and growth (Lea-Smith et al., 2016). In addition, alkanes were proven to be required for cyanobacterial tolerance to abiotic stresses including cold (Berla et al., 2015) and salt (Yamamori et al., 2018).
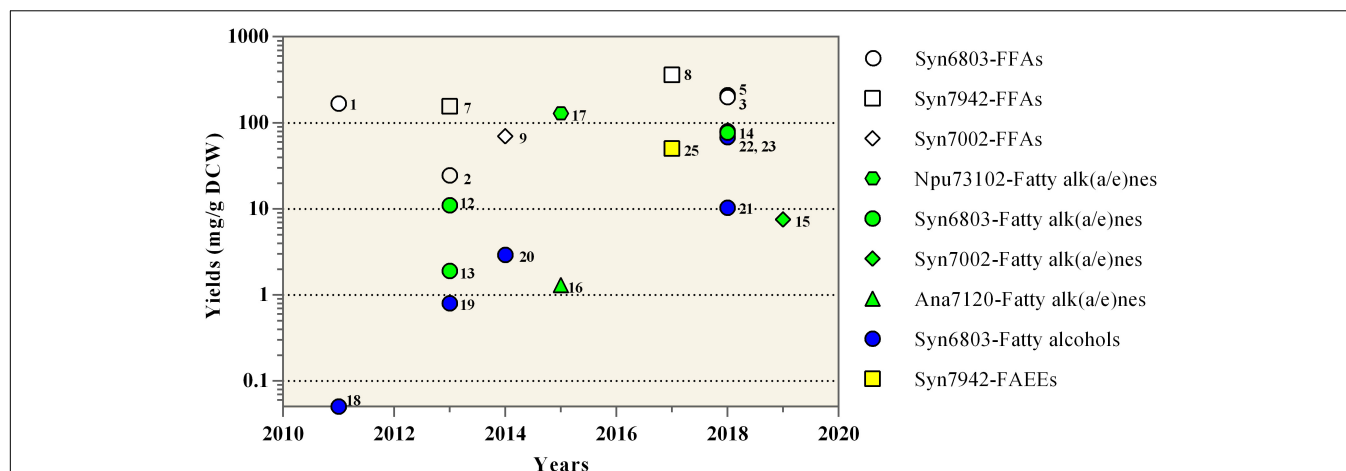
FIGURE 3 | Yields of oleochemicals reported from engineered cyanobacteria. Maximum oleochemical yields in each study are shown at the years in which the study has been published. The detailed information for each data point including the reference is listed in **Table 1** and can be retrieved using the adjacent number. The reported works on Syn6803, Syn7942, Syn7002, Npu73102, and Ana7120 are shown as circles, squares, diamonds, hexagons, and triangles, respectively. The work on the production of FFAs, Fatty alk(a/e)nes, Fatty alcohols, and FAEEs are shown as white, green, blue, and yellow colors, respectively.

ENGINEERING CYANOBACTERIA TO PRODUCE FATTY Alk(a/e)nes

With the carbon chain lengths ranging from C15 to 19, cyanobacterial alk(a/e)nes could be directly used in diesel and jet engines, and have attracted great attention from academics. On the one hand, chemical structures and profiles of cyanobacterial alk(a/e)nes were examined across a wide range of cyanobacterial species (Liu et al., 2013; Coates et al., 2014; Zhu et al., 2018). On the other hand, several model cyanobacterial species were metabolically engineered for improving their alk(a/e)nes production (Mendez-Perez et al., 2011; Hu et al., 2013; Wang et al., 2013; Kageyama et al., 2015; Peramuna et al., 2015) (**Table 1**). Among these engineering approaches, the over-expression of endogenous or heterogeneous alk(a/e)ne biosynthesis genes was widely used and proven to be successful (Xie et al., 2017). In addition, the cyanobacterial alk(a/e)ne production can be further improved by increasing the copy numbers of these genes through inserting them into different genomic loci (Wang et al., 2013).

As expected, the over-expression of the multi-subunit acetyl-CoA carboxylase, which catalyzes the first step of fatty acid biosynthesis, was proven to be an effective way to enhance the cyanobacterial alkane production (Tan et al., 2011; Wang et al., 2013). The cyanobacterial alkane production can also be improved by blocking the competing pathway, like the poly-β-hydroxybutyrate (PHB) pathway (Wang et al., 2013). Different from cyanobacterial FFA production, Aas is beneficial for cyanobacterial alkane production in the Aar-Ado pathway (Gao et al., 2012b). The over-expression of Aas promoted cyanobacterial alkane production, because the acyl-ACP precursors of the Aar-Ado pathway are mainly from the Aas-mediated reactivation of FFAs, which are from the hydrolysis of membrane lipids rather than the *de novo* fatty acid biosynthesis pathway (Gao et al., 2012b). Similarly,

the over-expression of lipolytic enzymes, which release the FFA by hydrolyzing membrane lipids, was also beneficial for alkane production (Wang et al., 2013; Peramuna et al., 2015). For example, the heptadecane production in Npu73102 was significantly improved by over-expression of Aar, Ado, and a lipase candidate Npun_F5141, together with the high light illumination, reaching 12.9% of dry cell weight (DCW) (Peramuna et al., 2015) (**Figure 3**).

In addition to the above engineering approaches on cyanobacterial native alkane biosynthesis pathways, several synthetic metabolic pathways were recently constructed and evaluated for alkane production in cyanobacteria (Yunus et al., 2018; Knoot and Pakrasi, 2019). In brief, three newly identified fatty acid decarboxylases were heterologously expressed in cyanobacteria successfully for converting FFA precursors to C_{n-1} alk(a/e)ne end-products (**Figure 2**), including UndA (Rui et al., 2014) and UndB (Rui et al., 2015) from *Pseudomonas fluorescens* Pf-5, together with fatty acid photodecarboxylase (FAP) from *Chlorella variabilis* (Sorigue et al., 2017). Although catalyzing similar reactions with another two 1-alkenes producing enzymes, namely Ols from *Synechococcus* (Mendez-Perez et al., 2011) and OleTJE from *Jeotgalicoccus* (Rude et al., 2011), both UndA and UndB prefer FFA substrates with medium-chain lengths (C10–C16) rather than long-chain substrates (Rui et al., 2014, 2015). Recently, a phenylalanine 239 to alanine mutation of UndA (UndA-F239A) increased its enzymatic activities toward long chain fatty acids and improved its compatibility with cyanobacterial fatty acid compositions (Knoot and Pakrasi, 2019). FAP from eukaryotic algae mediates the light-driven conversion of fatty acid substrates to alkanes, with a wide range of substrate chain lengths (C12 to C18) and a higher substrate specificity to hexadecanoic acid (Sorigue et al., 2017). When expressing in the *aas* mutant of Syn6803 harboring a truncated *E. coli* thioesterase 'TesA, FAP can markedly improve alkane production (Yunus et al., 2018). And the total alkane yield was further increased to

77.1 mg/g DCW through the removal of the chloroplast transit peptide of FAP ('FAP) and the increase of light illumination (Yunus et al., 2018) (**Table 1** and **Figure 3**).

ENGINEERING CYANOBACTERIA TO PRODUCE FATTY ALCOHOLS

Fatty alcohols can be used in the manufacture of cosmetics, detergents, lubricants, and potentially as biofuels (d'Espaux et al., 2017). Similar to several reports on microbial production of fatty alcohols (Steen et al., 2010; d'Espaux et al., 2017), cyanobacterial fatty alcohol production (**Table 1**) was mainly realized by heterologous expression of fatty acid reductases (Far) which utilize fatty acyl-ACP or acyl-CoA as substrates and NADH or NADPH as cofactors (Tan et al., 2011; Yao et al., 2014; Kaczmarzyk et al., 2018). All the Fars that worked well in cyanobacteria are from plants (Tan et al., 2011) and bacterium (Yao et al., 2014), whereas cyanobacteria expressing the Fars from mice failed to produce any fatty alcohols (Tan et al., 2011). The engineered strain of Syn6803 harboring the Far from jojoba (*Simmondsia chinensis*) produced only 0.05 mg fatty alcohols per gram DCW (Tan et al., 2011). Then, the cyanobacterial fatty alcohol yield was dramatically improved to 0.76 mg/g DCW by increasing the copy numbers of the plant Far-expressing cassettes and blocking both the PHB and glycogen biosynthesis pathways (Qi et al., 2013) (**Table 1**).

Compared to the plant Fars, a bacterial Far (Maqu_2220 from *Marinobacter aquaeolei* VT8) showed better substrate preferences to long-chain fatty acyl CoA/ACP (C16–C18) (Hofvander et al., 2011) and better performance in the engineered strain of Syn6803 for fatty alcohol production (Yao et al., 2014). For further improving fatty alcohol production of this Maqu_2220-expressing strain, the inactivation of the fatty alkane biosynthesis pathway which competes with Far for acyl-ACP precursors was shown to be effective, resulting in 2.9 mg/g DCW fatty alcohols (Yao et al., 2014). Recently, phosphate acyltransferase PlsX was identified as another key node in C18 fatty acyl-ACP consumption, and the fatty alcohol yield of Syn6803 was increased to 10.4 mg/g DCW by transcriptional inhibition of *plsX* using CRISPR-interference (CRISPRi) technique (Kaczmarzyk et al., 2018) (**Table 1**).

Besides Fars, carboxylic acid reductase (CAR) from *Mycobacterium marinum,* which can effectively convert FFAs into corresponding fatty aldehydes, was also used for cyanobacterial fatty alcohol production (Yunus and Jones, 2018; Yunus et al., 2018). In the presence of Ado, heterologous expression of CAR in the FFA-producing strain of *Syn*6803 unexpectedly led to the conversion of most of the FFA pool into corresponding fatty alcohols rather than fatty alk(a/e)nes, resulting in ~68 mg/g DCW fatty alcohols (Yunus et al., 2018). It was speculated that the Ado failed in competition with native aldehyde reductases (or alcohol dehydrogenases, Adh) (Yunus et al., 2018). For producing medium chain-length fatty alcohols, Tes3 from *A. tetradius*, rather than *E. coli* 'TesA was co-expressed in the *aas* mutant of Syn6803 together with CAR and its maturation protein Sfp. Through the optimization of promoters

and ribosomal binding sites and *in situ* product extraction with isopropyl myristate, the titers of 1-octanol and 1-decanol of the above mutant were increased to more than 100 mg/L, which is the highest titer of cyanobacterial fatty alcohols to date (Yunus and Jones, 2018) (**Table 1**). However, it should be noted that this 1-octanol producing strain displayed genetic instability and reduced 1-octanol production during continuous sub-culturing (Yunus and Jones, 2018).

ENGINEERING CYANOBACTERIA TO PRODUCE FATTY ACID ESTERS

Industrially, fatty acid esters are produced by transesterification of vegetable oils or animal fats with an alcohol in the presence of a suitable catalyst. For microbial production of fatty acid esters, the transesterification process is performed enzymatically by a multi-functional wax ester synthase/acyl-CoA:diacylglycerol acetyl transferase (WS/DGAT) (Janssen and Steinbuchel, 2014). Due to its wide substrate specificities to alcohols with various carbon lengths, the WS/DGAT (AtfA) from *Acinetobacter baylyi* ADP1 was normally used in cyanobacteria to mediate the combination of the activated fatty acids with ethanol (Lee et al., 2017) or fatty alcohols (Kaiser et al., 2013) (**Table 1**). To achieve wax ester production in Syn7942, AtfA was co-expressed with the Aar as well as a long-chain alcohol dehydrogenase from Syn6803 (Adh, Slr1192) or *A. bayli* (ACIAD3612) (Kaiser et al., 2013). Given the fact that FFA accumulates in the wax-producing strain as the byproduct (Kaiser et al., 2013), the endogenous *aldE* gene might be a candidate target for improving wax production by metabolic engineering.

For fatty acid ethyl esters (FAEEs) production in Syn7942, an ethanol biosynthesis pathway was firstly constructed by introducing both pyruvate decarboxylase and alcohol dehydrogenase from *Zymomonas mobilis* (Lee et al., 2017). The further expression of the AtfA in the ethanol-producing strain resulted in 40% ethanol reduction and the detection of trace concentrations of palmitic acid ethyl ester (Lee et al., 2017). A synthetic phosphoketolase pathway containing a phosphoketolase from *A. nidulans* (XpkA) and a phosphotransacetylase from *B. subtilis* (Pta) was then introduced to increase the acetyl-CoA pool and the FAEE production (Lee et al., 2017). The FAEE production was finally increased further to 50.0 mg/g DCW by culture optimization (Lee et al., 2017) (**Figure 3**). However, the low substrate specificities of the WS/DGAT to ethanol (Stoveken et al., 2005) could be a bottleneck for the FAEE production, considering the appearance of ethanol byproducts.

REDIRECTING CARBON FLUX TO OLEOCHEMICAL BIOSYNTHESIS PATHWAYS

Increasing the substrate supply and blocking the competitive pathways are routine strategies for improving the production of target products. Different from heterotrophic microbes,

autotrophic cyanobacteria utilize CO_2 rather than sugars as carbon sources, using the Calvin-Bassham-Benson (CBB) pathway. Thus, the RuBisCO, which is the key enzyme in the CBB cycle, has been considered as an ideal target for improving cyanobacterial carbon fixation. Over-expression of the RuBisCO from Syn7942 in Syn7002 led to a more than three-fold increase in FFA production (Ruffing, 2014). However, the same strategy did not work to improve the FFA production of Syn7942 (Ruffing, 2013a).

For directing the fixed carbon flux to *de novo* fatty acid biosynthesis pathway, a heterologous phosphoketolase pathway, which was discovered to be efficient for increasing the supply of acetyl-CoA precursor, was introduced into the FAEE-producing strain of Syn7942, and greatly enhanced the FAEE production (Lee et al., 2017). In addition, over-expression of the ACCase has been considered to direct the carbon flux into the fatty acid pathway, and was confirmed effective for FFA over-production in *E. coli* (Lu et al., 2008). However, it did not always work for improving FFA production of cyanobacteria. For example, the over-expression of an ACCase from *C. reinhardtii* led to a 20% increase of the specific FFA production in the FFA-producing strain of Syn7942 (Ruffing, 2013a) and a 56% increase of total fatty alk(a/e)nes in Syn6803 (Tan et al., 2011) (**Table 1**), but there was no significant change of FFA yield in Syn6803 using a similar approach (Liu et al., 2011a).

As a major carbon sink, glycogen can account for more than 50% of the DCW in some cyanobacteria treated by the stressed conditions (Song et al., 2016), which indicates that glycogen biosynthesis competes with the fatty acid biosynthesis pathway for carbon flux. However, the complete disruption of the glycogen biosynthesis always resulted in an increase of cyanobacterial susceptibility to stress conditions (Luan et al., 2019), and the reconfiguration of electron flow in photosynthesis (Work et al., 2015). Thus, the deletion of *glgC*, the key gene for cyanobacterial glycogen biosynthesis, showed a slight increase of fatty alcohol production in Syn6803 (Qi et al., 2013) and no increase of lauric acid production in Syn7002 (Work et al., 2015).

IMPROVING CYANOBACTERIAL TOLERANCE TO OLEOCHEMICALS

As hydrophobic compounds, oleochemicals inevitably interact with cell membranes, which are sites of photosynthesis and respiration, and will result in a series of physiological effects, including reduced photosynthetic yields, chlorophyll-a degradation, changes in the cellular localization of the light-harvesting pigments (Ruffing and Jones, 2012), increased reactive oxygen species (ROS), cell membrane permeability (Ruffing, 2013b), and impaired cell growth (Kamarainen et al., 2012; Ruffing and Jones, 2012).

A dozen candidate genes were identified by comparative transcriptome analyses with potentials to mitigate FFA toxicity. The disruption of two porins and the overexpression of ROS-degrading proteins were confirmed to be effective in reducing the toxic effects of FFA production and recovering cell growth (Ruffing, 2013b). Furthermore, transporters specific to

oleochemicals are promising candidates to secrete oleochemicals out of cells. The inactivation of Aas, which also functions as a FFA uptake transporter, was found to be able to alleviate FFA toxicity of cyanobacteria (von Berlepsch et al., 2012). Recently, it was found that the over-expression of a RND-type FFA exporter (RndA1B1) or the native or foreign (AcrAB) efflux systems for FFAs enhanced FFA secretion and cell growth (Kato et al., 2015; Bellefleur et al., 2019). Besides transporters, *in situ* removal of oleochemicals from the culture medium by some organic solutes was demonstrated to be able to significantly increase cyanobacterial cell tolerance to fatty alcohol, FFA, or FAEE (Kato et al., 2017; Lee et al., 2017; Yunus and Jones, 2018).

DISCUSSION

In the past decade, cyanobacteria were successively engineered to produce oleochemicals directly from CO_2, inspired by successful strategies on oleochemical production by *E. coli*. Although these efforts proved the concept of engineering cyanobacteria for oleochemical production, the engineered strains are far from being used for commercial applications mainly due to their poor production ability. To improve cyanobacterial oleochemical production, more intensive efforts are needed in the future.

As shown in **Figure 2**, all fatty acyl chains of oleochemicals are ultimately from the *de novo* fatty acid biosynthesis pathway in cyanobacteria. The fundamental works on the regulatory mechanisms of cyanobacterial fatty acid metabolism are required to identify potential targets for unlocking and boosting fatty acid biosynthesis. Excretion of oleochemicals out of cells by transporters can alleviate the toxicity of FFA and was shown to be an effective strategy for improving cyanobacterial FFA production (Kato et al., 2015, 2017). However, native or foreign transporters for other oleochemicals still need to be discovered and evaluated in cyanobacteria. In addition, it should be helpful for improving cyanobacterial production of oleochemicals by introducing a heterologous reductant regenerating system to balance the reductant generation and utilization, considering the fact that reductant is needed in in oleochemical biosynthesis pathways.

Moreover, genetic instability of engineered cyanobacteria (Jones, 2014) is another potential issue for future commercial application. Despite limited observation up to now (Takahama et al., 2003; Jones, 2014), genetic instability can randomly result in some mutations on the genes associated with the production traits. And these mutations will be enriched and finally lead to the failure of cyanobacterial production of some toxic compounds like oleochemicals, if they are beneficial for fitness. For reliable production of oleochemicals in engineered cyanobacteria, key genes associated with genetic fidelity should be identified though systematic genetic analysis and then modified in cyanobacteria to construct the chassis with a higher genetic stability. On the other hand, some inducible promoters or novel genome editing tools should be used to drive the oleochemical biosynthesis only when cyanobacterial cells stop to propagate, reducing the chance of spreading the mutation through the population in long-term culturing.

AUTHOR CONTRIBUTIONS

XT and SY conceived the outline and revised the manuscript. LW, LC, and XT drafted the manuscript. All authors read and approved the final manuscript.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Afrin, S., Khan, M. R. I., Zhang, W., Wang, Y., Zhang, W., He, L., et al. (2018). Membrane-located expression of thioesterase from *Acinetobacter baylyi* enhances free fatty acid production with decreased toxicity in *Synechocystis* sp, PCC6803. *Front. Microbiol.* 9, 2842. doi: 10.3389/fmicb.2018.02842

Ainsworth, E. A., Lemonnier, P., and Wedow, J. M. (2020). The influence of rising tropospheric carbon dioxide and ozone on plant productivity. *Plant Biol.* 22(Suppl. 1), 5–11. doi: 10.1111/plb.12973

Alper, H., and Stephanopoulos, G. (2009). Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nat. Rev. Microbiol.* 7, 715–723. doi: 10.1038/nrmicro2186

Angermayr, S. A., Hellingwerf, K. J., Lindblad, P., and Teixeira De Mattos, M. J. (2009). Energy biotechnology with cyanobacteria. *Curr. Opin. Biotechnol.* 20, 257–263. doi: 10.1016/j.copbio.2009.05.011

Angermayr, S. A., Rovira, A. G., and Hellingwerf, K. J. (2015). Metabolic engineering of cyanobacteria for the synthesis of commodity products. *Trends Biotechnol.* 33, 352–361. doi: 10.1016/j.tibtech.2015.03.009

Atsumi, S., Hanai, T., and Liao, J. C. (2008). Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* 451, 86–89. doi: 10.1038/nature06450

Atsumi, S., Higashide, W., and Liao, J. C. (2009). Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat. Biotechnol.* 27, 1177–1180. doi: 10.1038/nbt.1586

Bellefleur, M. P. A., Wanda, S. Y., and Curtiss, R. III (2019). Characterizing active transportation mechanisms for free fatty acids and antibiotics in *Synechocystis* sp. PCC 6803. *BMC Biotechnol.* 19:5. doi: 10.1186/s12896-019-0500-3

Berla, B. M., Saha, R., Maranas, C. D., and Pakrasi, H. B. (2015). Cyanobacterial alkanes modulate photosynthetic cyclic electron flow to assist growth under cold stress. *Sci. Rep.* 5:14894. doi: 10.1038/srep14894

Buijs, N. A., Siewers, V., and Nielsen, J. (2013). Advanced biofuel production by the yeast *Saccharomyces cerevisiae. Curr. Opin. Chem. Biol.* 17, 480–488. doi: 10.1016/j.cbpa.2013.03.036

Chen, G., Chen, J., He, Q. F., Zhang, Y., Peng, Z. Y., Fan, Z. X., et al. (2017). Functional expression of the *Arachis hypogaea* L. Acyl-ACP thioesterases *AhFatA* and *AhFatB* enhances fatty acid production in *Synechocystis* sp PCC6803. *Energies* 10:2093. doi: 10.3390/en10122093

Coates, R. C., Podell, S., Korobeynikov, A., Lapidus, A., Pevzner, P., Sherman, D. H., et al. (2014). Characterization of cyanobacterial hydrocarbon composition and distribution of biosynthetic pathways. *PLoS One* 9:e85140. doi: 10.1371/journal.pone.0085140

Cuellar-Bermudez, S. P., Romero-Ogawa, M. A., Vannela, R., Lai, Y. S., Rittmann, B. E., and Parra-Saldivar, R. (2015). Effects of light intensity and carbon dioxide on lipids and fatty acids produced by *Synechocystis* sp. PCC6803 during continuous flow. *Algal Res.* 12, 10–16. doi: 10.1016/j.algal.2015.07.018

Demirbas, A. (2009). Political, economic and environmental impacts of biofuels: a review. *Appl. Energy* 86, S108–S117.

d'Espaux, L., Ghosh, A., Runguphan, W., Wehrs, M., Xu, F., Konzock, O., et al. (2017). Engineering high-level production of fatty alcohols by *Saccharomyces cerevisiae* from lignocellulosic feedstocks. *Metab. Eng.* 42, 115–125. doi: 10.1016/j.ymben.2017.06.004

Dexter, J., and Fu, P. (2009). Metabolic engineering of cyanobacteria for ethanol production. *Energ. Environ. Sci.* 2, 857–864.

Gao, Q., Tan, X., and Lu, X. (2012a). [Characterization of a key gene in membrane lipid cycle in *Synechocystis* sp. PCC6803]. *Sheng Wu Gong Cheng Xue Bao* 28, 1473–1481.

Gao, Q., Wang, W., Zhao, H., and Lu, X. (2012b). Effects of fatty acid activation on photosynthetic production of fatty acid-based biofuels in *Synechocystis* sp. PCC6803. *Biotechnol. Biofuels* 5:17. doi: 10.1186/1754-6834-5-17

Gaurav, N., Sivasankari, S., Kiran, G. S., Ninawe, A., and Selvin, J. (2017). Utilization of bioresources for sustainable biofuels: a review. *Renew. Sustain. Energy Rev.* 73, 205–214. doi: 10.1016/j.rser.2017.01.070

Gombos, Z., Kanervo, E., Tsvetkova, N., Sakamoto, T., Aro, E. M., and Murata, N. (1997). Genetic enhancement of the ability to tolerate photoinhibition by introduction of unsaturated bonds into membrane glycerolipids. *Plant Physiol.* 115, 551–559. doi: 10.1104/pp.115.2.551

Graham-Rowe, D. (2011). Agriculture: beyond food versus fuel. *Nature* 474, S6–S8.

Gu, H., Jinkerson, R. E., Davies, F. K., Sisson, L. A., Schneider, P. E., and Posewitz, M. C. (2016). Modulation of medium-chain fatty acid synthesis in *Synechococcus* sp. PCC 7002 by replacing FabH with a chaetoceros ketoacyl-acp synthase. *Front. Plant Sci.* 7:690. doi: 10.3389/fpls.2016.00690

Hagemann, M., and Hess, W. R. (2018). Systems and synthetic biology for the biotechnological application of cyanobacteria. *Curr. Opin. Biotechnol.* 49, 94–99. doi: 10.1016/j.copbio.2017.07.008

Han, J., Mccarthy, E. D., Hoeven, W. V., Calvin, M., and Bradley, W. H. (1968). Organic geochemical studies, ii. A preliminary report on the distribution of aliphatic hydrocarbons in algae, in bacteria, and in a recent lake sediment. *Proc. Natl. Acad. Sci. U.S.A.* 59, 29–33. doi: 10.1073/pnas.59.1.29

Hauf, W., Schmid, K., Gerhardt, E. C. M., Huergo, L. F., and Forchhammer, K. (2016). Interaction of the nitrogen regulatory protein GlnB (PII) with biotin carboxyl carrier protein (BCCP) controls acetyl-coa levels in the cyanobacterium *Synechocystis* sp, PCC 6803. *Front. Microbiol.* 7:1700. doi: 10.3389/fmicb.2016.01700

Heath, R. J., and Rock, C. O. (1996a). Inhibition of beta-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in *Escherichia coli. J. Biol. Chem.* 271, 10996–11000. doi: 10.1074/jbc.271.18.10996

Heath, R. J., and Rock, C. O. (1996b). Regulation of fatty acid elongation and initiation by acyl-acyl carrier protein in *Escherichia coli. J. Biol. Chem.* 271, 1833–1836. doi: 10.1074/jbc.271.4.1833

Hofvander, P., Doan, T. T., and Hamberg, M. (2011). A prokaryotic acyl-CoA reductase performing reduction of fatty acyl-CoA to fatty alcohol. *FEBS Lett.* 585, 3538–3543. doi: 10.1016/j.febslet.2011.10.016

Hu, P., Borglin, S., Kamennaya, N. A., Chen, L., Park, H., Mahoney, L., et al. (2013). Metabolic phenotyping of the cyanobacterium *Synechocystis* 6803 engineered for production of alkanes and free fatty acids. *Appl. Energy* 102, 850–859. doi: 10.1016/j.apenergy.2012.08.047

Janssen, H. J., and Steinbuchel, A. (2014). Fatty acid synthesis in *Escherichia coli* and its applications towards the production of fatty acid based biofuels. *Biotechnol. Biofuels* 7:7. doi: 10.1186/1754-6834-7-7

Jones, P. R. (2014). Genetic instability in cyanobacteria - an elephant in the room? *Front. Bioeng. Biotechnol.* 2:12. doi: 10.3389/fbioe.2014.00012

Kaczmarzyk, D., Cengic, I., Yao, L., and Hudson, E. P. (2018). Diversion of the long-chain acyl-ACP pool in *Synechocystis* to fatty alcohols through CRISPRi

repression of the essential phosphate acyltransferase PlsX. *Metab. Eng.* 45, 59–66. doi: 10.1016/j.ymben.2017.11.014

Kaczmarzyk, D., and Fulda, M. (2010). Fatty acid activation in cyanobacteria mediated by acyl-acyl carrier protein synthetase enables fatty acid recycling. *Plant Physiol.* 152, 1598–1610. doi: 10.1104/pp.109.148007

Kageyama, H., Waditee-Sirisattha, R., Sirisattha, S., Tanaka, Y., Mahakhant, A., and Takabe, T. (2015). Improved alkane production in nitrogen-fixing and halotolerant cyanobacteria via abiotic stresses and genetic manipulation of alkane synthetic genes. *Curr. Microbiol.* 71, 115–120. doi: 10.1007/s00284-015-0833-7

Kaiser, B. K., Carleton, M., Hickman, J. W., Miller, C., Lawson, D., Budde, M., et al. (2013). Fatty aldehydes in cyanobacteria are a metabolically flexible precursor for a diversity of biofuel products. *PLoS One* 8:e58307. doi: 10.1371/journal.pone.0058307

Kamarainen, J., Knoop, H., Stanford, N. J., Guerrero, F., Akhtar, M. K., Aro, E. M., et al. (2012). Physiological tolerance and stoichiometric potential of cyanobacteria for hydrocarbon fuel production. *J. Biotechnol.* 162, 67–74. doi: 10.1016/j.jbiotec.2012.07.193

Kato, A., Takatani, N., Ikeda, K., Maeda, S. I., and Omata, T. (2017). Removal of the product from the culture medium strongly enhances free fatty acid production by genetically engineered *Synechococcus elongatus*. *Biotechnol. Biofuels* 10:141. doi: 10.1186/s13068-017-0831-z

Kato, A., Takatani, N., Use, K., Uesaka, K., Ikeda, K., Chang, Y., et al. (2015). Identification of a cyanobacterial RND-type efflux system involved in export of free fatty acids. *Plant Cell Physiol.* 56, 2467–2477. doi: 10.1093/pcp/pcv150

Kawahara, A., Sato, Y., Saito, Y., Kaneko, Y., Takimura, Y., Hagihara, H., et al. (2016). Free fatty acid production in the cyanobacterium *Synechocystis* sp. PCC 6803 is enhanced by deletion of the cyAbrB2 transcriptional regulator. *J. Biotechnol.* 220, 1–11. doi: 10.1016/j.jbiotec.2015.12.035

Keasling, J. D. (2012). Synthetic biology and the development of tools for metabolic engineering. *Metab. Eng.* 14, 189–195. doi: 10.1016/j.ymben.2012.01.004

Kizawa, A., Kawahara, A., Takashima, K., Takimura, Y., Nishiyama, Y., and Hihara, Y. (2017). The LexA transcription factor regulates fatty acid biosynthetic genes in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant J.* 92, 189–198. doi: 10.1111/tpj.13644

Klahn, S., Baumgartner, D., Pfreundt, U., Voigt, K., Schon, V., Steglich, C., et al. (2014). Alkane biosynthesis genes in cyanobacteria and their transcriptional organization. *Front. Bioeng. Biotechnol.* 2:24. doi: 10.3389/fbioe.2014.00024

Knoot, C. J., and Pakrasi, H. B. (2019). Diverse hydrocarbon biosynthetic enzymes can substitute for olefin synthase in the cyanobacterium *Synechococcus* sp. PCC 7002. *Sci. Rep.* 9:1360. doi: 10.1038/s41598-018-38124-y

Kuo, J., and Khosla, C. (2014). The initiation ketosynthase (FabH) is the sole rate-limiting enzyme of the fatty acid synthase of *Synechococcus* sp. PCC 7002. *Metab. Eng.* 22, 53–59. doi: 10.1016/j.ymben.2013.12.008

Lai, M. C., and Lan, E. I. (2015). Advances in metabolic engineering of cyanobacteria for photosynthetic biochemical production. *Metabolites* 5, 636–658. doi: 10.3390/metabo5040636

Lea-Smith, D. J., Ortiz-Suarez, M. L., Lenn, T., Nürnberg, D. J., Baers, L. L., Davey, M. P., et al. (2016). Hydrocarbons are essential for optimal cell size, division, and growth of cyanobacteria. *Plant Physiol.* 172, 1928–1940. doi: 10.1104/pp.16.01205

Lee, H. J., Choi, J., Lee, S.-M., Um, Y., Sim, S. J., Kim, Y., et al. (2017). Photosynthetic CO_2 conversion to fatty acid ethyl esters (FAEEs) using engineered cyanobacteria. *J. Agric. Food Chem.* 65, 1087–1092. doi: 10.1021/acs.jafc.7b00002

Lee, S. K., Chou, H., Ham, T. S., Lee, T. S., and Keasling, J. D. (2008). Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Curr. Opin. Biotechnol.* 19, 556–563. doi: 10.1016/j.copbio.2008.10.014

Li, F., Cai, B., Ye, Z., Wang, Z., Zhang, W., Zhou, P., et al. (2019). Changing patterns and determinants of transportation carbon emissions in Chinese cities. *Energy* 174, 562–575. doi: 10.1016/j.energy.2019.02.179

Lindberg, P., Park, S., and Melis, A. (2009). Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metab. Eng.* 12, 70–79. doi: 10.1016/j.ymben.2009.10.001

Liu, A., Zhu, T., Lu, X., and Song, L. (2013). Hydrocarbon profiles and phylogenetic analyses of diversified cyanobacterial species. *Appl. Energy* 111, 383–393. doi: 10.1016/j.apenergy.2013.05.008

Liu, X., Sheng, J., and Curtiss Iii, R. (2011a). Fatty acid production in genetically modified cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6899–6904. doi: 10.1073/pnas.1103014108

Liu, X., Fallon, S., Sheng, J., and Curtiss, R. (2011b). CO_2-limitation-inducible green recovery of fatty acids from cyanobacterial biomass. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6905–6908. doi: 10.1073/pnas.1103016108

Los, D. A., and Mironov, K. S. (2015). Modes of fatty acid desaturation in cyanobacteria: an update. *Life* 5, 554–567. doi: 10.3390/life5010554

Lu, X. (2010). A perspective: Photosynthetic production of fatty acid-based biofuels in genetically engineered cyanobacteria. *Biotechnol. Adv.* 28, 742–746. doi: 10.1016/j.biotechadv.2010.05.021

Lu, X., Vora, H., and Khosla, C. (2008). Overproduction of free fatty acids in *E. coli*: implications for biodiesel production. *Metab. Eng.* 10, 333–339. doi: 10.1016/j.ymben.2008.08.006

Luan, G., Zhang, S., Wang, M., and Lu, X. (2019). Progress and perspective on cyanobacterial glycogen metabolism engineering. *Biotechnol. Adv.* 37, 771–786. doi: 10.1016/j.biotechadv.2019.04.005

Marella, E. R., Holkenbrink, C., Siewers, V., and Borodina, I. (2018). Engineering microbial fatty acid metabolism for biofuels and biochemicals. *Curr. Opin. Biotechnol.* 50, 39–46. doi: 10.1016/j.copbio.2017.10.002

Mata, T. M., Martins, A. A., and Caetano, N. S. (2010). Microalgae for biodiesel production and other applications: a review. *Renew. Sustain. Energy Rev.* 14, 217–232. doi: 10.1016/j.rser.2009.07.020

Mendez-Perez, D., Begemann, M. B., and Pfleger, B. F. (2011). Modular synthase-encoding gene involved in alpha-olefin biosynthesis in *Synechococcus* sp Strain PCC 7002. *Appl. Environ. Microbiol.* 77, 4264–4267. doi: 10.1128/AEM.00467-11

Mikami, K., Kanesaki, Y., Suzuki, I., and Murata, N. (2002). The histidine kinase Hik33 perceives osmotic stress and cold stress in *Synechocystis* sp PCC 6803. *Mol. Microbiol.* 46, 905–915. doi: 10.1046/j.1365-2958.2002.03202.x

Murata, N., and Los, D. A. (2006). Histidine kinase Hik33 is an important participant in cold-signal transduction in cyanobacteria. *Physiol. Plant.* 126, 17–27. doi: 10.1111/j.1399-3054.2006.00608.x

Murata, N., and Wada, H. (1995). Acyl-lipid desaturases and their importance in the tolerance and acclimatization to cold of cyanobacteria. *Biochem. J.* 308(Pt 1), 1–8. doi: 10.1042/bj3080001

Nanjo, Y., Mizusawa, N., Wada, H., Slabas, A. R., Hayashi, H., and Nishiyama, Y. (2010). Synthesis of fatty acids *de novo* is required for photosynthetic acclimation of *Synechocystis* sp. PCC 6803 to high temperature. *Biochim. Biophys. Acta* 1797, 1483–1490. doi: 10.1016/j.bbabio.2010.03.014

Naoki, S., and Norio, M. (1982). Lipid biosynthesis in the blue-green alga, Anabaena variabilis: I. Lipid classes. *Biochim. Biophys. Acta* 710, 271–278. doi: 10.1016/0005-2760(82)90109-6

Okazaki, K., Sato, N., Tsuji, N., Tsuzuki, M., and Nishida, I. (2006). The significance of C16 fatty acids in the sn-2 positions of glycerolipids in the photosynthetic growth of *Synechocystis* sp. PCC6803. *Plant Physiol.* 141, 546–556. doi: 10.1104/pp.105.075796

Oliver, N. J., Rabinovitch-Deere, C. A., Carroll, A. L., Nozzi, N. E., Case, A. E., and Atsumi, S. (2016). Cyanobacterial metabolic engineering for biofuel and chemical production. *Curr. Opin. Chem. Biol.* 35, 43–50. doi: 10.1016/j.cbpa.2016.08.023

Peralta-Yahya, P. P., and Keasling, J. D. (2010). Advanced biofuel production in microbes. *Biotechnol. J.* 5, 147–162. doi: 10.1002/biot.200900220

Peramuna, A., Morton, R., and Summers, M. L. (2015). Enhancing alkane production in cyanobacterial lipid droplets: a model platform for industrially relevant compound production. *Life* 5, 1111–1126. doi: 10.3390/life5021111

Peramuna, A., and Summers, M. L. (2014). Composition and occurrence of lipid droplets in the cyanobacterium *Nostoc punctiforme*. *Arch. Microbiol.* 196, 881–890. doi: 10.1007/s00203-014-1027-6

Pfleger, B. F., Gossing, M., and Nielsen, J. (2015). Metabolic engineering strategies for microbial synthesis of oleochemicals. *Metab. Eng.* 29, 1–11. doi: 10.1016/j.ymben.2015.01.009

Qi, F., Yao, L., Tan, X., and Lu, X. (2013). Construction, characterization and application of molecular tools for metabolic engineering of *Synechocystis* sp. *Biotechnol. Lett.* 35, 1655–1661. doi: 10.1007/s10529-013-1252-0

Rexroth, S., Mullineaux, C. W., Ellinger, D., Sendtko, E., Rögner, M., and Koenig, F. (2011). The plasma membrane of the cyanobacterium *Gloeobacter violaceus* contains segregated bioenergetic domains. *Plant Cell* 23:2379. doi: 10.1105/tpc. 111.085779

Rude, M. A., Baron, T. S., Brubaker, S., Alibhai, M., Del Cardayre, S. B., and Schirmer, A. (2011). Terminal olefin (1-Alkene) biosynthesis by a novel P450 fatty acid decarboxylase from *Jeotgalicoccus* species. *Appl. Environ. Microbiol.* 77, 1718–1727. doi: 10.1128/AEM.02580-10

Ruffing, A. M. (2013a). Borrowing genes from *Chlamydomonas reinhardtii* for free fatty acid production in engineered cyanobacteria. *J. Appl. Phycol.* 25, 1495–1507. doi: 10.1007/s10811-013-9993-7

Ruffing, A. M. (2013b). RNA-Seq analysis and targeted mutagenesis for improved free fatty acid production in an engineered cyanobacterium. *Biotechnol. Biofuels* 6:15. doi: 10.1186/1754-6834-6-113

Ruffing, A. M. (2014). Improved free fatty acid production in cyanobacteria with *Synechococcus* sp. PCC 7002 as host. *Front. Bioeng. Biotechnol.* 2:17. doi: 10. 3389/fbioe.2014.00017

Ruffing, A. M., and Jones, H. D. (2012). Physiological effects of free fatty acid production in genetically engineered *Synechococcus elongatus* PCC 7942. *Biotechnol. Bioeng.* 109, 2190–2199. doi: 10.1002/bit.24509

Rui, Z., Harris, N. C., Zhu, X., Huang, W., and Zhang, W. (2015). Discovery of a family of desaturase-like enzymes for 1-Alkene biosynthesis. *ACS Catal.* 5, 7091–7094. doi: 10.1021/acscatal.5b01842

Rui, Z., Li, X., Zhu, X., Liu, J., Domigan, B., Barr, I., et al. (2014). Microbial biosynthesis of medium-chain 1-alkenes by a nonheme iron oxidase. *Proc. Natl. Acad. Sci. U.S.A.* 111, 18242–18242. doi: 10.1073/pnas.1419701112

Sakamoto, T., and Bryant, D. A. (2002). Synergistic effect of high-light and low temperature on cell growth of the Delta12 fatty acid desaturase mutant in *Synechococcus* sp. PCC 7002. *Photosynth. Res.* 72, 231–242.

Santos-Merino, M., Garcillán-Barcia, M. P., and De La Cruz, F. (2018). Engineering the fatty acid synthesis pathway in *Synechococcus elongatus* PCC 7942 improves omega-3 fatty acid production. *Biotechnol. Biofuels* 11:239. doi: 10.1186/s13068-018-1243-4

Sato, N., and Wada, H. (2009). "Lipid biosynthesis and its regulation in cyanobacteria," in *Lipids in Photosynthesis: Essential and Regulatory Functions*, eds H. Wada, and N. Murata, (Dordrecht: Springer), 157–177. doi: 10.1007/978-90-481-2863-1_8

Savakis, P., and Hellingwerf, K. J. (2015). Engineering cyanobacteria for direct biofuel production from CO_2. *Curr. Opin. Biotechnol.* 33, 8–14. doi: 10.1016/j.copbio.2014.09.007

Schirmer, A., Rude, M. A., Li, X., Popova, E., and Del Cardayre, S. B. (2010). Microbial biosynthesis of alkanes. *Science* 329, 559–562. doi: 10.1126/science. 1187936

Sheehan, J., Dunahay, T., Benemann, J., and Roessler, P. (1998). *A look back at the U.S. Department of Energy's Aquatic Species Program: Biodiesel from Algae. Close-Out Report*. Golden: National Renewable Energy Lab.

Solaymani, S. (2019). CO_2 emissions patterns in 7 top carbon emitter economies: the case of transport sector. *Energy* 168, 989–1001. doi: 10.1016/j.energy.2018. 11.145

Song, K., Tan, X., Liang, Y., and Lu, X. (2016). The potential of *Synechococcus elongatus* UTEX 2973 for sugar feedstock production. *Appl. Microbiol. Biotechnol.* 100, 7865–7875. doi: 10.1007/s00253-016-7510-z

Sorigue, D., Legeret, B., Cuine, S., Blangy, S., Moulin, S., Billon, E., et al. (2017). An algal photoenzyme converts fatty acids to hydrocarbons. *Science* 357, 903–907. doi: 10.1126/science.aan6349

Steen, E. J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., Mcclure, A., et al. (2010). Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* 463, 559–562. doi: 10.1038/nature0 8721

Stoveken, T., Kalscheuer, R., Malkus, U., Reichelt, R., and Steinbuchel, A. (2005). The wax ester synthase/acyl coenzyme a:diacylglycerol acyltransferase from *Acinetobacter* sp. Strain ADP1: characterization of a novel type of acyltransferase. *J. Bacteriol.* 187, 1369–1376. doi: 10.1128/jb.187.4.1369-1376. 2005

Suzuki, I., Los, D. A., Kanesaki, Y., Mikami, K., and Murata, N. (2000). The pathway for perception and transduction of low-temperature signals in *Synechocystis*. *EMBO J.* 19, 1327–1334. doi: 10.1093/emboj/19.6.1327

Takahama, K., Matsuoka, M., Nagahama, K., and Ogawa, T. (2003). Construction and analysis of a recombinant cyanobacterium expressing a chromosomally inserted gene for an ethylene-forming enzyme at the psbAI locus. *J. Biosci. Bioeng.* 95, 302–305. doi: 10.1016/s1389-1723(03)80034-8

Tan, X., Yao, L., Gao, Q., Wang, W., Qi, F., and Lu, X. (2011). Photosynthesis driven conversion of carbon dioxide to fatty alcohols and hydrocarbons in cyanobacteria. *Metab. Eng.* 13, 169–176. doi: 10.1016/j.ymben.2011.01.001

Torella, J. P., Ford, T. J., Kim, S. N., Chen, A. M., Way, J. C., and Silver, P. A. (2013). Tailored fatty acid synthesis via dynamic control of fatty acid elongation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11290–11295. doi: 10.1073/pnas.1307129110

Towijit, U., Songruk, N., Lindblad, P., Incharoensakdi, A., and Jantaro, S. (2018). Co-overexpression of native phospholipid-biosynthetic genes *plsX* and *plsC* enhances lipid production in *Synechocystis* sp. PCC 6803. *Sci. Rep.* 8:13510. doi: 10.1038/s41598-018-31789-5

Trautmann, D., Beyer, P., and Al-Babili, S. (2013). The ORF slr0091 of *Synechocystis* sp. PCC6803 encodes a high-light induced aldehyde dehydrogenase converting apocarotenals and alkanals. *FEBS J.* 280, 3685–3696. doi: 10.1111/febs.12361

Verma, E., Chakraborty, S., Tiwari, B., and Mishra, A. K. (2018). Transcriptional regulation of acetyl CoA and lipid synthesis by PII protein in *Synechococcus* PCC 7942. *J. Basic Microbiol.* 58, 187–197. doi: 10.1002/jobm.201700467

von Berlepsch, S., Kunz, H.-H., Brodesser, S., Fink, P., Marin, K., Flügge, U.-I., et al. (2012). The acyl-acyl carrier protein synthetase from *Synechocystis* sp. PCC 6803 mediates fatty acid import. *Plant Physiol.* 159, 606–617. doi: 10.1104/pp. 112.195263

Wada, H., Gombos, Z., and Murata, N. (1990). Enhancement of chilling tolerance of a cyanobacterium by genetic manipulation of fatty acid desaturation. *Nature* 347, 200–203. doi: 10.1038/347200a0

Wada, H., and Murata, N. (1990). Temperature-induced changes in the fatty acid composition of the cyanobacterium, *Synechocystis* PCC6803. *Plant Physiol.* 92, 1062–1069. doi: 10.1104/pp.92.4.1062

Wada, H., and Murata, N. (1998). "Membrane lipids in cyanobacteria," in *Lipids in Photosynthesis: Structure, Function and Genetics*, eds S. Paul-André, and M. Norio, (Dordrecht: Springer), 65–81. doi: 10.1007/0-306-48087-5_4

Wang, B., Eckert, C., Maness, P. C., and Yu, J. (2018). A genetic toolbox for modulating the expression of heterologous genes in the cyanobacterium *Synechocystis* sp. PCC 6803. *ACS Synth. Biol.* 7, 276–286. doi: 10.1021/acssynbio. 7b00297

Wang, W., Liu, X., and Lu, X. (2013). Engineering cyanobacteria to improve photosynthetic production of alka(e)nes. *Biotechnol. Biofuels* 6:69. doi: 10.1186/1754-6834-6-69

Weier, D., Muller, C., Gaspers, C., and Frentzen, M. (2005). Characterisation of acyltransferases from *Synechocystis* sp. PCC6803. *Biochem. Biophys. Res. Commun.* 334, 1127–1134. doi: 10.1016/j.bbrc.2005.06.197

Wijffels, R. H., and Barbosa, M. J. (2010). An outlook on microalgal biofuels. *Science* 329, 796–799. doi: 10.1126/science.1189003

Winters, K., Parker, P. L., and Van Baalen, C. (1969). Hydrocarbons of blue-green algae: Geochemical signficance. *Science* 163, 467–468. doi: 10.1126/science. 163.3866.467

Work, V. H., Melnicki, M. R., Hill, E. A., Davies, F. K., Kucek, L. A., Beliaev, A. S., et al. (2015). Lauric acid production in a glycogen-less strain of *Synechococcus* sp. PCC 7002. *Front. Bioeng. Biotechnol.* 3:48. doi: 10.3389/fbioe.2015.00048

Xie, M., Wang, W., Zhang, W., Chen, L., and Lu, X. (2017). Versatility of hydrocarbon production in cyanobacteria. *Appl. Microbiol. Biotechnol.* 101, 905–919. doi: 10.1007/s00253-016-8064-9

Xiong, W., Cano, M., Wang, B., Douchi, D., and Yu, J. (2017). The plasticity of cyanobacterial carbon metabolism. *Curr. Opin. Chem. Biol.* 41, 12–19. doi: 10.1016/j.cbpa.2017.09.004

Yamamori, T., Kageyama, H., Tanaka, Y., and Takabe, T. (2018). Requirement of alkanes for salt tolerance of cyanobacteria: Characterization of alkane synthesis genes from salt-sensitive *Synechococcus elongatus* PCC7942 and salt-tolerant *Aphanothece halophytica*. *Lett. Appl. Microbiol.* 67, 299–305. doi: 10.1111/lam. 13038

Yao, L., Qi, F., Tan, X., and Lu, X. (2014). Improved production of fatty alcohols in cyanobacteria by metabolic engineering. *Biotechnol. Biofuels* 7:94. doi: 10.1186/1754-6834-7-94

Yoshino, T., Liang, Y., Arai, D., Maeda, Y., Honda, T., Muto, M., et al. (2015). Alkane production by the marine cyanobacterium *Synechococcus*

sp. NKBG15041c possessing the alpha-olefin biosynthesis pathway. *Appl. Microbiol. Biotechnol.* 99, 1521–1529. doi: 10.1007/s00253-014-6286-2

Yu, A. Q., Pratomo Juwono, N. K., Leong, S. S., and Chang, M. W. (2014). Production of fatty acid-derived valuable chemicals in synthetic microbes. *Front. Bioeng. Biotechnol.* 2:78. doi: 10.3389/fbioe.2014.00078

Yu, X., Liu, T., Zhu, F., and Khosla, C. (2011). In vitro reconstitution and steady-state analysis of the fatty acid synthase from *Escherichia coli. Proc. Natl. Acad. Sci. U.S.A.* 108, 18643–18648. doi: 10.1073/pnas.1110852108

Yunus, I. S., and Jones, P. R. (2018). Photosynthesis-dependent biosynthesis of medium chain-length fatty acids and alcohols. *Metab. Eng.* 49, 59–68. doi: 10.1016/j.ymben.2018.07.015

Yunus, I. S., Wichmann, J., Wordenweber, R., Lauersen, K. J., Kruse, O., and Jones, P. R. (2018). Synthetic metabolic pathways for photobiological conversion of CO2 into hydrocarbon fuel. *Metab. Eng.* 49, 201–211. doi: 10.1016/j.ymben.2018.08.008

Zhou, J., and Li, Y. (2010). Engineering cyanobacteria for fuels and chemicals production. *Protein Cell* 1, 207–210. doi: 10.1007/s13238-010-0043-9

Zhu, T., Scalvenzi, T., Sassoon, N., Lu, X., and Gugger, M. (2018). Terminal olefin profiles and phylogenetic analyses of olefin synthases of diverse cyanobacterial species. *Appl. Environ. Microbiol.* 84:e00425-18. doi: 10.1128/AEM.00425-18

Type IV Pili-Independent Photocurrent Production by the Cyanobacterium *Synechocystis* sp. PCC 6803

Miyuki A. Thirumurthy[1†], Andrew Hitchcock[2†], Angelo Cereda[1†], Jiawei Liu[3], Marko S. Chavez[4], Bryant L. Doss[3], Robert Ros[3], Mohamed Y. El-Naggar[4,5,6], John T. Heap[7,8], Thomas S. Bibby[9*] and Anne K. Jones[1*]

[1] School of Molecular Sciences, Arizona State University, Tempe, AZ, United States, [2] Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield, United Kingdom, [3] Department of Physics, Arizona State University, Tempe, AZ, United States, [4] Department of Physics and Astronomy, University of Southern California, Los Angeles, CA, United States, [5] Department of Biological Sciences, University of Southern California, Los Angeles, CA, United States, [6] Department of Chemistry, University of Southern California, Los Angeles, CA, United States, [7] Imperial College Centre for Synthetic Biology, Department of Life Sciences, Imperial College London, London, United Kingdom, [8] School of Life Sciences, University of Nottingham, Nottingham, United Kingdom, [9] Ocean and Earth Science, University of Southampton, Southampton, United Kingdom

Biophotovoltaic devices utilize photosynthetic organisms such as the model cyanobacterium *Synechocystis* sp. PCC 6803 (*Synechocystis*) to generate current for power or hydrogen production from light. These devices have been improved by both architecture engineering and genetic engineering of the phototrophic organism. However, genetic approaches are limited by lack of understanding of cellular mechanisms of electron transfer from internal metabolism to the cell exterior. Type IV pili have been implicated in extracellular electron transfer (EET) in some species of heterotrophic bacteria. Furthermore, conductive cell surface filaments have been reported for cyanobacteria, including *Synechocystis*. However, it remains unclear whether these filaments are type IV pili and whether they are involved in EET. Herein, a mediatorless electrochemical setup is used to compare the electrogenic output of wild-type *Synechocystis* to that of a Δ*pilD* mutant that cannot produce type IV pili. No differences in photocurrent, i.e., current in response to illumination, are detectable. Furthermore, measurements of individual pili using conductive atomic force microscopy indicate these structures are not conductive. These results suggest that pili are not required for EET by *Synechocystis*, supporting a role for shuttling of electrons via soluble redox mediators or direct interactions between the cell surface and extracellular substrates.

Keywords: cyanobacteria, type IV pili, nanowire, photocurrent, biophotovoltaics, extracellular electron transfer

INTRODUCTION

Electron transfer and redox reactions form the foundation for energy transduction in biological systems (Marcus and Sutin, 1985). Some microbes have the capacity to transfer electrons beyond their cell wall to extracellular acceptors (Hernandez and Newman, 2001), a function that may be important in microbial ecology (Lis et al., 2015; Polyviou et al., 2018) and has been exploited

in bioelectronic applications. Although electron transfer between redox-active sites separated by less than 1.6 nm is well understood to occur via electron tunneling described by Marcus theory, little is known about the mechanisms of electron transfer over larger distances, i.e., nanometers to micrometers, observed in biological ecosystems (Gray and Winkler, 2005). Long-range electron transfer in various microbes may employ soluble redox mediators, conductive bacterial nanowires or pili (Reguera et al., 2005; Marsili et al., 2008; Brutinel and Gralnick, 2012; Kotloski and Gralnick, 2013; Yang et al., 2015; Ing et al., 2018; Heidary et al., 2020). Furthermore, an understanding of this activity forms the foundation for the development of microbial fuel cells and photobiological electrochemical systems, devices that employ microbes to generate electricity (Rabaey and Verstraete, 2005; Kracke et al., 2015).

Two distinct mechanisms have been hypothesized to account for extracellular electron transfer (EET) in anaerobic, heterotrophic bacteria: utilization of soluble, diffusing redox shuttles like flavins to transfer electrons from the cellular interior to the extracellular surface (Watanabe et al., 2009; Glasser et al., 2017) and direct interaction between a redox-active component on the cell surface and the extracellular target (Shi et al., 2009). The latter has been proposed to proceed via redox proteins on the cell surface (e.g., multiheme cytochromes) or via extracellular appendages that have come to be known as bacterial nanowires (Gorby et al., 2006; El-Naggar et al., 2010). The composition of these nanowires is hypothesized to vary between different organisms; recent work by El-Naggar and coworkers has shown that the nanowires of *Shewanella oneidensis* MR-1 are extensions of EET-protein-containing outer membrane that appear to form from chains of vesicles (Pirbadian et al., 2014). On the other hand, Lovley and coworkers reported that the nanowires of electrogenic *Geobacter* sp. are conductive pili (Reguera et al., 2005; Holmes et al., 2016), whereas recent studies have shown that *Geobacter sulfurreducens* produces OmcS cytochrome filaments that are distinct from type IV pili (Tfp) (Filman et al., 2019; Wang et al., 2019). For a recent review of *Geobacter* protein nanowires see Lovley and Walker (2019). However, details about the types of charge carriers and the exact mechanisms of interfacial electron transport within conductive appendages remain unclear.

Biophotovoltaic devices (BPVs) interconvert light and electrical energy using a photosynthetic organism. The most common devices employ oxygenic phototrophs to harvest light energy and transfer electrons produced by water oxidation to extracellular acceptors, generating power or hydrogen (Zou et al., 2009; Pisciotta et al., 2010; McCormick et al., 2011, 2015; Bradley et al., 2012; Lea-Smith et al., 2015; Saper et al., 2018; Tschörtner et al., 2019). Cyanobacteria, green algae, and plants have been used to generate power in BPVs, with much work performed using the model freshwater cyanobacterial species *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*). Current production in BPVs containing *Synechocystis* is largely dependent on illumination, and previous studies employing chemical and genetic inhibition indicate that water splitting by Photosystem II (PSII) provides the majority of electrons (Bombelli et al., 2011; Pisciotta et al., 2011; Cereda et al., 2014). Improvements of BPVs based on advances in device architecture, electrode material,

proton exchange membrane and use of mediators and biofilms have been reported (Thorne et al., 2011; Bombelli et al., 2012, 2015; Call et al., 2017; Rowden et al., 2018; Wenzel et al., 2018; Wey et al., 2019), but improvements arising from engineering of phototrophs have been limited to genetic removal of competing electron sinks (Bradley et al., 2013; McCormick et al., 2013; Saar et al., 2018) by lack of understanding of how photosynthetic electrons are transferred from the photosynthetic apparatus to extracellular acceptors.

Tfp are required for gliding motility, phototaxis, cell adhesion, flocculation, and natural transformation competency in *Synechocystis*, which produces morphologically distinct thick ($\sim$5–8 nm, >2 μm in length, form tufts) and thin ($\sim$3–4 nm, $\sim$1 μm, cover whole surface of cell) pili (Bhaya et al., 2000; Yoshihara et al., 2001; Schuergers and Wilde, 2015; Chen et al., 2020). Tfp have also been implicated as having a role in reductive iron (Kranzler et al., 2011; Lamb et al., 2014) and manganese uptake (Lamb and Hohmann-Marriott, 2017). *Synechocystis* has also been reported to produce conductive filaments under conditions of CO_2 limitation (Gorby et al., 2006), although whether these are Tfp is unclear. For detailed reviews of Tfp structure, biogenesis, and function in *Synechocystis*, see Schuergers and Wilde (2015) and Chen et al. (2020).

Synechocystis cannot produce pili in the absence of the leader peptidase/methylase, encoded by the *pilD* gene (Bhaya et al., 2000). Herein, the rates of EET by a $\Delta pilD$ mutant are compared to those of wild-type organisms by measuring photocurrent production in our previously described mediatorless bioelectrochemical cell (Cereda et al., 2014). Photocurrent production by the wild-type and $\Delta pilD$ cells is not significantly different, suggesting pili do not play a role in photocurrent generation or EET by *Synechocystis*, at least under the conditions investigated here. Additionally, conductivity measurements using atomic force microscopy (AFM) of wild-type *Synechocystis* pili found no evidence for conductivity in these structures. Our results support the hypothesis that redox mediator shuttling may be the major mechanism of photocurrent production by cyanobacteria (Saper et al., 2018; Wenzel et al., 2018).

MATERIALS AND METHODS

Growth of *Synechocystis* sp. PCC 6803

A glucose-tolerant (GT) strain of *Synechocystis* was used as the wild type in this study (see **Supplementary Table S1** for details). *Synechocystis* was cultured in BG11 media (Rippka et al., 1979) buffered with 10 mM N-[tris(hydroxymethyl) methyl]-2-aminoethanesulfonic acid (TES)-KOH pH 8.2 (BG11-TES). For photoautotrophic growth, 200 ml cultures contained within 250 ml flasks were bubbled with sterile air at 30°C under a constant illumination of approximately 50 μmol photons m^{-2} s^{-1}. For photomixotrophic growth, 5 mM glucose was added to the medium. For growth on plates, media was supplemented with 1.5% (w/v) agar and 0.3% (w/v) sodium thiosulphate; 34 μg/ml chloramphenicol (for $\Delta pilD$) or 20 μg/ml zeocin ($\Delta psbB$) were added where required.

Growth was monitored by measurement of the optical density at 750 nm (OD_{750}).

Deletion of *pilD* (slr1120)

For deletion of *pilD*, the central portion of the slr1120 open reading frame was replaced with a chloramphenicol acetyl transferase (*cat*) gene by allele exchange using a plasmid (pICJH4) constructed by Gibson assembly (Gibson et al., 2009) of three PCR products (two amplified from *Synechocystis* genomic DNA and the third from pACYC184) together with the 2.6 kb *Eco*RI–*Hin*dIII restriction fragment of pUC19. The allele exchange cassette comprised a first region of 685 bp of homology with the *Synechocystis* chromosome including upstream flanking sequence and the first 28 codons of *pilD* followed by two stop codons (amplified with primers *pilD*-us-F and *pilD*-us-R), the *cat* cassette (amplified with primers *cat*-F and *cat*-R), and a second region of 500 bp of homology with the *Synechocystis* chromosome beginning with the 12th-from-last codon of *pilD* followed by flanking downstream DNA (amplified with primers *pilD*-ds-F and *pilD*-ds-R) (see **Supplementary Table S2** for primer sequences). The pICJH4 plasmid was confirmed to be correctly assembled by automated DNA sequencing and introduced into wild-type *Synechocystis* by natural transformation. Recombinants were selected on plates containing 5 μg ml^{-1} chloramphenicol, and segregation of genome copies was achieved by sequentially increasing the chloramphenicol concentration (up to 40 μg ml^{-1}). Segregation at the *pilD* locus was confirmed by PCR with primer pair *pilD*-screen-F and *pilD*-screen-R.

RNA Isolation and RT-PCR

End-point reverse transcriptase PCR analysis of *Synechocystis* strains was performed as described previously for *Acaryochloris marina* (Chen et al., 2016). Briefly, *Synechocystis* cells were harvested at mid-log phase (OD_{750} = ∼0.6), and total RNA was isolated by the hot TRIzol method (Pinto et al., 2009). RNA was treated with the Ambion Turbo DNA-*free*™ Kit to remove contaminating genomic DNA, and 100 ng was used for cDNA synthesis and PCR, which were performed in a single reaction using the MyTaq one-step reverse transcription-PCR (RT-PCR) kit (Bioline). Gene-specific primer pairs *pilA1*-RT-F/R, *pilD*-RT-F/R or *rnpB*-RT-F/R were used to detect transcript of *pilA1* (124 bp), *pilD* (180 bp), and the reference gene *rnpB* (180 bp) (Polyviou et al., 2015), respectively. The reaction setup and thermocycling conditions were performed according to the manufacturer's instructions, and 10 μl of PCR product was analyzed on a 2% (w/v) agarose gel.

Immunodetection of PilA1

Denatured whole-cell extracts were separated by SDS-PAGE on 12% Bis-Tris gels (Invitrogen) and transferred to polyvinylidene difluoride membranes (Invitrogen). Membranes were incubated with an anti-PilA1 primary antibody raised against a synthetic peptide corresponding to PilA1 residues 147–160 as described previously (Linhartová et al., 2014) and then a secondary antibody conjugated with horseradish peroxidase (Sigma Aldrich). Chemiluminescence was detected using the WESTAR®

EtaC kit (Geneflow Ltd.) and an Amersham™ Imager 600 (GE Healthcare).

Oxygen Evolution and Determination of Chlorophyll Content

Oxygen evolution was measured as described in our previous work (Cereda et al., 2014). Chlorophyll was extracted from cell pellets from 1 ml of culture at OD_{750} ≈ 0.4 with 100% methanol and quantified spectrophotometrically according to Porra et al. (1989).

Electrochemical Measurements

Electrochemical measurements were made in a three-electrode cell with carbon cloth as working electrode as described previously (Cereda et al., 2014).

Atomic Force Microscopy Imaging of Wild-Type and Mutant Cells (Δ*pilD**)

Synechocystis wild-type and Δ*pilD** cells grown photoautotrophically in liquid BG11 or on BG11 agar plates were collected, washed three times, and resuspended in 1 ml deionized water (centrifugation speed 3,500 × *g*). Aliquots of 5 μl were spotted onto a mica support and air dried. After drying, samples were imaged using an Asylum Research MFP 3D (Santa Barbara, CA, United States) Atomic Force Microscope (AFM) in tapping mode using Tap300Al-G probes (with 40 N/m force constant, 300 kHz resonant frequency). The images were processed using Gwyddion software.

Scanning Electron Microscopy (SEM) Imaging

Wild-type *Synechocystis* and the Δ*pilD** strain were grown photoautotrophically and harvested via centrifugation (3,500 × *g*). Cells were transferred to the carbon cloth used for electrochemical measurements, fixed onto the cloth in 50 mM sodium phosphate buffer (pH 7.2) with 2% glutaraldehyde for 30 min at room temperature, and washed three times in the same buffer for a total of 30 min. After a second fixation step for 30 min at room temperature in the same buffer plus 0.5% (v/v) osmium tetroxide, samples were washed three times with deionized water. Samples were critical point dried with carbon dioxide (Balzers CPD020 unit), mounted on aluminum specimen stubs, and coated with approximately 15 nm of gold-palladium (Technics Hummer-II sputter-coater). Sample analysis was performed with a JEOL JSM-6300 SEM operated at 15 kV, and images were acquired with an IXRF Systems digital scanning unit.

AFM-Based Electrical Characterization of Pili

Glass coverslips (43 × 50 NO. 1 Thermo Scientific Gold Seal Cover Glass) coated with 5 nm titanium and then 100 nm gold via electron beam evaporation were used as conductive substrates. The Au-coated coverslips were rinsed with acetone, isopropanol, ethanol, and deionized water and then dried with nitrogen prior to use. *Synechocystis* cells were drop cast onto

the clean conductive substrates, rinsed with sterile water, and left to dry overnight. An Oxford Instruments Asylum Research Cypher ES AFM was used to make all electrical measurements. Dried samples were affixed and electrically connected to AFM disks with silver paint (TED PELLA, Inc). The sample disks were wired to the AFM upon loading. Si probes, with a Ti/Ir (5/20) coating, a resonant frequency of 75 kHz (58-97), a spring constant of 2.8 N/m (1.4-5.8), and a tip radius of 28 ± 10 nm, were used (Oxford Instruments AFM probe Model: ASYELEC.01-R2). Pili electrical characterization was performed using Oxford Instruments Asylum Research Fast Current Mapping (FCM). To generate FCM images, a bias is held between the probe and substrate while, for each pixel, current and force are measured with respect to the vertical distance of consecutive probe approaches and retractions over the sample. Each approach is terminated when a user-defined force is met (a force setpoint), and each retraction is terminated when a user-defined distance is met (a force distance). A bias of 5.00 V was used. A force setpoint of 49.34 nN and a force distance of 1000 nm were used for thick pili measurements. A force set point of 27.86 nN and a force distance of 750 nm were used for thin pili measurements.

RESULTS

Generation and Phenotypic Analysis of a ΔpilD Strain

The PilD protein is a bifunctional, membrane-bound leader peptidase/methylase that processes PilA precursors and N-methylates the amino acid at position 1 in the mature protein (Strom et al., 1993). PilD is absolutely required for pilus assembly, and a ΔpilD mutant in a motile strain of *Synechocystis* has been reported to be non-piliated, non-motile, and recalcitrant to transformation (Bhaya et al., 2000). Since *Synechocystis* contains multiple *pilA* genes (Yoshihara et al., 2001) but only a single copy of *pilD* (slr1120), we used a ΔpilD knockout mutant to investigate whether pili are required for EET in *Synechocystis*. The ΔpilD mutant generated herein has most of the open reading frame replaced with a chloramphenicol resistant cassette (**Figure 1A**) and was confirmed to be fully segregated by PCR (**Figure 1B**).

It should be noted that GT strains of *Synechocystis* are typically non-motile because of a frameshift mutation in the *spkA* (sll1574) gene, which in motile strains encodes a functional Ser/Thr protein kinase (Kamei et al., 2001). In the originally genome-sequenced Kazusa strain (Kaneko et al., 1996), a 1 bp insertion also results in a frameshift mutation in *pilC* (slr0162/3), preventing pilus assembly (Bhaya et al., 2000), which means this strain is non-competent for transformation with exogenous DNA (Ikeuchi and Tabata, 2001). The *pilC* mutation seems to be specific to the Kazusa strain as other GT strains contain an intact *pilC* gene (Tajima et al., 2011; Kanesaki et al., 2012; Trautmann et al., 2012; Morris et al., 2014; Ding et al., 2015), and the GT wild-type strain used in this study (**Supplementary Table S1**) is naturally transformable and thus must produce Tfp.

When first generated, the ΔpilD mutant displayed an obvious aggregation phenotype, with cells forming small clumps when grown photoheterotrophically in liquid medium. The cells were very difficult to collect with a loop from an agar plate, and the strain grew very poorly, if at all, under photoautotrophic conditions (**Table 1** and **Figure 1C**). Similar phenotypes were described for a ΔpilD mutant generated by Linhartová et al. (2014), who showed that the build-up of unprocessed PilA-prepilins triggered degradation of the essential membrane proteins SecY and YidC. Linhartová et al. (2014) isolated suppressor mutants that were able to grow photoautotrophically by prolonged growth in the absence of glucose or targeted deletion of the *pilA1* gene. Similarly, after continued sub-culturing on agar plates we also isolated suppressor mutants that were capable of photoautotrophic growth, and when cultures were well mixed by air bubbling or orbital shaking, these suppressor strains grew at rates comparable to the wild type without significant clumping (**Table 1** and **Figure 1C**). We will henceforth refer to the strain which can grow photoautrophically as ΔpilD*. Linhartová et al. (2014) showed that the loss of PilA1 pre-pilins in their ΔpilD* strain was at least partially responsible for the improvement in growth; conversely, we found that Pre-PilA1 is still present in our ΔpilD* strain, albeit to a lesser extent than in the originally isolated ΔpilD strain (**Figure 1D**). Another study found that the level of *pilA1* mRNA in a ΔpilD strain capable of phototrophic growth is similar to that of the wild-type organism (Bhaya et al., 2000); sequencing confirmed *pilA1* and its promoter are not mutated in our ΔpilD* strain, and we confirmed *pilA1* is expressed using end-point RT-PCR (**Figure 1E**), indicating that reduced transcription of the *pilA1* gene is unlikely to be the reason for the decrease in PilA production. Further investigation of the nature of the suppressor mutation(s) in ΔpilD* strains is beyond the scope of the present work and will be reported elsewhere (Linhartova, Sobtoka, et al. Unpublished).

The initially isolated ΔpilD mutant described by Linhartová et al. (2014) had impaired PSII activity. Because it has previously been shown that photocurrent from *Synechocystis* is largely dependent on the supply of electrons from water splitting by PSII (Bombelli et al., 2011; Pisciotta et al., 2011; Cereda et al., 2014), we measured the rate of oxygen evolution by wild-type or ΔpilD* cells. For both photoautotrophically and photoheterotrophically cultured cells, the growth rate, chlorophyll content, and oxygen evolution of the ΔpilD* was not significantly different to that of the wild-type organism (**Table 1**). This suggests that PSII activity and the photosynthetic capacity of the ΔpilD* strain are similar to the wild type, allowing direct electrochemical comparison of the two strains when the same number of cells is used (normalized by OD_{750}).

Electrochemical Properties of the ΔpilD* Strain

The light-dependent, EET capacity of the wild-type and ΔpilD* strains of *Synechocystis* was probed by measuring the photocurrent produced when a potential of +240 mV (vs. standard hydrogen electrode) was applied. This potential was chosen because it has previously been shown to be sufficiently oxidizing for the cells to transfer electrons to an external
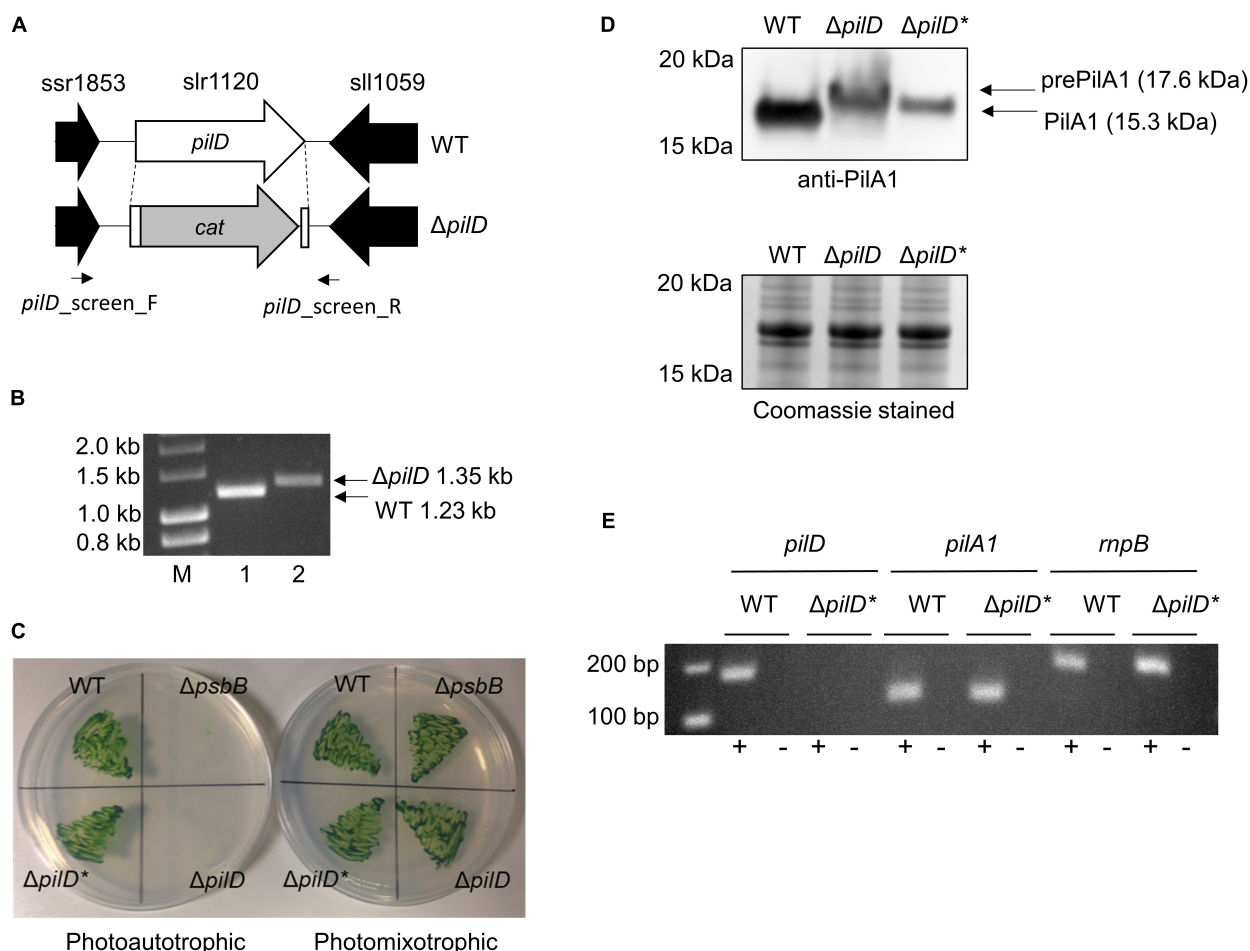
FIGURE 1 | Generation and phenotypic analysis of a Δ*pilD* mutant strain of *Synechocystis*. **(A)** Strategy for deletion of *pilD* (slr1120) by replacement with the chloramphenicol acetyl transferase (*cat*) cassette. The position of screening primers used in panel **(B)** is shown. **(B)** Agarose gel showing PCR products amplified with the *pilD*_screen_F/*pilD*_screen_R primer pair with wild type (WT, lane 1) or Δ*pilD* (lane 2) genomic DNA as template. A larger 1.35 kb PCR product is observed for the Δ*pilD* mutant compared to the 1.23 kb WT band. Lane M = HyperLadder™ 1 kb molecular weight marker (Bioline). **(C)** Growth of the WT, Δ*pilD* and Δ*pilD** (suppressor mutant capable of photoautotrophic growth) in the absence or presence of 5 mM glucose. The originally isolated Δ*pilD* mutant cannot grow under photoautotrophic conditions; a Δ*psbB* mutant that is also unable to grow photoautotrophically is included as a control. **(D)** Level of (pre)PilA1 in WT, Δ*pilD* and Δ*pilD** in photomixotrophically grown whole-cell extracts determined by immunodetection with anti-PilA1 antibodies (upper panel). The accumulation of prePilA1 in the original mutant is decreased in the suppressor strain. The predicted molecular weights of pre- and processed PilA1 are indicated. The lower panel shows a duplicate Coomassie-stained SDS-PAGE gel to demonstrate approximately equal protein loading of each sample. **(E)** End-point RT-PCR analysis of *pilA1* expression in WT and Δ*pilD** showing the transcript is present in both strains. As expected, *pilD* transcripts were absent from the mutant; the *rnpB* housekeeping gene is included as a control. Reactions were performed in the presence (+) or absence (–) of reverse transcriptase.

substrate (Cereda et al., 2014). As shown in **Figure 2A**, when Δ*pilD** cells are applied to the working electrode of a photo-bioelectrochemical cell followed by incubation for a few minutes at the desired electrochemical potential, photocurrent can be observed [red light with peak λ = 660 nm, maximum intensity 20 W m^{-2} (110 μmol photons m^{-2} s^{-1})]. The photocurrent produced by Δ*pilD** is similar to the photocurrent produced by wild type whether the cells were grown photoautotrophically or photomixotrophically (**Figure 2B**). For the Δ*pilD** strain, photocurrent increases linearly (R^2 = 0.99) with cell density to a magnitude (88 ± 15%) comparable to that produced by the wild type (100 ± 12%) (**Supplementary Figure S1**). This shows that the electrical output of both strains is directly related to the

concentration of *Synechocystis* cells present in the electrochemical cell. In short, photocurrent production by the two strains is not significantly different, suggesting that it is independent of Tfp.

Atomic Force Microscopy (AFM) Imaging of Wild-Type and Δ*pilD** Cells

Planktonic growth under rapidly mixed conditions has previously been reported to negatively impact pili stability via shearing action (Yoshihara et al., 2001; Lamb et al., 2014). To provide evidence that wild-type *Synechocystis* has Tfp under the growth conditions employed in this study, we visualized the cells by AFM. To ensure that the imaged cells

TABLE 1 | Growth rate, chlorophyll content, and oxygen evolution of WT, Δ*pilD* and Δ*pilD** *Synechocystis* cells.

Strain	Growth condition[a]	Doubling time (h)	Chl content (μg OD$_{750}$ unit^{-1})	O$_2$ evolution (nmol O$_2$ OD$_{750}$ unit^{-1} min^{-1})	O$_2$ evolution (μmol O$_2$ mg chl^{-1} h^{-1})[b]
WT	PM	12 ± 0.5	3.9 ± 0.3	41 ± 1	631
WT	PA	16 ± 0.5	4.2 ± 0.2	46 ± 5	657
Δ*pilD*	PM	20 ± 1.0[c]	3.5 ± 0.7[c]	32 ± 8[c]	549[c]
Δ*pilD**	PM	12 ± 0.5	3.8 ± 0.4	40 ± 4	632
Δ*pilD**	PA	16 ± 0.5	4.1 ± 0.1	46 ± 4	673

[a]*Growth under PM, photomixotrophic (plus 5 mM glucose) or PA, photoautotrophic conditions, as described in the section "Materials and Methods."* [b]*Calculated from Chl content of 1 OD$_{750}$ unit of cells and the oxygen evolution (nmol O$_2$ OD$_{750}$ unit^{-1} min^{-1}).* [c]*Accuracy of the growth rate, Chl content, and oxygen evolution is limited for this strain as a result of the clumping phenotype.*

are as morphologically like those used in the electrochemical measurements, samples were washed in deionized water prior to AFM visualization to remove contaminants, simulating the pretreatment conditions used for the electrochemical experiments. **Figure 3** shows representative images. Wild-type cells grown planktonically have hair-like pilus structures protruding from the cell surfaces (**Figure 3A**). Conversely, corresponding images of Δ*pilD** cells grown and treated in the same way reveal an almost complete lack of cell surface protrusions (**Figure 3B**).

Scanning Electron Microscopy (SEM) Imaging of *Synechocystis* Cells

Scanning electron microscopy was used to visualize the physical interaction between *Synechocystis* cells and the carbon electrode. SEM micrographs of both wild-type and Δ*pilD** cells confirm uniform adhesion of cells to the carbon cloth electrode surface. We note that sample preparation for SEM imaging can affect the total number of cells attached to the electrode and can underestimate the actual coverage. Nonetheless, in all images, cells appear to be in direct contact with the carbon cloth electrode. High-resolution images from wild-type cells clearly show structures consistent with being pili present between the cells and the carbon substrate (**Figures 4A–D**). Conversely, high-resolution images from the Δ*pilD** strain show a complete absence of any type of pilus-like structures (**Figures 4E–H**), suggesting some other mechanism for the physical interaction with the electrode surface.

Conductivity Measurements of Pili Using AFM

The Fast Current Mapping (FCM) mode of AFM was used to simultaneously generate topographical and current map images of *Synechocystis* pili overtop Au-coated glass coverslips. FCM was chosen for the conductivity measurements to minimize lateral tip-sample forces, which we observed to be damaging and disruptive to the filaments in contact mode conductive AFM. During FCM, current and force curves are generated at each pixel, while the AFM probe vertically approaches and
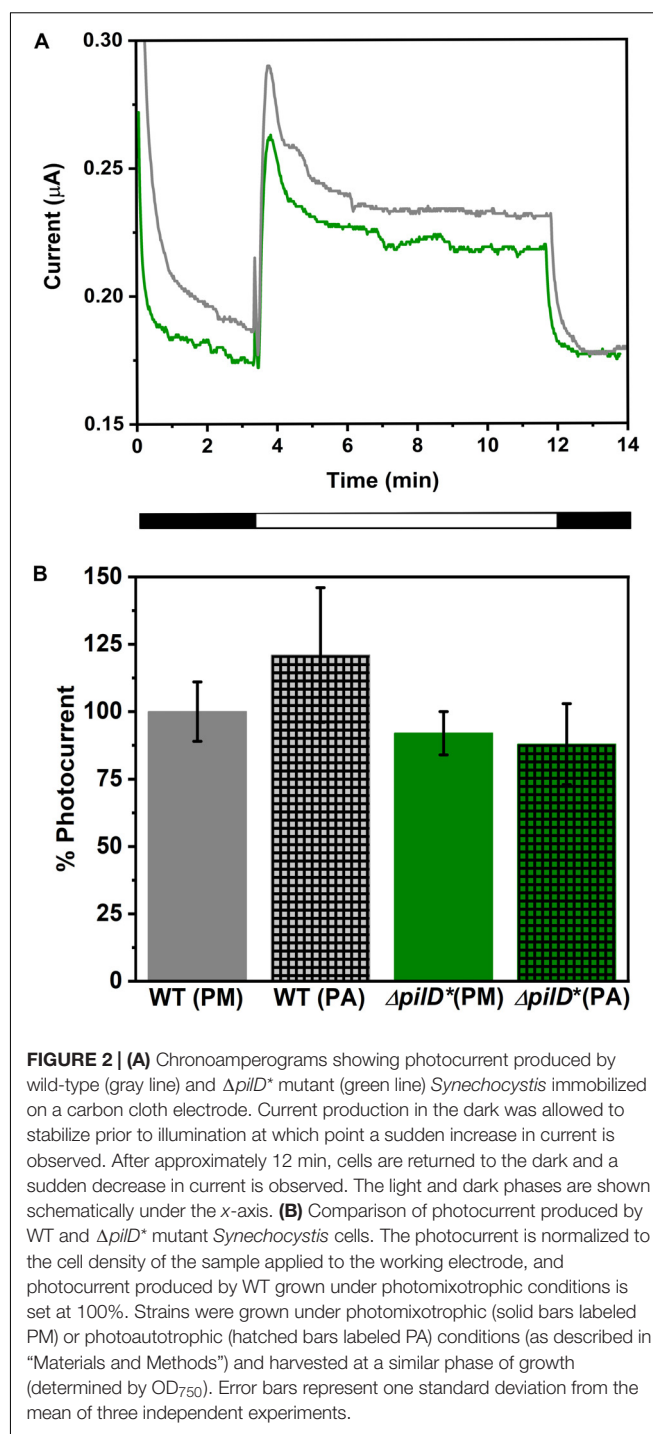
FIGURE 2 | **(A)** Chronoamperograms showing photocurrent produced by wild-type (gray line) and Δ*pilD** mutant (green line) *Synechocystis* immobilized on a carbon cloth electrode. Current production in the dark was allowed to stabilize prior to illumination at which point a sudden increase in current is observed. After approximately 12 min, cells are returned to the dark and a sudden decrease in current is observed. The light and dark phases are shown schematically under the x-axis. **(B)** Comparison of photocurrent produced by WT and Δ*pilD** mutant *Synechocystis* cells. The photocurrent is normalized to the cell density of the sample applied to the working electrode, and photocurrent produced by WT grown under photomixotrophic conditions is set at 100%. Strains were grown under photomixotrophic (solid bars labeled PM) or photoautotrophic (hatched bars labeled PA) conditions (as described in "Materials and Methods") and harvested at a similar phase of growth (determined by OD$_{750}$). Error bars represent one standard deviation from the mean of three independent experiments.

retracts from the sample. Thick and thin pili are clearly visible in the topographical images (**Figures 5A,B**). The diameters of the thin (**Figure 5A**) and thick (**Figure 5B**) pili were obtained from AFM height measurements as 3 and 6 nm, respectively. Note that the heights, rather than the apparent widths, were used to estimate the diameters, since AFM lateral measurements are subject to tip convolution artifacts resulting in a significant broadening of structures. There are no current
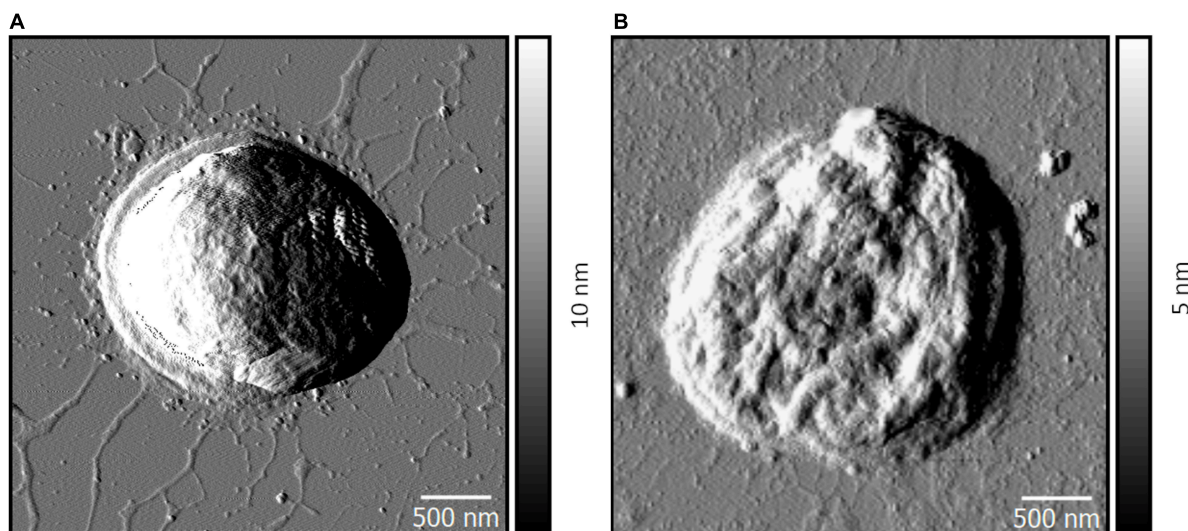
FIGURE 3 | Representative AFM amplitude images of wild-type **(A)** and Δ*pilD** **(B)** *Synechocystis* cells.
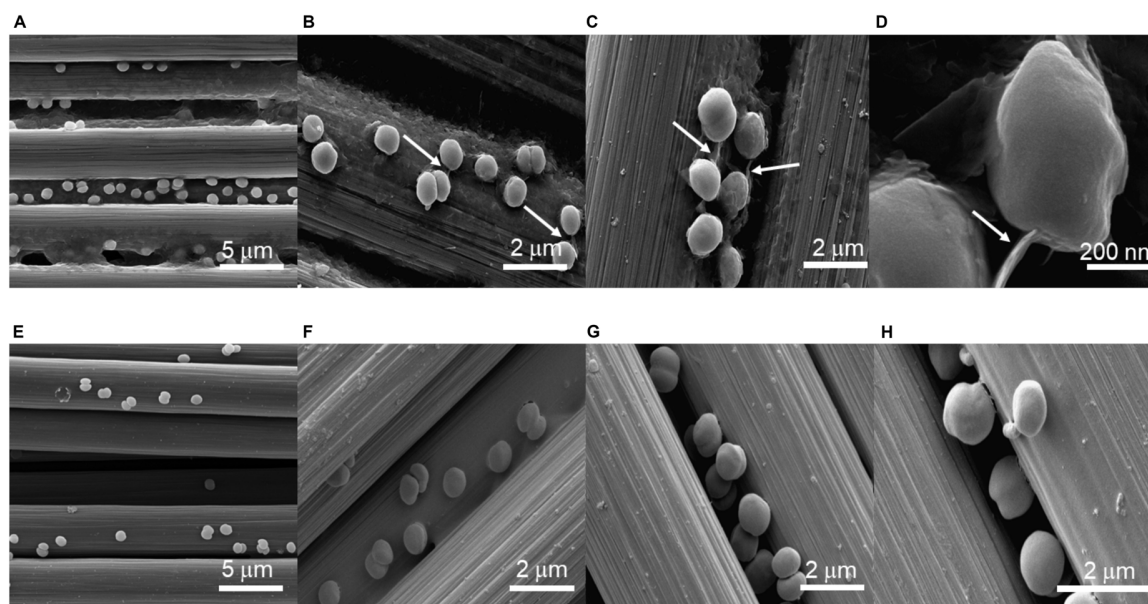
FIGURE 4 | Scanning electron micrographs of wild-type **(A–D)** and Δ*pilD** **(E–H)** *Synechocystis* cells immobilized on a carbon cloth electrode. Arrows in panels **(B–D)** point to structures consistent with pili.

readings along the lengths of pili in the current map images (**Figures 5C,D**). Representative point measurements of current during probe approach and retraction (**Figures 5E,F**) show pili current readings comparable to background values when the probe contacts the pili with the same force used to observe current readings from the Au substrate. Our results indicate that, within the sensitivity of our instrumentation, *Synechocystis* pili are not conductive. We note that AFM measurements were made with dried cells and conductivity may differ under other conditions.

DISCUSSION

Conductive pili are hypothesized to be important for long-range electron transport by various microorganisms including dissimilatory metal-reducing bacteria such as *G. sulfurreducens*. Gorby et al. (2006) reported scanning tunneling microscopy images suggesting that, under CO_2 limitation, *Synechocystis* also produces such conductive filaments. However, controversy exists as to whether the structures they observed are true Tfp assemblies. Lovley (2012) has suggested the diameter of the

FIGURE 5 | Topographical atomic force microscopy images of a thin **(A)** and a thick **(B)** *Synechocystis* pilus. Current maps of the thin **(C)** and the thick **(D)** *Synechocystis* pili shown in **(A,B)**. The current map in **(C)** shows a zoomed in region of the thin pilus, approximated by the box in **(A)**. Representative current versus time curves during probe approach and retraction over the pili (red) and over the Au substrate (black) for the thin **(E)** and thick **(F)** pili. Cross marks in the topographical images indicate the locations where the curves were measured over the pili (red) and over the Au substrates (black). The scale bars in **(A,B,D)** indicate 200 nm. The scale bar in **(C)** indicates 100 nm.

filaments is too large for Tfp. Furthermore, it is hypothesized that similar structures observed in *S. oneidensis* by Gorby et al. (2006) in the same study are filamentous extracellular polysaccharides that arise as an artifact of dehydration during sample preparation or imaging (Dohnalkova et al., 2011). Finally, although appendages produced by *S. oneidensis* have been shown to be conductive under dry conditions (Gorby et al., 2006; El-Naggar et al., 2010), additional work has shown that nanowires of *S. oneidensis* MR-1 are not pili but rather outer membrane extensions containing the multiheme cytochrome conduits of EET (Pirbadian et al., 2014). Consistent with these findings, experiments with mutant strains of *S. oneidensis* have shown that pili are not required for EET (Bouhenni et al., 2010). Thus, the potential role of pili in EET in cyanobacteria such as *Synechocystis* was ambiguous and warranted investigation.

The results herein show that our Δ*pilD** strain, which lacks the *pilD* gene and is unable to synthesize mature pili,

produces a similar amount of light-dependent current as wild-type *Synechocystis* in a mediatorless biophotovoltaic device. Given that the rate of photo-electron production by PSII was shown to be similar in the mutant and wild-type using oxygen evolution measurements, we conclude that, at least under the conditions used in this study, pili are not required for photocurrent production. In support of this conclusion, our AFM-based electrical measurements suggest that neither thick nor thin pili of *Synechocystis* are conductive. Microbial cell-to-electrode electron transfer by *Synechocystis* must therefore be facilitated by an alternative, i.e., non-pili-mediated, mechanism, either by direct transfer from some other cell surface electron transport proteins or by mediated-transfer via unknown redox-shuttles excreted into the extracellular environment/electrolyte (Saper et al., 2018; Wenzel et al., 2018). Secreted flavins have been detected in cultures of *Shewanella* and other bacteria and are believed to play a role in EET by serving as soluble redox mediators (Okamoto et al., 2013; Tian et al., 2019).

We confirmed direct contact between *Synechocystis* cells and the carbon cloth electrode with high-resolution SEM images. This demonstrates that the absence of pili in the Δ*pilD** mutant cells does not appear to affect the adhesion of the mutant cells to the electrode surface, and mediated electron transfer may be more important in cyanobacteria than electron transfer via direct contact between cells and the electrode. Wenzel et al. (2018) elegantly demonstrated that bio-anodes with mesopores large enough to accommodate cells, thereby providing an increase in the direct contact area between the bacteria and the electrode surface, showed only a small increase in current generation compared to nanoporous electrodes, which are not directly accessible to the relatively large cells but provide an increased surface area for interactions with soluble redox-carriers. Coupled with our demonstration that pili do not appear to be necessary for EET, it appears most likely that cyanobacteria use a redox shuttle-mediated mechanism for electron transfer from the bacteria to the electrode rather than a direct electron transfer, or both mechanisms may be important under different growth conditions or environmental stresses. Identifying the components responsible for the reduction of the extracellular environment by cyanobacteria is a crucial next step, both for exploiting cyanobacterial EET and determining the role of this phenomenon in natural systems.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

JH, TB, and AJ conceived the study and designed the research. AH and JH generated and characterized the Δ*pilD* mutant. MT, JL, BD, and RR performed or analyzed the atomic

force microscopy. MT and AJ performed or analyzed the scanning electron microscopy. MC and ME-N performed conductive AFM. MT, AC, and AJ performed or analyzed the electrochemical experiments. MT, AH, JH, TB, and AJ wrote the manuscript, which was edited and approved for submission by all the other authors.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb. 2020.01344/full#supplementary-material

REFERENCES

Bhaya, D., Bianco, N. R., Bryant, D., and Grossman, A. (2000). Type IV pilus biogenesis and motility in the cyanobacterium *Synechocystis* sp. PCC 6803. *Mol. Microbiol.* 37, 941–951. doi: 10.1046/j.1365-2958.2000.02068.x

Bombelli, P., Bradley, R. W., Scott, A. M., Philips, A. J., McCormick, A. J., Cruz, S. M., et al. (2011). Quantitative analysis of the factors limiting solar power transduction by *Synechocystis* sp. PCC 6803 in biological photovoltaic devices. *Energy Environ. Sci.* 4, 4690–4698. doi: 10.1039/C1EE02531G

Bombelli, P., Müller, T., Herling, T. W., Howe, C. J., and Knowles, T. P. (2015). A high power-density, mediator-free, microfluidic biophotovoltaic device for cyanobacterial cells. *Adv. Energy Mater.* 5:1401299. doi: 10.1002/aenm.201401299

Bombelli, P., Zarrouati, M., Thorne, R. J., Schneider, K., Rowden, S. J., Ali, A., et al. (2012). Surface morphology and surface energy of anode materials influence power outputs in a multi-channel mediatorless bio-photovoltaic (BPV) system. *Phys. Chem. Chem. Phys.* 14, 12221–12229. doi: 10.1039/c2cp42526b

Bouhenni, R. A., Vora, G. J., Biffinger, J. C., Shirodkar, S., Brockman, K., Ray, R., et al. (2010). The role of *Shewanella oneidensis* MR-1 outer surface structures in extracellular electron transfer. *Electroanalysis* 22, 856–864. doi: 10.1002/elan.200880006

Bradley, R. W., Bombelli, P., Lea-Smith, D. J., and Howe, C. J. (2013). Terminal oxidase mutants of the cyanobacterium *Synechocystis* sp. PCC 6803 show increased electrogenic activity in biological photo-voltaic systems. *Phys. Chem. Chem. Phys.* 15, 13611–13618. doi: 10.1039/c3cp52438h

Bradley, R. W., Bombelli, P., Rowden, S. J., and Howe, C. J. (2012). Biological photovoltaics: intra- and extra-cellular electron transport by cyanobacteria. *Biochem. Soc. Trans.* 40, 1302–1307. doi: 10.1042/BST20120118

Brutinel, E. D., and Gralnick, J. A. (2012). Shuttling happens: soluble flavin mediators of extracellular electron transfer in *Shewanella*. *Appl. Microbiol. Biotechnol.* 93, 41–48. doi: 10.1007/s00253-011-3653-0

Call, T. P., Carey, T., Bombelli, P., Lea-Smith, D. J., Hooper, P., Howe, C. J., et al. (2017). Platinum-free, graphene based anodes and air cathodes for single chamber microbial fuel cells. *J. Mater. Chem. A* 5, 2382–23886. doi: 10.1039/c7ta06895f

Cereda, A., Hitchcock, A., Symes, M. D., Cronin, L., Bibby, T. S., and Jones, A. K. (2014). A bioelectrochemical approach to characterize extracellular electron transfer by *Synechocystis* sp. PCC6803. *PLoS One* 9:e91484. doi: 10.1371/journal.pone.0091484

Chen, G. E., Hitchcock, A., Jackson, P. J., Chaudhuri, R. R., Dickman, M. J., Hunter, C. N., et al. (2016). Two unrelated 8-vinyl reductases ensure production of mature chlorophylls in *Acaryochloris marina*. *J. Bacteriol.* 198, 1393–1400. doi: 10.1128/JB.00925-15

Chen, Z., Li, X., Tan, X., Zhang, Y., and Wang, B. (2020). Recent advances in biological functions of thick pili in the cyanobacterium *Synechocystis* sp. PCC 6803. *Front. Plant Sci.* 11:241. doi: 10.3389/fpls.2020.00241

Ding, Q., Chen, G., Wang, Y., and Wei, D. (2015). Identification of specific variations in a non-motile strain of cyanobacterium *Synechocystis* sp. PCC 6803 originated from ATCC 27184 by whole genome resequencing. *Int. J. Mol. Sci.* 16, 24081–24093. doi: 10.3390/ijms161024081

Dohnalkova, A. C., Marshall, M. J., Arey, B. W., Williams, K. H., Buck, E. C., and Fredrickson, J. K. (2011). Imaging hydrated microbial extracellular polymers: comparative analysis by electron microscopy. *Appl. Environ. Microbiol.* 77, 1254–1262.

El-Naggar, M. Y., Wanger, G., Leung, K. M., Yuzvinsky, T. D., Southam, G., Yang, J., et al. (2010). Electrical transport along bacterial nanowires from *Shewanella oneidensis* MR-1. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18127–18131. doi: 10.1073/pnas.1004880107

Filman, D. J., Marino, S. F., Ward, J. E., Yang, L., Mester, Z., Bullitt, E., et al. (2019). Cryo-EM reveals the structural basis of long-range electron transport in a cytochrome-based bacterial nanowire. *Commun. Biol.* 2:219. doi: 10.1038/s42003-019-0448-9

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A. III, and Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6, 343–345. doi: 10.1038/nmeth.1318

Glasser, N. R., Saunders, S. H., and Newman, D. K. (2017). The colorful world of extracellular electron shuttles. *Annu. Rev. Microbiol.* 71, 731–751.

Gorby, Y. A., Yanina, S., McLean, J. S., Rosso, K. M., Moyles, D., Dohnalkova, A., et al. (2006). Electrically conductive bacterial nanowires produced by *Shewanella oneidensis* strain MR-1 and other microorganisms. *Proc. Natl. Acad. Sci. U.S.A.* 103, 11358–11363. doi: 10.1073/pnas.0604517103

Gray, H. B., and Winkler, J. R. (2005). Long-range electron transfer. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3534–3539. doi: 10.1073/pnas.0408029102

Heidary, N., Kornienko, N., Kalathil, S., Fang, X., Ly, K. H., Greer, H. F., et al. (2020). Disparity of cytochrome utilization in anodic and cathodic extracellular electron transfer pathways of *Geobacter sulfurreducens* biofilms. *J. Am. Chem. Soc.* 142, 5194–5203. doi: 10.1021/jacs.9b13077

Hernandez, M. E., and Newman, D. K. (2001). Extracellular electron transfer. *Cell. Mol. Life Sci.* 58, 1562–1571. doi: 10.1007/PL00000796

Holmes, D. E., Dang, Y., Walker, D. J., and Lovley, D. R. (2016). The electrically conductive pili of *Geobacter* species are a recently evolved feature for extracellular electron transfer. *Microb. Genom.* 2:e000072. doi: 10.1099/mgen.0.000072

Ikeuchi, M., and Tabata, S. (2001). *Synechocystis* sp. PCC 6803 — a useful tool in the study of the genetics of cyanobacteria. *Photosynth. Res.* 70, 73–83. doi: 10.1023/A:1013887908680

Ing, N. L., El-Naggar, M. Y., and Hochbaum, A. I. (2018). Going the distance: long-range conductivity in protein and peptide bioelectronic materials. *J. Phys. Chem. B* 122, 10403–10423. doi: 10.1021/acs.jpcb.8b07431

Kamei, A., Yuasa, T., Orikawa, K., Geng, X. X., and Ikeuchi, M. (2001). A eukaryotic-type protein kinase, SpkA, is required for normal motility of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* 183, 1505–1510. doi: 10.1128/JB.183.5.1505-1510.2001

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., et al. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* 30, 109–136. doi: 10.1093/dnares/3.3.109

Kanesaki, Y., Shiwa, Y., Tajima, N., Suzuki, M., Watanabe, S., Sato, N., et al. (2012). Identification of substrain-specific mutations by massively parallel whole-genome resequencing of *Synechocystis* sp. PCC 6803. *DNA Res.* 19, 67–79. doi: 10.1093/dnares/dsr042

Kotloski, N. J., and Gralnick, J. A. (2013). Flavin electron shuttles dominate extracellular electron transfer by *Shewanella oneidensis*. *mBio* 4:e00553-12. doi: 10.1128/mBio.00553-12

Kracke, F., Vassilev, I., and Krömer, J. O. (2015). Microbial electron transport and energy conservation - the foundation for optimizing bioelectrochemical systems. *Front. Microbiol.* 6:575. doi: 10.3389/fmicb.2015.00575

Kranzler, C., Lis, H., Shaked, Y., and Keren, N. (2011). The role of reduction in iron uptake processes in a unicellular, planktonic cyanobacterium. *Environ. Microbiol.* 13, 2990–2999. doi: 10.1111/j.1462-2920.2011.02572.x

Lamb, J. J., Hill, R. E., Eaton-Rye, J. J., and Hohmann-Marriott, M. F. (2014). Functional role of PilA in iron acquisition in the cyanobacterium *Synechocystis* sp. PCC 6803. *PLoS One* 9:e105761. doi: 10.1371/journal.pone.0105761

Lamb, J. J., and Hohmann-Marriott, M. F. (2017). Manganese acquisition is facilitated by PilA in the cyanobacterium *Synechocystis* sp. PCC 6803. *PLoS One* 12:e0184685. doi: 10.1371/journal.pone.0184685

Lea-Smith, D. J., Bombelli, P., Vasudevan, R., and Howe, C. J. (2015). Photosynthetic, respiratory and extracellular electron transport pathways in cyanobacteria. *Biochim. Biophys. Acta* 1857, 247–255. doi: 10.1016/j.bbabio.2015.10.007

Linhartová, M., Bučinská, L., Halada, P., Ječmen, T., Setlík, J., Komenda, J., et al. (2014). Accumulation of the type IV prepilin triggers degradation of SecY and YidC and inhibits synthesis of photosystem II proteins in the cyanobacterium *Synechocystis* PCC 6803. *Mol. Microbiol.* 93, 1207–1223. doi: 10.1111/mmi.12730

Lis, H., Kranzler, C., Keren, N., and Shaked, Y. (2015). A comparative study of iron uptake rates and mechanisms amongst marine and fresh water cyanobacteria: prevalence of reductive iron uptake. *Life* 5, 841–860. doi: 10.3390/life5010841

Lovley, D. R. (2012). Electromicrobiology. *Annu. Rev. Microbiol.* 66, 391–409. doi: 10.1146/annurev-micro-092611-150104

Lovley, D. R., and Walker, D. J. F. (2019). *Geobacter* protein nanowires. *Front. Microbiol.* 10:2078. doi: 10.3389/fmicb.2019.02078

Marcus, R. A., and Sutin, N. (1985). Electron transfer in chemistry and biology. *Biochim. Biophys. Acta* 811, 265–322. doi: 10.1016/0304-4173(85)90014-X

Marsili, E., Baron, D. B., Shikhare, I. D., Coursolle, D., Gralnick, J. A., and Bond, D. R. (2008). *Shewanella* secretes flavins that mediate extracellular electron transfer. *Proc. Natl. Acad. Sci. U.S.A.* 105, 3968–3973. doi: 10.1073/pnas.0710525105

McCormick, A. J., Bombelli, P., Bradley, R. W., Thorne, R., Wenzel, T., and Howe, C. J. (2015). Biophotovoltaics: oxygenic photosynthetic organisms in the world of bioelectrochemical systems. *Energy Environ. Sci.* 8, 1092–1109. doi: 10.1039/C4EE03875D

McCormick, A. J., Bombelli, P., Lea-Smith, D. J., Bradley, R. W., Scott, A. M., Fisher, A. C., et al. (2013). Hydrogen production through oxygenic photosynthesis using the cyanobacterium *Synechocystis* sp PCC 6803 in a bio-photoelectrolysis cell (BPE) system. *Energy Environ. Sci.* 6, 2682–2690. doi: 10.1039/C3EE40491A

McCormick, A. J., Bombelli, P., Scott, A. M., Phillips, A. J., Smith, A. G., Fischer, A. C., et al. (2011). Photosynthetic biofilms in pure culture harness solar energy in a mediatorless bio-photovoltaic cell (BPV) system. *Energy Environ. Sci.* 4, 4699–4709. doi: 10.1039/C1EE01965A

Morris, J. N., Crawford, T. S., Jeffs, A., Stockwell, P. A., Eaton-Rye, J. J., and Summerfield, T. C. (2014). Whole genome re-sequencing of two 'wild-type' strains of the model cyanobacterium *Synechocystis* sp. PCC 6803. *N. Z. J. Bot.* 52, 36–47. doi: 10.1080/0028825X.2013.846267

Okamoto, A., Hashimoto, K., Nealson, K. H., and Nakamura, R. (2013). Rate enhancement of bacterial extracellular electron transport involves bound flavin semiquinones. *Proc. Natl. Acad. Sci. U.S.A.* 110, 7856–7861. doi: 10.1073/pnas.1220823110

Pinto, F. L., Thapper, A., Sontheim, W., and Lindblad, P. (2009). Analysis of current and alternative phenol based RNA extraction methodologies for cyanobacteria. *BMC Mol. Biol.* 10:79. doi: 10.1186/1471-2199-10-79

Pirbadian, S., Barchinger, S. E., Leung, K. M., Byun, H. S., Jangir, Y., Bouhenni, R. A., et al. (2014). *Shewanella oneidensis* MR-1 nanowires are outer membrane and periplasmic extensions of the extracellular electron transport components. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12883–12888. doi: 10.1073/pnas.1410551111

Pisciotta, J. M., Zou, Y., and Baskakov, I. V. (2010). Light-dependent electrogenic activity of cyanobacteria. *PLoS One* 5:e10821. doi: 10.1371/journal.pone.0010821

Pisciotta, J. M., Zou, Y., and Baskakov, I. V. (2011). Role of the photosynthetic electron transfer chain in electrogenic activity of cyanobacteria. *Appl. Microbiol. Biotechnol.* 91, 377–385. doi: 10.1007/s00253-011-3239-x

Polyviou, D., Baylay, A. J., Hitchcock, A., Robidart, J., Moore, C. M., and Bibby, T. S. (2018). Desert dust as a source of iron to the globally important diazotroph *Trichodesmium*. *Front. Microbiol.* 8:2683. doi: 10.3389/fmicb.2017.02683

Polyviou, D., Hitchcock, A., Baylay, A. J., Moore, C. M., and Bibby, T. S. (2015). Phosphite utilization by the globally important marine diazotroph *Trichodesmium*. *Environ. Microbiol. Rep.* 7, 824–830. doi: 10.1111/1758-2229.12308

Porra, R. J., Thompson, W. A., and Kriedemann, P. E. (1989). Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls *a* and *b* extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim. Biophys. Acta* 975, 384–394. doi: 10.1016/S0005-2728(89)80347-0

Rabaey, K., and Verstraete, W. (2005). Microbial fuel cells: novel biotechnology for energy generation. *Trends Biotechnol.* 23, 291–298. doi: 10.1016/j.tibtech.2005.04.008

Reguera, G., McCarthy, K. D., Mehta, T., Nicoll, J. S., Tuominen, M. T., and Lovley, D. R. (2005). Extracellular electron transfer via microbial nanowires. *Nature* 435, 1098–1101. doi: 10.1038/nature03661

Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M., and Stanier, R. Y. (1979). Generic assignments, strain histories and properties of pure culture of cyanobacteria. *Microbiology* 111, 1–61. doi: 10.1099/00221287-111-1-1

Rowden, S. J. L., Bombelli, P., and Howe, C. J. (2018). Biophotovoltaics: design and study of bioelectrochemical systems for biotechnological applications and metabolic investigation. *Methods Mol. Biol.* 1770, 335–346. doi: 10.1007/978-1-4939-7786-4_20

Saar, K. L., Bombelli, P., Lea-Smith, D. J., Call, T., Aro, E. M., Müller, T., et al. (2018). Enhancing power density of biophotovoltaics by decoupling storage and power delivery. *Nat. Energy* 3, 75–81. doi: 10.1038/s41560-017-0073-0

Saper, G., Kallmann, D., Conzuelo, F., Zhao, F., Tóth, T. N., Liveanu, V., et al. (2018). Live cyanobacteria produce photocurrent and hydrogen using both the respiratory and photosynthetic systems. *Nat. Commun.* 9, 2168. doi: 10.1038/s41467-018-04613-x

Schuergers, N., and Wilde, A. (2015). Appendages of the cyanobacterial cell. *Life* 5, 700–715. doi: 10.3390/life5010700

Shi, L., Richardson, D. J., Wang, Z., Kerisit, S. N., Rosso, K. M., Zachara, J. M., et al. (2009). The roles of outer membrane cytochromes of *Shewanella* and *Geobacter* in extracellular electron transfer. *Environ. Microbiol. Rep.* 1, 220–227. doi: 10.1111/j.1758-2229.2009.00035.x

Strom, M. S., Nunn, D. N., and Lory, S. (1993). A single bifunctional enzyme, PilD, catalyzes cleavage and N-methylation of proteins belonging to the type IV pilin family. *Proc. Natl. Acad. Sci. U.S.A.* 90, 2404–2408. doi: 10.1073/pnas.90.6.2404

Tajima, N., Sato, S., Maruyama, F., Kaneko, T., Sasaki, N. V., Kurokawa, K., et al. (2011). Genomic structure of the cyanobacterium *Synechocystis* sp. PCC 6803 strain GT-S. *DNA Res.* 18, 393–399. doi: 10.1093/dnares/dsr026

Thorne, R., Hu, H., Schneider, K., Bombelli, P., Fischer, A., Peter, L. M., et al. (2011). Porous ceramic anode materials for photo-microbial fuel cells. *J. Mater. Chem.* 21, 18055–18060. doi: 10.1039/C1JM13058G

Tian, T., Fan, X., Feng, M., Su, L., Zhnag, W., Chi, H., et al. (2019). Flavin-mediated extracellular electron transfer in Gram-positive bacteria *Bacillus cereus* DIF1 and *Rhodococcus ruber* DIF2. *RSC Adv.* 9, 40903–40909. doi: 10.1039/C9RA08045G

Trautmann, D., Voß, B., Wilde, A., Al-Babili, S., and Hess, W. R. (2012). Microevolution in cyanobacteria: re-sequencing a motile substrain of *Synechocystis* sp. PCC 6803. *DNA Res.* 19, 435–448. doi: 10.1093/dnares/dss024

Tschörtner, J., Lai, B., and Krömer, J. O. (2019). Biophotovoltaics: green power generation from sunlight and water. *Front. Microbiol.* 10:866. doi: 10.3389/fmicb.2019.00866

Wang, F., Gu, Y., O'Brien, J. P., Yi, S. M., Yalcin, S. E., Srikanth, V., et al. (2019). Structure of microbial nanowires reveals stacked hemes that transport electrons over micrometers. *Cell* 177, 361–369. doi: 10.1016/j.cell.2019.03.029

Watanabe, K., Manefield, M., Lee, M., and Kouzuma, A. (2009). Electron shuttles in biotechnology. *Curr. Opin. Biotechnol.* 20, 633–641. doi: 10.1016/j.copbio.2009.09.006

Wenzel, T., Härtler, D., Bombelli, P., Howe, C. J., and Steiner, U. (2018). Porous translucent electrodes enhance current generation from photosynthetic biofilms. *Nat. Commun.* 9:1299. doi: 10.1038/s41467-018-03320-x

Wey, L. T., Bombelli, P., Chen, X., Lawrence, J. M., Rabideau, C. M., Rowden, S. J. L., et al. (2019). The development of biophotovoltaic systems for power generation and biological analysis. *ChemElectroChem* 6, 5375–5386. doi: 10.1002/celc.201900997

Yang, Y., Ding, Y., Hu, Y., Cao, B., Rice, S. A., Kjelleberg, S., et al. (2015). Enhancing bidirectional electron transfer of *Shewanella oneidensis* by a synthetic flavin pathway. *ACS Synth. Biol.* 4, 815–823. doi: 10.1021/sb500331x

Yoshihara, S., Geng, X., Okamoto, S., Yura, K., Murata, T., Go, M., et al. (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. doi: 10.1093/pcp/pce007

Zou, Y., Pisciotta, J., Billmyre, R. B., and Baskakov, I. V. (2009). Photosynthetic microbial fuel cells with positive light response. *Biotechnol. Bioeng.* 104, 939–946. doi: 10.1002/bit.22466

Systematic Identification of Target Genes for Cellular Morphology Engineering in *Synechococcus elongatus* PCC7942

Mingyi Zhang[1,2,3], Cuncun Qiao[1,2,3], Guodong Luan[1,2,3]*, Quan Luo[1,2]* and Xuefeng Lu[1,2,3,4,5]

[1] Key Laboratory of Biofuels, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, China, [2] Shandong Provincial Key Laboratory of Synthetic Biology, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao, China, [3] University of Chinese Academy of Sciences, Beijing, China, [4] Dalian National Laboratory for Clean Energy, Dalian, China, [5] Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China

Cyanobacteria are serving as promising microbial platforms for development of photosynthetic cell factories. For enhancing the economic competitiveness of the photosynthetic biomanufacturing technology, comprehensive improvements on industrial properties of the cyanobacteria chassis cells and engineered strains are required. Cellular morphology engineering is an up-and-coming strategy for development of microbial cell factories fitting the requirements of industrial application. In this work, we performed systematic evaluation of potential genes for cyanobacterial cellular morphology engineering. Twelve candidate genes participating in cell morphogenesis of an important model cyanobacteria strain, *Synechococcus elongatus* PCC7942, were knocked out/down and overexpressed, respectively, and the influences on cell sizes and cell shapes were imaged and calculated. Targeting the selected genes with potentials for cellular morphology engineering, the controllable cell lengthening machinery was also explored based on the application of sRNA approaches. The findings in this work not only provided many new targets for cellular morphology engineering in cyanobacteria, but also helped to further understand the cell division process and cell elongation process of *Synechococcus elongatus* PCC7942.

Keywords: cyanobacteria, cellular morphology, riboswitch, photosynthesis, cellular length

INTRODUCTION

Cyanobacteria emerged as the simplest and the most ancient oxygen-evolving phototrophs, paving the way for evolution of other aerobiont on the planet, and meantime contributing a large portion of the oxygen to the current biosphere environment (Flombaum et al., 2013; Rousseaux and Gregg, 2014). The flexible physiological and metabolic networks permit cyanobacteria significant potentials to acclimate to changeable environments and diverse ecosystems, including land, ocean, fresh water and polar regions (Waterbury et al., 1979; Demarsac and Houmard, 1993). By performing high efficient photosynthesis, cyanobacteria capture solar energy and carbon dioxide for production of diverse organic compounds, accounting for up to 20% of the primary production

within the scope of the global (Hagemann, 2011; Flombaum et al., 2013; Rousseaux and Gregg, 2014). Not only that, cyanobacteria also play important and active roles in the global cycle of other important elements such as nitrogen, phosphorus, and iron (Sohm et al., 2011; Fernandez-Juarez et al., 2019; Wang et al., 2019).

In recent years, due to the unique characteristics such as efficient photosynthesis, rapid growth, simple structure, and convenient genetic manipulations, cyanobacteria serve as promising microbial platforms for artificially designing, constructing, and controlling photosynthesis-driven routes for directional conversion of energy and materials (Lu, 2010; Desai and Atsumi, 2013). Based on the development and utilization of system biology technologies (multiple -omics approaches), massive information about the metabolic profiles and dynamics of cyanobacteria cells under stable and changing environments have been acquired (Aikawa et al., 2019; Lin et al., 2019). Notable improvements in developing efficient tools of synthetic biology and metabolic engineering over the last decade have permitted effective regulation and expansion of the photosynthetic metabolism network (Sun et al., 2018b; Santos-Merino et al., 2019). Through assembling and regulating the native, heterologous, or artificial metabolic pathways in cyanobacteria chassis cells, photosynthetic production of dozens of natural or non-natural metabolites utilizing solar energy and carbon dioxide has been achieved with diverse cyanobacteria cell factories (Desai and Atsumi, 2013). Since far, some of the cyanobacteria cell factories products could be synthesized and accumulated at levels of g/L, accounting for up to 70% of the intracellular photosynthetic carbon flow (Gao et al., 2012, 2016a; Liu et al., 2019). Besides the synthesis capacity of final products, there are some other important traits of the cyanobacteria cell factories influencing the economic competitiveness of the photosynthetic biomanufacturing technology, including the tolerance to environmental stresses, the resistance to biocontaminants, and the convenience for biomass harvesting (Luan and Lu, 2018). To remove the restrictions over practical applications of photosynthetic biomanufacturing, these complex industrial traits of the cyanobacterial cell factories are yet to be significantly improved, which would require comprehensive remodeling of the behaviors and characteristics of the cyanobacteria chassis cells.

Cellular morphology is a basic and essential characteristic of cyanobacteria, as well as other microorganisms, significantly determining some of the important industrial properties of the derived cell factories. Previously it has been reported that modifications on cell surface facilitated cyanobacteria cells to survive in grazing of predators or infection of cyanophages (Xu et al., 1997; Simkovsky et al., 2012). In addition, the size and shape of cyanobacteria cells also significantly influenced the grazing resistances (Young, 2006; Jezberova and Komarkova, 2007) and the recovery characteristics of the photosynthetic cell factories, which is of great significance for economic feasibilities of industrially leveled photosynthetic biomanufacturing (Zamalloa et al., 2011; Chisti, 2013). Engineering cell sizes or shapes through manipulating the node genes influencing or determining cell morphogenesis provided a promising approach to optimize

industrial properties of microbial cell factories (Jiang and Chen, 2016). In heterologous cell factories derived from *Escherichia coli* (*E. coli*) and *Halomonas campaniensis*, enlarged cell sizes and volumes significantly improved the yield of the PHB products and decreased the difficulties in biomass harvesting processes (Wang et al., 2014; Jiang et al., 2015, 2017). With *Synechococcus elongatus* PCC7942 (hereafter PCC7942 for short), a model strain of fresh water cyanobacteria, the concept of "morphology engineering" has also been confirmed. Through controllable expression of the components in Min system (which is participated in regulation of FtsZ protein and the Z-ring structure determining cell morphogenesis), cell lengths of PCC7942 could be extended from several micrometers to near millimeter levels. The elongated cells showed normal-gravity induced sedimentation behaviors and enhanced fragilities with mechanical treatments, which were both expected to enable convenient biomass harvesting and downstream processing in scaled cultivations (Jordan et al., 2017).

To facilitate more accurate and controllable editing of cellular morphology in cyanobacteria as required by an ideal robust industrial process, the identification of effective target genes is as important as the development of manipulation tools. In recent years, more and more accessible synthetic biology tools, including riboswitch (Ohbayashi et al., 2016), CRIPSRi (Yao et al., 2016), and microRNA tools (Sun et al., 2018a), have been developed and widely adopted in cyanobacteria engineering, permitting smart and rapid regulation of target genes. As compared, more effective node genes for morphology engineering in cyanobacteria are yet to be explored and evaluated. In this work, we performed systematic evaluation of the potential genes participating in cell morphogenesis of PCC7942. Combining knockout and overexpression manipulations, the influences of manipulating the corresponding genes on cell sizes and shapes were systematically evaluated and compared. The results provide useful information for designing cyanobacteria cell factories with smartly regulated morphology in future.

RESULTS AND DISCUSSION

Screening the Potential Genes Involved in Cell Morphogenesis of PCC7942

As for rod-shaped bacteria, represented by *E. coli* and *Bacillus subtilis* (*B. subtilis*), cellular morphology is influenced simultaneously by the activities and process of cell division and elongation (Wang et al., 2014; Jiang and Chen, 2016; Jordan et al., 2017). Thus, the genes participating in the two machineries might contribute to the process and outcome of cell morphogenesis, while the disturbance of the expression pattern of these genes might cause significant changes in cell sizes and shapes. Among the morphogenesis machineries of rod-shaped bacterial cells, FtsZ and MreB serve as the most important cellular skeleton proteins, functioning in recruiting and orchestrating the subsequent components in the divisome and elongasome, respectively (Rohs et al., 2018). In PCC7942, two highly conserved homologous of *ftsZ* (*Synpcc7942_2378*) and *mreB* (*Synpcc7942_0300*) have been annotated on the

chromosome, with amino acids sequence identities of as high as 49 and 56% to the homologous of *E. coli*, respectively (**Table 1**).

To facilitate the cellular division process, FtsZ, a tubulin-like GTPase protein, would polymerize to form the Z-ring structure as the skeleton and scaffold of the cellular divisome complex. In cells of typical rod-shaped bacteria, including *E. coli* and *B. subtilis*, multiple components, including FtsA/ZipA/ZapB/SepF, FtsE, FtsK, FtsQ, FtsL, FtsB, FtsW, FtsI, and FtsN, would be further recruited to activate the cellular division process (Errington et al., 2003; Marbouty et al., 2009). FtsA, as an important and conserved cytoplasmic actin-like protein in multiple bacterial species, is responsible for stabilizing the Z-ring structure and recruiting subsequent proteins (Pichoff and Lutkenhaus, 2002). In cyanobacteria, FtsA is missing, while another protein termed as ZipN (*Synpcc7942_1943* in PCC7942), was discovered to work as an FtsA-like orchestrator for divisome assembly (Koksharova and Wolk, 2002; Marbouty et al., 2009; Camargo et al., 2019). A Cdv2 protein (*Synpcc7942_2059*), with 32% sequence similarity to SepF of *B. subtilis* (**Table 1**), has also been identified on the chromosome of PCC7942, which might bring additional contribution to the stabilization of the Z-ring structure (Miyagishima et al., 2005). Some other conserved cellular divisome components, including FtsE (*Synpcc7942_1414*), FtsW (*Synpcc7942_0324*), and FtsI (*Synpcc7942_0482*), have also been annotated in PCC7942, while the other portion are not detected in cyanobacteria (Miyagishima et al., 2005). Previously, it has been reported that the deficiency of these genes participating in cellular divisome led to filamentation of the mutant cells (Miyagishima et al., 2005). Some negative factors, including MinCDE, DivlVA, EzrA, SulA, and Noc would also participate in regulating the cell division process in rod-shaped bacteria by directly interacting with FtsZ to regulate or position the Z-ring structure (Goehring and Beckwith, 2005; Margolin, 2005). The MinCDE system (*Synpcc7942_2001*, *Synpcc7942_0220*, and *Synpcc7942_0897*), and the SulA protein (*Synpcc7942_2477*) has been identified on the chromosome of PCC7942 (Miyagishima et al., 2005; Koksharova and Babykin, 2011). And the Cdv3 protein (*Synpcc7942_2006*), which was reported to be involved in PCC7942 cell division, shows sequence similarity to the DivlVA from *B. subtilis* (Miyagishima et al., 2005). In addition, a periplasm located protein Cdv1 (*Synpcc7942_0653*), with high similarity to peptidyl-prolyl *cis-trans* isomerase (PPlase) was also recognized as a factor related with cell division, although the detailed mechanisms are yet elucidated (Miyagishima et al., 2005). Besides, there are also two cyanobacteria-specific proteins, Ftn6 (*Synpcc7942_1707* in PCC7942) with unknown functions and CikA as the circadian input kinase, identified to influence cell division process (Koksharova and Wolk, 2002; Miyagishima et al., 2005).

In addition to the factors participating in constructing and regulating the cell division process, the synthesis of peptidoglycan, the main component of cell wall, is also an important contributor to cell morphogenesis in *E. coli*. The non-canonical transglycosylase protein RodA has been reported to be involved in the regulation of cell shapes and lengths in many bacterial species, through interaction with the MreB skeleton (Henriques et al., 1998; Arora et al., 2018; Rohs et al., 2018). In

PCC7942, the *Synpcc7942_1104* gene was annotated to encode the functional homolog of RodA.

As mentioned above, the effects of genetic manipulations of MinCDE on cellular morphology in PCC7942 have been elucidated in a previous research (Jordan et al., 2017). Thus, in this work we main evaluated the effects of manipulating the other genes on cell morphogenesis engineering. The detailed information about the candidate genes have been summarized in **Table 1** and the potential interrelationships among the corresponding proteins are presented in **Figure 1** based on previous results and hypotheses.

Exploring the Effects of Knocking Out/Down the Candidate Genes on Cell Morphogenesis of PCC7942

To explore promising targets for morphology engineering, we first tried to knock out the twelve candidate genes (as listed in **Figure 1** and **Table 1**) through homologous recombination in PCC7942 and to observe the changes of cellular morphology of the mutants with microscope. Due to the essential roles on cell division and cell elongation, the *ftsZ* and *mreB* genes could not be completely eliminated from the chromosome of PCC7942. Thus, homozygous transformants carrying complete disruption of *ftsZ* and *mreB* were not obtained through several attempts in this work. The result of quantitative PCR showed that the *ftsZ* and *mreB* mutants still remained the wild type genes (i.e., *ftsZ* and *mreB*) with ratios of about 76 and 39%, respectively (**Supplementary Figures S1A, S2**). The other ten mutants were successfully constructed as designed (**Table 1** and **Supplementary Figure S1A**). In consistence with the crucial role as skeleton bricks of the cell divisome, the deficiency of FtsZ in PCC7942 caused significant influence on cellular morphology, resulting in filamentous cells (**Table 1** and **Figure 2A**). Cells of the three other mutant strains deficient of subsequent cellular divisome components, including FtsI, FtsW, and ZipN, were also filamented, which could not be effectively calculated for cell lengths and areas under microscope (**Figure 2A**). In addition to the above four mutants with filamentous cells, the cell lengths and areas of the other eight mutants (ΔCdv1, ΔCdv2, ΔCdv3, ΔSulA, ΔRodA, ΔMreB, ΔFtn6, and ΔFtsE) were calculated and summarized in **Table 1** and **Figures 2B,C**. The Cdv2 was predicated to perform similar functions of ZipN by stabilizing the structure of Z-ring, while when *cdv2* was disrupted, the effects on cellular morphology of PCC7942 was not as significant as that of *zipN*. The cell length and area of the ΔCdv2 mutant was just increased by 1.7- and 2-fold, respectively than these of the wild type (**Table 1**). Knockout of *ftsE*, another potential factor involved in cell divisome, caused minor changes in cell size and cell shape. The general phenomenon of cell filamentation in the deficient mutants of the components participating in the cellular divisome suggested the essential role of cell division on maintaining the short rod shape of PCC7942 cells, by preventing the formation of excessively long cells from the horizontal axis.

As compared, RodA and MreB displayed more significant importance on maintaining of cell shape from vertical axis. When

TABLE 1 | Selection and engineering of candidate genes potentially participating in cell morphogenesis in *Synechococcus elongatus* PCC7942.

Proteins known to regulate cell shape			Homologous proteins in *Synechococcus elongatus* PCC7942									
Protein	Host strain	Identity (Query cover) with *S. elongatus*	Protein	Gene ID	Gene Knockout/Knockdown				Gene Overexpression			
					Cell length relative to WT (%)	Cell area relative to WT (%)	Cell shape	Growth relative to WT (% ± SD)	Cell length relative to WT (%)	Cell area relative to WT (%)	Cell shape	Growth relative to WT (% ± SD)
FtsZ	*E. coli*	49% (98%)	FtsZ	*Synpcc7942_2378*	Filamentous	Filamentous	Filamentous	55 ± 29	46	39	Short rod-shaped	48 ± 5
PPIA	*E. coli*	32% (77%)	Cdv1	*Synpcc7942_0653*	513	500	Elongated	96 ± 13	130	121	WT	85 ± 19
RodA	*E. coli*	33% (94%)	RodA	*Synpcc7942_1104*	53	102	Round	106 ± 6	130	118	WT	90 ± 20
MreB	*E. coli*	56% (96%)	MreB	*Synpcc7942_0300*	61	133	Round	83 ± 17	95	106	Spindle-shaped	70 ± 1
FtsE	*E. coli*	43% (97%)	FtsE	*Synpcc7942_1414*	95	99	WT	89 ± 3	133	118	WT	102 ± 7
SepF	*B. subtilis*	32% (90%)	Cdv2	*Synpcc7942_2059*	168	203	Elongated	98 ± 3	114	91	WT	92 ± 26
ZipN	*Synechocystis*	42% (72%)	ZipN	*Synpcc7942_1943*	Filamentous	Filamentous	Filamentous	98 ± 8	623	666	Elongated	104 ± 6
Cdv3	*Synechocystis*	35% (75%)	Cdv3	*Synpcc7942_2006*	542	732	Elongated	83 ± 2	576	760	Elongated	89 ± 3
SulA	*Synechocystis*	56% (99%)	SulA	*Synpcc7942_2477*	123	134	WT	95 ± 7	124	103	WT	98 ± 0
Ftn6	*Synechocystis*	37% (37%)	Ftn6	*Synpcc7942_1707*	855	1155	Filamentous	107 ± 2	141	115	WT	70 ± 2
FtsI	*Synechocystis*	46% (94%)	FtsI	*Synpcc7942_0482*	Filamentous	Filamentous	Filamentous	100 ± 9	130	128	WT	93 ± 12
FtsW	*Synechococcus sp.* PCC 7002	49% (93%)	FtsW	*Synpcc7942_0324*	Filamentous	Filamentous	Filamentous	29 ± 0	124	104	WT	68 ± 13

E. coli, Escherichia coli; Synechocystis, Synechocystis sp. PCC 6803; B. subtilis, Bacillus subtilis; Gene ID, Gene identifier in CyanoBase; Chla, Chlorophyll a [Log_2(Day 7 Chla/Day 3 Chla)];% identity (Query cover) with S. elongatus was calculated based on BLASTP program.

FIGURE 1 | Proposed working model for the factors potentially participating in cell morphogenesis of PCC7942. The interactions, locations, and components were selected and illustrated referencing previous discoveries in *Escherichia coli*, *Bacillus subtilis*, and *Synechocystis* sp. PCC6803.
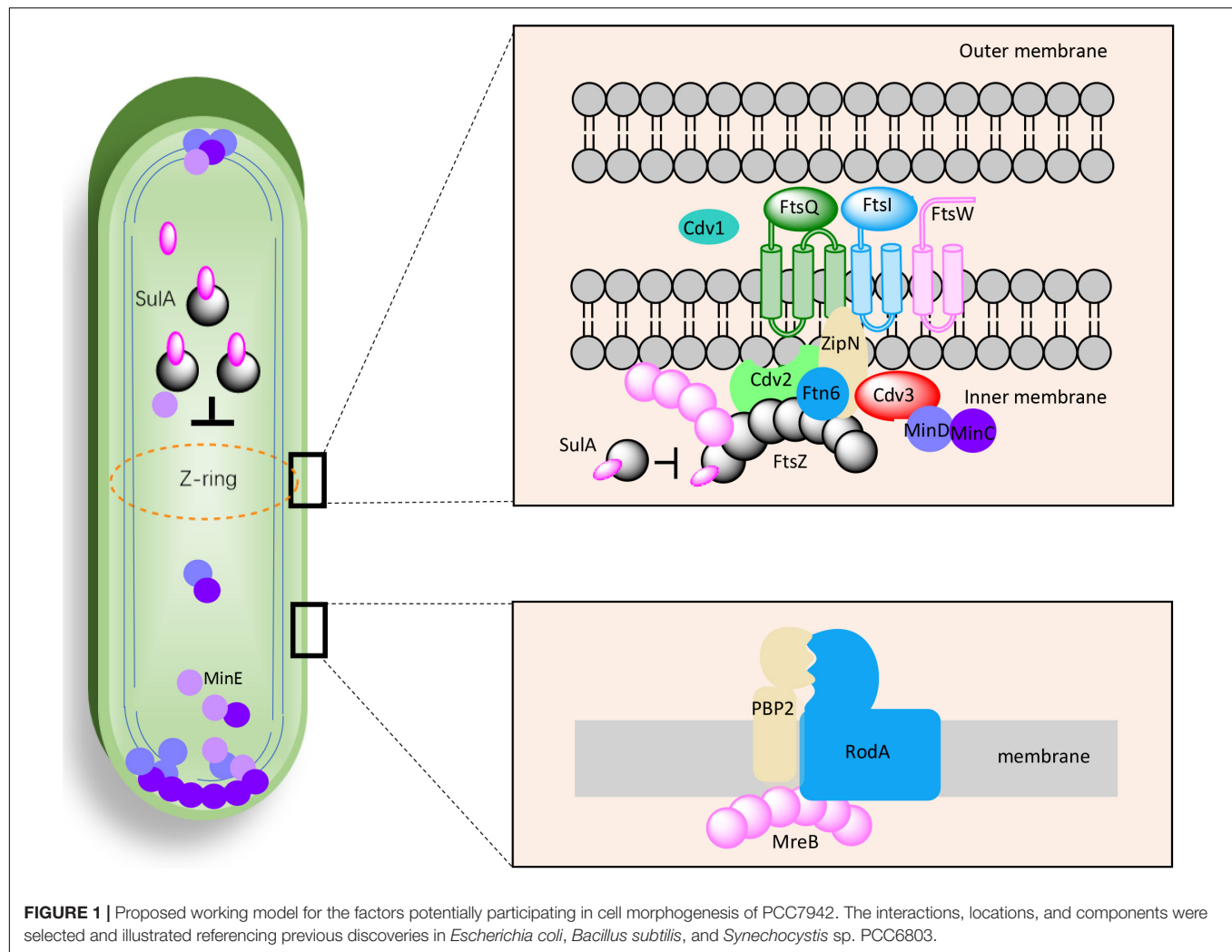
RodA and MreB was eliminated, respectively, in PCC7942, the cell shapes of the disruption mutants were remodeled from short rod to sphere (**Figure 2A**), which might resulted from the impaired synthesis of cell wall and the loss of vertical tension. In addition, although cell length of the ΔMreB mutant was reduced by 40% comparing with that of the wild type, the cell area was not adversely affected, but enlarged by 1.3-fold (**Table 1** and **Figure 2B**), which is similar with the phenotypes of *mreB* knocking-down mutants of *E. coli* (Kruse et al., 2003). The cell length of the ΔRodA mutant was reduced by 47% while the cell area of the mutant was still maintained on the same level of the wild type control (**Figure 2C**).

Moreover, the cells of PCC7942 mutants carrying ΔFtn6, ΔCdv3, and ΔCdv1 also showed significant elongation by 8. 6-, 5. 4-, and 5.1-fold of the wild type cell length, and the cell area were also enlarged by 11. 6-, 7. 3-, and 5.0-fold, respectively (**Table 1**), which is in consistence with previous reports (Miyagishima et al., 2005). A noteworthy point is that the elimination of Ftn6, Cdv3, and Cdv1 resulted in unequal division of the mutant cells (**Figure 2A**), suggesting that these proteins might contribute to the accurate positioning of Z-ring in cell membranes. The

cell lengths and areas of the mutant cells deficient in SulA, another potential inhibitor of Z-ring structure, were just slightly influenced, both increased by 1.3-fold than these of the wild type cells (**Table 1** and **Figures 2B,C**).

Deficiency or weakening of the genes participating in morphogenesis revealed different effects on cell growth of PCC7942. The ΔFtsZ and ΔFtsW mutants showed severely impaired growth, which decreased by 45 and 71% compared to that of the wild type strain (**Table 1**), indicating possible defection of cell divisions (Boyle et al., 1997; Sarcina and Mullineaux, 2000). In addition, the mutants carrying *mreB* and *cdv3* deficiencies also showed retarded growth rates of about 83% compared to that of wild type (**Table 1**), suggesting their potential roles in maintaining normal cell growth and division. As a key effector not only regulating shape determination but also patterning cell-wall growth, MreB is essential for cell viability, and *mreB* depletion resulted in loss of rod-shape and eventually cell lysis in *E. coli* (Gital et al., 2005; Kruse et al., 2005). Regarding Cdv3, previously it has been reported that partially knockout of this gene in *Synechocystis* reduced the growth rate by 50% (Marbouty et al., 2009). Other mutants did not show notable changes in cell

FIGURE 2 | The disruption effects of the candidate genes on cellular morphology of PCC7942. **(A)** Microscopic photo of the wild type PCC7942 and disruption mutants. **(B)** The distribution histogram and normal distribution curve of the cell length (left) and the cell area (right) in the mutant strains compared with the wild-type strain (blue line). At least 100 cells were evaluated for statistical analysis. The left and right Y axes represent the distribution probability of the mutants and the wild type, respectively. **(C)** The comparison of cell length (left) and cell area (right) between wild type and mutant strains based on the value of median with 95% CI. The Kruskal-Wallis test in One-way ANOVA was used to analyze the significant difference. * indicates <0.05; ** indicates <0.01; *** indicates <0.001; **** indicates <0.0001.

growth and photosynthesis compared to the wild type PCC7942 (**Table 1** and **Supplementary Figure S3**).

Exploring the Effects of Overexpressing the Candidate Genes on PCC7942 Cell Morphogenesis

In addition to the strategy of gene knockout, we also explored the effects of overexpressing each of the twelve candidate genes on cellular morphology of PCC7942. All the candidates were cloned and placed under control of a flexible gene expression system, consisting the *Ptrc*-promoter and a theophylline responsive riboswitch ENYC4 (Ohbayashi et al., 2016). The cassettes were subsequently integrated on the neutral site 2 (NS2) on the chromosome of PCC7942. Genotypes of the mutants were confirmed by PCR and DNA sequencing (**Supplementary Figure S1A**). As shown in **Supplementary Figure S1B**, cellular morphology of the PCC7942 wild type strain was not significantly influenced by the induction dose and induction time of theophylline, thus we calculated and compared the cell sizes and shapes of the overexpression strains after 3 days induction with 1 mM theophylline (**Table 1** and **Figure 3A**). Overexpression of FtsZ resulted in the formation of significantly minimized cells, with 54% reduced cell lengths and 61% reduced cell area. This result is inconsistent with the phenomenon previously reported that constitutive expression of FtsZ (by *Ptrc*-promoter) in PCC7942 led to the generation of filamented cells (Mori and Johnson, 2001; Cohen et al., 2018). Previously similar counterintuitive phenomenon has been discovered in *E. coli* cells that 2–7-folds enforced expression of FtsZ resulted in minimized cells, while when the expression level was further enhanced to 12-fold higher, the cells would be remodeled into filamentous pattern (Ward and Lutkenhaus, 1985). A possible explanation could be that excessive FtsZ in appropriate range (with sufficient subsequent factors) accelerated the assembly of Z-ring and over-activated cell division, which in turn promoted the generation of minimized cells. While when the expression level of FtsZ was further increased, the abundances of the other cellular divisome components would become relatively insufficient. This might subsequently impair the formation and function of cell divisome and elongate the mutant cells. A similar mechanism could also explain the phenomenon that the overexpression of Cdv3 and ZipN resulted in significantly increased cell size (5.8- and 6.2-folds increased lengths than that of the wild type control, as shown in **Table 1** and **Figures 3B,C**), because excessive abundances of specific factors for the cell division might competitively block and inhibit the interaction and affinity of the subsequent cells, subsequently impaired the normal cell division and generated elongated cells (Gao et al., 2017). An additional phenotype supporting this hypothesis is the increased frequency of uneven cell divisions (**Figure 3A**). The response mode of cellular morphology to theophylline induction in these two mutants were further explored. As shown in **Figures 4A,B**, when no theophylline was added, cell lengths of the mutants carrying additional copy of *zipN* or *cdv3* (NS2-Ptrc-ENYC4-ZipN and NS2-Ptrc-ENYC4-Cdc3) were maintained well (1.6- and 1.8-fold higher than that of the wild type), while when 1 mM theophylline

was added, cell lengths of the mutant cells kept increasing to about 4-fold of the wild type level. The addition of theophylline and the elongation of the cells did not caused significant influence on cell growth (**Figures 4A,B** inset charts). In addition, not only on the time scale, the cell length also exhibited a good response to the concentration gradient of theophylline (**Figure 4C**). Taking Enyc4-ZipN as an example, when the theophylline concentration was gradually increased, the cell length and the cell sedimentation rate were both synchronously increased, indicating the potential of this strategy for application in controllable biomass harvesting (Jordan et al., 2017).

Comparing with the FtsZ-overexpression strain, the strain overexpressing MreB, another important cell morphogenesis skeleton bricks, showed little difference in cell lengths and areas from that of the wild type control, except that the MreB-overexpression cells tended to be spindle (**Figure 3A**). Excessive accumulation of other factors potentially involved in cell morphogenesis, including Cdv1, RodA, FtsE, FtsI, FtsW, SulA, and Ftn6, caused similar effects on cell length, with increase ranging from 1.2- to 1.4-fold, while rod-shapes of the cells were maintained (**Table 1** and **Figure 3A**). Although it has been reported previously that overexpression of SepF in *B. subtilis* resulted in filamentation of the mutant cells and eventually cell deaths (Gao et al., 2017), the excessive abundance of the homologous protein Cdv2 in PCC7942 caused minor effects on cellular morphology (with 10% increased cell lengths and 10% decreased cell area, **Table 1**). This indicated that the Cdv2/SepF was not as essential as other components such as ZipN and Cdv3 for Z-ring stability and functionality in PCC7942.

Some of the overexpression mutants exhibited growth impairment phenotypes similar to the knockout mutants. For example, the growths rates of the strains overexpressing FtsZ, MreB, FtsW, and Ftn6, were significantly reduced by 52, 30, 32, and 30%, respectively, compared to the wild type control (**Table 1**). As mentioned above, the mismatch in concentrations and timing of different morphogenesis components might result in disorder and disturbance of the Z-ring assemble and functionalities. The subsequent impairments on formation and function of cell divisome might further cause the weakening of cell survival and proliferations (Chiu et al., 2008). It is noteworthy that at least for a portion of the potential genes participating in morphogenesis (FtsW, MreB, and FtsW), artificial (either up- or down-) regulation of the abundances resulted in significantly weakened cell growth (**Table 1**), and these effects should not be resulted from morphology changes, because growths of both the elongated (ΔFtsZ) and the shortened (ENYC4-FtsZ) were similarly reduced. The detailed mechanisms are yet to be disclosed, while possible influence on cyanobacteria cultivation process should be taken into consideration when the strategy of morphogenesis engineering is adopted.

Adopting Small RNA (sRNA) Based Gene Repression Approach to Regulate Cellular Morphology of PCC7942

By adopting the theophylline-responsive riboswitch approach, we achieved flexible regulation of cell length of PCC7942 by
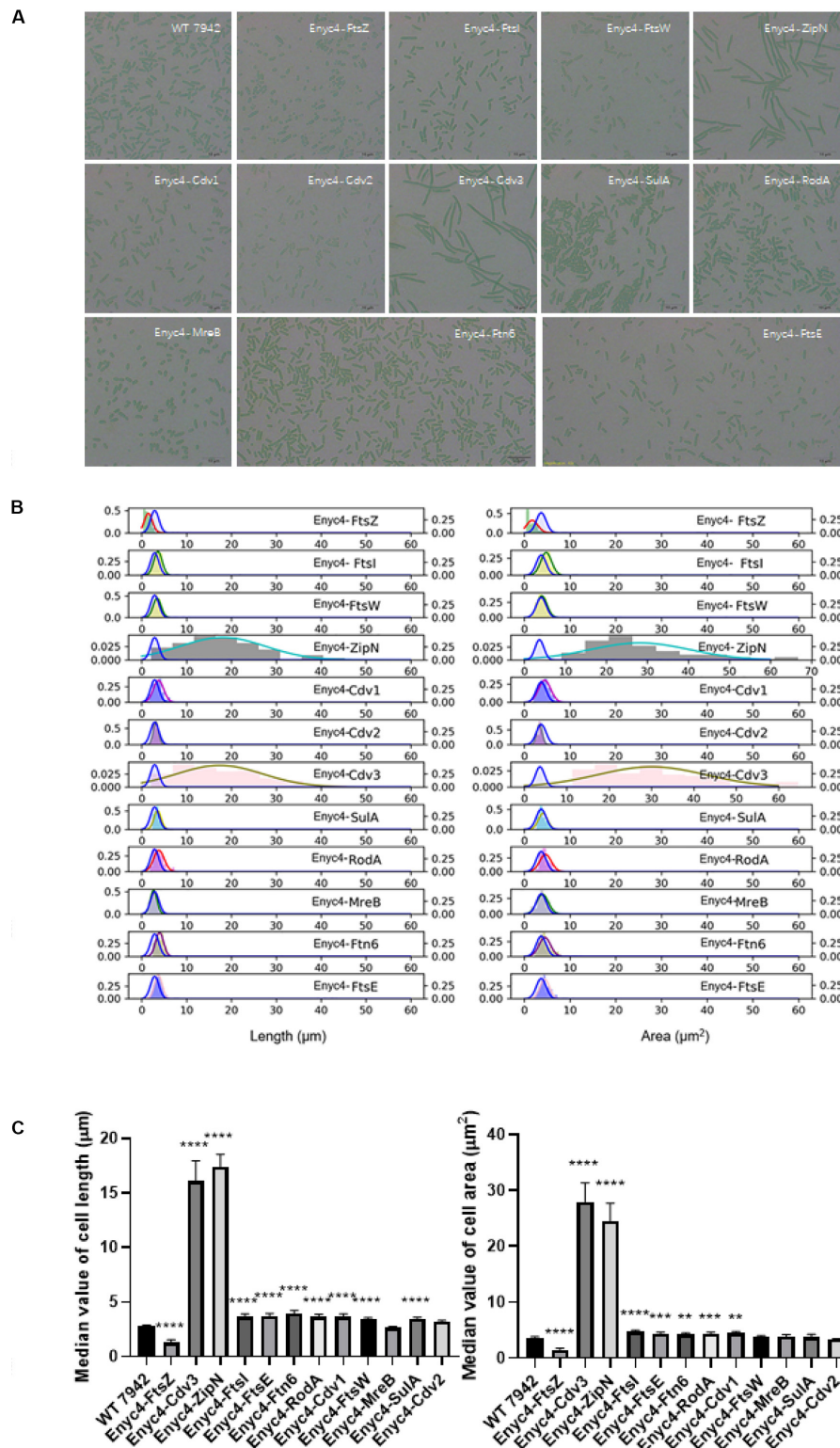
FIGURE 3 | The overexpression effects of the candidate genes on cellular morphology of PCC7942. **(A)** Microscopic photo of the wild type PCC7942 and overexpression mutants. **(B)** The distribution histogram and normal distribution curve of the cell length (left) and the cell area (right) in the mutant strains compared with the wild type strain (blue line). At least 100 cells were evaluated for statistical analysis. The right left and right Y axes indicates the probability of the mutants and the wild type, respectively. **(C)** The comparison of cell length (left) and cell area (right) between wild type and riboswitch-regulated strains based on the values of median and the SEC. The Kruskal-Wallis test in One-way ANOVA was used to analyze the significant difference. * indicates <0.05; ** indicates <0.01; *** indicates <0.001; **** indicates <0.0001.
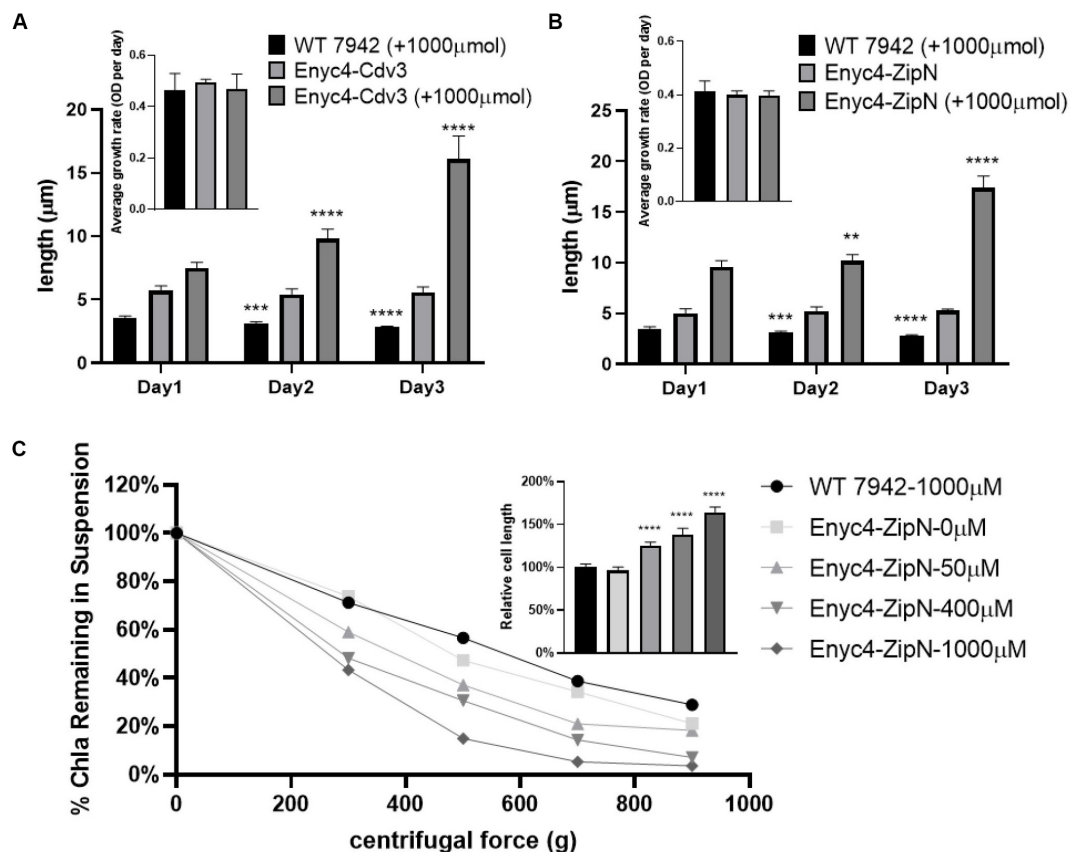
FIGURE 4 | Response of the cell lengths to theophylline induction in PCC7942 strains carrying riboswitch-regulated expression of Cdv3 and ZipN. The cell lengths of the PCC7942 mutants carrying Ptrc-ENYC4-Cdv3 **(A)** and Ptrc-ENYC4-ZipN **(B)** were calculated during 3 days cultivation with (1000 μM) or without th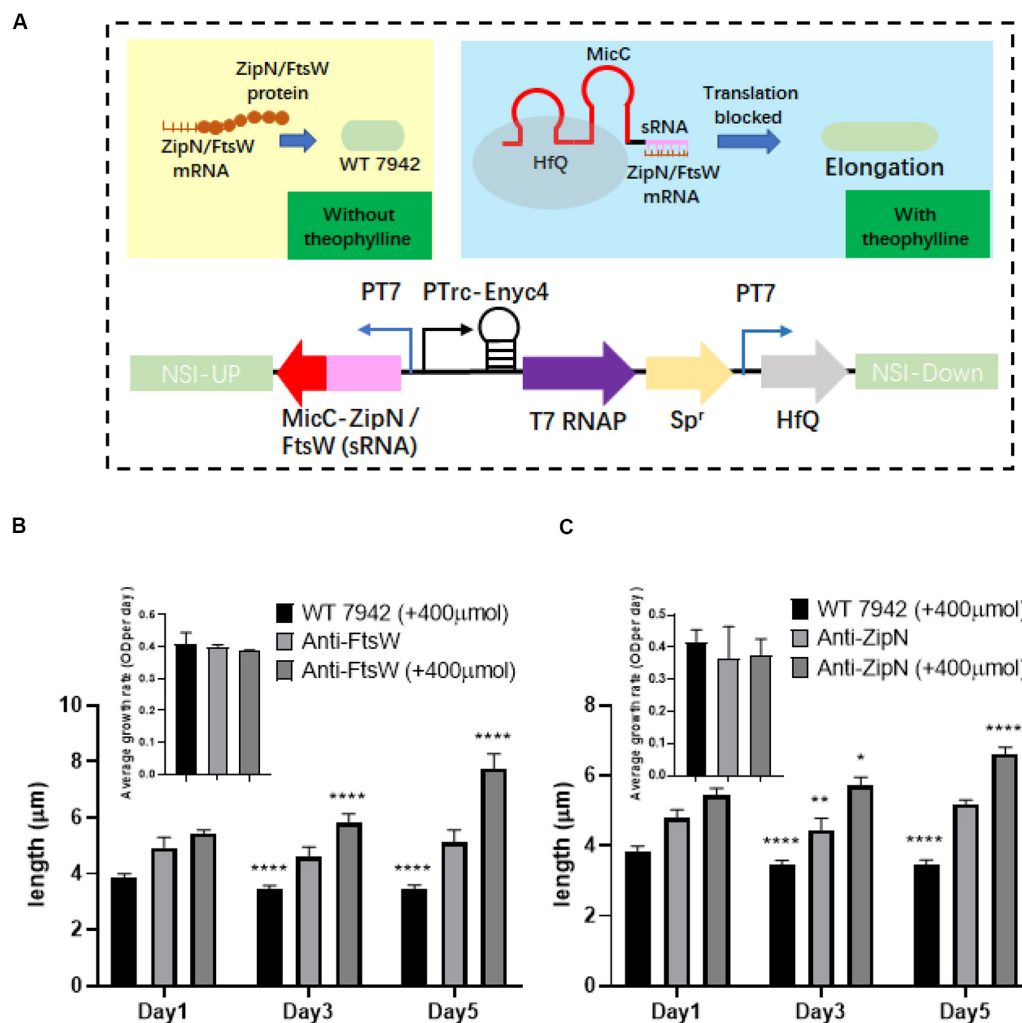eophylline induction. The inset figures show the growth rates of the wild type and mutant strains during the process. The Kruskal-Wallis test in One-way ANOVA was used to analyze the significant difference of cell length at Day 2 (and Day 3) compared with that of at Day 1 in each strain. * indicates <0.05; ** indicates <0.01; *** indicates <0.001; **** indicates <0.0001. The culture of wild-type control was supplemented with 1000 μM theophylline, which didn't affect its growth. **(C)** At Day 5 of induction by theophylline, Chla (Chlorophyll a) contents in the supernatant of wild type and different Enyc4-ZipN cultures were analyzed after centrifugation. Ratio of Chla content under the applied centrifugal force relative to the control condition (centrifugal force 0 g) was calculated for each culture. The embedded graph shows ratios of the cell lengths of the Enyc4-ZipN strain compared to the wild type strain.

inducible overexpression of important contributors involved in cell division. Subsequently, we also attempted to regulate cellular morphology of PCC7942 through controllable down-regulation of potential targets. sRNA regulatory tools are promising metabolic engineering approaches to repress the expression of both endogenous and exogenous target genes (Nakashima et al., 2006). In recent years, this approach has also been successfully developed and adopted in cyanobacteria for remodeling cellular metabolism (Li et al., 2018; Sun et al., 2018a). In this work, we aimed to adopt Hfq-MicC tool (Sun et al., 2018a) to regulate the expression of ZipN and FtsW. The components and working mechanism of the controllable gene repression system is illustrated in **Figure 5A**, in which sRNA-MicC cassette and the Hfq protein were placed under the control of the T7 promoter, while the heterologous T7 RNA polymerase was driven by the previously utilized theophylline-responsive expression system (Ptrc-ENYC4). When the expression of T7 RNA polymerase was induced by addition of theophylline, the sRNA-MicC RNA would be subsequently transcribed and bind to the mRNA of

the target gene, resulting in controllable gene silencing. Targeting ZipN and FtsW, two mutant strains were further constructed (termed as Anti-ZipN and Anti-FtsW, respectively). As shown in **Figure 5**, when 400 μM theophylline was added into the medium, cell lengths of the Anti-ZipN and Anti-FtsW strains were significantly increased by 1.3- and 1.5-fold than theoe of the cells without theophylline induction, respectively (**Figures 5B,C**). However, there are still two obvious drawbacks of this system. When no theophylline was supplemented, leak expression of T7 RNA polymerase or Hfq-MicC system might partially inhibit the expression of ZipN and FtsW, leading to slightly increased cell lengths of the mutant strains (by 1.3- and 1.4-fold higher than the wild type level). While when theophylline was added, the elongations of the mutant cells was not as significant as that in the disruption mutants (filamentous cells of the ΔZipN and ΔFtsW mutants). In future, the development of more stringent and smart gene repression approach enabling a wider regulatory space could be expected to bring in more desirable control system for cellular morphology.

FIGURE 5 | Regulation of PCC7942 cell lengths through the adoption of sRNA based gene repression approach. **(A)** Design of the Hfq-MicC sRNA approach for controllable repression of ZipN or FtsW to regulate cell lengths of PCC7942. **(B)** The dynamics of cell lengths of the PCC7942-Anti-FtsW **(B)** and PCC7942-Anti-ZipN **(C)** strains induced with 400 μM theophylline for 3 days. The inset figures represent the growths of the wild type control and the mutant strains. The Kruskal-Wallis test in One-way ANOVA was used to analyze the significant difference cell length at Day 2 (and Day 3) compared with that at Day 1 for each strain. * indicates <0.05; ** indicates <0.01; *** indicates <0.001; **** indicates <0.0001.

CONCLUSION

Cyanobacteria are promising microbial chassis for photosynthetic biomanufacturing in future, and optimization of industrial properties of cyanobacteria chassis cells and engineered strains are necessary for developing economically competitive photosynthetic cell factories (Gao et al., 2016b; Luan and Lu, 2018). Cellular morphology engineering is an up-and-coming strategy to improve complex phenotypes required by industrial application. In this work, we performed systematic exploration of promising target genes for engineering cellular morphology of an important cyanobacteria strain, *Synechococcus elongatus* PCC7942. Previously, the knockout strategy was adopted to identify the influence of target genes on cell morphogenesis, while the overexpression strategy was relatively rarely utilized. Aiming to get a more clear and comprehensive map of potential nodes in cellular morphology engineering, we combined the knockout/down and overexpression strategies targeting each of the twelve potential genes participating in cell division and/or elongation. The influence of elevated and decreased abundance of the targets on cell morphogenesis were systematically calculated and compared, illustrating a more clear and comprehensive map of cellular morphology engineering nodes. As the most important skeleton bricks for Z-ring structure, the expression level of FtsZ show strongly negatively regulatory effects on cell length of PCC7942. The cells of the FtsZ defective mutant were filamented and the overexpression of FtsZ resulted in generation of minimized cells. However, the disruption and overexpression of the other components involved in cell division both elongated cells of the respective mutants, which might be resulted from imbalance of the ratios among diverse components. As compared, MreB and RodA, the two factors contributing

to the cell wall synthesis, show more significant influence on cell shape of PCC7942. Shapes of the two disruption mutants (MreB and RodA defective strains) were remodeled from rod into sphere. In addition, the overexpression of MreB also led to the formation of spindle-shaped cells. Adopting a previously developed sRNA based expression regulation approach, we partially achieved flexible and controllable regulation of the PCC7942 cell lengths, and more desirable regulatory effects could be expected through development and application of more power synthetic biology toolbox.

MATERIALS AND METHODS

Chemicals and Reagents

Chemicals utilized in this work were purchased from Sinopharm Chemical Reagent Co., Ltd. (Shanghai, China) except theophylline from Sigma-Aldrich (St. Louis, MO, United States). *Taq* and *FastPfu* Fly DNA polymerases for PCR and pEASY-Blunt Cloning kits were obtained from Transgene Biotech (Beijing, China). Restriction enzymes and T4 DNA ligase were purchased from Thermo Fisher (Waltham, MA, United States). Oligonucleotides synthesis and DNA sequencing was processed by TsingKe (Qingdao, China).

Strain Construction and Cultivation

Escherichia coli DH5α was used as the host strain for plasmids construction and grown in LB media at 37°C. The wild-type strain of PCC7942 is a gift of Prof. Xudong Xu from Institute of Hydrobiology, Chinese Academy of Sciences. To construct the knockout plasmids, the respective upstream and downstream homologous fragments of each target gene were amplified by PCR, and subsequently fused with *aacC1* (Gentamicin-resistance gene, GmR, 1.2 kb) by fusion PCR. The fused fragment was then cloned into pEASY-blunt simple vector. The generated plasmids were transformed into the PCC7942 wild type cells. Gentamicin resistant transformants were obtained after 7 days cultivation on selective BG11 agar plates (containing 10 μg/ml gentamicin). Genotypes of the transformants were confirmed by PCR and DNA sequencing. To overexpress target genes, the backbone of plasmid was amplified from previously developed plasmid (Qiao et al., 2018), containing upstream and downstream homologous fragments of NS2 (neutral site 2 on the chromosome of PCC7942), the *Ptrc*-ENYC4 promoter, chloramphenicol resistant gene (CmR, 0.95 kb). Then the PCR amplicons of the target genes were digested with restriction enzyme (*PacI/PaeI*) and ligated into the backbone plasmid. Chloramphenicol-containing BG11 agar plates were used to select resistant transformants. The plasmid used for synthetic sRNA expression was obtained from Prof. Weiwen Zhang of Tianjin University. Synthetic sRNAs that recognize specific sequences of target genes were introduced into plasmid through site-directed mutagenesis as previously reported (Yoo et al., 2013). And here spectinomycin resistance was used as the phenotypes to isolate positive transformants. All information about strains, plasmids and oligonucleotides was presented in **Supplementary Table S1**. All of the PCC7942 derived strains were grown in BG11 medium in flasks that were incubated on a horizontal rotary shaker at 150 rpm at 30°C under constant white-light illumination of 30 μmol/m^2/s. Theophylline was added to liquid media when necessary.

Microscopy and Cell Length Measurements

All images were captured using Olympus BX51 microscope (100X/1.3 Oil Ph3) with an Olypus DP72 camera. For cell length measurements, three independent biology parallels with at least 250 cells per image were recorded and further measured by manual tools in DP2-BSW3 software (Olympus, Japan).

Growth, Chlorophyll *a* (Chl*a*) and Oxygen Evolution Measurements

Growths rates (shown in **Table 1**) were calculated based on equation Log [(Day 7 Chl*a*/Day 3 Chl*a*), 2], while growths in **Figures 4**, **5** were shown by OD$_{730}$ per day. In growth measurement for gene overexpression strains, 1 mM theophylline was added. Chl*a* was extracted by cell suspension in equal volume methyl alcohol overnight at 4°C. Then Chl*a* content was determined by equation 12.9447*(A665-A720) with methyl alcohol as control. Oxygen evolution was measured under light intensity of 143 μmol/m^2/s using a Clark-type oxygen electrode (Hansatech, British) connected to the Oxy Lab software, final data were divided by Chl*a* content of each culture.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

MZ performed the experiments and participated in manuscript preparation. CQ and QL participated in the research. QL, GL, and XL designed and directed the research. GL and XL wrote and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.01608/full#supplementary-material

REFERENCES

Aikawa, S., Nishida, A., Hasunuma, T., Chang, J. S., and Kondo, A. (2019). Short-term temporal metabolic behavior in halophilic cyanobacterium *Synechococcus* sp. strain PCC 7002 after salt shock. *Metabolites* 9:297. doi: 10.3390/metabo9120297

Arora, D., Chawla, Y., Malakar, B., Singh, A., and Nandicoori, V. K. (2018). The transpeptidase PbpA and noncanonical transglycosylase RodA of *Mycobacterium tuberculosis* play important roles in regulating bacterial cell lengths. *J. Biol. Chem.* 293, 6497–6516. doi: 10.1074/jbc.M117.811190

Boyle, D. S., Khattar, M. M., Addinall, S. G., Lutkenhaus, J., and Donachie, W. D. (1997). ftsW is an essential cell-division gene in *Escherichia coli*. *Mol. Microbiol.* 24, 1263–1273. doi: 10.1046/j.1365-2958.1997.4091773.x

Camargo, S., Picossi, S., Corrales-Guerrero, L., Valladares, A., Arevalo, S., and Herrero, A. (2019). ZipN is an essential FtsZ membrane tether and contributes to the septal localization of SepJ in the filamentous cyanobacterium *Anabaena*. *Sci. Rep.* 9:2744. doi: 10.1038/s41598-019-39336-6

Chisti, Y. (2013). Constraints to commercialization of algal fuels. *J. Biotechnol.* 167, 201–214. doi: 10.1016/j.jbiotec.2013.07.020

Chiu, S. W., Chen, S. Y., and Wong, H. C. (2008). Dynamic localization of MreB in *Vibrio parahaemolyticus* and in the ectopic host bacterium *Escherichia coli*. *Appl. Environ. Microbiol.* 74, 6739–6745. doi: 10.1128/AEM.01021-08

Cohen, S. E., McKnight, B. M., and Golden, S. S. (2018). Roles for ClpXP in regulating the circadian clock in *Synechococcus elongatus*. *Proc. Natl. Acad. Sci. U.S.A.* 115, E7805–E7813. doi: 10.1073/pnas.1800828115

Demarsac, N. T., and Houmard, J. (1993). Adaptation of cyanobacteria to environmental stimuli - new steps towards molecular mechanisms. *FEMS Microbiol. Lett.* 104, 119–189. doi: 10.1111/j.1574-6968.1993.tb05866.x

Desai, S. H., and Atsumi, S. (2013). Photosynthetic approaches to chemical biotechnology. *Curr. Opin. Biotechnol.* 24, 1031–1036. doi: 10.1016/j.copbio.2013.031015

Errington, J., Daniel, R. A., and Scheffers, D. J. (2003). Cytokinesis in bacteria. *Microbiol. Mol. Biol. Rev.* 67, 52–65. doi: 10.1128/mmbr.67.1.52-65.2003

Fernandez-Juarez, V., Bennasar-Figueras, A., Tovar-Sanchez, A., and Agawin, N. S. R. (2019). The role of iron in the P-acquisition mechanisms of the unicellular N2-fixing cyanobacteria *Halothece* sp., found in association with the mediterranean seagrass *Posidonia oceanica*. *Front. Microbiol.* 10:1903. doi: 10.3389/fmicb.2019.01903

Flombaum, P., Gallegos, J. L., Gordillo, R. A., Rincon, J., Zabala, L. L., Jiao, N. A. Z., et al. (2013). Present and future global distributions of the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 9824–9829. doi: 10.1073/pnas.1307701110

Gao, X., Gao, F., Liu, D., Zhang, H., Nie, X. Q., and Yang, C. (2016a). Engineering the methylerythritol phosphate pathway in cyanobacteria for photosynthetic isoprene production from CO_2. *Energ. Environ. Sci.* 9, 1400–1411. doi: 10.1039/c5ee03102h

Gao, X., Sun, T., Pei, G., Chen, L., and Zhang, W. (2016b). Cyanobacterial chassis engineering for enhancing production of biofuels and chemicals. *Appl. Microbiol. Biotechnol.* 100, 3401–3413. doi: 10.1007/s00253-016-7374-2

Gao, Y., Wenzel, M., Jonker, M. J., and Hamoen, L. W. (2017). Free SepF interferes with recruitment of late cell division proteins. *Sci. Rep.* 7:16928. doi: 10.1038/s41598-017-17155-x

Gao, Z. X., Zhao, H., Li, Z. M., Tan, X. M., and Lu, X. F. (2012). Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy. Environ. Sci.* 5, 9857–9865. doi: 10.1039/C2ee22675h

Gital, Z., Dye, N. A., Reisenauer, A., Wachi, M., and Shapiro, L. (2005). MreB actin-mediated segregation of a specific region of a bacterial chromosome. *Cell* 120, 329–341. doi: 10.1016/j.cell.2005.01.007

Goehring, N. W., and Beckwith, J. (2005). Diverse paths to midcell: assembly of the bacterial cell division machinery. *Curr. Biol.* 15, R514–R526. doi: 10.1016/j.cub.2005.06.038

Hagemann, M. (2011). Molecular biology of cyanobacterial salt acclimation. *FEMS Microbiol. Rev.* 35, 87–123. doi: 10.1111/j.1574-6976.2010.00234.x

Henriques, A. O., Glaser, P., Piggot, P. J., and Moran, C. P. Jr. (1998). Control of cell shape and elongation by the *rodA* gene in *Bacillus subtilis*. *Mol. Microbiol.* 28, 235–247. doi: 10.1046/j.1365-2958.1998.00766.x

Jezberova, J., and Komarkova, J. (2007). Morphological transformation in a freshwater *Cyanobium* sp. induced by grazers. *Environ. Microbiol.* 9, 1858–1862. doi: 10.1111/j.1462-2920.2007.01311.x

Jiang, X. R., and Chen, G. Q. (2016). Morphology engineering of bacteria for bio-production. *Biotechnol. Adv.* 34, 435–440. doi: 10.1016/j.biotechadv.2015.12.007

Jiang, X. R., Wang, H., Shen, R., and Chen, G. Q. (2015). Engineering the bacterial shapes for enhanced inclusion bodies accumulation. *Metab. Eng.* 29, 227–237. doi: 10.1016/j.ymben.2015.03.017

Jiang, X. R., Yao, Z. H., and Chen, G. Q. (2017). Controlling cell volume for efficient PHB production by *Halomonas*. *Metab. Eng.* 44, 30–37. doi: 10.1016/j.ymben.2017.09.004

Jordan, A., Chandler, J., MacCready, J. S., Huang, J., Osteryoung, K. W., and Ducat, D. C. (2017). Engineering cyanobacterial cell corphology for enhanced recovery and processing of biomass. *Appl. Environ. Microbiol.* 83:e00053-17. doi: 10.1128/AEM.00053-17

Koksharova, O. A., and Babykin, M. M. (2011). Cyanobacterial cell division: genetics and comparative genomics of cyanobacterial cell division. *Russ. J. Genet.* 47, 255–261. doi: 10.1134/s1022795411030070

Koksharova, O. A., and Wolk, C. P. (2002). A novel gene that bears a DnaJ motif influences cyanobacterial cell division. *J. Bacteriol.* 184, 5524–5528. doi: 10.1128/jb.184.19.5524-5528.2002

Kruse, T., Bork-Jensen, J., and Gerdes, K. (2005). The morphogenetic MreBCD proteins of *Escherichia coli* form an essential membrane-bound complex. *Mol. Microbiol.* 55, 78–89. doi: 10.1111/j.1365-2958.2004.04367.x

Kruse, T., Møller-Jensen, J., Løbner-Olesen, A., and Gerdes, K. (2003). Dysfunctional MreB inhibits chromosome segregation in *Escherichia coli*. *EMBO J.* 22, 5283–5292. doi: 10.1093/emboj/cdg504

Li, S., Sun, T., Xu, C., Chen, L., and Zhang, W. (2018). Development and optimization of genetic toolboxes for a fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Metab. Eng.* 48, 163–174. doi: 10.1016/j.ymben.2018.06.002

Lin, W. R., Tan, S. I., Hsiang, C. C., Sung, P. K., and Ng, I. S. (2019). Challenges and opportunity of recent genome editing and multi-omics in cyanobacteria and microalgae for biorefinery. *Bioresour. Technol.* 291:121932. doi: 10.1016/j.biortech.2019.121932

Liu, X. F., Miao, R., Lindberg, P., and Lindblad, P. (2019). Modular engineering for efficient photosynthetic biosynthesis of 1-butanol from CO_2 in cyanobacteria. *Energ. Environ. Sci.* 12, 2765–2777. doi: 10.1039/c9ee01214a

Lu, X. (2010). A perspective: photosynthetic production of fatty acid-based biofuels in genetically engineered cyanobacteria. *Biotechnol. Adv.* 28, 742–746. doi: 10.1016/j.biotechadv.2010.05.021

Luan, G., and Lu, X. (2018). Tailoring cyanobacterial cell factory for improved industrial properties. *Biotechnol. Adv.* 36, 430–442. doi: 10.1016/j.biotechadv.2018.01.005

Marbouty, M., Saguez, C., Cassier-Chauvat, C., and Chauvat, F. (2009). ZipN, an FtsA-like orchestrator of divisome assembly in the model cyanobacterium *Synechocystis* PCC6803. *Mol. Microbiol.* 74, 409–420. doi: 10.1111/j.1365-2958.2009.06873.x

Margolin, W. (2005). FtsZ and the division of prokaryotic cells and organelles. *Nat. Rev. Mol. Cell Biol.* 6, 862–871. doi: 10.1038/nrm1745

Miyagishima, S. Y., Wolk, C. P., and Osteryoung, K. W. (2005). Identification of cyanobacterial cell division genes by comparative and mutational analyses. *Mol. Microbiol.* 56, 126–143. doi: 10.1111/j.1365-2958.2005.04548.x

Mori, T., and Johnson, C. H. (2001). Independence of circadian timing from cell division in cyanobacteria. *J. Bacteriol.* 183, 2439–2444. doi: 10.1128/JB.183.8.2439-2444.2001

Nakashima, N., Tamura, T., and Good, L. (2006). Paired termini stabilize antisense RNAs and enhance conditional gene silencing in *Escherichia coli*. *Nucleic Acids Res.* 34:e138. doi: 10.1093/nar/gkl697

Ohbayashi, R., Akai, H., Yoshikawa, H., Hess, W. R., and Watanabe, S. (2016). A tightly inducible riboswitch system in *Synechocystis* sp. PCC 6803. *J. Gen. Appl. Microbiol.* 62, 154–159. doi: 10.2323/jgam.2016.02.002

Pichoff, S., and Lutkenhaus, J. (2002). Unique and overlapping roles for ZipA and FtsA in septal ring assembly in *Escherichia coli*. *EMBO J.* 21, 685–693. doi: 10.1093/emboj/21.4.685

Qiao, C., Duan, Y., Zhang, M., Hagemann, M., Luo, Q., and Lu, X. (2018). Effects of reduced and enhanced glycogen pools on salt-induced sucrose production in

a sucrose-secreting strain of *Synechococcus elongatus* PCC 7942. *Appl. Environ. Microbiol.* 84:e02023-17. doi: 10.1128/AEM.02023-17

Rohs, P. D. A., Buss, J., Sim, S. I., Squyres, G. R., Srisuknimit, V., Smith, M., et al. (2018). A central role for PBP2 in the activation of peptidoglycan polymerization by the bacterial cell elongation machinery. *PLoS Genet.* 14:e1007726. doi: 10.1371/journal.pgen.1007726

Rousseaux, C. S., and Gregg, W. W. (2014). Interannual variation in phytoplankton primary production at a global scale. *Remote Sens (Basel)* 6, 1–19. doi: 10.3390/rs6010001

Santos-Merino, M., Singh, A. K., and Ducat, D. C. (2019). New applications of synthetic biology tools for cyanobacterial metabolic engineering. *Front. Bioeng. Biotechnol.* 7:33. doi: 10.3389/fbioe.2019.00033

Sarcina, M., and Mullineaux, C. W. (2000). Effects of tubulin assembly inhibitors on cell division in prokaryotes in vivo. *FEMS Microbiol. Lett.* 191, 25–29. doi: 10.1016/S0378-1097(00)00365-7

Simkovsky, R., Daniels, E. F., Tang, K., Huynh, S. C., Golden, S. S., and Brahamsha, B. (2012). Impairment of O-antigen production confers resistance to grazing in a model amoeba-cyanobacterium predator-prey system. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16678–16683. doi: 10.1073/pnas.1214904109

Sohm, J. A., Webb, E. A., and Capone, D. G. (2011). Emerging patterns of marine nitrogen fixation. *Nat. Rev. Microbiol.* 9, 499–508. doi: 10.1038/nrmicro2594

Sun, T., Li, S., Song, X., Pei, G., Diao, J., Cui, J., et al. (2018a). Re-direction of carbon flux to key precursor malonyl-CoA via artificial small RNAs in photosynthetic *Synechocystis* sp. PCC 6803. *Biotechnol. Biofuels* 11:26. doi: 10.1186/s13068-018-1032-0

Sun, T., Li, S. B., Song, X. Y., Diao, J. J., Chen, L., and Zhang, W. W. (2018b). Toolboxes for cyanobacteria: recent advances and future direction. *Biotechnol. Adv.* 36, 1293–1307. doi: 10.1016/j.biotechadv.2018.04.007

Wang, Y., Wu, H., Jiang, X., and Chen, G. Q. (2014). Engineering *Escherichia coli* for enhanced production of poly(3-hydroxybutyrate-co-4-hydroxybutyrate) in larger cellular space. *Metab. Eng.* 25, 183–193. doi: 10.1016/j.ymben.2014.07.010

Wang, Z., Huang, S., and Li, D. (2019). Decomposition of cyanobacterial bloom contributes to the formation and distribution of iron-bound phosphorus (Fe-P): insight for cycling mechanism of internal phosphorus loading. *Sci. Total Environ.* 652, 696–708. doi: 10.1016/j.scitotenv.2018.10.260

Ward, J. E. Jr., and Lutkenhaus, J. (1985). Overproduction of FtsZ induces minicell formation in *E. coli*. *Cell* 42, 941–949. doi: 10.1016/0092-8674(85)90290-9

Waterbury, J. B., Watson, S. W., Guillard, R. R. L., and Brand, L. E. (1979). Widespread occurrence of a unicellular, marine, planktonic, cyanobacterium. *Nature* 277, 293–294. doi: 10.1038/277293a0

Xu, X. D., Khudyakov, I., and Wolk, C. P. (1997). Lipopolysaccharide dependence of cyanophage sensitivity and aerobic nitrogen fixation in *Anabaena* sp. strain PCC 7120. *J. Bacteriol.* 179, 2884–2891. doi: 10.1128/jb.179.9.2884-2891.1997

Yao, L., Cengic, I., Anfelt, J., and Hudson, E. P. (2016). Multiple gene repression in cyanobacteria using CRISPRi. *ACS Synth. Biol.* 5, 207–212. doi: 10.1021/acssynbio.5b00264

Yoo, S. M., Na, D., and Lee, S. Y. (2013). Design and use of synthetic regulatory small RNAs to control gene expression in *Escherichia coli*. *Nat. Protoc.* 8, 1694–1707. doi: 10.1038/nprot.2013.105

Young, K. D. (2006). The selective value of bacterial shape. *Microbiol. Mol. Biol. Rev.* 70, 660–703. doi: 10.1128/MMBR.00001-06

Zamalloa, C., Vulsteke, E., Albrecht, J., and Verstraete, W. (2011). The techno-economic potential of renewable energy through the anaerobic digestion of microalgae. *Bioresour. Technol.* 102, 1149–1158. doi: 10.1016/j.biortech.2010.09.017

Check for
updates

Expression of Formate-Tetrahydrofolate Ligase Did Not Improve Growth but Interferes With Nitrogen and Carbon Metabolism of *Synechocystis* sp. PCC 6803

Shanshan Song[1], Stefan Timm[1], Steffen N. Lindner[2], Viktoria Reimann[3], Wolfgang R. Hess[3], Martin Hagemann[1] and Eva-Maria Brouwer[1]*

[1] Plant Physiology Department, Institute of Biosciences, University of Rostock, Rostock, Germany, [2] Max Planck Institute of Molekular Plant Physiology, Potsdam-Golm, Germany, [3] Faculty of Biology, Genetics and Experimental Bioinformatics, University of Freiburg, Freiburg im Breisgau, Germany

The introduction of alternative CO_2-fixing pathways in photoautotrophic organism may improve the efficiency of biological carbon fixation such as minimizing the carbon loss due to photorespiration. Here, we analyzed the effects of creating a formate entry point into the primary metabolism of the cyanobacterium *Synechocystis* sp. PCC 6803. The formate-tetrahydrofolate ligase (FTL) from *Methylobacterium extorquens* AM1 was expressed in *Synechocystis* to enable formate assimilation and reducing the loss of fixed carbon in the photorespiratory pathway. Transgenic strains accumulated serine and 3-phosphoglycerate, and consumed more 2-phosphoglycolate and glycine, which seemed to reflect an efficient utilization of formate. However, labeling experiments showed that the serine accumulation was not due to the expected incorporation of formate. Subsequent DNA-microarray analysis revealed profound changes in transcript abundance due to *ftl* expression. Transcriptome changes were observed in relation to serine and glycine metabolism, C1-metabolism and particularly nitrogen assimilation. The data implied that *ftl* expression interfered with the signaling the carbon/nitrogen ratio in *Synechocystis*. Our results indicate that the expression of new enzymes could have a severe impact on the cellular regulatory network, which potentially hinders the establishment of newly designed pathways.

Keywords: C1 metabolism, cyanobacteria, carbon fixation, formate assimilation, metabolome, serine, transcriptomics

INTRODUCTION

Inorganic carbon fixation by photoautotrophic organisms via the Calvin-Benson-Bassham (CBB) cycle represents the biochemical process that supplies organic carbon for almost all living organisms on Earth. In nature, factors limiting the growth of photosynthetic organisms vary among species and habitats and include the availability of water, light, and nutrients e.g., combined nitrogen

sources (Evans, 1997). However, in modern agriculture, using fertilizers and often irrigation, CO_2 fixation became the rate-limiting factor of crop plant yield under ambient conditions due to the inefficiency of the key CO_2-fixing enzyme of the CBB cycle, the ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO) (Long et al., 2006; Tcherkez et al., 2006; Bar-Even et al., 2010). RubisCO possesses a slow catalytic rate and an oxygenation side reaction producing the toxic byproduct 2-phosphoglycolate (2PG). 2PG must be metabolized through the photorespiratory pathway, which releases previously fixed CO_2 and liberates NH_3 (Bauwe et al., 2010). Many attempts were undertaken to enhance photosynthetic carbon fixation, like engineering RubisCO toward higher catalytic efficiency and specificity for CO_2 (Whitney et al., 2011). Other attempts aimed at increasing the CO_2 concentration in close vicinity to RubisCO, which has been naturally achieved during evolution of different CO_2-concentrating mechanisms (CCMs) in cyanobacteria, algae and C4 plants (Leegood, 2002; Badger and Price, 2003; Giordano et al., 2005; Long et al., 2018). In recent years, it has also been tried to optimize photorespiration in plants by expressing different artificial bypass reactions to improve the recycling of 2PG (Kebeish et al., 2007; Hagemann and Bauwe, 2016; South et al., 2019). Another *in vitro* approach established a carbon-conserving photorespiration by converting glycolate via glycolyl-CoA and glycolaldehyde into CBB cycle intermediates (Trudeau et al., 2018).

As an alternative to the improvement of the CBB cycle and photorespiration, which are intimately linked to plant primary metabolism, the generation of entirely new synthetic CO_2-fixing pathways has been proposed. Shih et al. (2014) generated a synthetic photorespiratory CO_2-fixing bypass in cyanobacteria, which provided the basis for an alternative carbon fixation pathway in cyanobacteria, algae and plants. Schwander et al. (2016) were able to design and prove an *in vitro* CO_2 fixing pathway, the CETCH [(CoA)/ethylmalonyl-CoA/hydroxybutyryl-CoA] cycle that involves 11 enzymatic steps. The direct assembly of this synthetic pathway in living organisms is challenging due to limited understanding of the complex interplay among the different enzymes used in this synthetic network. Furthermore, the interference of the synthetic networks with the complex metabolic and regulatory background of the host organism can lead to undesired side reactions and toxicity (Schwander et al., 2016).

Recently, formate has been proposed as an ideal feedstock for bio-economy, because it can be produced at relatively high efficiency from multiple available resources such as the electrochemical reduction of CO_2 and oxidization of natural gas (Bar-Even et al., 2013). Furthermore, formate is soluble and of low toxicity. Many methylotrophic organisms can grow with formate as sole carbon source (Marx et al., 2003). The establishment of additional CO_2 reduction into formate in photoautotrophic organisms such as crop plants was proposed to support CO_2 fixation via the CBB cycle (Bar-Even, 2018). The most valuable entry point of formate into primary carbon metabolism is via conversion into 10-formyl-tetrahydrofolate (formyl-THF) by the formyl-THF ligase (FTL) (Bar-Even, 2016). FTL catalyzes an ATP-dependent kinase reaction that gives rise to the intermediate formyl-phosphate and the activated

formyl-group is then transferred on THF to give formyl-THF (Mejillano et al., 1989). FTL does not directly generate a carbon–carbon bond but it activates formate, making it a good electrophile for downstream reactions with a nucleophilic carbon atom. FTL is the only known naturally occurring formate-fixing reaction that supports formatotrophic growth (Bar-Even, 2018). In most organisms, formyl-THF naturally participates in the synthesis of purines and also takes part in the formylation of initiator methionyl-tRNA^{Met} in bacteria, mitochondria and chloroplasts. It can also be converted to methylene-THF via the bi-functional methylene-THF dehydrogenase/methenyl-THF cyclohydrolase (FolD) (Hanson and Roje, 2001). Subsequently, methylene-THF can, together with glycine, serve for serine biosynthesis via the serine-hydroxymethyltransferase (SHMT), which represents an important step in the C1-metabolism of most organisms (**Figure 1**). In plants and other oxygenic phototrophs, the CO_2-releasing step via glycine cleavage in the photorespiratory pathway produces high amounts of methylene-THF, which is then used by SHMT to synthesize serine on the expense of a second glycine molecule. It has been discussed that an increased pool of methylene-THF due to efficient formate incorporation could turn photorespiration into less CO_2-releasing or even CO_2-fixing, when the glycine-decarboxylase reaction is reversed Recently, the formate-assimilation pathway including a reversed glycine decarboxylase flux was successively established in *E. coli*, proving the afore calculated kinetical feasibility and functionality of the designed CO_2-fixing shunt (Bar-Even et al., 2010; Yishai et al., 2017; Bang and Lee, 2018; Döring et al., 2018; Kim et al., 2020).

Here, we aimed to establish formate assimilation in cyanobacteria, which use light energy for oxygenic photosynthesis and the CBB cycle for CO_2 assimilation
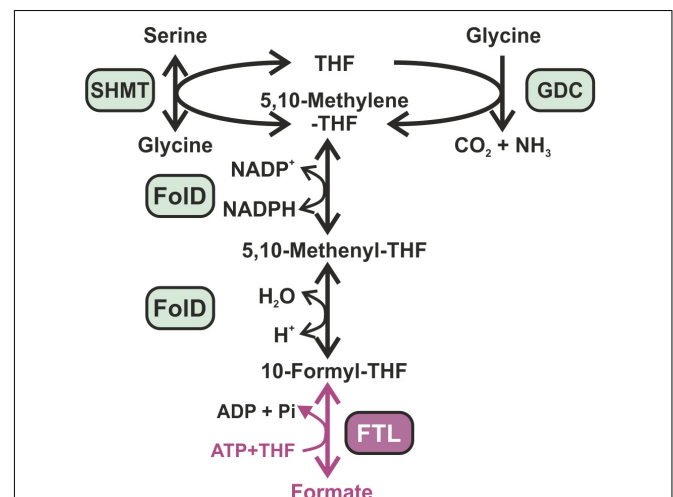
FIGURE 1 | Synthetic formate assimilation (FA) pathway in *Synechocystis* sp. PCC 6803. Enzymes present in the *Synechocystis* host cell are marked in green, while the additional enzyme necessary for formate incorporation is marked in pink. FTL, formate-THF ligase; FolD, bifunctional methylene-THF dehydrogenase/methenyl-THF cyclohydrolase; GDC, glycine decarboxylase complex; SHMT, serine hydroxymethyltransferase.

similar to plants (Hohmann-Marriott and Blankenship, 2011). We selected the model organism *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*) for the expression of *ftl*, which should enable formate utilization and its conversion into biomass via photorespiratory 2PG metabolism (**Figure 1**). The study was initiated to test, if formate assimilation can be established in a prokaryotic oxygenic phototroph before making the next step of crop plant engineering. In contrast to our expectations, the knowledge from *E. coli* could be not directly transferred to *Synechocystis*, because instead of improved growth on formate we found marked alteration of cellular C/N metabolism in the *ftl*-expressing strain.

MATERIALS AND METHODS

Strains and Culture Conditions

The cyanobacterial strains used in this work are listed in **Supplementary Table S1**. The glucose-tolerant strain of *Synechocystis* sp. PCC 6803 served as wild type (WT). Cultivation of mutants and transgenic strains were performed at 50 μg/ml erythromycin (Ery). Axenic cultures of *Synechocystis* were maintained on agar plates (BG 11, pH 8, solidified by 0.9% Kobe agar) at 30°C under continuous illumination of 50 μmol photons m^{-2} s^{-1}. The drop-dilution assay was performed with serial dilution of 2 μl cell suspensions at $OD_{750nm} = 1$ spotting on agar plates without antibiotics and with different supplements as given in the text. Liquid cultures were grown in the batch mode using BG 11 medium. Cells suspensions were sparked either with ambient air (0.04% CO_2) or air-enriched with CO_2 [5% (v/v)] at 30°C under continuous illumination of 100 μmol photons m^{-2} s^{-1}. Contamination by heterotrophic bacteria was evaluated by spreading of 0.2 ml of the culture on LB agar plates.

The *E. coli* DH5α cultured in LB medium at 37°C was used for routine DNA manipulations.

Generation of Transgenic Cyanobacterial Strains

Genes for overexpression and mutation were PCR amplified using gene specific primers (see **Supplementary Table S1**). All PCR products were ligated with pGEM®-T (Promega, Walldorf, Germany) and verified by sequencing (Microsynth Seqlab, Göttingen, Germany). For the overexpression of FTL in *Synechocystis*, *ftl* from *Methylobacterium extorquens* AM1 (Yishai et al., 2017) was amplified with primers adding *Bgl*II and *Mun*I sites at its 5' and 3' ends and inserted under control of the strong light-induced P_{psbAII} promoter into plasmid pAII carrying an erythromycin resistance cassette (Lagarde et al., 2000).

Protein Extraction From Synechocystis and Western Blot

Twenty ml of *Synechocystis* cells ($OD_{750nm} = 1$) were collected by centrifugation at 6000 × g for 10 min and immediately frozen in liquid nitrogen and stored at −80°C for further protein extraction. Frozen cells were resuspended in 200 μl homogenization buffer [75 mM Tris−HCl pH 7.5, 1.5 mM EDTA, 1.5 mm PMSF, 1.5 mM NaHSO₃, 0.15 mM Pefabloc (Merck, Darmstadt, Germany)]. Samples were supplemented with glass beads (diameter 0.5 mm) and subjected to 5 freeze-thaw cycles. Protein quantification was done with Amidoblack (Schulz et al., 1994). The calibration curve was done with different concentration of bovine serum albumin.

SDS−PAGE and Western Blot were done according to standard protocols (Laemmli, 1970; Towbin et al., 1979). The FTL antibody was raised in rabbit against recombinant the generated His-tagged FTL by Davids Biotechnology GmbH (Regensburg, Germany).

Enzyme Assays

The N-terminal His_6-tagged *ftl* was obtained after ligation of a *Sac*I/*Kpn*I fragment into pBAD/HisA. The recombinant FTL was purified from cells of *E. coli* strain BL21 (DE3). The pre-cultures were inoculated in fresh LB-medium to an OD_{600nm} of 0.1 and incubated at 37°C to OD_{600nm} of 0.6 to 0.8 before induction of *ftl* expression with 0.02% L-arabinose. Expression was carried out for 4 h at 37°C. Cells were harvested by centrifugation at 6000 × g for 10 min and washed with lysis buffer [20 mM Tris−HCl pH 7.8, 50 mM NaCl, 10 mM imidazole]. Cells were suspended in lysis buffer supplemented with 1 mg/ml lysozyme and incubated on ice for 30 min. The resulting suspension was subsequently sonicated for 3 × 30 s at maximal power. Lysate was cleared by centrifugation at 14000 × g for 30 min at 4°C.

His-tagged proteins were purified via IMAC according to the manufactures protocol (QIAexpressionist, Qiagen) in the gravity flow mode. Lysate passed the Ni-NTA three times, followed by three washing steps with 20 batch volumes washing buffer [20 mM Tris−HCl pH 7.8, 1 M NaCl, 40 mM imidazole]. Elution was done with one batch volume of elution buffer [20 mM sodium phosphate pH 7.8, 500 mM NaCl, 300 mM imidazole] and repeated up to 3 times if desired. Pure recombinant FTL of elution fraction 2 was used for biochemical assays or antibody production.

The FTL activity assay measures the conversion of THF and formate into 10-formyl-THF, which was then quantitatively converted into methenyl-THF by the addition of acid as described (Marx et al., 2003). The assay was performed at 25°C for up to 10 min. Methenyl-THF was determined spectrophotometrically by its characteristic absorption maximum at 350 nm. The 1 ml standard assay mixture contained 0.1 M Tris buffer (pH 8.0), 2 mM tetrahydrofolate (THF) (Merck, Darmstadt, Germany), 10 mM $MgCl_2$, 5 mM ATP, 200 mM sodium formate, and 50 μg cell protein extract. The reaction was stopped at different time points (1, 5, and 10 min) by the addition of 2 ml of 0.36 N HCl. The assay was done under low oxygen condition established by a stream of N_2 to minimize oxidative degradation of the co-substrate THF, whereas the enzyme FTL (EC 6.3.4.3) itself does not contain an oxygen-sensitive cofactor. The absorbance of methenyl-THF was then determined at 350 nm.

Enzyme assay was performed with three technical replicates and given are the mean value ± SD.

Quantification of Soluble Amino Acids and Organic Acids

Pre-cultures had been cultivated under constant illumination and aerated with 5% CO_2 in BG11 medium. The cells were diluted to OD_{750nm} = 1 and shifted to ambient air bubbling either with or without 10 mM sodium formate under constant illumination for 24 h. Free amino acids and organic acids were extracted from frozen *Synechocystis* cell pellets of 10 ml of cultures at OD_{750nm} = 1 using 80% ethanol at 65°C for 3 h. Cell suspensions were mixed thoroughly by shaking every 30 min. Cell debris were removed by centrifugation at $6000 \times g$ for 15 min. the supernatant was lyophilized and re-dissolved in 1 ml MS-grade water (Carl Roth, Karlsruhe, Germany). Amino acids and organic acids were separated through liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) with Discovery H5 F5 HPLC column (Merck, Darmstadt, Germany) as described in Reinholdt et al. (2019).

All assays were repeated 3 times with independent cell cultivation and three technical replicates each. Pair-wise *t*-test was applied for the statistical comparison of mean values of all 9 data sets.

RNA-Isolation and Microarray

For transcriptomics, cells were cultured with or without addition of 10 mM sodium formate for 3 days under constant illumination and ambient air bubbling. Cells from 10 ml of cell suspension were harvested by quick centrifugation at 4°C. The cell pellets were frozen in liquid N_2 and stored at −80°C. RNA isolation, direct RNA labeling and DNA-microarray hybridization were performed as previously described (Gärtner et al., 2019). A high-resolution microarray manufactured by Agilent (Design ID 075764, format 8×60 K; slide layout = IS-62976-8-V2) was used for transcriptomic analysis. The array design allows the direct hybridization of total RNA without conversion into cDNA and covers the probes for all annotated genes as well as other transcripts identified in the course of comprehensive RNA sequencing studies. Before labeling, total RNA was incubated with Turbo DNase (Invitrogen) according to the manufacturer's protocol and precipitated with ethanol/sodium acetate. Further details of the labeling and hybridization protocol can be found in Voß and Hess (2014).

Raw data were further processed with the R package Limma. Median signal intensities were background corrected and quantile normalized. The microarray hybridization was performed with two biological replicates for each treatment. The used array design contained three technical replicates for each single probe and almost all features were covered by several independent probes. Mean values for all probes of a given feature were used for the final calculation of relative transcript ratios normalized to untreated WT. For statistical evaluation, i.e., the *p*-value calculation, the Benjamini–Hochberg procedure was used. Further details of data processing and statistical evaluation using the R software were described previously (Georg et al., 2009). The full array data have been deposited in the GEO database under the accession number GSE143785.

^{14}C-Formate Uptake

^{14}C-Labeled sodium formate was purchased from Merck (Darmstadt, Germany). Pre-cultures had been cultivated under constant illumination and aerated with ambient air and diluted to OD_{750nm} = 1 prior to the experiment. Given amounts of sodium formate containing 5% w/w ^{14}C-labeled sodium formate were added and 1 ml cell suspension were filtered via nitrocellulose membranes ($45\,\mu m$) and immediately washed with 20 ml BG11 medium at given time points. The membranes were transferred into 5 ml scintillation cocktail (Ultima Gold, PerkinElmer) and analyzed in a scintillation counter (Tri-Carb 2810TS, PerkinElmer). Concentration had been calculated from an individually calibration curve for each experiment. All assays were repeated three times with independent cell cultivation.

^{13}C-Labeling Pattern Analysis

For stationary isotope tracing of proteinogenic amino acids, cells were pre-cultivated with CO_2-enriched air (5% CO_2) in BG11 medium. Cells of the WT or strain exFTL were shifted to ambient air starting with OD_{750nm} = 0.2 under continuous light and were cultivated for 5 d in the presence of either ^{13}C-labeled or unlabeled sodium formate. 2 ml of cells (OD_{750nm} = 1) were harvested by centrifugation for 5 min at $11\,000 \times g$. The pellet was hydrolyzed by incubation with 1 ml of 6N hydrochloric acid for 24 h at 95°C. The acid was evaporated by heating to 95°C. Hydrolyzed amino acids were separated and analyzed as described by Yishai et al. (2017). Hydrolyzed amino acids were separated through ultraperformance liquid chromatography (Acquity, Waters, Milford, MA, United States) using a C_{18}-reversed-phase column (Waters) according to previous description. Mass spectra were acquired using an Exactive^{TM} mass spectrometer (Thermo Fisher Scientific). Standards of authentic amino acids (Merck, Darmstadt, Germany) were analyzed under the same conditions in order to determine typical retention times. The program package Xcalibur (Thermo Fisher Scientific) was used for data analysis.

All assays were performed with three biological replicates for each treatment. Representative result shown in here.

RESULTS AND DISCUSSION

Impact of Externally Supplied Formate on *Synechocystis* WT

Prior to establish formate assimilation, we tested the ability of *Synechocystis* to import external formate into the cell. To this end, uptake assays using ^{14}C-labeled sodium formate were performed to verify formate uptake into *Synechocystis* (**Figure 2A**). The rapid initial formate accumulation in the cells was followed by saturation after 20^-30 min. From the uptake measurements we calculated an initial formate uptake rate of 50 nmol formate h^{-1} ml^{-1} OD_{750nm}^{-1} at pH 8. The initial uptake rate increased to 65 nmol formate h^{-1} ml^{-1} OD_{750nm}^{-1}, when the assay was performed at pH 7 instead of pH 8. Under more acidic conditions, formate (pKa of 3.74

FIGURE 2 | Uptake of sodium formate and its effect on Synechocystis cells. **(A)** Uptake assay using 10 mM sodium formate containing 5% w/w ^{14}C-labeled sodium formate with WT cells grown in BG11 at pH 8 and 7, respectively. Given are means of three independent replicates ± SD **(B)** Initial uptake rate of cells incubated for 15 min with 1–200 mM ^{14}C-labeled formate at pH 8. Given are means of three independent replicates ± SD **(C)** Tolerance of WT cells toward formate. Serial dilutions of cell suspension (OD$_{750nm}$ = 1) were spotted on BG11 agar plates, pH 8, supplemented with different concentrations of sodium formate (1–200 mM). Pictures show representative results after 7 days under continuous illumination. **(D)** Photosynthetic O$_2$ evolution rates in the presence of different concentrations of formate (0–50 mM) and light intensities (25–100 µmol photons m^{-2} s^{-1}). Given are means of three independent replicates ± SD (*$p < 0.05$).

for formic acid) is less charged what obviously promoted its initial uptake rate. Further analysis of the initial uptake rate of cells supplemented with different concentration of sodium formate revealed a concentration-dependent increase and did not end up in saturation (**Figure 2B**). Formate, whether dissociated or not, should traverse the outer membrane easily via porins and uptake is rather limited by the permeability of the plasma membrane. Assuming an active transport of the formate anion, a concentration of 200 mM exogenous sodium formate should exceed the affinity of a putative transporter (e.g., FocA in *E. coli* with a K$_m$ 119 mM, Wiechert and Beitz, 2017) and end up in no further increase in its uptake rate. However, a linear decrease of the uptake rate correlating with the applied exogenous concentration was observed indicating that non-dissociated sodium formate is entering the cell rather via diffusion than trough specific transporters. Furthermore, a formate-nitrite transporter family, facilitating formate import and export in proteobacteria (Suppmann and Sawers, 1994), was not yet identified in cyanobacteria (Hahn and Schleiff, 2014).

It has been shown that formate can be toxic for oxygenic phototrophs at higher concentrations by interfering with the

bicarbonate-binding site at photosystem II (Stemler and Radmer, 1975; Semin et al., 1990; Shevela et al., 2007). Hence, the impact of endogenous formate on *Synechocystis* wild type (WT) was investigated. First, the formate tolerance of WT cells was studied on agar plates supplemented with formate concentrations ranging from 0 to 200 mM (**Figure 2C**). The growth of *Synechocystis* was unchanged up to 50 mM formate and became somewhat reduced if formate concentrations exceeded 100 mM. Nevertheless, cells survived formate concentrations of up to 200 mM. Second, the growth rates of WT cells were evaluated in liquid media to characterize the long-term impact of formate. No significantly different growth 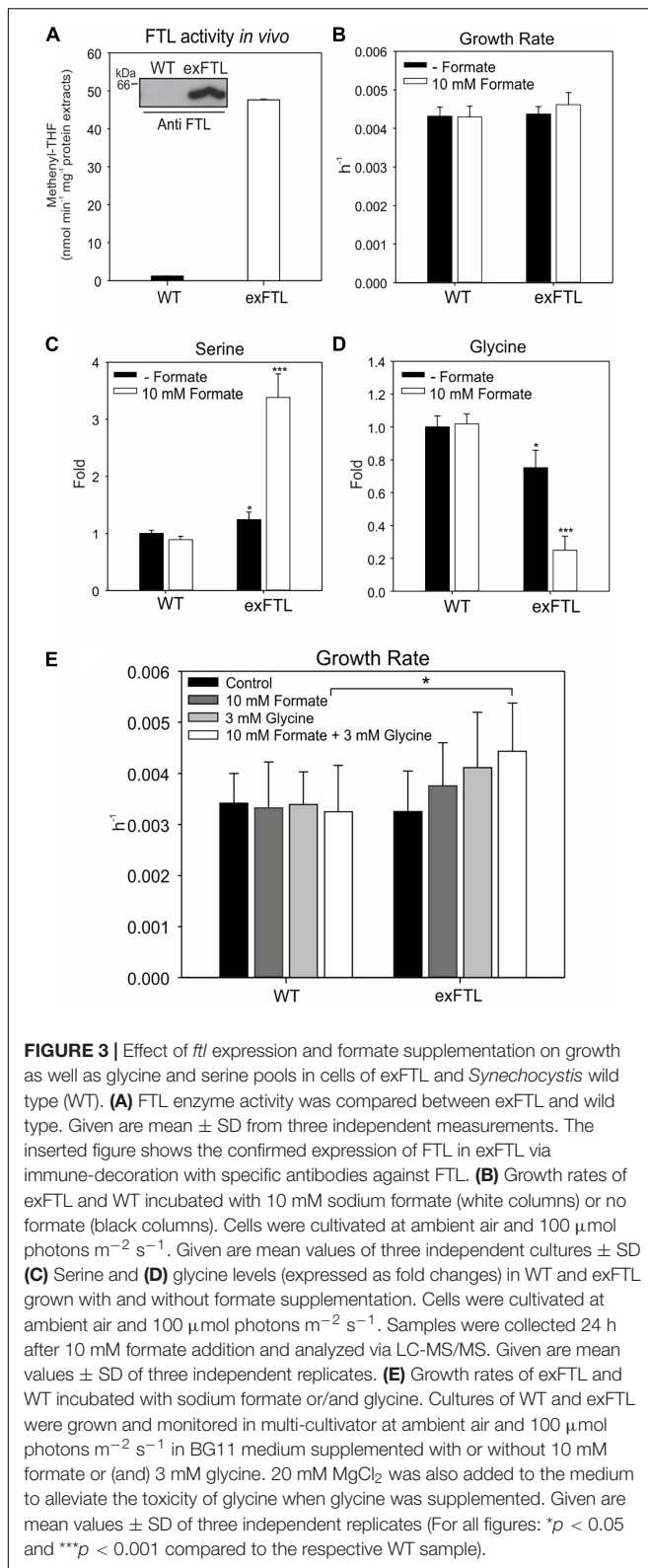rate was observed in the presence of 10 and 20 mM formate compared to non-treated cells over a time of 7 days. Similar observations were made under different light intensities (50, 100, and 200 µmol photons m^{-2} s^{-1}) and different inorganic carbon concentrations (0.04% or 5% CO$_2$), respectively (**Supplementary Table S2**). Third, effects of formate on photosynthesis were studied using *Synechocystis* WT cells exposed to different formate concentrations and light intensities during measuring photosynthetic oxygen evolution. Oxygen evolution was unaffected at 10 mM formate under all tested light intensities (25, 50, or 100 µmol photons m^{-2} s^{-1}), but severely inhibited at 50 mM formate at light intensities of 50 and 100 µmol photons m^{-2} s^{-1} (**Figure 2D**).

Collectively, these data indicated that formate is entering the cells probably via diffusion and low concentrations up to 20 mM of formate were well tolerated by *Synechocystis*, whilst higher concentrations (>50 mM) had a negative impact on photosynthesis and growth. As demonstrated for *E. coli*, the supplementation of 10 mM formate provided sufficient C1 units for serine synthesis via the formate assimilation pathway in a serine auxotrophic strain (Yishai et al., 2017; Kim et al., 2019). Therefore, 10 mM formate was used for subsequent experiments. Generally, the above described experiments verified that *Synechocystis* provides a suitable chassis to implement the formate assimilation pathway.

Effects of Formate-Tetrahydrofolate Ligase (FTL) Expression on Synechocystis

The heterologous expression of *ftl* should be sufficient to complete a formate assimilation pathway, because all other necessary enzymes are annotated in the *Synechocystis* genome (gene bank accession NC_000911). The FTL from *Methylobacterium extorquens* AM1 was chosen for expression in *Synechocystis,* as the expression of this gene successfully supported formate assimilation in *E. coli* via the desired pathway (Yishai et al., 2018; Kim et al., 2020).

The *ftl* gene was stably inserted into the *psbA2*-site on the *Synechocystis* chromosome, thus, its expression is controlled by the strong, light-induced promoter P$_{psbA2}$ (Lagarde et al., 2000). Genotype and expression of *ftl* in the resulting exFTL strain were confirmed via PCR, Coomassie-staining, and Western-blotting. Furthermore, FTL enzyme activity was detected in exFTL but not in the WT (**Figure 3A**). As expected, the expression of *ftl* driven by the *psbA2* promoter resulted in strong protein accumulation

FIGURE 3 | Effect of *ftl* expression and formate supplementation on growth as well as glycine and serine pools in cells of exFTL and *Synechocystis* wild type (WT). **(A)** FTL enzyme activity was compared between exFTL and wild type. Given are mean ± SD from three independent measurements. The inserted figure shows the confirmed expression of FTL in exFTL via immune-decoration with specific antibodies against FTL. **(B)** Growth rates of exFTL and WT incubated with 10 mM sodium formate (white columns) or no formate (black columns). Cells were cultivated at ambient air and $100\ \mu$mol photons $m^{-2}\ s^{-1}$. Given are mean values of three independent cultures ± SD **(C)** Serine and **(D)** glycine levels (expressed as fold changes) in WT and exFTL grown with and without formate supplementation. Cells were cultivated at ambient air and $100\ \mu$mol photons $m^{-2}\ s^{-1}$. Samples were collected 24 h after 10 mM formate addition and analyzed via LC-MS/MS. Given are mean values ± SD of three independent replicates. **(E)** Growth rates of exFTL and WT incubated with sodium formate or/and glycine. Cultures of WT and exFTL were grown and monitored in multi-cultivator at ambient air and $100\ \mu$mol photons $m^{-2}\ s^{-1}$ in BG11 medium supplemented with or without 10 mM formate or (and) 3 mM glycine. 20 mM $MgCl_2$ was also added to the medium to alleviate the toxicity of glycine when glycine was supplemented. Given are mean values ± SD of three independent replicates (For all figures: $*p < 0.05$ and $***p < 0.001$ compared to the respective WT sample).

in exFTL (**Supplementary Figure S1**). Phenotyping of exFTL revealed a growth rate similar to WT without additional external formate, while the addition of 10 mM sodium formate led to a minor stimulating effect on exFTL (**Figure 3B**).

Metabolic Consequences of *ftl* Expression in *Synechocystis*

Next, we compared the metabolome of WT and exFTL in the presence and absence of 10 mM formate. Significant differences, particularly in serine and glycine pools (**Figures 3C,D**) but also in many other metabolites (**Figure 4A**) were observed upon formate addition. The addition of formate resulted in approximately 3-fold higher serine levels in exFTL compared to WT, while in the absence of formate only a slight difference was found (**Figure 3C**). In contrast, glycine decreased approximately 3-fold in exFTL upon formate supplementation, whereas only small differences appeared without formate addition (**Figure 3D**). The decreased glycine content implied that the amount of glycine, as a precursor of serine in the SHMT reaction, might be the limiting factor for higher formate incorporation into serine. Therefore, growth experiments were performed, in which the medium was supplemented with 3 mM glycine together with 10 mM sodium formate. In addition, 20 mM $MgCl_2$ were added to alleviate the toxicity of glycine to *Synechocystis* (Eisenhut et al., 2007). Consistent with our assumption, exFTL showed significantly higher growth rate than WT in the presence of both formate and glycine, whereas only minor, non-significant difference appeared when only glycine or formate was added (**Figure 3E**).

However, the changes in serine and glycine pools were part of more general metabolic alterations in exFTL, whereas the metabolome of WT was almost unaffected by the addition of formate (**Figure 4A** and **Supplementary Figure S2**). For example, the level of the RubisCO carboxylation product 3PGA increased, while the steady state amount of the RubisCO oxygenation product 2PG decreased in exFTL. Furthermore, the expression of *ftl* in *Synechocystis* caused significant decreased contents of 2-oxoglutarate (2OG) and other intermediates of the tricarboxylic acid (TCA) cycle already in the absence of formate. Similar observations were made for the amino acids leucine, proline, histidine, valine and phenylalanine. Some of these changes were intensified by addition of 10 mM sodium formate (**Figure 4A** and **Supplementary Figure S2**). In case of valine, the addition of formate rescued the initial decrease to a WT-like level and histidine even exceeded the WT level by 2-fold. For lysine, threonine and asparagine, all originating from oxaloacetate, the contents increased upon formate addition in comparison to non-treated exFTL and WT, respectively.

The two amino acids most directly linked to the C1-pool via their THF-derivatives mediated biosynthesis – methionine and histidine – showed distinct regulations upon formate addition. Whereas the histidine level in exFTL clearly increased upon formate addition, methionine seemed to be unaffected under all growth conditions (**Figure 4A** and **Supplementary Figure S2**). Interestingly, the ratio between glutamine and glutamate changed in exFTL upon formate addition (**Figure 4A**). These changes are usually related to C_i-limiting conditions and in line with changed 2PG amounts (Eisenhut et al., 2008). Whereas the alterations of other metabolites like citrate, succinate as well as serine and glycine are consistent with a C_i-limited phenotype, 2PG, 3PGA and 2OG reacted completely oppositional to this hypothesis. Furthermore, not all of these metabolites were affected solely by expression of *ftl* independent of formate addition.
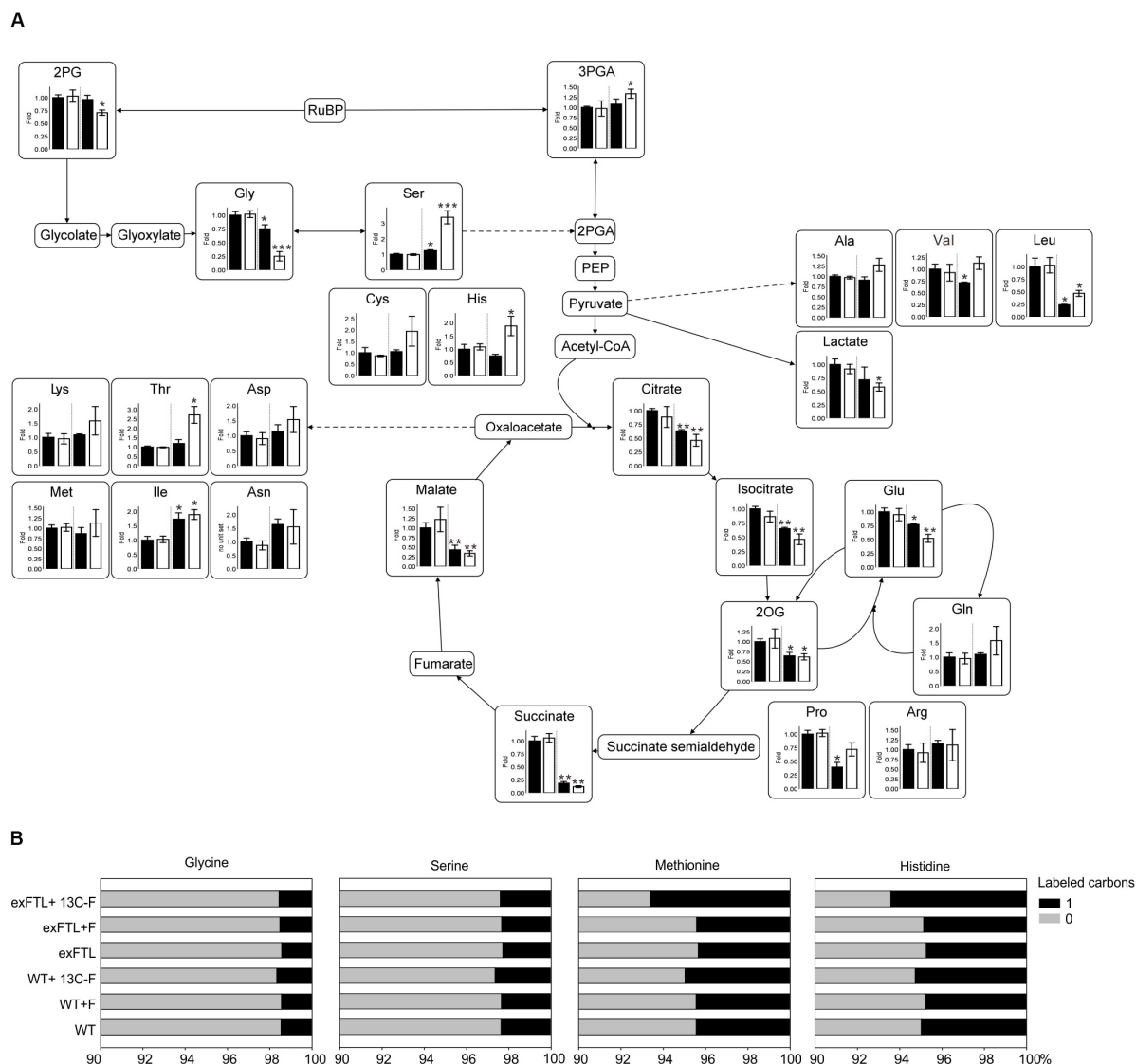
FIGURE 4 | Relative changes of selected metabolites and ^{13}C labeling pattern in exFTL compared to *Synechocystis* wild type (WT). **(A)** Cells of WT or exFTL were collected 24 h after addition of 10 mM sodium formate (white columns) or without (black columns) incubated at ambient air and 100 µmol photons m^{-2} s^{-1}. Metabolite contents are mean values ± SD relative to the WT mean value from three independent biological replicates.(*$p < 0.05$; **$p < 0.01$; and ***$p < 0.001$ compared to the respective WT sample) 2PG, 2-phosphoglycolate; 3-PGA, 3-phosphoglycerate; Gly, glycine; Ser, serine; Cys, cysteine; His, histidine; Ala, alanine; Val, valine; Leu, leucine; Lys, lysine; Thr, threonine; Asp, aspartate; Asn, asparagine; Met, methionine; Ile, isoleucine; Glu, glutamate; Gln, glutamine; Pro, proline; Arg, arginine; 2OG, 2-oxoglutarate. **(B)** *Synechocystis* WT and exFTL were pre-cultivated under high carbon condition and thereafter transferred to ambient air condition incubated with either unlabeled formate, ^{13}C-labeled formate or without formate under ambient air and 100 µmol photons m^{-2} s^{-1} for 5 days. Given are results from one typical experiment.

Among all detected metabolites, the alpha aminobutyric acid (AABA) showed the highest relative change. Its levels increased up to 8-fold in exFTL upon formate supplementation (**Supplementary Figure S2**). This metabolite might originate from serine breakdown to cysteine via oxobutanoate (**Supplementary Figure S3**), which could also explain the formate-induced increase in cysteine.

To verify whether externally supplied formate was incorporated into cellular biomass, the ^{13}C-labeling pattern of proteinogenic amino acids was evaluated in cells grown in the presence of ^{13}C-labeled sodium formate for 5 days. The ^{13}C-incorporation into the amino acids methionine, histidine, glycine, and serine was analyzed to elucidate whether the C1-building blocks for their biosynthesis derived from ^{13}C-formate in exFTL. As expected, glycine was unlabeled (**Figure 4B**), what proofs that formate oxidation did not occur in *Synechocystis* as it would cause labeling of all amino acids. However, serine also appeared completely unlabeled in exFTL despite its massive accumulation in the presence of formate. Only methionine and histidine were slightly more ^{13}C-labeled in exFTL compared

to incubation with non-labeled formate (**Figure 4B**). These results indicated that a rather small amount of formate–derived C1-units was used for methionine and histidine synthesis but not for serine production. Unlike previous studies with designed *E. coli* strains, supplied formate was not the source for the enhanced serine pool in *Synechocystis* (Yishai et al., 2017, 2018; Bang and Lee, 2018; Döring et al., 2018; Kim et al., 2019).

Analysis of Transcriptome Changes in exFTL

The ^{13}C-labeling results revealed that the enhanced serine accumulation in exFTL did not result from significant FTL-mediated formate incorporation. Therefore, the changed serine/glycine ratio likely originated from some regulatory impact of *ftl* expression in *Synechocystis*. To verify this assumption, transcriptomic analyses were preformed using a DNA microarray with RNA isolated from cells of the WT and exFTL, cultured with and without formate. Significant differences in the gene expression between WT and exFTL were detected in the absence of formate, whereas the addition of formate changed the expression pattern in exFTL further (**Supplementary Figure S4**).

In contrast, formate addition had only a minor impact on gene expression in WT cells, which is consistent with the small changes regarding growth and metabolome reported before. Basically, only the *hliB-lilA* (*ssr2595-slr1544*) operon, which encode the small chlorophyll a-binding-like protein ScpD (also called HliB) and LilA that is a member of the extended light-harvesting-like protein family (Kufryk et al., 2008), exceeded the threshold of 2-fold induction. Furthermore, *cmpB* (*slr0041*) and *cmpC* (*slr0043*) encoding subunits of the ABC-type bicarbonate transporter BCT1 showed more than 2-fold lowered expression (the other genes for the BCT1 transporter were also down-regulated but below significance level, complete data set available under accession number GSE143785).

The global comparison of gene expression revealed that 272 transcripts became significantly ($p < 0.05$) more strongly expressed whereas 232 were more lowly expressed in exFTL compared to WT under both conditions (threshold for significance was 2-fold). The upregulated genes belong to many different categories, comprising many genes for ribosomal proteins and enzymes of the C1, nitrogen and carbon metabolism (for examples see **Tables 1**, **2**). Significant changes in expression

TABLE 1 | Expression of genes encoding proteins involved in photosynthesis, photorespiration, as well as primary carbon, amino acid and purine metabolism in exFTL compared to wild type (WT).

Cyano Base Acc.	Gene	Description	WT	WT + formate	exFTL	exFTL + formate
Phosphoserine pathway						
slr1124	psp	phosphoserine phosphatase	1	1.04	1.54	1.35
sll1908	serA	3PGA dehydrogenase	1	1.07	2.00*	1.96
sll1559	pstA	phosphoserine & Gly_Ala transaminase	1	1.07	2.43*	2.13
Photosynthesis						
slr0009	rbcL	ribulose bisphosphate carboxylase large subunit	1	0.82	0.91	0.35
slr0012	rbcS	ribulose bisphosphate carboxylase small subunit	1	0.77	0.60	0.21
slr1840	slr1840	glycerate kinase	1	1.04	2.27*	2.00
Photorespiration						
sll1931	glyA	serine hydroxymethyltransferase	1	0.92	1.16	0.77
slr0879	gcvH	glycine decarboxylase complex H-protein	1	0.88	3.44*	3.77*
Glycolysis/TCA cycle/CBB						
sll1196	pfk	phosphofructokinase	1	0.91	0.61	1.23
slr0783	tpiA	triosephosphate isomerase	1	0.98	1.85	2.24
slr0884	gap1	glyceraldehyde 3-phosphate dehydrogenase 1 (NAD$^+$)	1	1.02	0.65	1.33
slr1945	yibO	2.3-bisphosphoglycerate-ind. P-glycerate mutase	1	1.09	2.34*	2.02
slr1934	slr1934	pyruvate dehydrogenase E1 component. α subunit	1	0.99	2.04*	2.18*
slr1289	icd	isocitrate dehydrogenase (NADP +)	1	0.88	0.48	0.88
Methionine/histidine metabolism						
sll0900	hisG	ATP phosphoribosyltransferase	1	1.18	2.24*	2.39*
slr1705	aspA	aspartoacylase	1	1.22	3.12*	1.39
Purine metabolism						
sll1823	purA	adenylosuccinate synthetase	1	1.03	3.14*	3.97*
slr0520	purL	phosphoribosyl formylglycinamidine synthase	1	0.97	1.90	1.94
sll1852	ndkR	nucleoside diphosphate kinase	1	1.16	3.75*	3.63*
sll1815	adk	adenylate kinase	1	0.83	2.28*	1.53
slr0597	purH	phosphoribosylaminoimidazolecarboxamide formyltransferase	1	1.10	1.21	1.26
slr0477	purN	phosphoribosylglycinamide formyltransferase	1	1.03	1.44	1.36

*Synechocystis cultures of WT and exFTL were cultivated at ambient air supplemented with 10 mM sodium formate or without. Samples from WT and exFTL for RNA isolation were taken after 3 days. Results are given as relative transcript levels normalized to WT without formate addition. Changes that are exceeding the threshold are marked in red. Genes with significant (p < 0.05) formate-dependent changes in exFTL are highlighted in yellow (*p < 0.05 to the respective WT sample).*

TABLE 2 | Expression of genes related to folate synthesis and nitrogen metabolism in exFTL compared to wild type (WT).

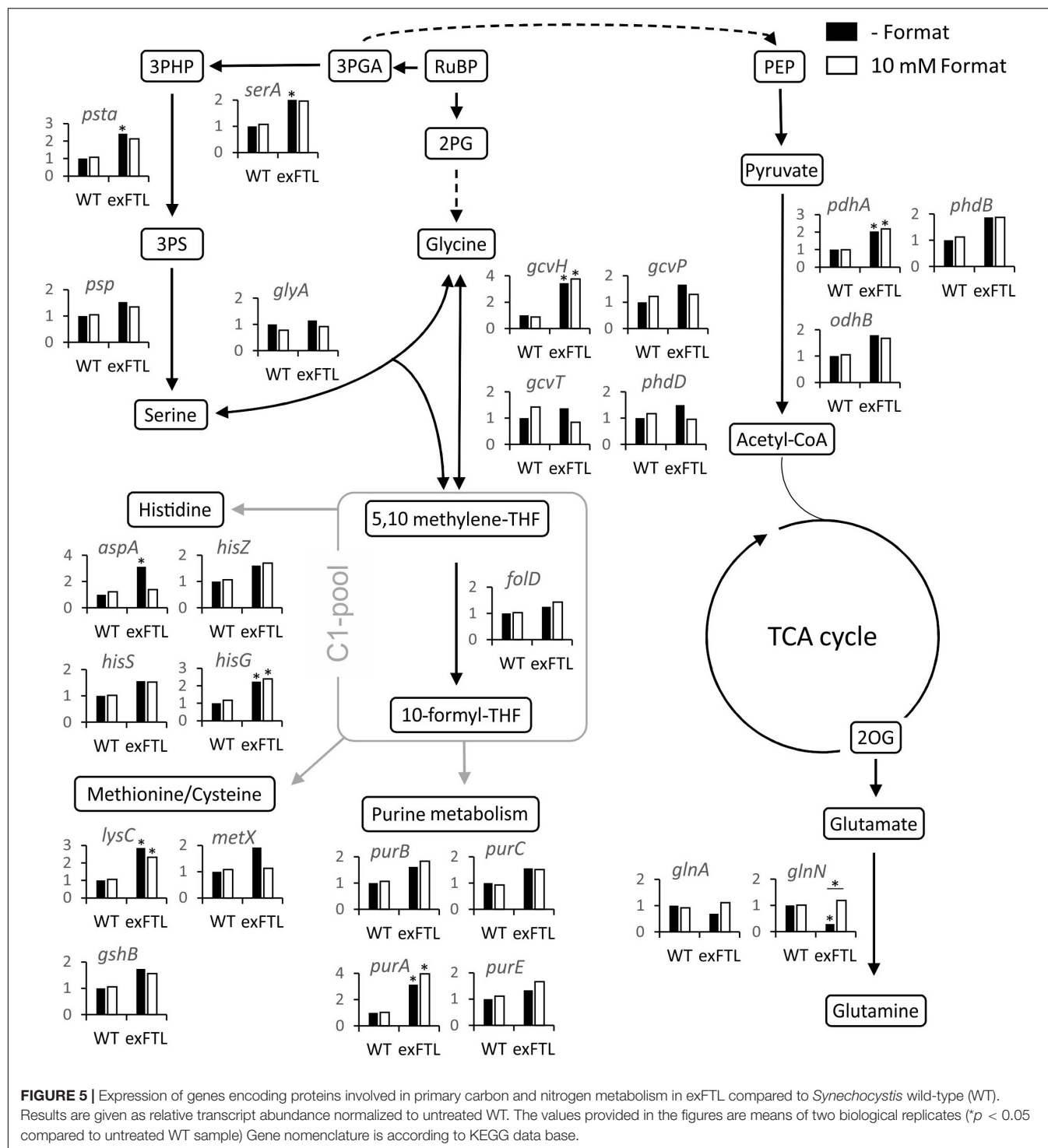| Cyano Base Acc.	Gene	Description	WT	WT + formate	exFTL	exFTL + formate
Folate biosynthesis						
slr0900	moaA	molybdopterin biosynthesis protein A	1	0.96	0.21*	0.52*
slr0901	moeA	molybdopterin biosynthesis MoeA protein	1	0.97	0.30*	0.45
slr0902	moaC	molybdenum cofactor biosynthesis protein C	1	0.91	0.37*	0.48
slr0903	moaE	molybdopterin (MPT) converting factor. subunit 2	1	0.87	0.39*	0.46
slr1626	slr1626	dihydroneopterin aldolase	1	0.95	1.70	2.39*
Nitrogen metabolism						
sll1450	nrtA	nitrate/nitrite transporter substrate-binding protein	1	0.63	0.14*	0.52
sll1451	nrtB	nitrate/nitrite transporter permease protein	1	0.72	0.12*	0.45
sll1452	nrtC	nitrate/nitrite transporter ATP-binding protein	1	0.77	0.14*	0.44
sll1453	nrtD	nitrate/nitrite transporter ATP-binding protein	1	0.71	0.12*	0.27*
sll1454	narB	ferredoxin-nitrate reductase	1	0.76	0.21*	0.30*
sll0450	norB	cytochrome b subunit of nitric oxide reductase	1	0.67	0.33*	0.23*
slr0898	nirA	ferredoxin–nitrite reductase	1	0.90	0.14*	0.99
slr0288	glnN	glutamate–ammonia ligase	1	1.01	0.29*	1.19
slr0899	cynS	cyanate lyase	1	0.81	0.14*	0.60
sll1454	narB	ferredoxin-nitrate reductase	1	0.76	0.21*	0.30*
ssl1911	gifA	glutamine synthetase inactivating factor IF7	1	0.92	25.62*	2.56*
sll1515	gifB	glutamine synthetase inactivating factor IF17	1	1.39	15.31*	1.88
ncl0540	NsiR4	sRNA, translational inhibitor of IF7	1	1.06	0.28*	1.18
sll1270	glnH/glnP	periplasmic substrate-binding protein of Bgt	1	0.94	0.59	2.07*
sll0108	amt1	ammonium/methylammonium permease	1	0.80	0.17*	0.87
sll1017	amt2	ammonium/methylammonium permease	1	1.11	0.34*	1.42
ssl0707	glnB	nitrogen regulatory protein P-II	1	0.95	0.56	2.09*
sll1423	ntcA	global nitrogen regulator	1	1.07	1.99	2.12

Synechocystis cultures of WT and exFTL were cultivated at ambient air supplemented with 10 mM sodium formate or without. Samples from WT and exFTL for RNA isolation were taken after 3 days. Results are given as relative transcript levels normalized to WT without formate addition. Changes that are exceeding the threshold are marked in red. Genes that show strong formate-dependent changes in exFTL are marked in yellow ($p < 0.05$ to the respective WT sample).*

were also found for genes involved in serine biosynthesis and related processes, for example photorespiration (**Figure 5**). To evaluate the FTL-mediated effect on transcript abundance, we initially focused on genes coding for enzymes closely related to FTL activity.

Serine can be synthesized by the photorespiratory 2PG metabolism and by the phosphoserine pathway in *Synechocystis* (Klemke et al., 2015). In the latter pathway serine originates from 3PGA and is made by three enzymatic reactions: 3-phosphoglycerate dehydrogenase (SerA), phosphoserine transaminase (PSTA) and phosphoserine phosphatase (PSP). The genes *serA* and *pstA* were significantly upregulated in exFTL compared to WT, whereas *psp* only showed a slight, non-significant increase in transcript abundance. Upon formate addition expression of none of the three genes was further enhanced, but rather slightly repressed (**Figure 5** and **Table 1**). The serine pool is connected to the glycine pool by the reversible action of SHMT (encoded by *glyA*) that converts serine into glycine and methylene-THF or *vice versa*. Glycine can also derive from the photorespiratory 2PG metabolism, where it is decarboxylated by glycine decarboxylase (GDC). Expression of all four genes for the GDC were coordinately enhanced in exFTL, but only the *gcvH* gene encoding the H protein subunit was significantly up-regulated (genes for P-, T- and L-protein subunits were around 1.5-times higher expressed). However, the
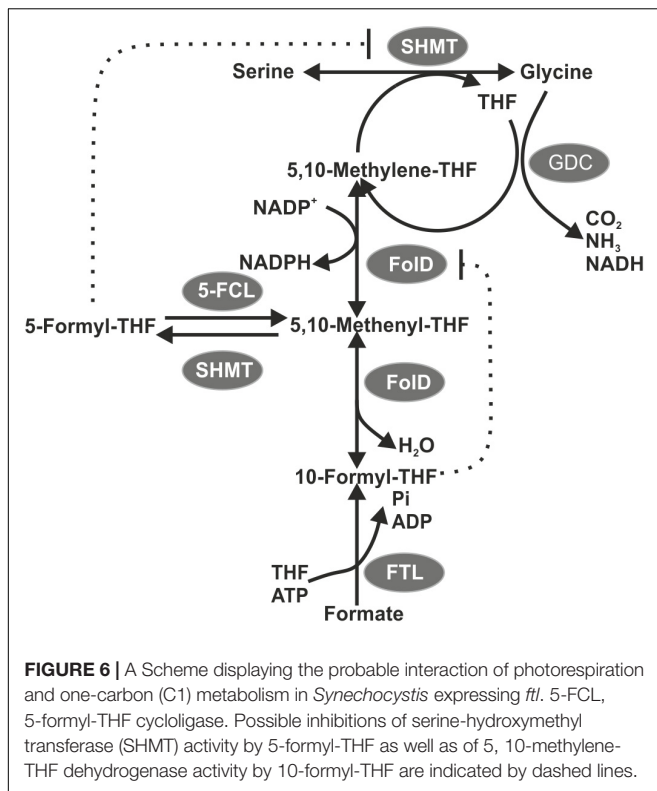
addition of formate reduced the expression of genes for the GDC subunits L and P to a WT-like level, which makes it unlikely that the decreased amount of glycine in formate-supplemented exFTL was due to altered GDC expression. Only the expression of *glyA* encoding SHMT was slightly down-regulated upon formate addition. However, expression changes were not consistent with the observed serine accumulation and glycine consumption under formate-supplemented conditions in exFTL, since the strongest transcript changes were observed if this strain was grown under formate-free conditions (**Table 1**). The slight downregulation of *glyA* expression upon formate addition in exFTL could partially explain the serine accumulation, when we assume that SHMT activity is mostly used to synthetize C1-units by converting serine into glycine and 5,10-methylene THF and to minor extend converts 5,10-methylene-THF and glycine into serine (photorespiratory direction). Furthermore, it is known that SHMT activity is inhibited by 5-formyl-THF (**Figure 6**), which can be produced from 5,10-methenyl-THF by SHMT itself (Goyer et al., 2005; Collakova et al., 2008). Hence, the combined small downregulation of its expression and possible inhibition of SHMT activity are consistent with the observed serine accumulation upon formate addition in exFTL (**Figure 6**).

In addition, enzymes involved in purine, methionine and histidine biosynthesis were more highly expressed under formate-free conditions in exFTL than in WT (**Table 1**), which

FIGURE 5 | Expression of genes encoding proteins involved in primary carbon and nitrogen metabolism in exFTL compared to *Synechocystis* wild-type (WT). Results are given as relative transcript abundance normalized to untreated WT. The values provided in the figures are means of two biological replicates (*$p < 0.05$ compared to untreated WT sample) Gene nomenclature is according to KEGG data base.

together with the changed expression of genes involved in folate biosynthesis (**Table 2**) indicates that the expression of *ftl* interfered with the C1-metabolism of *Synechocystis*. For example, slightly increased expression of the folate-dependent formyltransferases *purN* and *purH* were observed (**Table 1**). The formyltransferase transfers formyl group from 10-formyl-THF and releases THF during purine biosynthesis, therefore,

their increased activity would reduce the 10-formyl-THF pool in exFTL, which could have an impact on other shunts of the C1 metabolism as well. Most probably, the heterologous FTL is changing the balance of the different C1-intermediates bound to THF (see **Figure 6**) in *Synechocystis*, which is somehow sensed by the cell leading to the observed changes in the transcriptome and metabolome. Furthermore, it has been reported that the

FIGURE 6 | A Scheme displaying the probable interaction of photorespiration and one-carbon (C1) metabolism in *Synechocystis* expressing *ftl*. 5-FCL, 5-formyl-THF cycloligase. Possible inhibitions of serine-hydroxymethyl transferase (SHMT) activity by 5-formyl-THF as well as of 5, 10-methylene-THF dehydrogenase activity by 10-formyl-THF are indicated by dashed lines.

dehydrogenase activity of FolD is inhibited by an enhanced 10-formyl-THF pool in *E. coli* FolD (Dev and Harvey, 1978). This FolD inhibition could affect the equilibrium between the C1 units 5,10-methylene-THF and 5,10-methenyl-THF. Furthermore, it was shown with human cell lines that the FolD homolog C-1-tetrahydrofolate synthase is responsible for maintaining the redox balance between $NADP^+$ and NADPH (Fan et al., 2014). A similar effect was assumed for a *ftl*-supported $\Delta folD$ *E. coli* strain showing a glycine and purine auxotrophy (Sah et al., 2015). Hence, our results indicate that the bifunctional enzyme FolD in *Synechocystis* might be the limiting factor for the assimilation of formate into serine. Most likely, the FTL-related imbalance in the primary and C1 metabolism prevented efficient formate flux into the metabolites of *Synechocystis*.

In contrast to our expectations that FTL expression, especially upon formate addition, mostly influences genes for enzymes associated with C1 metabolism and purine biosynthesis, the strongest changes in transcript abundance were observed for genes connected to nitrogen uptake and assimilation (**Table 2**). Most pronounced differences were found for the glutamine synthase inactivating factors IF17 and IF7, with a 15-fold and 25-fold enhanced transcript level in exFTL compared to WT (**Table 2**). In addition, the gene encoding NsiR4, a regulatory sRNA inhibiting translation of IF7 (Klähn et al., 2015) was repressed in exFTL compared to WT, which is consistent with the enhanced mRNA level of IF7. Several other genes that are closely related to nitrogen metabolism in *Synechocystis* (Giner-Lamia et al., 2017) were significantly down regulated in exFTL. For example, this was overserved for the genes

encoding the ABC-type nitrate/nitrite transporter (*nrtABCD*), the ammonium/methylammonium permease (*amt1/2*) or the glutamate ammonia ligase (*glnN*) (**Table 2**). All these genes are regulated by NtcA (*ntcA*: 2-fold enhanced in exFTL, **Table 2**), the global nitrogen regulator, which senses the 2OG level as a measure of the cellular nitrogen status (Giner-Lamia et al., 2017). 2OG is synthesized in the oxidative branch of the TCA cycle, which is mostly open among cyanobacteria due to the lack of 2OG dehydrogenase complex. Therefore, 2OG is the final product of oxidative degradation of organic carbon and is mainly used as carbon precursor for ammonia assimilation in the glutamine synthase-glutamate synthase (GS/GOGAT) cycle. Hence, 2OG directly links carbon and nitrogen metabolism making it the major signal molecule to sense the C/N ratio *in vivo* (Zhang et al., 2018). The activity of the transcriptional factor NtcA is strongly regulated by the PII protein (*glnB*; 0.6-fold expression in exFTL, **Table 2**), which in dependence on the amount of 2OG and ATP regulates NtcA activity via an adaptor protein PipX. For example, at high N/C ratios with a lowered 2OG level, PII efficiently binds to PipX and NtcA free of PipX is inactivated, while under low N/C ratios with high 2OG amounts, NtcA-2OG is activated due to strong PipX-binding (Forchhammer and Selim, 2020). According to this model, the decreased 2OG level in exFTL indicates a relative nitrogen-rich status in exFTL (Forchhammer and Lüddecke, 2016). However, formate addition did not change the 2OG level in exFTL but reversed the expression of low-nitrogen-induced genes, for example the genes involved in ammonium uptake and nitrogen assimilation increased to a WT-like level and the transcripts for GS-inactivating factors decreased in comparison to the non-treated exFTL (**Figure 3** and **Table 2**). Hence, the contrary regulation of gene expression and 2OG content in the presence or absence of formate makes a direct impact of 2OG-signaling unlikely, analogous to the C_i-limited phenotype indicated by the metabolite profile changes described above. In this regard it is interesting to note that *ftl* expression causes a formate-independent downregulation of genes involved molybdopterin synthesis which is slightly stimulated in exFTL upon formate addition (**Table 2**). Molybdopterin is an essential cofactor for the function of nitrate reductase, the first enzyme initiating the assimilation of the inorganic nitrogen (Berks et al., 1995; Lin and Stewart, 1998). Hence, the reversible regulation of genes for molybdopterin biosynthesis in exFTL with or without formate could impact the nitrate assimilation and explain the strong effects of *ftl* expression on the overall expression of N-related genes.

Compared to genes related to N-assimilation, genes related to carbon fixation (*rbcL, rbcS*) and of the TCA cycle and glycolysis were less affected by the expression of FTL than by the addition of formate (**Table 1**). The genes encoding RubisCO (*rbcL, rbcS*) were significantly lower expressed in exFTL compared to WT in the presence of formate. Some genes for enzymes of the primary carbon metabolism such as triosephosphate isomerase, phosphofructokinase and were stimulated in exFTL upon addition of formate (**Table 1**). All of them are involved in the breakdown of organic carbon. Therefore, their enhanced expression could be linked to a need for enhanced mobilization of organic carbon reserves due to the lowered RubisCO

transcript abundance (e.g., Shinde et al., 2019). The only genes encoding enzymes involved in C metabolism that show an opposite regulation in exFTL upon formate addition were 2,3-diphosphoglycerate-independent phosphoglycerate mutase and the α-subunit of the pyruvate dehydrogenase E1 (**Table 2**). These findings indicate that formate is probably not used nor sensed as an additional carbon source by *Synechocystis*.

CONCLUSION

We aimed to establish a formate assimilation pathway in the cyanobacterium *Synechocystis* to analyze its potential contribution to CO_2 assimilation via the CBB cycle. Our experiments confirmed that this cyanobacterium represents a suitable chassis for such an attempt, given that external sodium formate is taken up and low amounts of formate are well tolerated without significant effects on growth, photosynthesis, metabolome and transcriptome of WT cells. However, the expression of *ftl* caused defects in sensing the carbon and nitrogen status of the cells. Our results indicate that there is the possibility of an additional, not yet described signal molecule involved in this sensing mechanism. Despite the observed changes in transcript abundances, the exFTL strain grew similarly well as WT under our standard conditions as well as in the presence of formate. These findings show that *Synechocystis*' physiology can compensate rather large metabolic changes due to transcriptional remodeling without significant effects on growth.

Many other cyanobacteria have been successfully engineered to produce a great variety of biofuels and chemical feedstock, but the observed production titers were usually low (reviewed in Hagemann and Hess, 2018). These limitations could be due to unintended side effects as illustrated here, genetic instability or other yet unknown reasons (Jones, 2014). To minimize these phenomena, comprehensive genetic and metabolic changes of cyanobacteria are necessary as recently shown in the case study on butanol production with *Synechocystis* (Liu et al., 2019). Our study provides an advance toward the possible impacts of metabolic engineering on cyanobacteria.

AUTHOR CONTRIBUTIONS

MH conceptualized the project. SS performed the physiological and biochemical experiments. E-MB performed ^{14}C-formate uptake experiments and RNA isolation. SL performed analysis of ^{13}C-labeled proteinogenic amino acids and data evaluation. ST performed LC/MS-MS measurements. VR and WH performed the microarray analysis and data evaluation. SS, MH, and E-MB made the final data evaluation and wrote the manuscript with contributions from all other authors. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.01650/full#supplementary-material

REFERENCES

Badger, M. R., and Price, G. D. (2003). CO2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J. Exp. Bot.* 54, 609–622. doi: 10.1093/jxb/erg076

Bang, J., and Lee, S. Y. (2018). Assimilation of formic acid and CO2 by engineered *Escherichia coli* equipped with reconstructed one-carbon assimilation pathways. *Proc. Natl. Acad. Sci. U.S.A.* 115, E9271–E9279.

Bar-Even, A. (2016). Formate assimilation: the metabolic architecture of natural and synthetic pathways. *Biochemistry* 55, 3851–3863. doi: 10.1021/acs.biochem.6b00495

Bar-Even, A. (2018). Daring metabolic designs for enhanced plant carbon fixation. *Plant Sci.* 273, 71–83. doi: 10.1016/j.plantsci.2017.12.007

Bar-Even, A., Noor, E., Flamholz, A., and Milo, R. (2013). Design and analysis of metabolic pathways supporting formatotrophic growth for electricity-dependent cultivation of microbes. *Biochim. Biophys. Acta* 1827, 1039–1047. doi: 10.1016/j.bbabio.2012.10.013

Bar-Even, A., Noor, E., Lewis, N. E., and Milo, R. (2010). Design and analysis of synthetic carbon fixation pathways. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8889–8894. doi: 10.1073/pnas.0907176107

Bauwe, H., Hagemann, M., and Fernie, A. R. (2010). Photorespiration: players, partners and origin. *Trends Plant Sci.* 15, 330–336. doi: 10.1016/j.tplants.2010.03.006

Berks, B. C., Ferguson, S. J., Moir, J. W. B., and Richardson, D. J. (1995). Enzymes and associated electron transport systems that catalyse the respiratory reduction of nitrogen oxides and oxyanions. *Biochim. Biophys. Acta Bioenerg.* 1232, 97–173. doi: 10.1016/0005-2728(95)00092-5

Collakova, E., Goyer, A., Naponelli, V., Krassovskaya, I., Gregory, J. F., Hanson, A. D., et al. (2008). Arabidopsis 10-formyl tetrahydrofolate deformylases are essential for photorespiration. *Plant Cell* 20, 1818–1832. doi: 10.1105/tpc.108.058701

Dev, I. K., and Harvey, R. J. (1978). N10-Formyltetrahydrofolate is the formyl donor for glycinamide ribotide transformylase in *Escherichia coli*. *J. Biol. Chem.* 253, 4242–4244.

Döring, V., Darii, E., Yishai, O., Bar-Even, A., and Bouzon, M. (2018). Implementation of a reductive route of one-carbon assimilation in *Escherichia coli* through directed evolution. *ACS Synth. Biol.* 7, 2029–2036. doi: 10.1021/acssynbio.8b00167

Eisenhut, M., Bauwe, H., and Hagemann, M. (2007). Glycine accumulation is toxic for the cyanobacterium *Synechocystis* sp. strain PCC 6803, but can be compensated by supplementation with magnesium ions. *FEMS Microbiol. Lett.* 277, 232–237. doi: 10.1111/j.1574-6968.2007.00960.x

Eisenhut, M., Huege, J., Schwarz, D., Bauwe, H., Kopka, J., and Hagemann, M. (2008). Metabolome phenotyping of inorganic carbon limitation in cells of the wild type and photorespiratory mutants of the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Plant Physiol.* 148, 2109–2120. doi: 10.1104/pp.108.129403

Evans, L. T. (1997). Adapting and improving crops: the endless task. *Phil. Trans. R. Soc. Lond. B* 352, 901–906. doi: 10.1098/rstb.1997.0069

Fan, J., Ye, J., Kamphorst, J. J., Shlomi, T., Thompson, C. B., and Rabinowitz, J. D. (2014). Quantitative flux analysis reveals folate-dependent NADPH production. *Nature* 510, 298–302. doi: 10.1038/nature13236

Forchhammer, K., and Lüddecke, J. (2016). Sensory properties of the PII signalling protein family. *FEBS J.* 283, 425–437. doi: 10.1111/febs.13584

Forchhammer, K., and Selim, K. A. (2020). Carbon/nitrogen homeostasis control in cyanobacteria. *FEMS Microbiol. Rev.* 44, 33–53. doi: 10.1093/femsre/fuz025

Gärtner, K., Klähn, S., Watanabe, S., Mikkat, S., Scholz, I., Hess, W. R., et al. (2019). Cytosine N4-Methylation via M.Ssp6803II is involved in the regulation of transcription, fine- tuning of DNA replication and DNA repair in the cyanobacterium *Synechocystis* sp. PCC 6803. *Front. Microbiol.* 10:1233. doi: 10.3389/fmicb.2019.01233

Georg, J., Voss, B., Scholz, I., Mitschke, J., Wilde, A., and Hess, W. R. (2009). Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. *Mol. Syst. Biol.* 5:305. doi: 10.1038/msb.2009.63

Giner-Lamia, J., Robles-Rengel, R., Hernández-Prieto, M. A., Muro-Pastor, M. I., Florencio, F. J., and Futschik, M. E. (2017). Identification of the direct regulon of NtcA during early acclimation to nitrogen starvation in the cyanobacterium *Synechocystis* sp, PCC 6803. *Nucleic Acids Res* 45, 11800–11820. doi: 10.1093/nar/gkx860

Giordano, M., Beardall, J., and Raven, J. A. (2005). CO2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annu. Rev. Plant Biol.* 56, 99–131. doi: 10.1146/annurev.arplant.56.032604.144052

Goyer, A., Collakova, E., La Díaz de Garza, R., Quinlivan, E. P., Williamson, J., Gregory, J. F., et al. (2005). 5-Formyltetrahydrofolate is an inhibitory but well tolerated metabolite in *Arabidopsis* leaves. *J. Biol. Chem.* 280, 26137–26142. doi: 10.1074/jbc.m503106200

Hagemann, M., and Bauwe, H. (2016). Photorespiration and the potential to improve photosynthesis. *Curr. Opin. Chem. Biol.* 35, 109–116. doi: 10.1016/j.cbpa.2016.09.014

Hagemann, M., and Hess, W. R. (2018). Systems and synthetic biology for the biotechnological application of cyanobacteria. *Curr. Opin. Biotechnol.* 49, 94–99. doi: 10.1016/j.copbio.2017.07.008

Hahn, A., and Schleiff, E. (2014). "The Cell Envelope," in *The Cell Biology of Cyanobacteria*, eds F. Enrique, and H. Antonia, (Norfolk: Caister Academic Press), 29–87.

Hanson, A. D., and Roje, S. (2001). One-carbon metabolism in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 52, 119–137.

Hohmann-Marriott, M. F., and Blankenship, R. E. (2011). Evolution of photosynthesis. *Annu. Rev. Plant Biol.* 62, 515–548.

Jones, P. R. (2014). Genetic instability in cyanobacteria – an elephant in the room? *Front. Bioeng. Biotechnol.* 2:12. doi: 10.3389/fbioe.2014.00012

Kebeish, R., Niessen, M., Thiruveedhi, K., Bari, R., Hirsch, H.-J., Rosenkranz, R., et al. (2007). Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in *Arabidopsis thaliana*. *Nat. Biotechnol.* 25, 593–599. doi: 10.1038/nbt1299

Kim, S., Lindner, S. N., Aslan, S., Yishai, O., Wenk, S., Schann, K., et al. (2020). Growth of E. coli on formate and methanol via the reductive glycine pathway. *Nat. Chem. Biol.* 16, 538–545. doi: 10.1038/s41589-020-0473-5

Kim, S.-J., Yoon, J., Im, D.-K., Kim, Y. H., and Oh, M.-K. (2019). Adaptively evolved *Escherichia coli* for improved ability of formate utilization as a carbon source in sugar-free conditions. *Biotechnol. Biofuels* 12, 1–12.

Klähn, S., Schaal, C., Georg, J., Baumgartner, D., Knippen, G., Hagemann, M., et al. (2015). The sRNA NsiR4 is involved in nitrogen assimilation control in cyanobacteria by targeting glutamine synthetase inactivating factor IF7. *Proc. Natl. Acad. Sci. U.S.A.* 112, E6243–E6252.

Klemke, F., Baier, A., Knoop, H., Kern, R., Jablonsky, J., Beyer, G., et al. (2015). Identification of the light-independent phosphoserine pathway as an additional source of serine in the cyanobacterium *Synechocystis* sp. PCC 6803. *Microbiology* 161, 1050–1060. doi: 10.1099/mic.0.000055

Kufryk, G., Hernandez-Prieto, M. A., Kieselbach, T., Miranda, H., Vermaas, W., and Funk, C. (2008). Association of small CAB-like proteins (SCPs) of *Synechocystis* sp. PCC 6803 with Photosystem II. *Photosynthesis Res.* 95, 135–145. doi: 10.1007/s11120-007-9244-3

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0

Lagarde, D., Beuf, L., and Vermaas, W. (2000). Increased production of zeaxanthin and other pigments by application of genetic engineering techniques to *Synechocystis* sp. strain PCC 6803. *Appl. Environ. Microbiol.* 66, 64–72. doi: 10.1128/aem.66.1.64-72.2000

Leegood, R. C. (2002). C(4) photosynthesis: principles of CO(2) concentration and prospects for its introduction into C(3) plants. *J. Exp. Bot.* 53, 581–590. doi: 10.1093/jexbot/53.369.581

Lin, J. T., and Stewart, V. (1998). "Nitrate assimilation by bacteria," in *Advances in microbial physiology*, ed. R. K. Poole, (San Diego, CA: Academic Press), 1–30. doi: 10.1016/s0065-2911(08)60014-4

Liu, X., Miao, R., Lindberg, P., and Lindblad, P. (2019). Modular engineering for efficient photosynthetic biosynthesis of 1-butanol from CO 2 in cyanobacteria. *Energy Environ. Sci.* 12, 2765–2777. doi: 10.1039/c9ee01214a

Long, B. M., Hee, W. Y., Sharwood, R. E., Rae, B. D., Kaines, S., Lim, Y.-L., et al. (2018). Carboxysome encapsulation of the CO2-fixing enzyme Rubisco in tobacco chloroplasts. *Nat. Commun.* 9:3570.

Long, S. P., Zhu, X.-G., Naidu, S. L., and Ort, D. R. (2006). Can improvement in photosynthesis increase crop yields? *Plant Cell Environ.* 29, 315–330. doi: 10.1111/j.1365-3040.2005.01493.x

Marx, C. J., Laukel, M., Vorholt, J. A., and Lidstrom, M. E. (2003). Purification of the formate-tetrahydrofolate ligase from *Methylobacterium extorquens* AM1 and demonstration of its requirement for methylotrophic growth. *J. Bacteriol.* 185, 7169–7175. doi: 10.1128/jb.185.24.7169-7175.2003

Mejillano, M. R., Jahansouz, H., Matsunaga, T. O., Kenyon, G. L., and Himes, R. H. (1989). Formation and utilization of formyl phosphate by N10-formyltetrahydrofolate synthetase: evidence for formyl phosphate as an intermediate in the reaction. *Biochemistry* 28, 5136–5145. doi: 10.1021/bi00438a034

Reinholdt, O., Schwab, S., Zhang, Y., Reichheld, J.-P., Fernie, A. R., Hagemann, M., et al. (2019). Redox-regulation of photorespiration through mitochondrial thioredoxin o1. *Plant Physiol.* 181, 442–457. doi: 10.1104/pp.19.00559

Sah, S., Aluri, S., Rex, K., and Varshney, U. (2015). One-carbon metabolic pathway rewiring in *Escherichia coli* reveals an evolutionary advantage of 10-formyltetrahydrofolate synthetase (Fhs) in survival under hypoxia. *J. Bacteriol.* 197, 717–726. doi: 10.1128/jb.02365-14

Schulz, J., Dettlaff, S., Fritzsche, U., Harms, U., Schiebel, H., Derer, W., et al. (1994). The amido black assay: a simple and quantitative multipurpose test of adhesion, proliferation, and cytotoxicity in microplate cultures of keratinocytes (HaCaT) and other cell types growing adherently or in suspension. *J. Immunol. Methods* 167, 1–13. doi: 10.1016/0022-1759(94)90069-8

Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N. S., and Erb, T. J. (2016). A synthetic pathway for the fixation of carbon dioxide *in vitro*. *Science* 354, 900–904. doi: 10.1126/science.aah5237

Semin, B. K., Loviagina, E. R., Aleksandrov, A. Y., Kaurov, Y. N., and Novakova, A. A. (1990). Effect of formate on Mössbauer parameters of the non-heme iron of PS II particles of cyanobacteria. *FEBS Lett.* 270, 184–186. doi: 10.1016/0014-5793(90)81263-n

Shevela, D., Klimov, V., and Messinger, J. (2007). Interactions of photosystem II with bicarbonate, formate and acetate. *Photosynthesis Res.* 94, 247–264. doi: 10.1007/s11120-007-9200-2

Shih, P. M., Zarzycki, J., Niyogi, K. K., and Kerfeld, C. A. (2014). Introduction of a synthetic CO₂-fixing photorespiratory bypass into a cyanobacterium. *J. Biol. Chem.* 289, 9493–9500. doi: 10.1074/jbc.C113.543132

Shinde, S., Singapuri, S. P., Zhang, X., Kalra, I., Liu, X., Morgan-Kiss, R. M., et al. (2019). Glycogen metabolism jump-starts photosynthesis through the oxidative pentose phosphate pathway (OPPP) in cyanobacteria. *bioRxiv*. [Preprint]. doi: 10.1101/657304

South, P. F., Cavanagh, A. P., Liu, H. W., and Ort, D. R. (2019). Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science* 363:eaat9077. doi: 10.1126/science.aat9077

Stemler, A., and Radmer, R. (1975). Source of photosynthetic oxygen in bicarbonate-stimulated hill reaction. *Science* 190, 457–458. doi: 10.1126/science.190.4213.457

Suppmann, B., and Sawers, G. (1994). Isolation and characterization of hypophosphite–resistant mutants of *Escherichia coli*: identification of the FocA protein, encoded by the pfl operon, as a putative formate transporter. *Mol. Microbiol.* 11, 965–982. doi: 10.1111/j.1365-2958.1994.tb00375.x

Tcherkez, G. G. B., Farquhar, G. D., and Andrews, T. J. (2006). Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7246–7251. doi: 10.1073/pnas.0600605103

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. U.S.A.* 76, 4350–4354. doi: 10.1073/pnas.76.9.4350

Trudeau, D. L., Edlich-Muth, C., Zarzycki, J., Scheffen, M., Goldsmith, M., Khersonsky, O., et al. (2018). Design and in vitro realization of carbon-conserving photorespiration. *Proc. Natl. Acad. Sci. U.S.A.* 115, E11455–E11464.

Voß, B., and Hess, W. R. (2014). "The Identification of Bacterial Non-coding RNAs through Complementary Approaches," in *Handbook of RNA biochemistry*, eds R. K. Hartmann, A. Bindereif, A. Schön, and E. Westhof, (Weinheim: Wiley-VCH), 787–800. doi: 10.1002/9783527647064.ch34

Whitney, S. M., Houtz, R. L., and Alonso, H. (2011). Advancing our understanding and capacity to engineer nature's CO2-sequestering enzyme, Rubisco. *Plant Physiol.* 155, 27–35. doi: 10.1104/pp.110.164814

Wiechert, M., and Beitz, E. (2017). Mechanism of formate-nitrite transporters by dielectric shift of substrate acidity. *EMBO J.* 36, 949–958. doi: 10.15252/embj.201695776

Yishai, O., Bouzon, M., Döring, V., and Bar-Even, A. (2018). In vivo assimilation of one-Carbon via a synthetic reductive glycine pathway in *Escherichia coli*. *ACS Synth. Biol.* 7, 2023–2028. doi: 10.1021/acssynbio.8b00131

Yishai, O., Goldbach, L., Tenenboim, H., Lindner, S. N., and Bar-Even, A. (2017). Engineered assimilation of exogenous and endogenous formate in *Escherichia coli*. *ACS Synth. Biol.* 6, 1722–1731. doi: 10.1021/acssynbio.7b00086

Zhang, C.-C., Zhou, C.-Z., Burnap, R. L., and Peng, L. (2018). Carbon/nitrogen metabolic balance: lessons from *Cyanobacteria*. *Trends Plant Sci.* 23, 1116–1130. doi: 10.1016/j.tplants.2018.09.008

Light-Driven Biosynthesis of *myo*-Inositol Directly From CO$_2$ in *Synechocystis* sp. PCC 6803

Xiaoshuai Wang[1,2,3], Lei Chen[1,2,3], Jing Liu[4], Tao Sun[1,2,5]* and Weiwen Zhang[1,2,3,5]*

[1] *Laboratory of Synthetic Microbiology, School of Chemical Engineering and Technology, Tianjin University, Tianjin, China,* [2] *Frontier Science Center for Synthetic Biology and Key Laboratory of Systems Bioengineering, Ministry of Education of China, Tianjin, China,* [3] *Collaborative Innovation Center of Chemical Science and Engineering, Tianjin, China,* [4] *School of Life Sciences, Tianjin University, Tianjin, China,* [5] *Center for Biosafety Research and Strategy, Tianjin University, Tianjin, China*

myo-inositol (MI) is an essential growth factor, nutritional source, and important precursor for many derivatives like D-*chiro*-inositol. In this study, attempts were made to achieve the "green biosynthesis" of MI in a model photosynthetic cyanobacterium *Synechocystis* sp. PCC 6803. First, several genes encoding *myo*-inositol-1-phosphate synthases and *myo*-inositol-1-monophosphatase, catalyzing the first or the second step of MI synthesis, were introduced, respectively, into *Synechocystis*. The results showed that the engineered strain carrying *myo*-inositol-1-phosphate synthase gene from *Saccharomyces cerevisiae* was able to produce MI at 0.97 mg L^{-1}. Second, the combined overexpression of genes related to the two catalyzing processes increased the production up to 1.42 mg L^{-1}. Third, to re-direct more cellular carbon flux into MI synthesis, an inducible small RNA regulatory tool, based on MicC-Hfq, was utilized to control the competing pathways of MI biosynthesis, resulting in MI production of ∼7.93 mg L^{-1}. Finally, by optimizing the cultivation condition via supplying bicarbonate to enhance carbon fixation, a final MI production up to 12.72 mg L^{-1} was achieved, representing a ∼12-fold increase compared with the initial MI-producing strain. This study provides a light-driven green synthetic strategy for MI directly from CO$_2$ in cyanobacterial chassis and represents a renewable alternative that may deserve further optimization in the future.

Keywords: *myo*-inositol, cyanobacteria, photosynthetic cell factory, small RNA tools, synthetic biology

INTRODUCTION

Inositol, known as cyclohexanehexol, is a vital growth factor previously identified in bacteria, fungi, higher plants, and animals. It has nine isomers (i.e., *myo*-, *cis*-, *epi*-, *allo*-, *muco*-, *neo*-, L-*chiro*-, D-*chiro*-, and *scyllo*-), and five of them have been found in nature, namely, D-*chiro*-inositol, L-*chiro*-inositol, *myo*-inositol, *neo*-inositol, and *scyllo*-inositol (Thomas et al., 2016). Among them, *myo*-inositol (*cis*-1, 2, 3, 5-*trans*-4, 6-cyclohexanehexol, hereafter MI) and its derivatives are the most abundant in nature and have attracted significant attention in recent years due to their wide applications in functional food and pharmaceutical industry (You et al., 2017). For example, MI was reported to be effective in restoring spontaneous ovarian activity, consequently improving

the fertility of most patients with polycystic ovary syndrome (Regidor et al., 2018; Januszewski et al., 2019). In addition, MI serves as a precursor for many valuable chemicals, further generating numerous important chemicals participating in maintaining homeostasis, such as inositol-1, 4, 5-trisphosphate (IP3) that functions as a Ca^{2+}-mobilizing second messenger in regulating many cellular processes (Berridge, 2009). Moreover, MI can also be converted to *scyllo*- and D-*chiro*–inositol, both of which have potential roles in the medicine industry in curing Alzheimer's disease and hyperglycemia (Ma et al., 2012; Cheng et al., 2019). It is thus valuable to develop cost-efficient strategies for MI production.

Several strategies have been so far reported for MI synthesis. Among all chemical approaches, it is difficult to operate and is less environmentally friendly due to its harsh chemical conditions, such as low pH, high temperature, and high pressure. Recently, the microbial production of MI through synthetic biology has attracted increasing attention (Fujisawa et al., 2017; You et al., 2017; Lu et al., 2018). For example, Lu et al. (2018) recently reported a novel pathway to produce MI from glucose through a trienzymatic cascade system *in vitro*, achieving a productivity of 45.2 mM within 24 h. By dynamically modulating the key enzyme phosphofructokinase-I (Pfk-I) in *Escherichia coli*, recently a level of MI production at 1.31 g L^{-1} was achieved (Brockman and Prather, 2015). In addition, Tanaka et al. (2013) constructed a pathway starting from MI to *scyllo*-inositol in *Bacillus subtilis*, resulting in *scyllo*-inositol productivity of 10 g L^{-1} after 48 h. More recently, by introducing *Mycobacterium tuberculosis ino1* gene encoding *myo*-inositol-1-phosphate synthase and overexpressing intrinsic inositol monophosphatase, *YktC*, as well as an artificial pathway converting *myo*-inositol to *scyllo*-inositol in *Bacillus subtilis*, Michon et al. (2020) achieved a production of 2 g L^{-1} *scyllo*-inositol using 20 g L^{-1} glucose. Nevertheless, even with all the exciting progresses, a new, renewable, and cost-efficient alternative for MI production remains to be developed.

Due to the ability of utilizing sunlight and CO_2 as sole energy and carbon sources, respectively, cyanobacteria are considered as promising green chassis for producing chemicals. Up to now, several dozens of biofuels and chemicals have been successfully synthesized directly from CO_2 in cyanobacteria, such as ethylene, ethanol, fatty acids, D-lactic acid, 3-hydroxypropionic acid, etc. (Gao et al., 2016). As a model cyanobacterium, *Synechocystis* sp. PCC 6803 (hereafter *Synechocystis*) has the advantages of a simple genetic background and feasible genetic tools for metabolic engineering and synthetic biology (Sun et al., 2018a). Given that cyanobacteria could directly use CO_2 to produce chemicals driven by sunlight, attempts were made in this study to construct green synthesis strategy for MI in *Synechocystis* chassis.

In this study (**Figure 1**), to achieve the green synthesis of MI, we first constructed an exogenous metabolic route to convert glucose-6-phosphate to MI in *Synechocystis* by, respectively, introducing *myo*-inositol-1-phosphate synthase from *Saccharomyces cerevisiae* or *Corynebacterium glutamicum* as well as the native genes (*sll1329* and *sll1383*) encoding myo-inositol-1-monophosphatase. The results showed that the engineered *Synechocystis* carrying *INO1* (*myo*-inositol-1-phosphate synthases from *S. cerevisiae*) performed the best,

with MI production of 0.97 mg L^{-1}. Second, the combined overexpression of *INO1*, *sll1329*, and *sll1383* further improved MI production. Third, to drive more carbon flux into MI synthesis, endogenous gene *zwf* (encoding glucose-6-phosphate dehydrogenase), *pgi* (encoding glucose-6-phosphate isomerase), and *pfkA* (encoding phosphofructokinase) were knocked down, respectively, and combined, using a theophylline-inducible small RNA (sRNA) regulatory tool based on MicC-Hfq, leading to MI production of up to 7.93 mg L^{-1}. Finally, by supplying bicarbonate to enhance carbon fixation, a final MI production up to 12.72 mg L^{-1} was achieved, representing a ~12-fold increase compared with the initial MI-producing strain.

MATERIALS AND METHODS

Chemicals and Reagents

MI standard was purchased from Rhawn Chemical Technology Co., Ltd. (Shanghai, China). The other chemicals used in this study were purchased from Sigma-Aldrich (MO, United States). T4 Polynucleotide Kinase, T4 DNA ligase, and all restriction enzymes were purchased from Thermo Fisher Scientific (MA, United States). Phanta Super-Fidelity DNA Polymerase, ChamQ SYBR qPCR Master Mix, and HiScript Q RT SuperMix for qPCR were obtained from Vazyme Biotech Co., Ltd. (Nanjing, China). The Plasmid Mini Kit I and Cycle Pure Kit used were purchased from Omega Bio-Tek (GA, United States). Synthesis of DNA oligonucleotide primers and Sanger sequencing were provided by Genewiz (Suzhou, China).
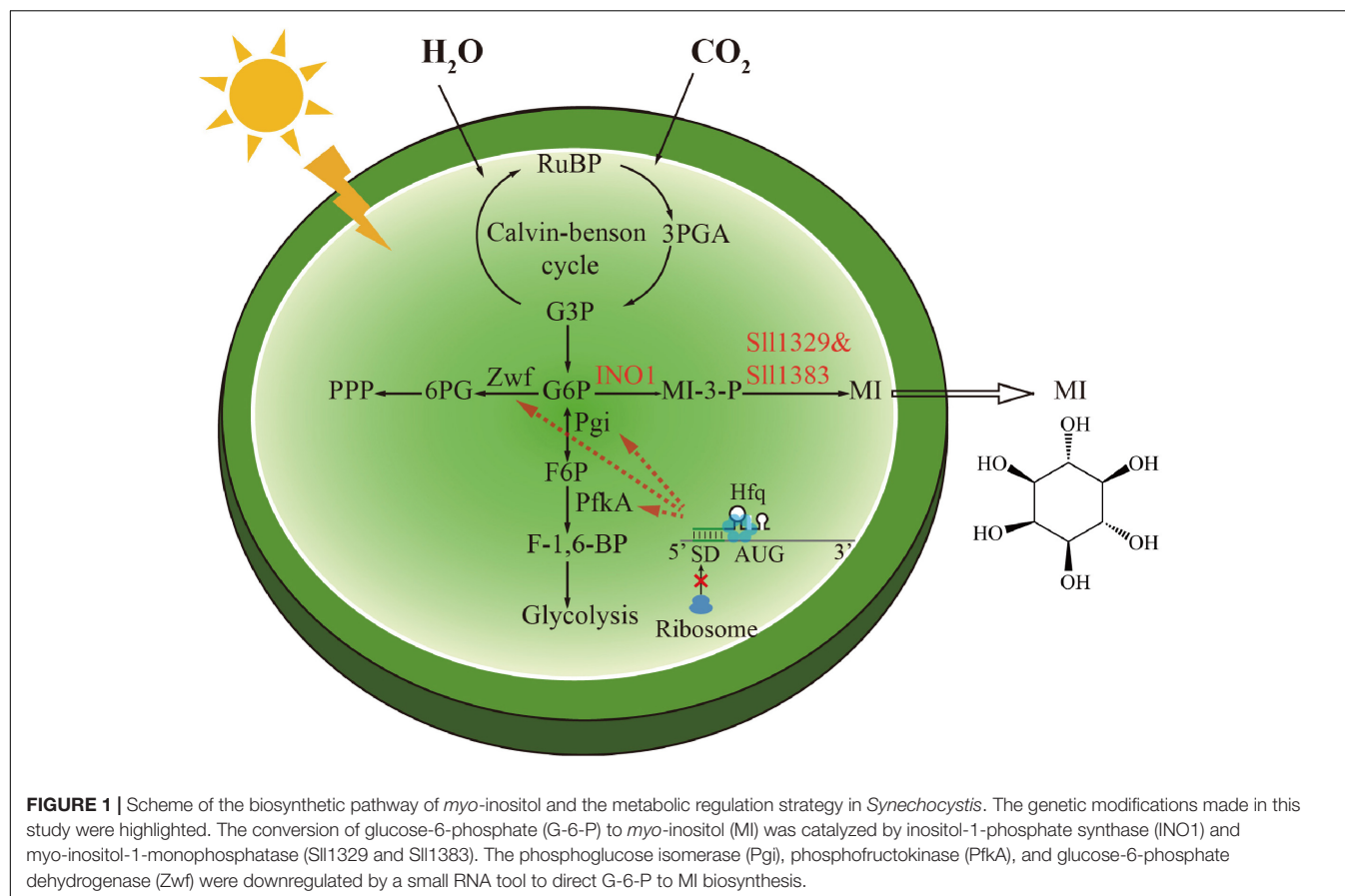
Culture Conditions

The wild-type (WT) and all engineered strains of *Synechocystis* were grown at 30°C in BG-11 liquid medium or on solid BG-11 agar plate at a light intensity of ~50 μmol photons $m^{-2} s^{-1}$ in an incubator (SPX-250B-G, Boxun, Shanghai, China) or illuminating shaking incubator (HNY-211B, Honour, Tianjin, China) at 130 rpm, respectively. Appropriate antibiotic(s) was added into the BG-11 growth medium as required (i.e., 20 μg ml^{-1} chloramphenicol or 20 μg ml^{-1} spectinomycin). The growth of the cells was monitored by measuring their optical density at 730 nm (OD_{730}) with a UV-1750 spectrophotometer (Shimadzu, Kyoto, Japan). *E. coli* Trans 5α was used as a host for constructing all recombinant plasmids, which were grown on Luria–Bertani solid agar plates or in a medium with appropriate antibiotic(s) to maintain the plasmids (i.e., 50 μg ml^{-1} chloramphenicol or 50 μg ml^{-1} spectinomycin) at 37°C in an incubator or a shaking incubator (HNY-100B, Honour, Tianjin, China) at 200 rpm, respectively.

Strain and Plasmid Construction

E. coli Trans 5α was used as a host for plasmid construction and amplification. In this study, two suicide plasmids, p3031 and p0168, that could replicate in *E. coli* and integrate into the genome of *Synechocystis* (between *slr2030* and *slr2031* for p3031 or within *slr0168* for p0168, respectively) via homologous recombination were utilized to express the related genes. The *INO1* and *cgl2996* genes were amplified using *S. cerevisiae* and

FIGURE 1 | Scheme of the biosynthetic pathway of *myo*-inositol and the metabolic regulation strategy in *Synechocystis*. The genetic modifications made in this study were highlighted. The conversion of glucose-6-phosphate (G-6-P) to *myo*-inositol (MI) was catalyzed by inositol-1-phosphate synthase (INO1) and myo-inositol-1-monophosphatase (Sll1329 and Sll1383). The phosphoglucose isomerase (Pgi), phosphofructokinase (PfkA), and glucose-6-phosphate dehydrogenase (Zwf) were downregulated by a small RNA tool to direct G-6-P to MI biosynthesis.

C. glutamicum genomic DNA as templates, respectively. Then *INO1* and *cgl2996* were, respectively, ligated into pCP3031 (**Supplementary Figure S1**), resulting in plasmids p3031I and p3031C, respectively. After being confirmed by DNA sequencing, these two genes were, respectively, introduced into WT, generating the strains WT-INO1 and WT-cgl. The *sll1329* and *sll1383* genes were amplified using *Synechocystis* genomic DNA as template and fused into one fragment linked by a ribosome binding site (RBS) via overlapping PCR. The fused fragment was then ligated into pCP3031, resulting in plasmid p3031SS. The p3031SS was introduced into WT, generating the strain WT-SS. In addition, the fragments of *sll1329* and *sll1383* were further fused with a strong promoter, Pcpc560, and inserted after the cassette of *INO1* on p3031I for transformation, generating the plasmid p3031S and the *Synechocystis* strain WT-INO1SS, respectively. The construction of sRNA-expressing plasmids was conducted as reported previously (Sun et al., 2018b). First, a light-induced promoter PpsbA2M (Ppsba2 without RBS) was utilized to express the sRNA scaffold *micC*, while a theophylline-induced riboswitch was used to control the expression of *hfq*, respectively (**Supplementary Figure S2**). Second, the synthetic 24-bp sRNA sequence (*aszwf*) targeting the translational starting site of the *zwf* was added into the location between PpsbA2M and MicC, leading to p0168Z, with a PpsbA2M-*aszwf*-*micC*-TrbcL-Ptrc-riboswitch-*hfq*-TrbcL cassette. Similarly, the plasmids with sRNA-expressing

cassette targeting the *pgi* and the *pfkA* (*aspgi* and *aspfkA*) were constructed independently, leading to p0168P and p0168PF, respectively. The p0168Z, p0168P, and p0168PF plasmids were, respectively, transferred into *Synechocystis* WT-INO1SS through natural transformation, generating the strains WT-INO1SS-ASZWF, WT-INO1SS-ASPGI, and WT-INO1SS-ASPFKA, respectively. Finally, two expressing cassettes, PpsbA2M-*aspgi*-*micC*-TrbcL and PpsbA2M-*aspfkA*-*micC*-TrbcL, were ligated into p0168Z, generating p0168ZP or p0168ZF, respectively targeting two genes (i.e., targeting *zwf* and *pgi* or *zwf* and *pfkA*). They were then introduced into WT-INO1SS to generate the strains WT-INO1SSZP and WT-INO1SSZF, respectively. All the strains and the plasmids used and constructed in this study are listed in **Table 1**.

Transformation of *Synechocystis*

Natural transformation of *Synechocystis* was performed according to the method published previously. Briefly, when *Synechocystis* grew to exponential phase ($OD_{730} \approx 0.5$), cells were collected by centrifugation ($3,000 \times g$, 13 min, 4°C) and washed with fresh BG-11 medium. The cells were then resuspended in fresh BG-11, and ~10 μg of corresponding plasmid DNA was added to the suspension. The cell and plasmid mixture was incubated at 30°C for at least 5 h under luminous intensity of ~50 μmol photons m^{-2} s^{-1}, followed by spreading onto BG-11 agar plates with appropriate antibiotic(s) (e.g., 20 μg

TABLE 1 | Strains used in this study.

Strains	Genotype	References
Escherichia coli Trans5α	F−, φ80d *lacZ*ΔM15, Δ(*lacZYA-arg*F) U169, *end*A1, *rec*A1, *hsd*R17 (rk−, mk+), *sup*E44λ, *thi*-1, *gyr*A96, *rel*A1, *pho*A	Stratagene
Cyanobacteria strains		
Synechocystis sp. PCC 6803	WT	ATCC 27184
WT-INO1	*slr2030_slr2031*:Pcpc560-*INO1*-TrbcL; Spe[R] in WT	This study
WT-cgl	*slr2030_slr2031*:Pcpc560-*cgl2996*-TrbcL; Spe[R] in WT	This study
WT-INO1SS	*slr2030_slr2031*:Pcpc560-*INO1*-TrbcL-Pcpc560-*sll1329*-rbs-*sll1383*-TrbcL; Spe[R] in WT	This study
WT-SS	*slr2030_slr2031*:Pcpc560-*sll1329*-rbs-*sll1383*-TrbcL; Spe[R] in WT	This study
WT-INO1SS-ASZWF	*slr0168*:PpsbA2M-*aszwf-micC*-TrbcL-Ptrc-riboswitch-*hfq*-TrbcL; Cm[R] in WT-INO1SS	This study
WT-INO1SS-ASPGI	*slr0168*:PpsbA2M-*aspgi-micC*-TrbcL-Ptrc-riboswitch-*hfq*-TrbcL; Cm[R] in WT-INO1SS	This study
WT-INO1SS-ASPFKA	*slr0168*:PpsbA2M-*aspfkA-micC*-TrbcL-Ptrc-riboswitch-*hfq*-TrbcL; Cm[R] in WT-INO1SS	This study
WT-INO1SSZP	*slr0168*:PpsbA2M-*aszwf-micC*-TrbcL- PpsbA2M-*aspgi-micC*-TrbcL-Ptrc-riboswitch-*hfq*-TrbcL; Cm[R] in WT-INO1SS	This study
WT-INO1SSZF	*slr0168*:PpsbA2M-*aszwf-micC*-TrbcL-PpsbA2M-*aspfkA-micC*-TrbcL Ptrc-riboswitch-*hfq*-TrbcL; Cm[R] in WT-INO1SS	This study

ml^{-1} chloramphenicol and/or 20 μg ml^{-1} spectinomycin). After incubation of ∼2 weeks, colonies were observed. After validation by colony PCR and sequencing, positive colonies would be transferred to liquid BG11 medium for growth and further examination.

MI Quantification

For *Synechocystis* samples, 1 ml of fresh cultures of *Synechocystis* was collected on the third day by centrifugation at 12,000 × *g* for 5 min at room temperature (Eppendorf 5430R, Hamburg, Germany). The MI content in the sample pellets and the supernatant were, respectively, measured after performing pre-column derivatization according to the two-stage technique described previously (Roessner et al., 2001). Meanwhile, the stock solution of MI was prepared in ddH₂O at a final concentration of 1 g L^{-1}. The MI standard curve was plotted using different concentrations of MI solution (**Supplementary Figure S4**). MI levels were quantified on a gas chromatography–mass spectrometry system—GC 7890 coupled to MSD 5975 (Agilent Technologies, Inc., Santa Clara, CA) equipped with a HP-5MS capillary column (30 m × 250 mm id).

Theophylline Treatment

The stock solution of theophylline was prepared by dissolving theophylline (Aladdin; Shanghai; China) in BG-11 medium at a final concentration of 10 mM. For theophylline-induced assays, all *Synechocystis* samples were collected by centrifugation at

3,000 × *g* and 4°C for 12 min and then re-suspended using fresh BG-11 medium with stock solution of theophylline at a final concentration of 2 mM (Sun et al., 2018b).

## Quantitative Real-Time PCR Analysis

Synechocystis samples were collected at 24 h after 2 mM theophylline induction. ∼5 ml of samples (OD₇₃₀ = 1.0) was collected by centrifugation at 3,000 × *g* and 4°C for 12 min. The supernatant was removed, and the cell pellet was used for RNA extraction. Total RNA extraction was achieved through a Direct-zol™ RNA MiniPrep Kit (Zymo, CA, United States), and cDNAs were synthesized using HiScript Q RT SuperMix for qPCR (Vazyme Biotech Co., Ltd., Nanjing, China). The 10-μl qRT-PCR reaction included 5 μl ChamQ SYBR qPCR Master Mix (Vazyme Biotech Co., Ltd., Nanjing, China), 3 μl ddH₂O, 1 μl diluted template cDNA, and 1 μl of each PCR primer (0.5 μl forward primer and 0.5 μl reverse primer). The reaction was conducted in the StepOne™ Real-Time PCR System (Applied Biosystems, CA, United States). 16S rRNA was selected as the reference gene, and the primers of the specific genes used are listed in **Supplementary Table S1**. Data analysis was carried out by using $2^{-\Delta\Delta CT}$ method as reported previously (Livak and Schmittgen, 2001).

RESULTS

Construction of MI-Producing *Synechocystis*

The bioconversion from glucose to MI involves three steps which are catalyzed by three sequentially acting enzymes (Fujisawa et al., 2017; You et al., 2017; Lu et al., 2018). First, with the aid of hexokinase, glucose was phosphorylated to glucose-6-phosphate (G6P). Second, G6P was converted into *myo*-inositol 3-phosphate (I3P), catalyzed by inositol-1-phosphate synthase (IPS), which was responsible for the committed step of inositol synthesis. Third, I3P was dephosphorylated to generate MI by inositol-1-monophosphatase. Previously, native genes potentially related to MI synthesis have been identified in *Synechocystis* (Chatterjee et al., 2004, 2006; Patra et al., 2007). In this study, to detect the production of MI in the *Synechocystis* WT strain, we measured both intracellular and extracellular MI, and the results showed that no detectable MI could be observed even after 7 days of cultivation, demonstrating that the MI produced natively was below the detection limit. Previously, studies showed that the *INO1* gene encoding inositol-1-phosphate synthase from *S. cerevisiae* could perform well in *E. coli* (Gupta et al., 2017). In addition, an IPS gene of the same function, *cgl2996*, was also identified in *C. glutamicum*. Therefore, *INO1* from *S. cerevisiae* and *cgl2996* from *C. glutamicum* were, respectively, introduced and evaluated in *Synechocystis*, generating the strains WT-INO1 and WT-cgl (**Table 1**).

The growth comparison between WT, WT-INO1, and WT-cgl suggested that overexpression of neither *INOI* nor *cgl2996* affected the growth of the engineered strains (**Figure 2A**). The potential production of MI was measured in both WT-INO1 and WT-cgl strains. As shown in **Figure 2B**, no MI was detected by GC-MS in the first 2 days as they were probably
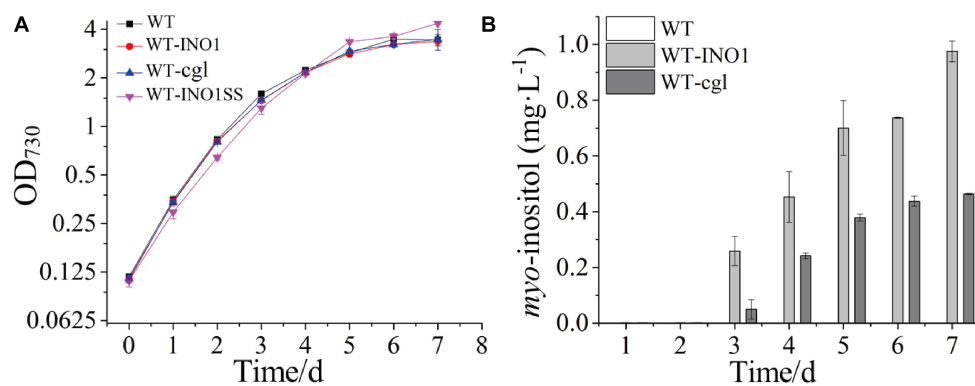
FIGURE 2 | Growth curves and *myo*-inositol quantitation in wild-type (WT) and engineered *Synechocystis* strains. The error bar represents the standard deviation of three biological replicates for each sample. **(A)** Growth curves of WT, WT-INO1, WT-cgl, and WT-INO1SS. **(B)** *myo*-inositol quantitation in WT and WT-cgl.

below the detection limit; however, MI could be observed in both intracellular and extracellular samples from the 3rd day (**Supplementary Figure S3**) and accumulated steadily until the 7th day (stationary phase) (**Figure 2B**). Finally, after cultivation for 7 days, the production of MI reached ∼975.5 µg L^{-1} in WT-INO1, while it only reached ∼463.8 µg L^{-1} in WT-cgl, respectively. These results demonstrated that the introduction of exogenous IPS could enable the MI biosynthesis in *Synechocystis*. In addition, the IPS from *S. cerevisiae* seemed to function better for MI biosynthesis than that from *C. glutamicum* in *Synechocystis*.

Enhancement of MI Production by Overexpressing Key Genes

myo-inositol-1-monophosphatase (IMP) is another crucial enzyme for MI synthesis that catalyzes the production of *myo*-inositol from *myo*-inositol 3-phosphate. This enzyme is putatively encoded by *sll1329* or *sll1383* in *Synechocystis* according to a previous study (Patra et al., 2007) and the KEGG pathway annotation[1]. In order to further improve the MI production, we simultaneously overexpressed the genes *sll1329* and *sll1383* (isozyme genes, both encoding IMP) using a strong promoter Pcpc560 in WT-INO1, resulting in the strain WT-INO1SS (**Table 1**). As a control, strain WT-SS with only overexpressed *sll1329* and *sll1383* was also constructed (**Table 1**). To evaluate their expression, the transcriptional levels of *sll1329* and *sll1383* in WT-INO1SS or WT-SS were quantified via qRT-PCR. As illustrated in **Figure 3A**, the transcriptional levels of *sll1329* and *sll1383* were, respectively, increased by more than 15- and 80-folds in both WT-INO1SS and WT-SS compared to that of WT, suggesting a successful overexpression. The growth of WT and WT-INO1SS was comparatively investigated, and the results showed that the overexpression of *sll1329* and *sll1383* caused no visible growth inhibition (**Figure 2A**). Moreover, the production of MI reached 1.42 mg L^{-1} in WT-INO1SS after cultivation for 7 days, achieving ∼45.6% increase compared to that in WT-INO1 (**Figure 3B**). In the control WT-SS strain where
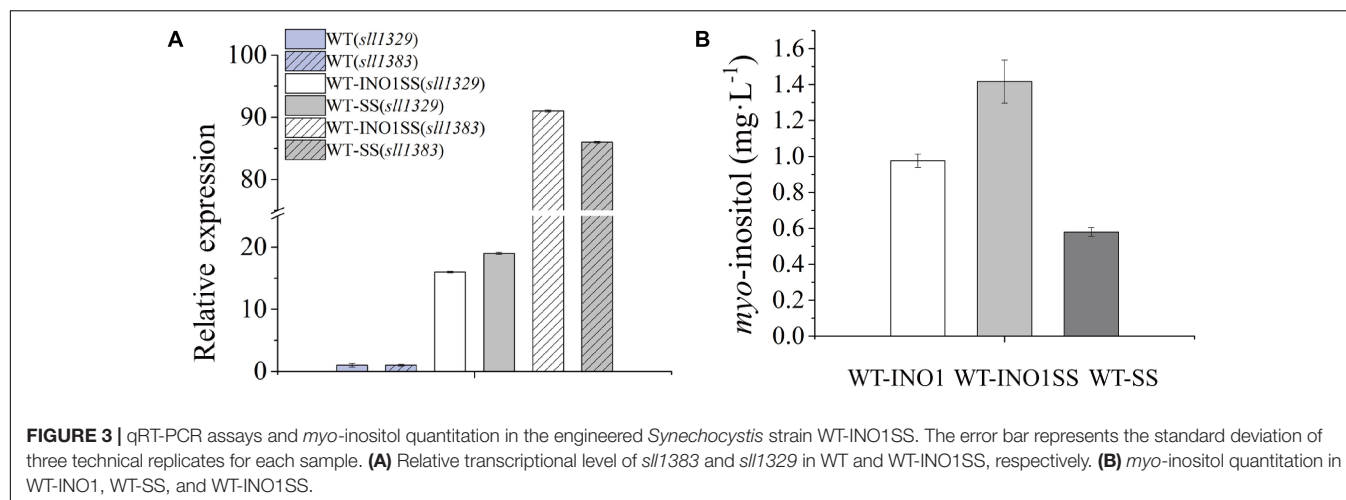
only *sll1329* and *sll1383* were expressed, MI analysis showed that it can produce MI production at 579.6 µg L^{-1}, confirming the existence of a native *INO1* gene encoding inositol-1-phosphate synthase in *Synechocystis* (Chatterjee et al., 2004, 2006; Patra et al., 2007), although its catalytic activity might be well lower than INO1 from *S. cerevisiae* and Cgl2996 from *C. glutamicum*. The results supported the report that the overexpression of *sll1329* and *sll1383* could improve the production of MI.

Re-direction of Carbon Flux Toward MI Production

Glycolysis and pentose phosphate pathway are the main carbon source competing pathways for MI synthesis (**Figure 1**; Hansen et al., 1999). However, total blocking of these essential pathways would cause severe growth inhibition or even a lethal phenotype. In order to drive more carbon flux from the competing pathways to MI biosynthesis, a sRNA tool MicC–Hfq, developed previously in *Synechocystis*, that allows "gene knock-down" was adopted (Sun et al., 2018b). In detail, the Hfq–MicC tool is composed of a chaperone protein Hfq and a well-studied sRNA scaffold named MicC from *E. coli*. With a designed target-binding region fused into the MicC scaffold, the fragment could regulate the expression of target genes effectively via altering their translation with the aid of the Hfq chaperone.

Glucose-6-phosphate (G6P) is not only a metabolic branch point but also a substrate for inositol synthesis in cells, as it could be routed into native cellular metabolism through both glycolysis and the oxidative pentose phosphate pathway, as well as into the heterologous biosynthetic pathway of MI production. Thus, *pgi* (encoding phosphoglucose isomerase), *pfkA* (encoding phosphofructokinase), and *zwf* (encoding glucose-6-phosphate dehydrogenase) were chosen as the target genes, and the related sRNA expressing systems were constructed. The synthetic sRNA sequences targeting the translational starting site of the *zwf*, *pgi*, or *pfkA* were fused, respectively, into the MicC scaffold and driven by PpsbA2M (PpsbA2 without RBS), while the expression of *hfq* was controlled by the Ptrc containing a theophylline-induced riboswitch. In addition, the related strains WT-INO1SS-ASPGI, WT-INO1SS-ASPFKA, and WT-INO1SS-ASZWF were,

---

[1]www.genome.jp/kegg-bin/show_pathway?syn00562

FIGURE 3 | qRT-PCR assays and *myo*-inositol quantitation in the engineered *Synechocystis* strain WT-INO1SS. The error bar represents the standard deviation of three technical replicates for each sample. **(A)** Relative transcriptional level of *sll1383* and *sll1329* in WT and WT-INO1SS, respectively. **(B)** *myo*-inositol quantitation in WT-INO1, WT-SS, and WT-INO1SS.

respectively, achieved via introducing the MicC–Hfq-expressing cassettes (**Table 1**). After induction with 2 mM theophylline, qRT-PCR analysis, growth curves, and *myo*-inositol production determination of the three strains were performed. First, qRT-PCR analysis was performed to validate the knockdown effect of the MicC–Hfq tool on all three target genes. As shown in **Figures 4A–C**, when compared to that in WT-INO1SS, 33, 45, and 39% down-regulation for *pgi*, *pfkA*, and *zwf* were achieved via the synthetic sRNA in WT-INO1SS-ASPGI, WT-INO1SS-ASPFKA, and WT-INO1SS-ASZWF, respectively. Second, the growth rate of the three strains was comparatively monitored, and only a slight inhibition was observed for WT-INO1SS-ASPGI, WT-INO1SS-ASPFKA, and WT-INO1SS-ASZWF compared to WT-INO1SS (**Figure 5A**). Third, the production of MI in three strains was determined, and the results showed that MI production was increased to 3.00, 3.06, and 3.06 mg L^{-1} in strains WT-INO1SS-ASPGI, WT-INO1SS-ASZWF, and WT-INO1SS-ASPFKA, respectively (**Figure 5B**), suggesting that the knock-down of *pgi*, *zwf*, and *pfkA* by the synthetic sRNAs was able to efficiently increase the metabolic flux from glucose-6-phosphate into MI.

Given that *pgi* and *pfkA* genes are both involved in the same pathway, a combined regulation for *zwf* and *pgi* or *zwf* and *pfkA* was carried out to evaluate whether it can further increase MI production. Accordingly, two strains, WT-INO1SSZP and WT-INO1SSZF, carrying the synthetic sRNAs targeting both *zwf* and *pgi* or both *zwf* and *pfkA*, respectively, were constructed. As illustrated in **Supplementary Figure S5**, the expression level of *zwf* was decreased by 42 and 41% in WT-INO1SSZP and WT-INO1SSZF; *pfkA* in WT-INO1SSZF and *pgi* in WT-INO1SSZP were decreased as well by 39 and 37% compared to that in WT after induction with 2 mM theophylline. The growth curves and the MI yields of the two strains were then measured after addition of 2 mM theophylline (**Figure 6A**); however, an obvious retardation in growth was observed for both strains, possibly due to the increased partitioning of carbon sources toward *myo*-inositol, which is consistent with the increased MI production in WT-INO1SSZP and WT-INO1SSZF even with the decreased growth (**Figure 6B**). The results showed that the MI production

reached 7.93 and 5.54 mg L^{-1} in WT-INO1SSZP and WT-INO1SSZF, respectively, indicating that the combined regulation of *zwf* and *pgi* was more effective for increasing MI synthesis. Between the two engineered strains, WT-INO1SSZP grew slower than WT-INO1SSZF (**Figure 6A**), which may be due to the fact that more glucose-6-phosphate was directed into the MI biosynthesis pathway from glycolysis and pentose phosphate pathway with the aid of sRNA tools in WT-INO1SSZP.

Cultivation Optimization to Enhance MI Production

Early studies have demonstrated that NaHCO$_3$ supplementation to cyanobacterial culture is an effective strategy for higher biomass and more production of target chemicals (Johnson et al., 2016; Wang et al., 2016). Thus, the effects of increased NaHCO$_3$ supply on MI production in the engineered strain were evaluated. Given that the strain WT-INO1SSZP showed the highest MI production capacity, it was chosen as the target for cultivation optimization. After supplementing 0.5 ml 1.0 M NaHCO$_3$ every 24 h into the BG-11, the engineered *Synechocystis* strain WT-INO1SSZP showed a faster growth rate compared with their corresponding strains without NaHCO$_3$ supplementation (**Figure 7A**). Meanwhile, the MI production in the strain WT-INO1SSZP was increased to 12.72 mg L^{-1} (**Figure 7B**), demonstrating that the enhanced carbon supply could significantly increase the MI production.

DISCUSSION

Early studies have shown the feasibility of directly converting light energy and CO$_2$ into green fuels and chemicals in *Synechocystis* (Gao et al., 2016; Wang et al., 2016). In this study, we engineered a photoautotrophic cyanobacterial system for the production of MI directly from CO$_2$. Previously, *INO1* and *myo*-inositol 1-phosphate synthase were overexpressed from *S. cerevisiae* in *E. coli* (Brockman and Prather, 2015), successfully achieving a heterologous production of MI. Consistently, we found that the overexpression of *INO1*
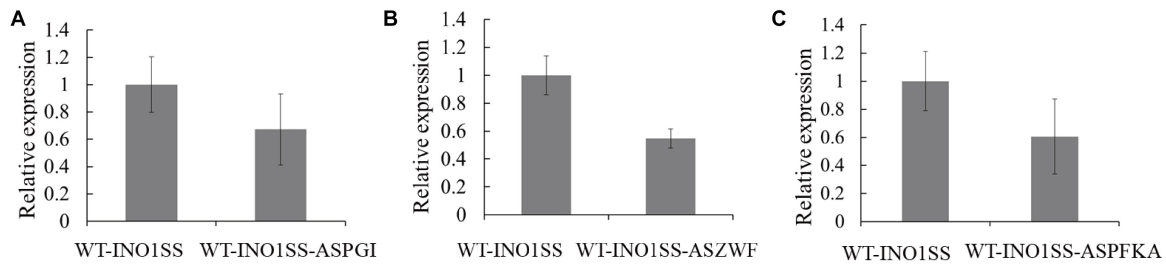
FIGURE 4 | qRT-PCR assays in the engineered *Synechocystis* strains WT-INO1SS-ASPGI, WT-INO1SS-ASZWF, and WT-INO1SS-ASPFKA after providing the supplement of 2 mM theophylline. The error bar represents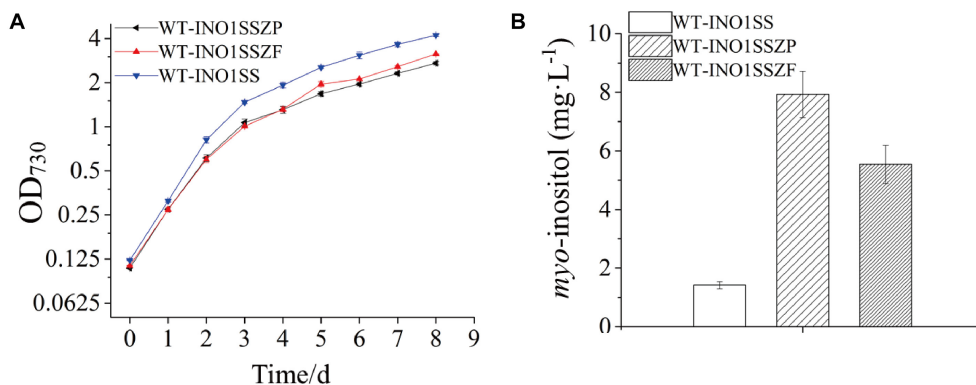 the standard deviation of three technical replicates for each sample. **(A)** Relative transcriptional level of *pgi* in WT-INO1SS and WT-INO1SS-ASPGI. **(B)** Relative transcriptional level of *zwf* in WT-INO1SS and WT-INO1SS-ASZWF. **(C)** Relative transcriptional level of *pfkA* in WT-INO1SS and WT-INO1SS-ASPFKA.
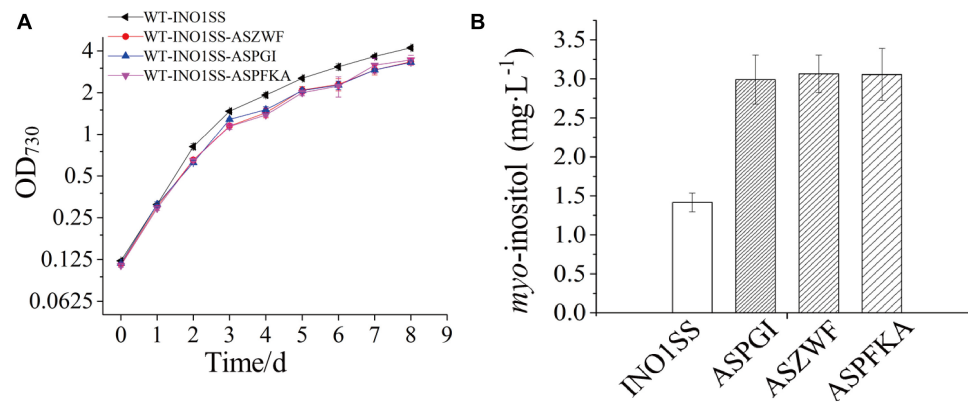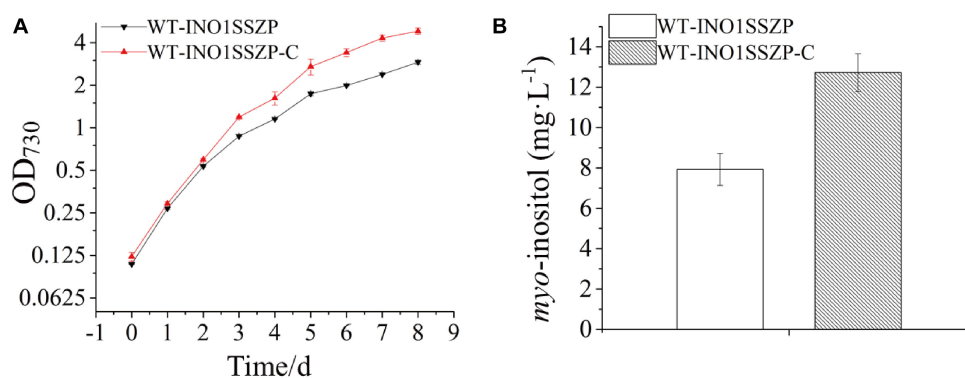

FIGURE 5 | Growth curves and *myo*-inositol quantitation in the engineered *Synechocystis* strains WT-INO1SS-ASPGI, WT-INO1SS-ASZWF, and WT-INO1SS-ASPFKA. The error bar represents the standard deviation of three biological replicates for each sample. INO1SS, strain WT-INO1SS; ASPGI, strain WT-INO1SS-ASPGI; ASZWF, strain WT-INO1SS-ASZWF; ASPFKA, strain WT-INO1SS-ASPFKA. **(A)** Growth curves of WT-INO1SS, WT-INO1SS-ASPGI, WT-INO1SS-ASZWF, and WT-INO1SS-ASPFKA. **(B)** *myo*-inositol quantitation in WT-INO1SS, WT-INO1SS-ASPGI, WT-INO1SS-ASZWF, and WT-INO1SS-ASPFKA after 8 days of cultivation.

FIGURE 6 | Growth curves and *myo*-inositol quantitation in the engineered *Synechocystis* strains WT-INO1SSZP and WT-INO1SSZF. The error bar represents the standard deviation of three biological replicates for each sample. **(A)** Growth curves of WT-INO1SS, WT-INO1SSZP, and WT-INO1SSZF. **(B)** *myo*-inositol quantitation in WT-INO1SS, WT-INO1SSZP, and WT-INO1SSZF after 8 days of cultivation.

led to detectable MI biosynthesis in *Synechocystis*. After the overexpression of *sll1329* and *sll1383* (encoding *myo*-inositol-1-monophosphatase), intracellular MI concentration was slightly increased by ∼45.6% in WT-INO1SS than that in WT-INO1. suggesting that IMP might not be the limiting step for MI synthesis in *Synechocystis*.

A significant overexpression for *sll1329* and *sll1383* on a transcriptional level was demonstrated via qRT-PCR (**Figure 3**), suggesting that IMP might not be the limiting step for MI synthesis in *Synechocystis*.

FIGURE 7 | *myo*-inositol quantitation and growth curves in the engineered *Synechocystis* strain WT-INO1SSZP, cultivated with or without NaHCO$_3$. The error bars represent the standard deviation of three biological replicates for each sample. WT-INO1SSZP represents cultivation without NaHCO$_3$, while WT-INO1SSZP-C represents cultivation with NaHCO$_3$. **(A)** Growth curves of WT-INO1SSZP with or without NaHCO$_3$. **(B)** *myo*-inositol quantitation in WT-INO1SSZP after cultivation for 8 days with or without NaHCO$_3$.

G6P was the direct precursor for MI synthesis; meanwhile, it is a fundamental metabolite to support microbial survival. Though the manipulation of carbon flux toward the G6P pool has previously been demonstrated to be an effective strategy to enhance the production of its derivatives in various microorganisms (Brockman and Prather, 2015; Gupta et al., 2017), deletion of the pentose phosphate pathway (PPP)-related gene *zwf* could totally block the essential pathways and cause severe growth inhibition. Thus, a suitable and efficient genetic tool for gene knockdown is valuable. Previously, the small RNA regulatory tool was demonstrated as feasible and efficient in regulating genes, especially essential genes, such as redirecting the carbon flux to the key precursor malonyl-CoA in *Synechocystis* (Sun et al., 2018b). In this work, the sRNA tool was utilized to decrease the flux to glycolysis and pentose phosphate pathway based on the theophylline-inducible riboswitch. Interestingly, the down-regulation of either PPP (*zwf*) or glycolysis pathway (*pgi* or *pfk*A) led to an ~2-fold increase in MI production compared to that of WT-INO1SS, while the combined regulation of the two pathways realized a synergetic effect with ~5.58-fold increase of MI production. The results demonstrated that control of the competing pathways and driving more carbon into MI biosynthesis were important for MI production. In the future, attempts could be made to target more "carbon-consuming pathways" like glycogen, fatty acids, as well as acetate synthesis to direct more carbon into MI synthesis (Zhou et al., 2014). Meanwhile, in this study, the control for competing pathways was achieved using an inducible riboswitch, which needs additional inducers at specific time points. As artificial quorum sensing systems allowed cell growth at low cell density and induced specific gene expressions at high cell density automatically (Kim et al., 2017; Gu et al., 2020), it may represent a more suitable switch to control the essential pathways and is worth investigating in the future.

Limiting the carbon flux into other pathways was efficient for enhancing MI synthesis, while improving the total carbon fixation could also be important as it provides more carbon

precursors. Previously, the overexpression of genes encoding ribulose-1,5-bisphosphate carboxylase/oxygenase or extra bicarbonate transporters were both demonstrated as feasible for enhancing carbon fixation and biomass accumulation in *Synechocystis* (Atsumi et al., 2009; Kamennaya et al., 2015; Liang and Lindblad, 2016). In addition, supplementation of inorganic carbon like CO$_2$ or bicarbonate has been considered as a more direct strategy for carbon fixation reinforcement and production improvement (Wang et al., 2013). Consistently, supplementation of bicarbonate for the cultivation of WT-INO1SSZP further reached a ~1.6-fold increase in MI production. Nevertheless, the final production in *Synechocystis* is still much lower than that in other heterotrophic microorganisms like *B. subtilis* and *E. coli* (Brockman and Prather, 2015; Michon et al., 2020), suggesting that less than enough carbon sources and precursors were fed into the synthetic pathway. In the future, cultivation supplemented with organic carbon sources like glucose, used for heterotrophic organisms or photomixotrophic, could also be a feasible strategy (Vassilev et al., 2017; Qian et al., 2020). In this work, our preliminary results showed that supplementation of 5 mM glucose could significantly improve the growth and MI production of WT-INO1SS-ZP, finally reaching a production at 21.74 mg L^{-1} after 8 days of cultivation (**Supplementary Figure S6**), which also supported the argument that more carbon sources and precursors are needed to achieve high MI production.

The relatively lower growth rate of *Synechocystis* could also be an important limiting factor for MI production. Previously, a fast-growing cyanobacterium named *Synechococcus elongatus* UTEX 2973 (Yu et al., 2015; Lin et al., 2020) was isolated, whose shortest doubling time can reach 1.5 h at 41°C under continuous 1,500 μmol photons m^{-2} s^{-1} white light with 5% CO$_2$, close to that of *S. cerevisiae* (1.67 h). More importantly, the potential of *S. elongatus* UTEX 2973 for products like sucrose was found to be 6- to 26-fold compared with those in traditional cyanobacterial chassis like *Synechocystis*, *Anabaena* sp. PCC 7120, and *S. elongatus* PCC 7942 (Song et al., 2016), suggesting its

application potential for chemical synthesis. Similarly, other fast-growing cyanobacteria like *S. elongatus* PCC 11802 and *Synechococcus* sp. PCC 11901 were identified recently, offering more candidates as chassis for MI production in the future (Jaiswal et al., 2020; Wlodarczyk et al., 2020).

CONCLUSION

In this study, we engineered the model cyanobacterium *Synechocystis* for the sustainable production of MI. With the expression of IPS and IMP genes, simultaneous knockdown of three genes related to competing pathways, and cultivation optimization, photosynthetic production of MI directly from CO_2 was achieved, with a production of up to 12.72 mg L^{-1} after cultivation for 8 days, which represents an increase of ~12 times compared with initial MI-producing WT-INO1. The study presented here demonstrated the feasibility of converting CO_2 directly into MI in cyanobacterial chassis.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

REFERENCES

Atsumi, S., Higashide, W., and Liao, J. C. (2009). Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat. Biotechnol.* 27, 1177–1180. doi: 10.1038/nbt.1586

Berridge, M. J. (2009). Inositol trisphosphate and calcium signalling mechanisms. *Biochim. Biophys. Acta* 1793, 933–940. doi: 10.1016/j.bbamcr.2008.10.005

Brockman, I. M., and Prather, K. L. J. (2015). Dynamic knockdown of *E. coli* central metabolism for redirecting fluxes of primary metabolites. *Metab. Eng.* 28, 104–113. doi: 10.1016/j.ymben.2014.12.005

Chatterjee, A., Dastidar, K. G., Maitra, S., Das-Chatterjee, A., Dihazi, H., Eschrich, K., et al. (2006). sll1981, an acetolactate synthase homologue of *Synechocystis* sp. PCC6803, functions as L-myo-inositol 1-phosphate synthase. *Planta* 224, 367–379. doi: 10.1007/s00425-006-0221-4

Chatterjee, A., Majee, M., Ghosh, S., and Majumder, A. L. (2004). sll1722, an unassigned open reading frame of *Synechocystis* PCC 6803, codes for L-myo-inositol 1-phosphate synthase. *Planta* 218, 989–998. doi: 10.1007/s00425-003-1190-5

Cheng, F., Han, L., Xiao, Y., Pan, C., Li, Y., Ge, X., et al. (2019). d- chiro-inositol ameliorates high fat diet-induced hepatic steatosis and insulin resistance via PKCepsilon-PI3K/AKT pathway. *J. Agric. Food Chem.* 67, 5957–5967. doi: 10.1021/acs.jafc.9b01253

Fujisawa, T., Fujinaga, S., and Atomi, H. (2017). An in vitro enzyme system for the production of myo-inositol from starch. *Appl. Environ. Microbiol.* 83:e00550-17. doi: 10.1128/AEM.00550-17

Gao, X., Sun, T., Pei, G., Chen, L., and Zhang, W. (2016). Cyanobacterial chassis engineering for enhancing production of biofuels and chemicals. *Appl. Microbiol. Biotechnol.* 100, 3401–3413. doi: 10.1007/s00253-016-7374-2

Gu, F., Jiang, W., Mu, Y., Huang, H., Su, T., Luo, Y., et al. (2020). Quorum sensing-based dual-function switch and its application in solving two key metabolic engineering problems. *ACS Synth. Biol.* 9, 209–217. doi: 10.1021/acssynbio.9b00290

Gupta, A., Reizman, I. M., Reisch, C. R., and Prather, K. L. (2017). Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat. Biotechnol.* 35, 273–279. doi: 10.1038/nbt.3796

AUTHOR CONTRIBUTIONS

XW conducted the experiments, analyzed the data, and wrote the manuscript. LC and TS designed the research and revised the manuscript. JL helped with some of the experiments. WZ designed the research and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.566117/full#supplementary-material

Hansen, C. A., Dean, A. B., Draths, K. M., and Frost, J. W. (1999). Synthesis of 1,2,3,4-tetrahydroxybenzene from d-glucose: exploiting myo-inositol as a precursor to aromatic chemicals. *J. Am. Chem. Soc.* 121, 3799–3800. doi: 10.1021/ja9840293

Jaiswal, D., Sengupta, A., Sengupta, S., Madhu, S., Pakrasi, H. B., and Wangikar, P. P. (2020). A novel cyanobacterium *Synechococcus elongatus* PCC 11802 has distinct genomic and metabolomic characteristics compared to its neighbor PCC 11801. *Sci. Rep.* 10:191. doi: 10.1038/s41598-019-57051-0

Januszewski, M., Issat, T., Jakimiuk, A. A., Santor-Zaczynska, M., and Jakimiuk, A. J. (2019). Metabolic and hormonal effects of a combined Myo-inositol and d-chiro-inositol therapy on patients with polycystic ovary syndrome (PCOS). *Ginekol. Pol.* 90, 7–10. doi: 10.5603/GP.2019.0002

Johnson, T. J., Zahler, J. D., Baldwin, E. L., Zhou, R., and Gibbons, W. R. (2016). Optimizing cyanobacteria growth conditions in a sealed environment to enable chemical inhibition tests with volatile chemicals. *J. Microbiol. Methods* 126, 54–59. doi: 10.1016/j.mimet.2016.05.011

Kamennaya, N. A., Ahn, S., Park, H., Bartal, R., Sasaki, K. A., Holman, H. Y., et al. (2015). Installing extra bicarbonate transporters in the cyanobacterium *Synechocystis* sp. PCC6803 enhances biomass production. *Metab. Eng.* 29, 76–85. doi: 10.1016/j.ymben.2015.03.002

Kim, E. M., Woo, H. M., Tian, T., Yilmaz, S., Javidpour, P., Keasling, J. D., et al. (2017). Autonomous control of metabolic state by a quorum sensing (QS)-mediated regulator for bisabolene production in engineered *E. coli*. *Metab. Eng.* 44, 325–336. doi: 10.1016/j.ymben.2017.11.004

Liang, F., and Lindblad, P. (2016). Effects of overexpressing photosynthetic carbon flux control enzymes in the cyanobacterium *Synechocystis* PCC 6803. *Metab. Eng.* 38, 56–64. doi: 10.1016/j.ymben.2016.06.005

Lin, P. C., Zhang, F., and Pakrasi, H. B. (2020). Enhanced production of sucrose in the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Sci. Rep.* 10:390. doi: 10.1038/s41598-019-57319-5

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Lu, Y., Wang, L., Teng, F., Zhang, J., Hu, M., and Tao, Y. (2018). Production of myo-inositol from glucose by a novel trienzymatic cascade of polyphosphate

glucokinase, inositol 1-phosphate synthase and inositol monophosphatase. *Enzyme Microb. Technol.* 112, 1–5. doi: 10.1016/j.enzmictec.2018.01.006

Ma, K., Thomason, L. A., and McLaurin, J. (2012). scyllo-Inositol, preclinical, and clinical data for Alzheimer's disease. *Adv. Pharmacol.* 64, 177–212. doi: 10.1016/b978-0-12-394816-8.00006-4

Michon, C., Kang, C. M., Karpenko, S., Tanaka, K., Ishikawa, S., and Yoshida, K. I. (2020). A bacterial cell factory converting glucose into scyllo-inositol, a therapeutic agent for Alzheimer's disease. *Commun. Biol.* 3:93. doi: 10.1038/s42003-020-0814-7

Patra, B., Ghosh Dastidar, K., Maitra, S., Bhattacharyya, J., and Majumder, A. L. (2007). Functional identification of sll1383 from *Synechocystis* sp PCC 6803 as L-myo-inositol 1-phosphate phosphatase (EC 3.1.3.25): molecular cloning, expression and characterization. *Planta* 225, 1547–1558. doi: 10.1007/s00425-006-0441-7

Qian, D. K., Geng, Z. Q., Sun, T., Dai, K., Zhang, W., Jianxiong Zeng, R., et al. (2020). Caproate production from xylose by mesophilic mixed culture fermentation. *Bioresour. Technol.* 308:123318. doi: 10.1016/j.biortech.2020.123318

Regidor, P. A., Schindler, A. E., Lesoine, B., and Druckman, R. (2018). Management of women with PCOS using myo-inositol and folic acid. New clinical data and review of the literature. *Horm. Mol. Biol. Clin. Investig.* 34:20170067. doi: 10.1515/hmbci-2017-0067

Roessner, U., Luedemann, A., Brust, D., Fiehn, O., Linke, T., Willmitzer, L., et al. (2001). Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *Plant Cell* 13, 11–29. doi: 10.1105/tpc.13.1.11

Song, K., Tan, X., Liang, Y., and Lu, X. (2016). The potential of *Synechococcus elongatus* UTEX 2973 for sugar feedstock production. *Appl. Microbiol. Biotechnol.* 100, 7865–7875. doi: 10.1007/s00253-016-7510-z

Sun, T., Li, S., Song, X., Diao, J., Chen, L., and Zhang, W. (2018a). Toolboxes for cyanobacteria: recent advances and future direction. *Biotechnol. Adv.* 36, 1293–1307. doi: 10.1016/j.biotechadv.2018.04.007

Sun, T., Li, S., Song, X., Pei, G., Diao, J., Cui, J., et al. (2018b). Re-direction of carbon flux to key precursor malonyl-CoA via artificial small RNAs in photosynthetic *Synechocystis* sp. PCC 6803. *Biotechnol. Biofuels* 11:26. doi: 10.1186/s13068-018-1032-0

Tanaka, K., Tajima, S., Takenaka, S., and Yoshida, K. (2013). An improved *Bacillus subtilis* cell factory for producing scyllo-inositol, a promising therapeutic agent for Alzheimer's disease. *Microb. Cell Fact.* 12:124. doi: 10.1186/1475-2859-12-124

Thomas, M. P., Mills, S. J., and Potter, B. V. (2016). The "Other" inositols and their phosphates: synthesis, biology, and medicine (with recent advances in myo-inositol chemistry). *Angew. Chem. Int. Ed. Engl.* 55, 1614–1650. doi: 10.1002/anie.201502227

Vassilev, N., Malusa, E., Requena, A. R., Martos, V., Lopez, A., Maksimovic, I., et al. (2017). Potential application of glycerol in the production of plant beneficial microorganisms. *J. Ind. Microbiol. Biotechnol.* 44, 735–743. doi: 10.1007/s10295-016-1810-2

Wang, B., Pugh, S., Nielsen, D. R., Zhang, W., and Meldrum, D. R. (2013). Engineering cyanobacteria for photosynthetic production of 3-hydroxybutyrate directly from CO2. *Metab. Eng.* 16, 68–77. doi: 10.1016/j.ymben.2013.01.001

Wang, Y., Sun, T., Gao, X., Shi, M., Wu, L., Chen, L., et al. (2016). Biosynthesis of platform chemical 3-hydroxypropionic acid (3-HP) directly from CO2 in cyanobacterium *Synechocystis* sp. PCC 6803. *Metab. Eng.* 34, 60–70. doi: 10.1016/j.ymben.2015.10.008

Wlodarczyk, A., Selao, T. T., Norling, B., and Nixon, P. J. (2020). Newly discovered *Synechococcus* sp. PCC 11901 is a robust cyanobacterial strain for high biomass production. *Commun. Biol.* 3:215. doi: 10.1038/s42003-020-0910-8

You, C., Shi, T., Li, Y., Han, P., Zhou, X., and Zhang, Y. P. (2017). An in vitro synthetic biology platform for the industrial biomanufacturing of myo-inositol from starch. *Biotechnol. Bioeng.* 114, 1855–1864. doi: 10.1002/bit.26314

Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., et al. (2015). *Synechococcus elongatus* UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO(2). *Sci. Rep.* 5:8132. doi: 10.1038/srep08132

Zhou, J., Zhang, H., Meng, H., Zhang, Y., and Li, Y. (2014). Production of optically pure d-lactate from CO2 by blocking the PHB and acetate pathways and expressing d-lactate dehydrogenase in cyanobacterium *Synechocystis* sp. PCC 6803. *Process Biochem.* 49, 2071–2077. doi: 10.1016/j.procbio.2014.09.007

Phototrophic Co-cultures From Extreme Environments: Community Structure and Potential Value for Fundamental and Applied Research

Claire Shaw[1], Charles Brooke[1], Erik Hawley[2], Morgan P. Connolly[3], Javier A. Garcia[4], Miranda Harmon-Smith[5], Nicole Shapiro[5], Michael Barton[5], Susannah G. Tringe[5], Tijana Glavina del Rio[5], David E. Culley[6], Richard Castenholz[7] and Matthias Hess[1]*

[1] Systems Microbiology and Natural Products Laboratory, University of California, Davis, Davis, CA, United States, [2] Bayer, Pittsburg, PA, United States, [3] Microbiology Graduate Group, University of California, Davis, Davis, CA, United States, [4] Biochemistry, Molecular, Cellular, and Developmental Biology Graduate Group, University of California, Davis, Davis, CA, United States, [5] Department of Energy, Joint Genome Institute, Berkeley, CA, United States, [6] Greenlight Biosciences, Inc., Medford, MA, United States, [7] Department of Biology, University of Oregon, Eugene, OR, United States

Cyanobacteria are found in most illuminated environments and are key players in global carbon and nitrogen cycling. Although significant efforts have been made to advance our understanding of this important phylum, still little is known about how members of the cyanobacteria affect and respond to changes in complex biological systems. This lack of knowledge is in part due to our dependence on pure cultures when determining the metabolism and function of a microorganism. We took advantage of the Culture Collection of Microorganisms from Extreme Environments (CCMEE), a collection of more than 1,000 publicly available photosynthetic co-cultures maintained at the Pacific Northwest National Laboratory, and assessed via 16S rRNA amplicon sequencing if samples readily available from public culture collection could be used in the future to generate new insights into the role of microbial communities in global and local carbon and nitrogen cycling. Results from this work support the existing notion that culture depositories in general hold the potential to advance fundamental and applied research. Although it remains to be seen if co-cultures can be used at large scale to infer roles of individual organisms, samples that are publicly available from existing co-cultures depositories, such as the CCMEE, might be an economical starting point for such studies. Access to archived biological samples, without the need for costly field work, might in some circumstances be one of the few remaining ways to advance the field and to generate new insights into the biology of ecosystems that are not easily accessible. The current COVID-19 pandemic, which makes sampling expeditions almost impossible without putting the health of the participating scientists on the line, is a very timely example.

Keywords: biodiversity, biotechnology, culture collection, cyanobacteria, extreme environments, fundamental research, microbial ecology and diversity

INTRODUCTION

Cyanobacteria are photosynthetic prokaryotes that are found in the majority of illuminated habitats and are known to be some of the most morphologically diverse prokaryotes on our planet (Whitton and Potts, 2000). The global cyanobacterial biomass is estimated to total $\sim 3 \times 10^{14}$ g of carbon (Garcia-Pichel et al., 2003) and cyanobacteria may account for 20–30% of Earth's primary photosynthetic productivity (Pisciotta et al., 2010). The efficient photosynthetic machinery of cyanobacteria has inspired growing interest in the utilization of cyanobacteria and cyanobacteria containing co-cultures in microbial fuel cells (Zhao et al., 2012; Gajda et al., 2015). In addition to having a global effect on the carbon cycle, cyanobacteria-mediated nitrogen fixation has been estimated to supply 20–50% of the nitrogen input in some marine environments (Karl et al., 1997). A detailed comprehension of cyanobacteria and their role in global carbon and nitrogen cycling is therefore indispensable for a multi-scalar and holistic understanding of these globally important nutrient cycles.

Besides their ecological relevance, cyanobacteria have potential applications in biotechnology: cyanobacteria facilitate the assimilation of carbon dioxide, a cheap and abundant substrate, to synthesize a variety of value-added compounds with industrial relevance (Al-Haj et al., 2016). Although monocultures have dominated in microbial biomanufacturing, controlled co-cultures have been recognized as valuable alternatives, due to their potential for reducing the risk of costly contaminations and in some cases enabling increasing product yield (Wang et al., 2015; Yen et al., 2015; Padmaperuma et al., 2018). Numerous cyanobacterial strains have been investigated for their potential to produce bioactive compounds, biofertilizer, biofuels, and bioplastics (Abed et al., 2009; Woo and Lee, 2017; Miao et al., 2018); and co-expression of non-cyanobacterial genes as well as co-cultivation of cyanobacteria with non-photosynthetic bacteria has resulted in self-sustained systems and improved desirable cyanobacterial phenotypes (de-Bashan et al., 2002; Subashchandrabose et al., 2011; Formighieri and Melis, 2016). Genes coding for enzymes capable of catalyzing reactions that result in unique products, such as modified trichamide, a cyclic peptide suggested to protect the bloom-forming *Trichodesmium erythraeum* against predation (Sudek et al., 2006); and prochlorosins, a family of lanthipeptides with diverse functions that are synthesized by various strains of *Prochlorococcus* and *Synechococcus* (Li et al., 2010; Cubillos-Ruiz et al., 2017), have been identified from cyanobacterial genomes (Zarzycki et al., 2013; Kleigrewe et al., 2016). It is very likely that *de novo* genome assembly from metagenomic data will facilitate the discovery of novel enzymes from cyanobacteria that are recalcitrant to current isolation and cultivation techniques. Although metagenome-derived genomes hold great potential to enhance our knowledge about genomic dark matter, improved techniques to isolate and enable axenic culturing of microorganisms that are currently considered as "unculturable," as well as new genetic tools to study non-axenic cultures will be necessary in order to fully access the biotechnological potential of cyanobacteria.

Culture collections provide the possibility of preserving microbial isolates over extended periods of time without introducing significant genetic changes (McCluskey, 2017) and they provide easy access to these isolates and their associated metadata (Boundy-Mills et al., 2015). Although culture collections hold enormous potential for capturing and preserving microbial biodiversity, there are numerous challenges in maintaining these biological depositories. With recent advances in DNA sequencing technologies and the accessibility of 16S rRNA gene-based microbial community profiling, we are now well positioned to re-inventory, and standardize existing culture collections, which will be essential for preserving and cataloging the planet's microbial biodiversity.

To explore the potential of culture collections, specifically those that maintain samples of microbial co-cultures, we reexamined the biodiversity of 26 historical phototrophic samples from the Culture Collection of Microorganisms from Extreme Environments (CCMEE). While some of the samples, selected for this project were studied previously (**Supplementary Table 1**) using cloned-based 16S rRNA profiling and morphological characterization (Camacho et al., 1996; Miller and Castenholz, 2000; Nadeau and Castenholz, 2000; Nadeau et al., 2001; Dillon et al., 2002; Dillon and Castenholz, 2003; Norris and Castenholz, 2006; Toplin et al., 2008), the diversity and the overall community assemblage of these co-cultures have not yet been characterized. We selected samples from environments with distinct and extreme physical properties from across the globe, suggesting each co-culture would yield a unique microbial consortium. Although reasonable to assume that these consortia have changed over time in composition and function (due to their cultivation), it is very likely that results obtained during this work will still provide insights into the microbial biodiversity of extreme habitats, some of which may no longer be accessible.

MATERIALS AND METHODS

Sample Collection and Sample Description

Co-cultures selected for this study are part of a larger culture collection and were collected from different locations (**Table 1**) between 1988 and 2002. Isolates were collected using sterile techniques, kept in the dark and stored on ice as soon as possible. Samples were transported to the laboratory where aliquots were prepared preservation at $-80°C$ and cultivation. For this study, co-cultures were selected from the CCMEE to cover a variety of geographical locations (**Supplementary Figure 1**) as well as a range of different ecosystems (**Table 1**). Due to the lack of a consistent usage of terminology to describe the sampling sites, we categorized co-cultures according to the geographical location (e.g., Antarctica, Bermuda, Denmark, Mexico and Spain) and based on the general description of the ecosystems (i.e., creek, crust, freshwater, hot spring, marine, saline pond, terrestrial, travertine, and tree bark) from where the co-cultures were collected. In addition, we used the growth medium and temperature (i.e., 12, 23, 40, 45, and 55°C) at which available

TABLE 1 | Summary of photosynthetic co-cultures for which 16S rRNA gene profiles were generated.

FECB ID[a]	CCMEE ID[a]	Growth temperature [°C]	Sample location [country]	Sample location [region]	Habitat
FECB1	5011	12	Antarctica	McMurdo Ice Shelf; Bratina Island	Pond (saline)
FECB2	5019	12	Antarctica	McMurdo Ice Shelf; Bratina Island	Pond (freshwater)
FECB3	5034	12	Antarctica	McMurdo Ice Shelf; Bratina Island	Pond (brackish)
FECB4	5047	23	Spain	Lake Arcas	Lake (freshwater)
FECB5	5049	23	Spain	Lake Arcas	Lake (freshwater)
FECB6	5051	23	United States	Yellowstone National Park; Pott's Basin	Hot Spring
FECB10	5056	23	Mexico	Viscaino Desert	Terrestrial (epilithic)
FECB12	5062-A	23	United States	University of Oregon; Eugene, Oregon	Terrestrial (concrete)
FECB14	5093	23	United States	Yellowstone National Park; Pott's Geyser	Hot Spring
FECB15	5083	23	United States	Yellowstone National Park; Rabbit Creek	Warm Creek
FECB17	5085	23	United States	Yellowstone National Park; Rabbit Creek	Warm Creek
FECB19	5091	23	United States	Yellowstone National Park; Shoshone Geyser Basin	Hot Spring
FECB21	5093-B	23	United States	Yellowstone National Park; Pott's Geyser	Hot Spring
FECB22	5097	23	United States	Hawaii	Terrestrial (tree trunk)
FECB24	5098	23	Antarctica	McMurdo Ice Shelf; Bratina Island	Pond (freshwater)
FECB26	5099	23	Bermuda	Sommerset	Terrestrial (wooden fence)
FECB28	5102	23	Antarctica	McMurdo Ice Shelf; Bratina Island	Pond (saline)
FECB30	5108-A	23	Denmark	Limfjord Shallows, Limfjord	Marine
FECB32	6031	23	United States	Yellowstone National Park; Mammoth; Narrow Gauge	Travertine (endolithic)
FECB34	6069	23	United States	Yellowstone National Park; Lower Geyser Basin; Great Fountain	Crust
FECB36	6076	23	United States	Yellow Stone National Park; Lower Geyser Basin; Sentinel Spring	Hot Spring (silica crust)
FECB38	6083	23	United States	Yellowstone National Park; Lower Geyser Basin; Mushroom Spring	Crust
FECB52	5506	40	United States	Yellowstone National Park; Norris Geyser Basin	Terrestrial (acid crust)
FECB53	5610	40	United States	Yellowstone National Park; Sylvan Springs	Hot Spring (acid pool)
FECB58	5216	45	United States	Hunter's Hot Spring, Oregon	Hot Spring
FECB68	5240	55	United States	Hunter's Hot Spring, Oregon	Hot Spring

[a]Co-cultures used in this work are maintained in the CCMEE and can be requested from Dr. Cady (Sherry.Cady@pnnl.gov).

co-cultures have been maintained by the CCMEE curators to categorize the selected co-cultures.

To facilitate future work, we assigned new unique FECB identifiers (FECB for Functional Encyclopedia of Cyanobacteria) to the selected co-cultures (**Table 1**). When requesting aliquots for future work from the CCMEE, these new FECB identifiers should be used.

FECB1 (CCMEE ID 5011) and FECB3 (CCMEE ID 5034) were collected from saline and brackish melt ponds and were dominated by phototrophic cyanobacteria previously classified as *Oscillatoria* sp. (Nadeau et al., 2001). FECB2 (CCMEE ID 5019) was collected from a freshwater pond and its composition was not investigated prior to our efforts. FECB4 (CCMEE ID 5047; AP1) and FECB5 (CCMEE ID 5049; AO21) were also isolated from freshwater and the dominant photosynthetic organisms within these samples were classified previously by 16S rRNA sequence analysis as relatives of *Pseudanabaena limnetica* and *Oscillatoria* cf. *tenuis*, respectively (Camacho et al., 1996). FECB6 (CCMEE ID 5051), FECB14 (CCMEE ID 5093; WT-97 Cal), FECB15 (CCMEE ID 5083), and FECB19 (CCMEE ID 5091; Y-97) were collected from diverse hot springs within Yellowstone National

Park (YNP) (**Table 1**). FECB10 [CCMEE ID 5056; M88-VD (1)] was collected as epiliths (Dillon et al., 2002). FECB17 (CCMEE ID 5085; RC-97 Cal) and FECB36 (CCMEE ID 6076) were isolated from Rabbit Creek and a crust in the Sentinel Spring Meadows in YNP, respectively, and dominant phototrophs of these co-cultures were characterized previously as *Calothrix* spp. (Dillon and Castenholz, 2003). FECB22 (CCMEE ID 5097; HW-91) and FECB26 (CCMEE ID 5099; B77-scy, j) were collected from a tree trunk and a wooded fence, respectively. FECB24 (CCMEE ID 5098; AN-90) was collected from a shallow melt pond ($\sim$10 m²) in the Victoria Valley, Antarctica, whereas FECB28 (CCMEE ID 5102) was collected from a saline melt pond on Bratina Island, Antarctica (Nadeau and Castenholz, 2000). FECB32 (CCMEE ID 6031), FECB34 (CCMEE ID 6069) and FECB38 (CCMEE ID 6083) were endoliths collected from subsurface (1–5 mm depths) travertine deposits in YNP (Norris and Castenholz, 2006). FECB53 (CCMEE ID 5610) was collected from Sylvan Springs in YNP. Temperature and pH at FECB53's sampling site were determined to be 40°C and pH4, conditions which are considered to be too harsh to actively support growth of cyanobacteria, and Toplin et al. (2008) reported the thermo-acidophilic red algae

Cyanidioschyzon as a highly abundant phototropic strain in this sample. FECB58 (CCMEE ID 5216; OH-9-45C) and FECB68 (CCMEE ID 5240; OH-2-55C) were collected from Hunter's Hot Spring in Oregon and in 2000 Miller and Castenholz reported the isolation of several thermophilic clones belonging to the genus *Synechococcus* from these samples (Miller and Castenholz, 2000).

Growth of Co-cultures

To obtain sufficient biomass for subsequent DNA analysis, 100 μL of each co-culture were transferred to 25 mL of sterile BG11 media (Allen and Stanier, 1968). For FECB52 and FECB53 BG11 was substituted by Cyanidium medium (Castenholz, 1981). Co-cultures were subjected to a 12 h diurnal light/dark cycle while grown over 28 days at a temperature similar to the temperature that was measured at the location where the sample was collected. Growth temperature for each sample is indicated in **Table 1**.

DNA Extraction and 16S rRNA Gene Amplification

Total microbial DNA was extracted from 500 μL of each photosynthetic co-culture using the FastDNA SPIN Kit for Soil (MP Biomedical, Solon, OH, United States) according to the manufacturer's instructions. Extracted DNA was quantified via fluorescence (Qubit; Thermo Scientific, United States) and the hypervariable V4 region of the 16S rRNA gene was amplified from extracted DNA using the primer set 515F/805R (515F: 5′-GTGCCAGCMGCCGCGGTAA-3′ and 805R: 5′-GGACTACHVGGGTWTCTAAT-3′). The forward primer included an 11 bp barcode to allow multiplexing of samples during sequencing. The barcode sequence for each sample is listed in **Supplementary Table 2**. Subsequent PCR reactions were performed using the 5PRIME HotMasterMix amplification mix (QIAGEN, Beverly, MA, United States) with the following PCR conditions: initial denaturation for 90 s at 94°C, followed by 30 amplification cycles (45 s at 94°C, 60 s at 60°C, and 90 s at 72°C) followed by a final extension step of 72°C for 10 min. Amplification products were cooled to 4°C. Samples were sequenced at the Department of Energy's Joint Genome Institute (JGI; [1]) according to JGI's standard operating procedure using Illumina's MiSeq platform and v3 chemistry.

Sequence Data Analysis

Raw sequencing data were downloaded from the JGI's Genome Portal[2] were they are deposited and accessible under the project ID 1032475. Data were decompressed and de-interleaved using the 7-zip software[3] and an in-house script, respectively. De-interleaved files were subsequently processed using MOTHUR version 1.38.1 (Schloss et al., 2009; Kozich et al., 2013). Paired-end reads were combined using the *make.contigs* command. Sequences with ambiguous base calls and sequences longer than 325 bp were removed using *screen.seqs*. Duplicate sequences were merged using

unique.seqs, and the resulting unique sequences were aligned to the V4 region of the SILVA database (v123) (Quast et al., 2013). Chimeras were removed using UCHIME (Edgar et al., 2011) and quality filtered sequences were taxonomically classified at 80% confidence to the GreenGenes reference taxonomy (release gg_13_5_99) (McDonald et al., 2012). Non-prokaryotic sequences were removed and the *dist.seqs* command was used to calculate pairwise distances between the aligned sequences. The resulting pairwise distance matrix was used to cluster sequences into operational taxonomic units (OTUs) with a 97% sequence identity cut-off using UCLUST (Edgar, 2010). The most abundant sequence of each OTU was picked as the representative sequence. OTUs were taxonomically classified using the *classify.otu* command using the GreenGenes reference taxonomy (release gg_13_5_99). Shannon and Simpson estimators were calculated in MOTHUR (Schloss et al., 2009).

In order to visualize the overall compositional differences between the co-cultures, an uncorrected pairwise distance matrix was generated using the *dist.seqs* command in MOTHUR and a tree was generated using *Clearcut* (version 1.0.9) (Evans et al., 2006). A cladogram from the resulting tree file was constructed and visualized using iTOL (Letunic and Bork, 2016). Cluster designations were assigned at a branch length of 0.05, with branch length indicating the (number of differences/overall length of branches) between two samples. Samples whose branches split at a distance >0.05 were considered as part of the same cluster (**Figure 1**).

Availability of Data and Material

Co-cultures subject to this study are publicly available through the CCMEE upon request by contacting Sherry Cady[4] using the corresponding FECB ID or CCMEE ID (**Table 1**). Co-cultures can also be obtained from the Hess Lab at UC Davis. Sequences generated during this project have been deposited and are publicly available at NCBI's SRA under the BioProject ID PRJNA401502. All other data is included in this published article and its **Supplementary Information** files.

The CCMEE is now maintained at the Northwest National Laboratory by Dr. Cady. The CCMEE is comprised of >1,200 co-cultures, including the co-cultures that were studied in the work presented here, and has been established to provide a valuable resource to the scientific community. Cultures that are part of the CCMEE can be requested from Sherry Cady[4].

RESULTS AND DISCUSSION

Prokaryotic Diversity and Eukaryotic Population of Co-cultures

A total of 3,357,905 raw reads [mean (SD) = 129,150 (±15,845) reads per sample] were generated from the V4 region of the 16S rRNA gene (**Table 2**). Quality filtering removed ∼3.8% (±0.57%) of the raw reads from each sample due to insufficient quality. The remaining reads were assigned to a total of 5,785 distinct

---

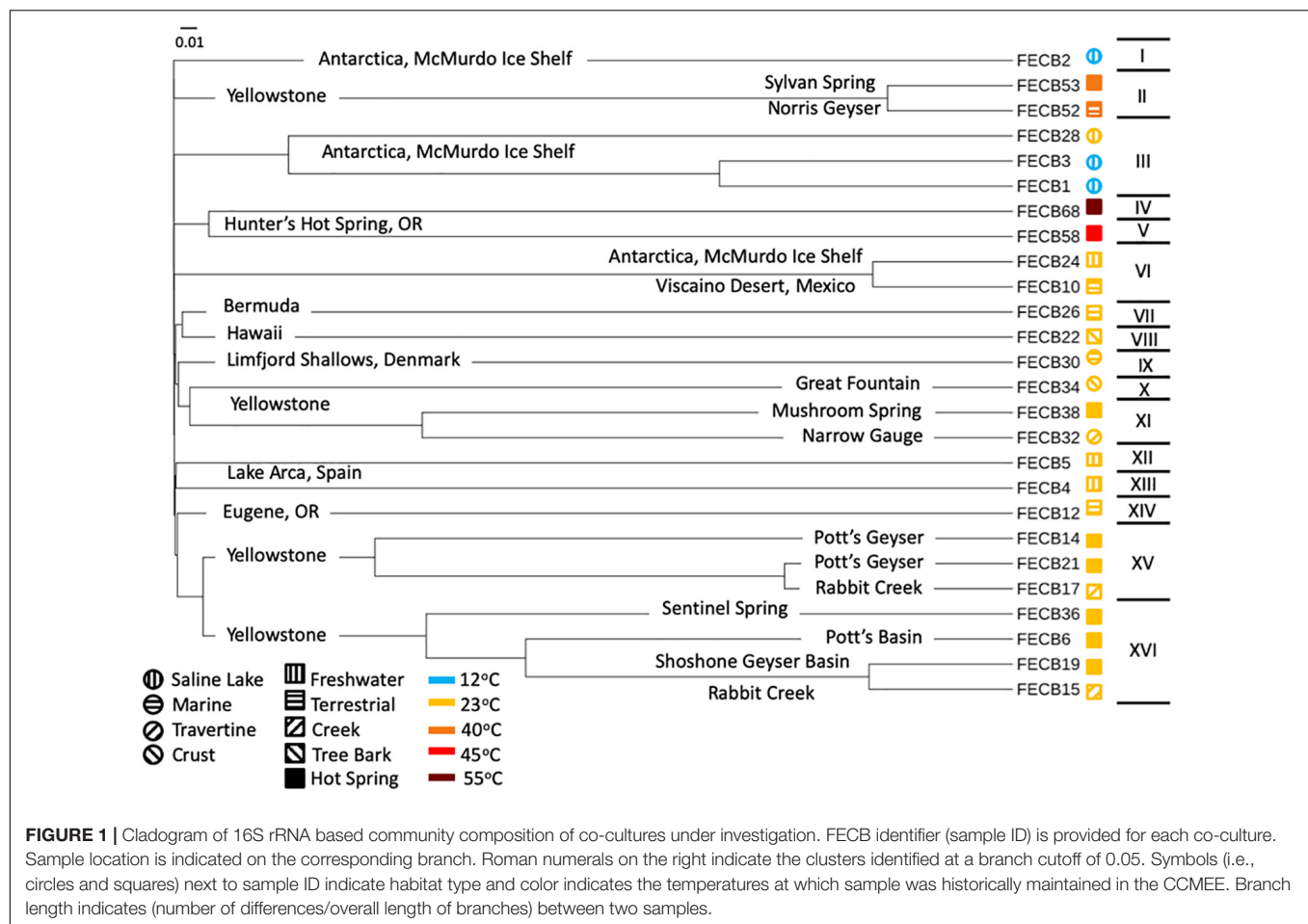[1]http://www.jgi.doe.gov
[2]http://genome.jgi.doe.gov/
[3]www.7-zip.org

[4]sherry.cady@pnnl.gov

FIGURE 1 | Cladogram of 16S rRNA based community composition of co-cultures under investigation. FECB identifier (sample ID) is provided for each co-culture. Sample location is indicated on the corresponding branch. Roman numerals on the right indicate the clusters identified at a branch cutoff of 0.05. Symbols (i.e., circles and squares) next to sample ID indicate habitat type and color indicates the temperatures at which sample was historically maintained in the CCMEE. Branch length indicates (number of differences/overall length of branches) between two samples.

Operational Taxonomic Units (OTUs) based on 97% sequence identity (**Supplementary Table 3**).

To estimate the microbial diversity within each sample, rarefaction analyses were performed (**Supplementary Figure 2**) and diversity indices were calculated (**Table 2**). The Inverse Simpson index (Simpson, 1949; Morris et al., 2014) of the samples ranged between 1.21 and 9.24 with the lowest and highest indices calculated for FECB53 and FECB32, respectively (**Table 2**), illustrating that co-cultures investigated during this project represented co-cultures from a wide range of diversity. Not surprisingly, the diversity in the co-cultures under investigation appeared to be negatively correlated with the proportion of reads recruited by the dominant OTU of each sample (Pearson $r = -0.8806$; $p < 0.01$). Although samples ranked slightly differently based on their diversity, when diversity was calculated using the Shannon index (Kim et al., 2017), the overall trend remained the same (**Table 2**).

The presence of eukaryotic microorganisms in each co-culture was estimated using mitochondrial reads resulting from the 16S rRNA sequencing (**Supplementary Table 4**). FECB52 had the greatest percentage of mitochondrial DNA at 0.29% of total reads. FECB2, FECB4, FECB6, FECB14, FECB17, FECB26, FECB30, and FECB68 all contained no mitochondrial reads. Average mitochondrial reads as a percentage of all reads averaged 0.02% across all 26 samples.

Ubiquity of Cyanobacteria and Proteobacteria Within Photosynthetic Co-cultures

While the microbial communities of the co-cultures under investigation varied greatly, cyanobacteria and proteobacteria co-occurred in all 26 of the community assemblages. Community composition analysis revealed that each of the co-cultures contained at least one OTU [mean (SD) = 2 ($\pm$1.23)] that recruited ($>0.1\%$ of the co-culture specific reads and that was classified as Cyanobacteria (**Table 3**). The only other phylum present in each of the individual 26 co-cultures and represented by at least one OTU recruiting ($>0.1\%$ of the reads was the *Proteobacteria* phylum (**Table 3**). In contrast, only three samples, namely FECB5, FECB30, and FECB68, contained OTUs that recruited $>0.1\%$ of the sample specific reads and that could not be classified at the phylum level or at a higher taxonomic resolution (**Table 3**). It is possible that the relatively high abundance of non-classified phyla might contribute to the separation of these samples into distinct clusters (i.e., cluster XII, IX, and IV; **Figure 1**). In addition to their ubiquity, *Cyanobacteria* and

TABLE 2 | Read statistics and diversity Index for co-cultures investigated in this study.

Sample ID	Total raw reads	Total quality filtered reads	Total OTUs observed	OTUs recruiting >0.1% of reads	% Reads represented by OTUs recruiting >0.1% of reads	% reads recruited by dominant OTU	Inverse Simpson index	Shannon index
FECB1	154,251	139,666	349	13	99.6%	47%	3.23	1.49
FECB2	128,172	116,166	197	3	99.8%	51%	2.04	0.77
FECB3	139,819	128,879	274	8	99.6%	80%	1.52	0.76
FECB4	94,960	84,983	288	17	99.4%	44%	3.92	1.77
FECB5	122,414	108,882	401	18	99.4%	40%	4.63	2.00
FECB6	153,355	136,199	549	28	99.4%	29%	6.5	2.39
FECB10	114,324	102,534	299	10	99.7%	72%	1.78	0.96
FECB12	151,629	141,953	335	19	99.5%	72%	1.87	1.23
FECB14	145,839	98,588	206	4	99.6%	67%	1.83	0.75
FECB15	105,256	96,030	278	9	99.6%	63%	2.18	1.14
FECB17	109,474	97,391	317	7	99.5%	59%	2.47	1.33
FECB19	118,003	103,477	300	10	99.5%	47%	3.27	1.50
FECB21	127,371	113,281	401	15	99.2%	58%	2.38	1.27
FECB22	129,934	117,651	399	23	99.4%	40%	4.46	2.08
FECB24	115,275	93,362	265	12	99.9%	65%	2.12	1.17
FECB26	121,303	103,597	265	16	99.5%	63%	2.29	1.32
FECB28	152,974	134,318	332	10	99.4%	37%	3.22	1.36
FECB30	140,681	118,256	546	22	98.7%	33%	4.61	2.00
FECB32	116,889	72,130	336	29	98.3%	19%	9.24	2.68
FECB34	138,723	125,710	288	14	99.9%	76%	1.69	1.11
FECB36	137,395	124,885	360	12	99.6%	52%	2.38	1.16
FECB38	139,001	124,379	479	17	99.1%	32%	5.22	2.01
FECB52	117,261	75,433	367	7	99.9%	71%	1.85	1.02
FECB53	119,380	86,565	303	5	99.1%	91%	1.21	0.44
FECB58	135,683	124,119	351	13	99.3%	38%	3.65	1.61
FECB68	128,539	114,032	402	11	99.7%	26%	5.27	1.90
Total	3,357,905	2,882,466	N/A	N/A	N/A	N/A	N/A	N/A
Average	129,150	110,864	342	14	99%	51%	N/A	N/A
Stdev	15,845	19,275	87.2	6	0%	17%	N/A	N/A

Proteobacteria also recruited the majority of the reads in all but four (i.e., FECB2, FECB12, FECB58, and FECB68) of the samples under investigation (**Figure 2** and **Supplementary Table 5**). In FECB2 and FECB12 the majority of the reads were recruited by OTUs classified as members of the phylum *Bacteroidetes* (recruiting 50.6 and 72% of the reads, respectively), whereas within FECB58 and FECB68, *Armatimonadetes* (38.3%) and *Chloroflexi* (25.9%) were identified as the most abundant phyla (**Figure 2** and **Supplementary Table 5**). The fact that these samples were dominated by phyla other than the *Cyanobacteria* or *Proteobacteria* may also help to explain why these samples form distinct clusters (cluster I, XIV and V, IV, respectively; **Figure 1**).

Firmicutes Dominate Photosynthetic Co-cultures From Hot Springs

Firmicutes abundances calculated for co-cultures from hot spring samples were higher compared to those calculated for co-cultures from other environments. OTUs assigned to the *Firmicutes* phylum were detected above the applied cut-off level of 0.1% in only five of the twenty-six co-cultures

under investigation (**Table 3**). Interestingly, these samples (i.e., FECB32, FECB34, FECB52, FECB58, and FECB68) are co-cultures collected from hot springs or from deposits within hot springs, with FECB52, FECB58, and FECB68 being maintained in culture at temperatures >40°C. OTU000073 (classified as *Alicyclobacillus tolerans*), OTU00082 (classified as members of the genus *Paenibacillus*), OTU000154 (classified as *Geobacillus vulcani*), and OTU000158 (classified as a member of the *Bacillaceae* family) recruited 5.9, 3.4, 0.5, and 0.4% of the reads generated from FECB52, FECB34, FECB68, and FECB58, respectively (**Supplementary Table 3**). *Alicyclobacillus tolerans* and *Geobacillus vulcani* have been described previously as aerobic spore-forming thermophiles and have been isolated from lead–zinc ores (Karavaiko et al., 2005) and hot springs (Nazina et al., 2004) located in Russia, respectively. Members of the genus *Paenibacillus* have been isolated from a wide variety of environments and some *Paenibacillus* species have been found to promote crop growth directly via biological nitrogen fixation, phosphate solubilization, production of the phytohormone indole-3-acetic acid; and they have been identified as a potential source of novel antimicrobial agents (Grady et al., 2016).

TABLE 3 | Count and phylogenetic classification of identified OTUs at the phylum level.

Phylum	FECB 1	FECB 2	FECB 3	FECB 4	FECB 5	FECB 6	FECB 10	FECB 12	FECB 14	FECB 15	FECB 17	FECB 19	FECB 21	FECB 22	FECB 24	FECB 26	FECB 28	FECB 30	FECB 32	FECB 34	FECB 36	FECB 38	FECB 52	FECB 53	FECB 58	FECB 68
Cyanobacteria	2	1	2	1	1	3	1	2	2	2	1	2	6	5	1	1	1	2	3	1	2	2	3	1	1	2
Proteobacteria	10	1	5	13	11	18	7	12	2	7	5	5	6	16	10	14	8	16	20	11	9	14	3	2	6	2
Actinobacteria	1	0	1	3	1	1	2	3	0	0	1	2	1	2	0	0	1	0	3	0	1	0	0	2	0	0
Bacteroidetes	0	1	0	0	1	6	0	1	0	0	0	0	2	0	1	1	0	0	1	1	0	1	0	0	1	0
Armatimonadetes	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
Unclassified Bacteria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2
Firmicutes	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1
Thermi	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Chloroflexi	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
Acidobacteria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0
Planctomycetes	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	1	0	0	0	0	0	0	0
Total OTU count	13	3	8	17	18	28	10	19	4	9	7	10	15	23	12	16	10	22	29	14	12	17	7	5	13	11

Only OTUs recruiting >0.1% of the co-culture specific reads are shown.

Photosynthetic Co-cultures From Antarctica and YNP to Study Adaptation to Increased Radiation, Low Temperatures and Oligotrophic Growth Conditions

Microbial adaptation to extreme environments and the molecular framework that enable microorganisms to survive and thrive in the presence of increased rates of radiation, low temperatures and in the absence of nutrients has fascinated the scientific community for decades and remains poorly understood. In an attempt to provide a better basis of the taxonomic make-up of co-cultures that were collected from ecosystems that are characterized by these extremes, we included co-cultures from Antarctica and YNP in this study (**Table 1**). OTU-based comparison of Antarctica and YNP co-cultures revealed between 197 (FECB2) and 549 (FECB6) distinct OTUs [mean (SD) 342 (87.2) OTUs] based on 97% sequence similarity (**Table 2**). The number of OTUs that recruited >0.1% of all reads ranged from 3 to 29 OTUs, with FECB2 and FECB32 having the lowest and highest OTU count, respectively (**Table 2**). FECB2 was dominated by an OTU classified as *Hymenobacter*, which recruited all *Bacteroidetes*-specific reads generated from this sample (**Tables 3**, **4**). The genus *Hymenobacter* contains several pigmented bacteria that have been isolated from Antarctica and have been reported to possess increased resistance to radiation (Oh et al., 2016; Marizcurrena et al., 2017), which might explain their increased abundance in FECB2, a co-culture isolated from an environment known to possess increased levels of UV radiation. Taking this into consideration, FECB2 and its individual community members could be a potential target for future studies to enhance our understanding of processes that infer resistance to radiation and DNA damage. The second most abundant OTU in FECB2, recruiting 48% of the generated samples, was classified as *Phormidium* sp. (**Supplementary Table 3**), a cyanobacterial genus that has been reported to dominate aquatic microbial mats from Antarctica (Jungblut et al., 2005; Strunecky et al., 2012). Representative isolates from this genus have been proposed previously as cost-effective options for industrial carotenoid production (Shukla and Kashyap, 2003), suggesting that FECB2 may hold the potential for this process.

FECB32 is a mixed culture isolated from an ancient travertine at Mammoth in YNP. Our analysis indicated that FECB32 contained 29 OTUs that each accounted for >0.1% of the reads generated (**Table 2**). Fifteen of these OTUs recruited >1% of all reads and four OTUs collectively accounted for ~60% of the reads generated from this co-culture (**Supplementary Table 5**). These four OTUs were classified as *Sphingopyxis alaskensis*, *Chelativorans* sp., and as members of the *Chitinophagaceae* and *Comamonadaceae* families, recruiting ~19, 13, 17, and 11% of the reads, respectively (**Supplementary Tables 3,4**). *S. alaskensis* is a Gram-negative bacterium found in relatively high abundance in oligotrophic regions of the ocean (Vancanneyt et al., 2001; Cavicchioli et al., 2003) and it has been studied in great detail as a model system for marine bacteria, specifically to understand microbial adaptation to cold or oligotrophic environments
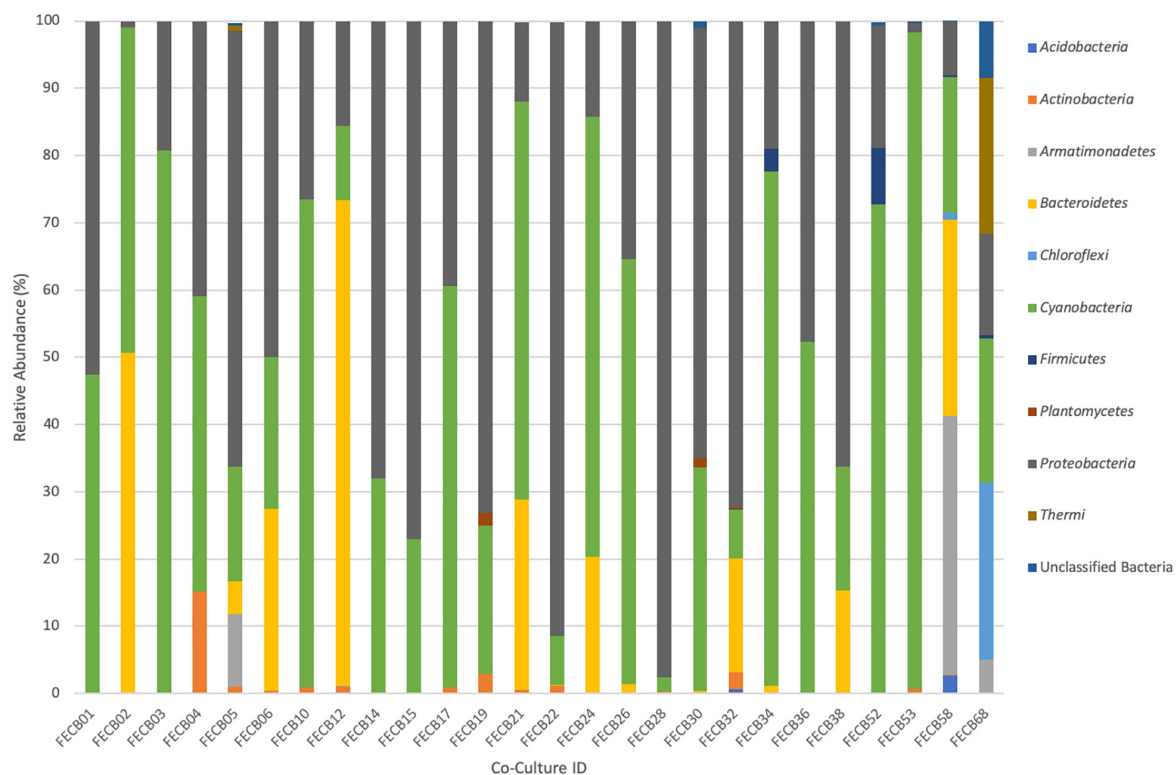
FIGURE 2 | Relative abundance of phyla associated with phototrophic co-cultures. 16S rRNA based community profile. Only phyla recruiting >0.1% of the reads in at least one of the co-cultures are shown.

(Lauro et al., 2009; Ting et al., 2010). The *Chitinophagaceae* family contains a wide phylogenetic diversity with many of its members being mesophilic. However, *Chitinophagaceae* have been reported to grow optimally at temperatures of 55°C and higher (Anders et al., 2014; Hanada et al., 2014).

Photosynthetic Co-cultures Containing the Deep-Branching Candidate Phylum Melainabacteria

Extreme environments similar to those on early Earth are often proposed to hold critical information about the historical progression of life on our planet and a niche that encompasses those physical stresses is the endolithic environment of rocks (Norris and Castenholz, 2006). Phylogenetic analysis of the heterotrophic population associated with FECB32, which was isolated from travertine deposited by hot springs in YNP, found that sequences from MLE-12 (OTU000109) recruited ~2% of the sample specific sequences (**Supplementary Table 3**). This rendered MLE-12, previously assigned to the deep-branching candidate phylum *Melainabacteria* (Di Rienzi et al., 2013), as the eleventh most abundant organism in this photosynthetic co-culture. It has been proposed previously that *Melainabacteria*, which is commonly found in aquatic habitats, separated from the cyanobacteria before the latter acquired photosynthetic capabilities (Di Rienzi et al., 2013). Hence FECB32 might be a particularly valuable co-culture to generate new insights into

the evolution of and relationship between the phylogenetically closely related *Cyanobacteria* and *Melainabacteria*.

Interestingly, OTU000109 was also detected in FECB36 and FECB38 (**Supplementary Table 3**), although at significantly lower abundance (<0.001%). FECB36 and FECB38 were similar to FECB32 in that they were isolated from sites in YNP. Interestingly, FECB32 and FECB38 cluster together (cluster IX) suggesting similar overall microbial community profiles, but separately from FECB36 (**Figure 1**). The only additional samples that contained OTUs classified as *Melainabacteria*, recruiting >0.1% of the generated reads, were FECB58 and FECB68 with ~0.9 and ~0.2% of their reads to this deeply branched phylum, respectively (**Supplementary Table 5**). It seems noteworthy that FECB58 and FECB68 were also isolated from hot springs and clustered closely together based on their overall microbiome composition (Clusters V and IV, respectively; **Figure 1**).

The McMurdo Dry Valley Lake System, a Physically Highly Stable Lacustrine System

The McMurdo Dry Valley (MDV) is one of the most extreme deserts on Earth, and although the importance of microbial communities for biogeochemical cycles of this region is widely accepted, the microbial ecology of the MDV remains poorly understood (Chan et al., 2013). FECB3, originating from a brackish pond on Bratina Island, was dominated by OTU000003,

TABLE 4 | Taxonomy relative abundance of dominant OTU identified in each co-culture.

Sample	Taxonomic assignment	% Reads recruited
FECB1	Cyanobacteria; Oscillatoriophycideae; Oscillatoriales; Phormidiaceae; Phormidium; pseudopriestleyi	47
FECB2	Bacteroidetes; Cytophagia; Cytophagales; Cytophagaceae; Hymenobacter	51
FECB3	Cyanobacteria; Oscillatoriophycideae; Oscillatoriales; Phormidiaceae; Phormidium; pseudopriestleyi	80
FECB4	Cyanobacteria; Oscillatoriophycideae; Oscillatoriales; Phormidiaceae; Planktothrix	44
FECB5	Proteobacteria; Alphaproteobacteria; Sphingomonadales; unclassified	40
FECB6	Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas	29
FECB10	Cyanobacteria; Nostocophycideae; Nostocales; Nostocaceae; unclassified	70
FECB12	Bacteroidetes; Saprospirae; Saprospirales; Chitinophagaceae; Sediminibacterium	72
FECB14	Proteobacteria; Alphaproteobacteria; Rhodospirillales; Rhodospirillacaea; Phaeospirillum; fulvum	67
FECB15	Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas	63
FECB17	Cyanobacteria; Nostocphycideae; Stigonematales; Rivulariaceae; Rivularia	59
FECB19	Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas	47
FECB21	Cyanobacteria; Nostocophycideae; Stigonematales; Rivulariaceae; Rivularia	58
FECB22	Proteobacteria; Alphabacteria; Caulobacterales; Caulobacteraceae; Mycoplana	40
FECB24	Cyanobacteria; Nostocphycideae; Nostocales; Nostocaceae; unclassified	65
FECB26	Cyanobacteria; Nostocophycideae; Nostocales; unclassified	63
FECB28	Proteobacteria; Gammaproteobacteria; Altermonadales; Alteromonadaceae; Marinobacter; hydrocarbonoclasticus	37
FECB30	Cyanobacteria; Oscillatoriophycideae; Oscillatoriales; Phormidiaceae; Geitlerinema	33
FECB32	Proteobacteria; Alphaproteobacteria; Sphingomonadales; Sphingomonadaceae; Sphingopyxis;alaskensis	19
FECB34	Cyanobacteria; Nostocophycideae; Nostocales; Nostocaceae; unclassified	76
FECB36	Cyanobacteria; Oscillatoriophycideae; unclassified	52
FECB38	Proteobacteria; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Stenotrophomonas; geniculata	32
FECB52	Cyanobacteria; unclassified	71
FECB53	Cyanobacteria; unclassified	91
FECB58	Armatimonadetes; OS-L; unclassified	38
FECB68	Chloroflexi; Chloroflexi; Chloroflexales; Chloroflexaceae; Chloroflexus	26

which recruited 80.3% of all reads (**Supplementary Table 5**). OTU000003 was classified as the cyanobacterium *Phormidium pseudopriestleyi*, previously reported to dominate microbial mats of the anoxic zone of Lake Fryxell, Antarctica (Jungblut et al., 2015). The second and third most abundant OTUs in FECB3 were OTU000015 and OTU000061, respectively (**Supplementary Table 5**). Both OTU000015 and OTU000061 were classified as Rhodobacteriaceae and recruited 9.2 and 8.2% of the reads generated for FECB3. Whereas a taxonomic classification of OTU000015 was not possible beyond the family level, OTU000061 was classified as *Paracoccus marcusii*, a Gram-negative organism that displays a bright orange color due to the synthesis of carotenoids such as astaxanthin (Harker et al., 1998).

Photosynthetic Co-cultures From Hunter's Hot Spring, Oregon

FECB58 and FECB68 were both isolated from Hunters Hot Spring in Oregon, United States and they shared similar microbial community members. Despite their similar community profile, abundances of the dominant OTUs associated with these two hot spring co-cultures were remarkably different. FECB58 was dominated by three OTUs (OTU000014, OTU000024, and OTU000033). OTU000014 was classified as OS-L, an uncultured representative of the phylum *Armatimonadetes*, OTU000024 which was classified as belonging to the *Bacteroidetes* phylum, and OTU000033 which was classified as *Thermosynechococcus*. These OTUs contributed 38, 29, and 20% of the reads generated from FECB58, respectively. Whereas OTU000014 recruited ∼4.9% of all reads generated from FECB68, representing the sixth most abundant OTU in the FECB68 community, OTU000024 and OTU000033 were only present at an abundance <0.0001% in FECB68 (**Supplementary Table 3**).

FECB68 was dominated by 6 OTUs (i.e., OTU000028, OTU000030, OTU000036, OTU000049, OTU000065, and OTU000014) recruiting ∼25.7, 23.1, 20.4, 14.3, 7.6, and 4.9% of

the reads, respectively. OTU000028 was classified as belonging to the genus *Chloroflexus*, whereas OTU000030 and OTU000036 were classified as representative of the genus *Meiothermus* and *Gloeobacter*, respectively. *Chloroflexus* is an anoxygenic phototrophic bacterium that grows at temperatures up to 70°C (Castenholz, 2015) and forms yellow-orange-greenish mats in association with cyanobacteria (Hanada, 2014). Members of the cyanobacterial genus *Gloeobacter* lack thylakoids, and have been proposed to host the earliest ancestors, or a missing link, in the cyanobacteria lineage (Saw et al., 2013). Thus, FECB68 offers a unique opportunity to investigate interspecies interaction between a member of these basal cyanobacteria and the thermophilic phototroph *Chloroflexus*, represented by OTU000028 in this co-culture. As outlined in a recent review by Castenholz (2015), Hunter's Hot Spring located in Oregon is one of the most studied hot springs in the world and a large repertoire of work has been conducted on this habitat over the last 40 years. However, most of this work was performed prior to the advent of recent molecular and -omics techniques.

Photosynthetic Co-cultures From Lignocellulosic Surfaces With Potential to Fix Nitrogen and Degrade Aromatic Compounds

FECB22 and FECB26 are mesophilic co-cultures collected from similar habitats (i.e., from tree bark and a wooden fence) from two locations (i.e., Hawaii and Bermuda) approximately 9,000 kilometers apart from each other (**Supplementary Figure 1** and **Table 1**). Diversity index calculation placed these two samples in the mid-range of the diversity spectrum of the 26 co-cultures analyzed for this study. The inverse Simpson and Shannon index was calculated at 4.46 and 2.08 for FECB22 and 2.29 and 1.32 for FECB26, respectively (**Table 2**). Within FECB22, 23 OTUs were identified as individually recruiting more than 0.1% of the generated reads. In contrast, FECB26 contained only 16 OTUs that recruited more than 0.1% of the reads each (**Supplementary Table 5**). FECB22, scraped from tree bark in Hawaii, was dominated by 11 OTUs, each recruiting (>1% of the reads. The most abundant OTU (OTU000017) was classified as a member of the Mycoplana, a genus that contains bacteria capable of aromatic compound degradation (Urakami et al., 1990), and it recruited 40.2% of the reads. OTU000042 (classified as *Rhizobium leguminosarum*), OTU000045 (classified as *Acetobacteraceae*), and OTU000072 (classified as *Cyanobacteria*), were the next most abundant OTUs, recruiting 17.1, 16.3, and 5.5% of the reads generated from FECB22, respectively. *Rhizobium leguminosarum* is a well-studied α-proteobacterium capable of N_2-fixation and "rhizobia" have been suggested repeatedly to facilitate more sustainable agricultural practices through their symbiosis with legumes, reducing the need for nitrogen fertilizer (Marek-Kozaczuk et al., 2017). It remains to be seen if OTU000042 provides N_2 to the other organisms in this co-culture or if it consumes all of the fixed N_2 itself. *Acetobacteraceae* are α-proteobacteria often associated with low pH environments and are known for their ability

to efficiently synthesize biological cellulose (Rozenberga et al., 2016; Semjonovs et al., 2017). Furthermore, *Acetobacteraceae* have been reported before as some of the dominant players in photosynthetic consortia during soil formation (Mapelli et al., 2011). It would be interesting to explore the agricultural and chemical potential of a minimalistic co-culture composed of the four OTUs (i.e., OTU000017, OTU000042, OTU000045, and OTU000072) that dominated FECB22, as they may combine the ability to degrade aromatic compounds and synthesize cellulose while removing nitrogen from the atmosphere. FECB26, on the other hand, was dominated by OTU000010, which recruited 63.2% of the reads generated and it was identified as an unclassified member of the *Nostocales*; a phylogenetic group known for their functional and morphological diversity. Members of the *Sphingomonadaceae* (i.e., OTU000041 and OTU000062), phototropic α-proteobacteria often found in high abundance in environments previously thought to support mostly the growth of cyanobacteria (Tahon and Willems, 2017), contributed to a total of 25.6% of the generated reads. Most interestingly, OTU000017 was also detected within FECB26 recruiting ~1.6% of the reads. It is possible that OTU000017 facilitates a metabolic reaction in which aromatic compounds typically associated with the decomposition of woody material under aerobic conditions are utilized.

CONCLUSION

Culture collections can provide easy access to biological samples without the need for extensive resources by the requesting individual, subsequently facilitating new studies and ultimately advancing our understanding of phylogenetic and functional biodiversity. While these collections present increased access to typically hard to acquire samples, there is lost diversity due to cultivation bias, but it remains to be understood exactly how prevalent and consistent the loss of diversity is sample to sample. Although care is taken to mimic the native environmental conditions of each sample in the cultivation process, there are real world factors that either cannot be mimicked in a lab setting or are unknown to researchers. More work is needed to assess this cultivation bias and to develop techniques to minimize the effects. Although some of the diversity of the original microbial community might have been lost due to a cultivation bias, the 16S rRNA based community fingerprints of the 26 photosynthetic co-cultures described here provide a first in-depth glimpse into the taxonomic and functional diversity of communities from extreme environments that were considered for a long time as too harsh to support the growth of complex microbial communities. The extreme conditions that are associated with the habitats from which these co-cultures were collected offer the unique opportunity to study the molecular mechanisms that support the growth of these extremophilic co-cultures and their role in global carbon and nitrogen cycling. Co-cultures from the CCMEE, and data presented here, also provide a first opening to enhance our understanding of the origin of oxygenic photosynthesis and aerobic respiration in *Cyanobacteria*, an area that is currently still poorly understood (Soo et al., 2017). Furthermore, an in-depth

understanding of these co-cultures holds the potential to discover novel microbial proteins that might render current agricultural, industrial and medical processes more economical and sustainable, for example by promoting or inhibiting plant and microbial growth.

The heterogeneity of the physical parameters reported for the sites where the samples presented in this work were collected, highlights a major challenge (i.e., standardization of protocols) associated with environmental samples and their corresponding metadata (i.e., data describing conditions at each sampling site), specifically when collected during independent sampling efforts. Fortunately, with recent advances in data technologies, the task of data acquisition and dissemination has become less of a challenge. In order to make the best use of these technologies defining a set of minimal information parameters to be recorded during the collection of an environmental sample is of great importance. Similar efforts have been successfully implemented by the Genomic Standards Consortium (GSC) for microbial genomes and metagenomes in the form of the "minimum information about a genome sequence" (MIGS) (Field et al., 2008) and are enforced when describing a novel microbial species (Kampfer et al., 2003).

The identification of Minimum Information about a Co-Culture Sample (MICCS) would be a significant step in standardizing sample acquisition and maintenance, increasing the value of current and future microbial samples collected from the environment. Developing MICCS and applying them to co-cultures currently available from existing culture depositories is beyond the scope of the work presented here, but we hope that the results presented here will contribute to the initiation of this process and stimulate broad involvement and support from the scientific community and various funding agencies.

In summary, we encourage the scientific community to take advantage of the CCMEE and the data we generated during this pilot study. Both data and samples from which these data were generated are publicly available from the CCMEE for further in-depth analyzes and investigations. Future work which might provide a more detailed picture of the microbe-microbe interactions in these co-cultures and their role in the global carbon and nitrogen cycle.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/genbank/, PRJNA401502.

AUTHOR CONTRIBUTIONS

CS, CB, RC, DC, MH, and ST wrote the manuscript. RC and MH designed the experiment. EH and MH performed the experiment. MB, TG, MH-S, EH, MH, NS, and ST generated the data. MB, CS, CB, MC, DC, JG, TG, MH-S, EH, MH, and NS analyzed the data. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.572131/full#supplementary-material

REFERENCES

Abed, R. M., Dobretsov, S., and Sudesh, K. (2009). Applications of cyanobacteria in biotechnology. *J. Appl. Microbiol.* 106, 1–12. doi: 10.1111/j.1365-2672.2008.03918.x

Al-Haj, L., Lui, Y. T., Abed, R. M., Gomaa, M. A., and Purton, S. (2016). Cyanobacteria as chassis for industrial biotechnology: progress and prospects. *Life (Basel)* 6:42. doi: 10.3390/life6040042

Allen, M. M., and Stanier, R. Y. (1968). Selective isolation of blue-green algae from water and soil. *J. Gen. Microbiol.* 51, 203–209. doi: 10.1099/00221287-51-2-203

Anders, H., Dunfield, P. F., Lagutin, K., Houghton, K. M., Power, J. F., Mackenzie, A. D., et al. (2014). Thermoflavifilum aggregans gen. nov., sp. nov., a thermophilic and slightly halophilic filamentous bacterium from the phylum *Bacteroidetes*. *Int. J. Syst. Evol. Microbiol.* 64(Pt 4), 1264–1270. doi: 10.1099/ijs.0.057463-0

Boundy-Mills, K., Hess, M., Bennett, A. R., Ryan, M., Kang, S., Nobles, D., et al. (2015). The United States culture collection network (USCCN): enhancing microbial genomics research through living microbe culture collections. *Appl. Environ. Microbiol.* 81, 5671–5674. doi: 10.1128/AEM.01176-15

Camacho, A., Garcia-Pichel, F., Vicente, E., and Castenholz, R. W. (1996). Adaptation to sulfide and to the underwater light field in three cyanobacterial

isolates from Lake Arcas (Spain). *FEMS Microbiol. Ecol.* 21, 293–301. doi: 10.1111/j.1574-6941.1996.tb00126.x

Castenholz, R. W. (1981). "Isolation and cultivation of thermophilic cyanobacteria," in *Isolation and Identification of Bacteria*, eds M. P. Starr, H. Stolp, H. G. Trüper, A. Balows, and H. G. Schlegel (Berlin: Springer), 236–246. doi: 10.1007/978-3-662-13187-9_11

Castenholz, R. W. (2015). Portrait of a geothermal spring, hunter's hot springs, oregon. *Life (Basel)* 5, 332–347. doi: 10.3390/life5010332

Cavicchioli, R., Ostrowski, M., Fegatella, F., Goodchild, A., and Guixa-Boixereu, N. (2003). Life under nutrient limitation in oligotrophic marine environments: an eco/physiological perspective of *Sphingopyxis alaskensis* (formerly *Sphingomonas alaskensis*). *Microb. Ecol.* 45, 203–217. doi: 10.1007/s00248-002-3008-6

Chan, Y., Van Nostrand, J. D., Zhou, J., Pointing, S. B., and Farrell, R. L. (2013). Functional ecology of an Antarctic Dry Valley. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8990–8995. doi: 10.1073/pnas.1300643110

Cubillos-Ruiz, A., Berta-Thompson, J. W., Becker, J. W., van der Donk, W. A., and Chisholm, S. W. (2017). Evolutionary radiation of lanthipeptides in marine cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 114, E5424–E5433. doi: 10.1073/pnas.1700990114

de-Bashan, L. E., Bashan, Y., Moreno, M., Lebsky, V. K., and Bustillos, J. J. (2002). Increased pigment and lipid content, lipid variety, and cell and population size of the microalgae *Chlorella* spp. when co-immobilized in alginate beads with the microalgae-growth-promoting bacterium *Azospirillum brasilense*. *Can. J. Microbiol.* 48, 514–521. doi: 10.1139/w02-051

Di Rienzi, S. C., Sharon, I., Wrighton, K. C., Koren, O., Hug, L. A., Thomas, B. C., et al. (2013). The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. *eLife* 2:e01102. doi: 10.7554/eLife.01102

Dillon, J. G., and Castenholz, R. W. (2003). The synthesis of the UV-screening pigment, scytonemin, and photosynthetic performance in isolates from closely related natural populations of cyanobacteria (*Calothrix* sp.). *Environ. Microbiol.* 5, 484–491. doi: 10.1046/j.1462-2920.2003.00436.x

Dillon, J. G., Tatsumi, C. M., Tandingan, P. G., and Castenholz, R. W. (2002). Effect of environmental factors on the synthesis of scytonemin, a UV-screening pigment, in a cyanobacterium (*Chroococcidiopsis* sp.). *Arch. Microbiol.* 177, 322–331. doi: 10.1007/s00203-001-0395-x

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Evans, J., Sheneman, L., and Foster, J. (2006). Relaxed neighbor joining: a fast distance-based phylogenetic tree construction method. *J. Mol. Evol.* 62, 785–792. doi: 10.1007/s00239-005-0176-2

Field, D., Garrity, G., Gray, T., Morrison, N., Selengut, J., Sterk, P., et al. (2008). The minimum information about a genome sequence (MIGS) specification. *Nat. Biotechnol.* 26, 541–547. doi: 10.1038/nbt1360

Formighieri, C., and Melis, A. (2016). Sustainable heterologous production of terpene hydrocarbons in cyanobacteria. *Photosynthesis Res.* 130, 123–135. doi: 10.1007/s11120-016-0233-2

Gajda, I., Greenman, J., Melhuish, C., and Ieropoulos, I. (2015). Self-sustainable electricity production from algae grown in a microbial fuel cell system. *Biomass Bioenergy* 82, 87–93. doi: 10.1016/j.biombioe.2015.05.017

Garcia-Pichel, F., Belnap, J., Neuer, S., and Schanz, F. (2003). Estimates of global cyanobacterial biomass and its distribution. *Algological Stud.* 109, 213–227. doi: 10.1127/1864-1318/2003/0109-0213

Grady, E. N., MacDonald, J., Liu, L., Richman, A., and Yuan, Z. C. (2016). Current knowledge and perspectives of *Paenibacillus*: a review. *Microb. Cell Fact.* 15:203. doi: 10.1186/s12934-016-0603-7

Hanada, S. (2014). "The phylum chloroflexi, the family *Chloroflexaceae*, and the related phototrophic families *Oscillochloridaceae* and *Roseiflexaceae*," in *Other Major Lineages of Bacteria and the Archaea*, eds E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin, Heidelberg: Springer Berlin Heidelberg), 515–532. doi: 10.1007/978-3-642-38954-2_165

Hanada, S., Tamaki, H., Nakamura, K., and Kamagata, Y. (2014). *Crenotalea thermophila* gen. nov., sp. nov., a member of the family *Chitinophagaceae*

isolated from a hot spring. *Int. J. Syst. Evol. Microbiol.* 64(Pt 4), 1359–1364. doi: 10.1099/ijs.0.058594-0

Harker, M., Hirschberg, J., and Oren, A. (1998). *Paracoccus marcusii* sp. nov., an orange gram-negative coccus. *Int. J. Syst. Bacteriol.* 48(Pt 2), 543–548. doi: 10.1099/00207713-48-2-543

Jungblut, A. D., Hawes, I., Mackey, T. J., Krusor, M., Doran, P. T., Sumner, D. Y., et al. (2015). Microbial mat communities along an oxygen gradient in a perennially ice-covered Antarctic lake. *Appl. Environ. Microbiol.* 82, 620–630. doi: 10.1128/AEM.02699-15

Jungblut, A.-D., Hawes, I., Mountfort, D., Hitzfeld, B., Dietrich, D. R., Burns, B. P., et al. (2005). Diversity within cyanobacterial mat communities in variable salinity meltwater ponds of McMurdo Ice Shelf, Antarctica. *Environ. Microbiol.* 7, 519–529. doi: 10.1111/j.1462-2920.2005.00717.x

Kampfer, P., Buczolits, S., Albrecht, A., Busse, H. J., and Stackebrandt, E. (2003). Towards a standardized format for the description of a novel species (of an established genus): *Ochrobactrum gallinifaecis* sp. nov. *Int. J. Syst. Evol. Microbiol.* 53(Pt 3), 893–896. doi: 10.1099/ijs.0.02710-0

Karavaiko, G. I., Bogdanova, T. I., Tourova, T. P., Kondrat'eva, T. F., Tsaplina, I. A., Egorova, M. A., et al. (2005). Reclassification of 'Sulfobacillus thermosulfidooxidans subsp. thermotolerans' strain K1 as *Alicyclobacillus tolerans* sp. nov. and *Sulfobacillus disulfidooxidans* Dufresne et al. 1996 as *Alicyclobacillus disulfidooxidans* comb. nov., and emended description of the genus *Alicyclobacillus*. *Int. J. Syst. Evol. Microbiol.* 55(Pt 2), 941–947. doi: 10.1099/ijs.0.63300-0

Karl, D., Letelier, R., Tupas, L., Dore, J., Christian, J., and Hebel, D. (1997). The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. *Nature* 388, 533–538. doi: 10.1038/41474

Kim, B. R., Shin, J., Guevarra, R., Lee, J. H., Kim, D. W., Seol, K. H., et al. (2017). Deciphering diversity indices for a better understanding of microbial communities. *J. Microbiol. Biotechnol.* 27, 2089–2093. doi: 10.4014/jmb.1709.09027

Kleigrewe, K., Gerwick, L., Sherman, D. H., and Gerwick, W. H. (2016). Unique marine derived cyanobacterial biosynthetic genes for chemical diversity. *Nat. Prod. Rep.* 33, 348–364. doi: 10.1039/c5np00097a

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* 79, 5112–5120. doi: 10.1128/AEM.01043-13

Lauro, F. M., McDougald, D., Thomas, T., Williams, T. J., Egan, S., Rice, S., et al. (2009). The genomic basis of trophic strategy in marine bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15527–15533. doi: 10.1073/pnas.0903507106

Letunic, I., and Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44, W242–W245. doi: 10.1093/nar/gkw290

Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P. J., et al. (2010). Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10430–10435. doi: 10.1073/pnas.0913677107

Mapelli, F., Marasco, R., Rizzi, A., Baldi, F., Ventura, S., Daffonchio, D., et al. (2011). Bacterial communities involved in soil formation and plant establishment triggered by pyrite bioweathering on arctic moraines. *Microb. Ecol.* 61, 438–447. doi: 10.1007/s00248-010-9758-7

Marek-Kozaczuk, M., Wdowiak-Wrobel, S., Kalita, M., Chernetskyy, M., Derylo, K., Tchorzewski, M., et al. (2017). Host-dependent symbiotic efficiency of *Rhizobium leguminosarum* bv. trifolii strains isolated from nodules of *Trifolium rubens*. *Antonie Van Leeuwenhoek* 110, 1729–1744. doi: 10.1007/s10482-017-0922-7

Marizcurrena, J. J., Morel, M. A., Brana, V., Morales, D., Martinez-Lopez, W., and Castro-Sowinski, S. (2017). Searching for novel photolyases in UVC-resistant Antarctic bacteria. *Extremophiles* 21, 409–418. doi: 10.1007/s00792-016-0914-y

McCluskey, K. (2017). A review of living collections with special emphasis on sustainability and its impact on research across multiple disciplines. *Biopreserv. Biobank* 15, 20–30. doi: 10.1089/bio.2016.0066

McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., DeSantis, T. Z., Probst, A., et al. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* 6, 610–618. doi: 10.1038/ismej.2011.139

Miao, R., Xie, H., and Lindblad, P. (2018). Enhancement of photosynthetic isobutanol production in engineered cells of *Synechocystis* PCC 6803. *Biotechnol. Biofuels* 11:267. doi: 10.1186/s13068-018-1268-8

Miller, S. R., and Castenholz, R. W. (2000). Evolution of thermotolerance in hot spring cyanobacteria of the genus *Synechococcus*. *Appl. Environ. Microbiol.* 66, 4222–4229. doi: 10.1128/aem.66.10.4222-4229.2000

Morris, E. K., Caruso, T., Buscot, F., Fischer, M., Hancock, C., Maier, T. S., et al. (2014). Choosing and using diversity indices: insights for ecological applications from the German biodiversity exploratories. *Ecol. Evol.* 4, 3514–3524. doi: 10.1002/ece3.1155

Nadeau, T.-L., and Castenholz, R. W. (2000). Characterization of psychrophilic oscillatorians (cyanobacteria) from antarctic meltwater ponds. *J. Phycol.* 36, 914–923. doi: 10.1046/j.1529-8817.2000.99201.x

Nadeau, T.-L., Milbrandt, E. C., and Castenholz, R. W. (2001). Evolutionary relationships of cultivated antarctic oscillatorians (cyanobacteria). *J. Phycol.* 37, 650–654. doi: 10.1046/j.1529-8817.2001.037004650.x

Nazina, T. N., Lebedeva, E. V., Poltaraus, A. B., Tourova, T. P., Grigoryan, A. A., Sokolova, D., et al. (2004). *Geobacillus gargensis* sp. nov., a novel thermophile from a hot spring, and the reclassification of Bacillus vulcani as *Geobacillus vulcani* comb. nov. *Int. J. Syst. Evol. Microbiol.* 54(Pt 6), 2019–2024. doi: 10.1099/ijs.0.02932-0

Norris, T. B., and Castenholz, R. W. (2006). Endolithic photosynthetic communities within ancient and recent travertine deposits in Yellowstone National Park. *FEMS Microbiol. Ecol.* 57, 470–483. doi: 10.1111/j.1574-6941.2006.00134.x

Oh, T. J., Han, S. R., Ahn, D. H., Park, H., and Kim, A. Y. (2016). Complete genome sequence of *Hymenobacter* sp. strain PAMC26554, an ionizing radiation-resistant bacterium isolated from an Antarctic lichen. *J. Biotechnol.* 227, 19–20. doi: 10.1016/j.jbiotec.2016.04.011

Padmaperuma, G., Kapoore, R. V., Gilmour, D. J., and Vaidyanathan, S. (2018). Microbial consortia: a critical look at microalgae co-cultures for enhanced biomanufacturing. *Crit. Rev. Biotechnol.* 38, 690–703. doi: 10.1080/07388551.2017.1390728

Pisciotta, J. M., Zou, Y., and Baskakov, I. V. (2010). Light-dependent electrogenic activity of cyanobacteria. *PLoS One* 5:e10821. doi: 10.1371/journal.pone.0010821

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596. doi: 10.1093/nar/gks1219

Rozenberga, L., Skute, M., Belkova, L., Sable, I., Vikele, L., Semjonovs, P., et al. (2016). Characterisation of films and nanopaper obtained from cellulose synthesised by *acetic acid bacteria*. *Carbohydr. Polym.* 144, 33–40. doi: 10.1016/j.carbpol.2016.02.025

Saw, J. H., Schatz, M., Brown, M. V., Kunkel, D. D., Foster, J. S., Shick, H., et al. (2013). Cultivation and complete genome sequencing of *Gloeobacter kilaueensis* sp. nov., from a lava cave in Kilauea Caldera, Hawai'i. *PLoS One* 8:e76376. doi: 10.1371/journal.pone.0076376

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

Semjonovs, P., Ruklisha, M., Paegle, L., Saka, M., Treimane, R., Skute, M., et al. (2017). Cellulose synthesis by *Komagataeibacter rhaeticus* strain P 1463 isolated from Kombucha. *Appl. Microbiol. Biotechnol.* 101, 1003–1012. doi: 10.1007/s00253-016-7761-8

Shaw, C., Brooke, C. G., Connolly, M. P., Garcia, J. A., Harmon-Smith, M., Shapiro, N., et al. (2020). Phototrophic co-cultures from extreme environments: community structure and potential value for fundamental and applied research. *bioRxiv* [Preprint]. doi: 10.1101/427211

Shukla, S. P., and Kashyap, A. K. (2003). An assessment of biopotential of three cyanobacterial isolates from Antarctic for carotenoid production. *Indian J. Biochem. Biophys.* 40, 362–366.

Simpson, E. H. (1949). Measurement of diversity. *Nature* 163, 688–688. doi: 10.1038/163688a0

Soo, R. M., Hemp, J., Parks, D. H., Fischer, W. W., and Hugenholtz, P. (2017). On the origins of oxygenic photosynthesis and aerobic respiration in Cyanobacteria. *Science* 355, 1436–1440. doi: 10.1126/science.aal3794

Strunecky, O., Elster, J., and Komarek, J. (2012). Molecular clock evidence for survival of Antarctic cyanobacteria (Oscillatoriales, *Phormidium autumnale*) from Paleozoic times. *FEMS Microbiol. Ecol.* 82, 482–490. doi: 10.1111/j.1574-6941.2012.01426.x

Subashchandrabose, S. R., Ramakrishnan, B., Megharaj, M., Venkateswarlu, K., and Naidu, R. (2011). Consortia of cyanobacteria/microalgae and bacteria: biotechnological potential. *Biotechnol. Adv.* 29, 896–907. doi: 10.1016/j.biotechadv.2011.07.009

Sudek, S., Haygood, M. G., Youssef, D. T., and Schmidt, E. W. (2006). Structure of trichamide, a cyclic peptide from the bloom-forming cyanobacterium *Trichodesmium erythraeum*, predicted from the genome sequence. *Appl. Environ. Microbiol.* 72, 4382–4387. doi: 10.1128/AEM.00380-06

Tahon, G., and Willems, A. (2017). Isolation and characterization of aerobic anoxygenic phototrophs from exposed soils from the Sor Rondane Mountains, East Antarctica. *Syst. Appl. Microbiol.* 40, 357–369. doi: 10.1016/j.syapm.2017.05.007

Ting, L., Williams, T. J., Cowley, M. J., Lauro, F. M., Guilhaus, M., Raftery, M. J., et al. (2010). Cold adaptation in the marine bacterium, *Sphingopyxis alaskensis*, assessed using quantitative proteomics. *Environ. Microbiol.* 12, 2658–2676. doi: 10.1111/j.1462-2920.2010.02235.x

Toplin, J. A., Norris, T. B., Lehr, C. R., McDermott, T. R., and Castenholz, R. W. (2008). Biogeographic and phylogenetic diversity of thermoacidophilic cyanidiales in Yellowstone National Park, Japan, and New Zealand. *Appl. Environ. Microbiol.* 74, 2822–2833. doi: 10.1128/AEM.02741-07

Urakami, T., Oyanagi, H., Araki, H., Suzuki, K.-I., and Komagata, K. (1990). Recharacterization and emended description of the genus *Mycoplana* and description of two new species, *Mycoplana ramosa* and *Mycoplana segnis*. *Int. J. Syst. Evol. Microbiol.* 40, 434–442. doi: 10.1099/00207713-40-4-434

Vancanneyt, M., Schut, F., Snauwaert, C., Goris, J., Swings, J., and Gottschal, J. C. (2001). *Sphingomonas alaskensis* sp. nov., a dominant bacterium from a marine oligotrophic environment. *Int. J. Syst. Evol. Microbiol.* 51(Pt 1), 73–79. doi: 10.1099/00207713-51-1-73

Wang, R., Xue, S., Zhang, D., Zhang, Q., Wen, S., Kong, D., et al. (2015). Construction and characteristics of artificial consortia of *Scenedesmus obliquus*-bacteria for *S. obliquus* growth and lipid production. *Algal Res.* 12, 436–445. doi: 10.1016/j.algal.2015.10.002

Whitton, B. A., and Potts, M. (2000). *The Ecology of Cyanobacteria: their Diversity in time and Space*. Alphen aan den Rijn: Kluwer Academic.

Woo, H. M., and Lee, H. J. (2017). Toward solar biodiesel production from CO2 using engineered cyanobacteria. *FEMS Microbiol. Lett.* 364:fnx066. doi: 10.1093/femsle/fnx066

Yen, H.-W., Chen, P.-W., and Chen, L.-J. (2015). The synergistic effects for the co-cultivation of oleaginous yeast-*Rhodotorula glutinis* and microalgae-*Scenedesmus obliquus* on the biomass and total lipids accumulation. *Bioresource Technol.* 184, 148–152. doi: 10.1016/j.biortech.2014.09.113

Zarzycki, J., Axen, S. D., Kinney, J. N., and Kerfeld, C. A. (2013). Cyanobacterial-based approaches to improving photosynthesis in plants. *J. Exp. Bot.* 64, 787–798. doi: 10.1093/jxb/ers294

Zhao, J., Li, X. F., Ren, Y. P., Wang, X. H., and Jian, C. (2012). Electricity generation from Taihu Lake cyanobacteria by sediment microbial fuel cells. *J. Chem. Technol. Biotechnol.* 87, 1567–1573. doi: 10.1002/jctb.3794

Evaluation and Comparison of the Efficiency of Transcription Terminators in Different Cyanobacterial Species

Grant A. R. Gale[1,2,3], Baojun Wang[2,3] and Alistair J. McCormick[1,2]*

[1] School of Biological Sciences, Institute of Molecular Plant Sciences, University of Edinburgh, Edinburgh, United Kingdom, [2] Centre for Synthetic and Systems Biology, University of Edinburgh, Edinburgh, United Kingdom, [3] School of Biological Sciences, Institute of Quantitative Biology, Biochemistry and Biotechnology, University of Edinburgh, Edinburgh, United Kingdom

Cyanobacteria utilize sunlight to convert carbon dioxide into a wide variety of secondary metabolites and show great potential for green biotechnology applications. Although cyanobacterial synthetic biology is less mature than for other heterotrophic model organisms, there are now a range of molecular tools available to modulate and control gene expression. One area of gene regulation that still lags behind other model organisms is the modulation of gene transcription, particularly transcription termination. A vast number of intrinsic transcription terminators are now available in heterotrophs, but only a small number have been investigated in cyanobacteria. As artificial gene expression systems become larger and more complex, with short stretches of DNA harboring strong promoters and multiple gene expression cassettes, the need to stop transcription efficiently and insulate downstream regions from unwanted interference is becoming more important. In this study, we adapted a dual reporter tool for use with the CyanoGate MoClo Assembly system that can quantify and compare the efficiency of terminator sequences within and between different species. We characterized 34 intrinsic terminators in *Escherichia coli*, *Synechocystis* sp. PCC 6803, and *Synechococcus elongatus* UTEX 2973 and observed significant differences in termination efficiencies. However, we also identified five terminators with termination efficiencies of >96% in all three species, indicating that some terminators can behave consistently in both heterotrophic species and cyanobacteria.

Keywords: CyanoGate, *Escherichia coli*, Golden Gate, intrinsic terminator, MoClo, *Synechococcus elongatus* UTEX 2973, *Synechocystis* sp. PCC 6803, synthetic biology

INTRODUCTION

Cyanobacteria comprises a large and diverse phylum of photoautotrophic bacteria that can capture and convert inorganic carbon (e.g., CO_2) into a wide variety of secondary metabolites (Huang and Zimba, 2019). Many cyanobacterial species are genetically tractable and show great potential for green biotechnology applications, such as the sustainable production of biofuels and high value biomolecules (Lin et al., 2017; Knoot et al., 2018; Eungrasamee et al., 2019; Lin and Pakrasi, 2019; Włodarczyk et al., 2019). Much of the recent progress in engineering cyanobacteria has been driven by the uptake of synthetic biology approaches. One major aim of cyanobacterial synthetic biology is the development of new tools and strategies to facilitate stringent and precise control of gene

expression. A wide variety of new molecular tools and genetic parts to tune gene expression are now available for use by the research community (Englund et al., 2016; Kim et al., 2017; Ferreira et al., 2018; Kelly et al., 2018; Vasudevan et al., 2019; Yao et al., 2020). The increase in availability of well-characterized genetic parts has allowed rational design, a core process to the synthetic biology paradigm, to be more routinely employed in the engineering of new cyanobacterial strains. Nevertheless, the majority of synthetic biology work in cyanobacteria has thus far concentrated on characterizing genetic elements that control gene transcription (e.g., promoters, CRISPRi) or translation modulation (e.g., ribosomal binding sites (RBS), riboswitches, small RNAs) (Huang and Lindblad, 2013; Camsund et al., 2014; Ma et al., 2014; Immethun et al., 2017; Kelly et al., 2018; Sun et al., 2018; Behle et al., 2020; Yao et al., 2020). Transcription terminators are also key transcriptional control elements, but far fewer studies have examined their roles in regulating gene expression in cyanobacteria.

The rational design of efficient gene expression cassettes (and more advanced gene circuits) requires the use of genetic parts with well-characterized and predictable function (Moser et al., 2018). For instance, strong terminators attenuate transcription and isolate downstream genetic sequences, which can prevent interference and disruption of function from unwanted transcriptional readthrough (Kelly et al., 2019). This is particularly important when considering synthetic gene constructs, where several gene expression cassettes driven by strong promoters may occupy a short stretch of DNA. Furthermore, many prokaryotes (including cyanobacteria) are prone to homologous recombination. Homologous regions as small as 23–27 bp have been demonstrated to lead to recombination in *Escherichia coli*, so multiple distinct terminators are generally preferable for multi-gene expression systems and gene circuits (Shen and Huang, 1986; Sleight et al., 2010; Chen et al., 2013). As with other genetic parts, an understanding of terminator performance and robustness between species is also important. Promoters have been shown to drive gene expression differently in cyanobacteria compared to heterotrophic species (e.g., *Escherichia coli*) and between cyanobacterial species (Camsund et al., 2014; Vasudevan et al., 2019). In contrast, potential differences in behavior between cyanobacterial species has not yet been investigated for transcription terminators.

In prokaryotes, transcription is terminated by two distinct terminator types: (i) Rho-dependent terminators that rely on a Rho transcription factor, and (ii) Rho-independent, or intrinsic terminators, which do not require a transcription factor. In *E. coli*, approximately 20% of terminators are Rho-dependent (Peters et al., 2009). However, Rho transcription factors appear to be absent in cyanobacteria, such that all transcription termination events are thought to rely on intrinsic termination (Vijayan et al., 2011). Intrinsic terminators are defined by a sequence motif that forms a hairpin loop secondary structure in the nascent RNA transcript. The hairpin loop is comprised of a GC-rich stem (8–12 nucleotides) (nt) and a loop (3–6 nt). Upstream of the hairpin loop is an adenine-rich region (the

A-tract) typically 6–8 nt in length, while downstream is a uracil-rich region of 7–12 nt in length (the U-tract). Intrinsic termination depends upon the differential binding affinities between nucleotides. The interaction between U and A is weak, such that transcription of the U-tract results in a pause in transcription that allows the hairpin loop to form. The presence of the hairpin loop in the RNA polymerase (RNAP) exit channel, causes a ratcheting action and subsequent disruption of RNA-DNA binding. This leads to dissociation of RNAP from the DNA template and the subsequent release of the nascent RNA transcript (Wilson and Von Hippel, 1995; Herbert et al., 2008; Peters et al., 2011). In *E. coli*, many terminators have been assessed for termination efficiency (TE), which is typically calculated as a percentage estimate of the RNAP transcription elongation complexes prevented from continuing transcription passed a given sequence (i.e., a terminator) (Cambray et al., 2013; Chen et al., 2013). Importantly, a "no terminator" control was included to determine a normalized value for TE in those studies.

Characterization studies of terminators in cyanobacteria are currently limited to the model species *Synechocystis* sp. PCC 6803 (PCC 6803). Liu and Pakrasi (2018) evaluated the relative strengths of seven native terminators using a dual fluorescent reporter system similar to that used by Chen et al. (2013). More recently, Kelly et al. (2019) evaluated 19 synthetic and heterologous intrinsic terminators ported from *E. coli*, with the aim of identifying terminators able to insulate a specific genomic locus in PCC 6803 from native promoter readthrough originating from upstream of the insertion site. Each terminator sequence was inserted between the transcription start site (TSS) and RBS of an inducible promoter driving *YFP*, and following induction, twelve terminators were shown to efficiently block transcription indicating a potential efficiency of nearly100%. These studies have provided valuable insights into terminator function in PCC 6803. But if comparisons in performance between different strains are to be achieved, a normalized quantitative parameter, such as TE, should be calculated.

In this study we assembled a set of 34 intrinsic terminators from PCC 6803, and *E. coli* and synthetic libraries that have previously demonstrated a wide range of TE values in *E. coli* (Chen et al., 2013). We re-designed an established dual fluorescent reporter system to be compatible with the CyanoGate MoClo Assembly system, which allowed for increased cloning throughput (Liu and Pakrasi, 2018; Vasudevan et al., 2019). Importantly, all assays included a "no terminator" control vector as a reference to calculate a normalized TE value for each terminator, such that the TE values could be compared between different experiments and species irrespective of the instrument or gain settings used. We first validated and benchmarked our testing system by comparing TE values from the literature with our results in *E. coli*. Then we tested the performance of the terminators in two different cyanobacterial species: PCC 6803 and the recently described high-light tolerant *Synechococcus elongatus* UTEX 2973 (UTEX 2973) (Williams, 1988; Yu et al., 2015).

MATERIALS AND METHODS

Cyanobacterial Culture Conditions

The *Synechocystis* sp. PCC 6803 glucose tolerant (GT) strain (obtained from the Lea-Smith lab at the University of East-Anglia, United Kingdom) (Zavøel et al., 2017) and UTEX 2973 were maintained on 1.5% (w/v) agar plates containing BG11 medium (Lea-Smith et al., 2016). Liquid cultures were grown in BG11 (supplemented with 10 mM $NaHCO_3$) in 100 ml Erlenmeyer flasks. Liquid cultures were shaken at 100 rpm and aerated with filter-sterilized, water-saturated air. PCC 6803 and UTEX 2973 transconjugants were cultured in BG11 medium and on BG11 agar plates, supplemented with 50 μg/ml kanamycin (BG11 + Kan50). Strains were grown under continuous light with PCC 6803 grown at 30°C, 100 μmol photons $m^{-2} s^{-1}$ and UTEX 2973 at 40°C, 300 μmol photons $m^{-2} s^{-1}$ in a Multitron Pro incubator supplied with warm white LED lighting (Infors HT).

Vector Construction and Parts Assembly

All cloning was performed in OneShot TOP10 *E. coli* cells. Transformed cells were cultured in LB medium and on 1.5% (w/v) LB agar plates supplemented with either 100 μg/ml spectinomycin or 50 μg/ml kanamycin as required. *E. coli* strain MC1061 was cultured in LB medium supplemented with 100 μg/ml ampicillin and 25 μg/ml chloramphenicol. All *E. coli* strains were grown at 37°C with shaking at 225 rpm.

pPMQAK1-T (pCAT.000) from the CyanoGate toolkit was modified to generate pDUOTK1-L1 (pCA1.332, Addgene vector ID 162351)[1] (**Supplementary Information S1**) (Vasudevan et al., 2019). To assemble pDUOTK1-L1, pPMQAK1-T was first digested with *Bpi*I and *Bsa*I (Thermo Fisher Scientific). The linearized backbone was gel purified using a Monarch DNA Gel Extraction Kit (NEB). Sequences encoding P_{trc10}-*eYFP* from the CyanoGate vector pCAT.262, the *LacZ* expression cassette from the Plant MoClo level 1 acceptor vector pICH47732 and *mTagBFP*-T_{rrnB} (from an available vector containing BBa_K592100)[2] fused at the 5′ end to the RBS-associated sequence used by Chen et al. (2013) (BBa_B0034) were amplified using Q5 High-Fidelity DNA Polymerase (NEB) (**Supplementary Table S1**). Finally, the three amplicons and the linearized pPMQAK1-T backbone were assembled together using Golden Gate assembly (Vasudevan et al., 2019). pDUOTK1-L1 contains *Bsa*I restriction sites flanking *LacZ* that generate overhangs GCTT-CGCT, such that level 0 terminator parts can be assembled directly and screened using blue-white selection.

Terminator parts were generated by overlap extension PCR using two synthesized oligonucleotides (Integrated DNA Technology) (**Supplementary Table S1**), and the resulting amplicons were assembled into the level 0 (3U + Ter) acceptor vector pICH41276 (**Supplementary Information S1**) (Engler et al., 2014). New level 0 terminator parts and existing parts from CyanoGate toolkit (Addgene Kit #1000000146)[3] were

assembled into pDUOTK1-L1 to give vectors pC1.342 to pC1.375 (**Supplementary Table S2**).

Two "no terminator" control vectors were generated to determine 0% TE (i.e., the maximum ratio of mTagBFP relative to eYFP). pC1.376 was assembled as pDUOTK1-L1 above, but without inclusion of *LacZ* (**Supplementary Information S1**). For pC1.377, the spacer sequence rd1.2 (5′-cgcccccggaggctttcccggggcaaatca-3′) from Cambray et al. (2013) was generated using overlap extension PCR (**Supplementary Table S1**), and the PCR product was assembled into pDUOTK1-L1 using Golden Gate assembly.

Cyanobacterial Conjugation

Genetic modification by conjugation in PCC 6803 and UTEX 2973 was facilitated by *E. coli* strain MC1061 carrying the mobilizer vector pRK24[4] and helper vector pRL528[5] (Tsinoremas et al., 1994; Gale et al., 2019). Conjugal transfer was performed as in Gale et al. (2019).

Fluorescence Assays

To measure fluorescence in *E. coli*, transformants were first inoculated into 5 ml LB medium supplemented with 50 μg/ml kanamycin and grown overnight at 37°C with constant shaking at 225 rpm. To initiate the assay, overnight cultures were diluted 1:1000 into a black 96 well flat bottom plate (F-Bottom (Chimney Well) μCLEAR®, Greiner Bio-One) containing fresh LB medium supplemented with 50 μg/ml kanamycin to a final volume of 200 μl. The plates were incubated at 37°C with constant shaking at 600 rpm and culture density (OD_{600}) was measured hourly using a FLUOstar OMEGA microplate reader (BMG Labtech). At early exponential phase (*ca.* 4.5 h following inoculation), eYFP and mTagBFP fluorescence levels were measured for individual cells by flow cytometry (minimum 10,000 cells per culture) with a FACSCanto II with HTS Flow Cytometer (Becton Dickinson). Cells were gated using forward and side scatter. Median eYFP and mTagBFP fluorescence levels were calculated from excitation/emission wavelengths 488 nm/530/30 nm and 407 nm/450/50 nm, respectively. An "empty" pPMQAK1-T vector (i.e., with no eYFP or mTagBFP expression cassettes) was included as a base line control. Fluorescence values for the latter control were subtracted from transconjugant strain measurements.

To measure fluorescence in cyanobacteria, PCC 6803 or UTEX 2973 transconjugants maintained on BG11 + Kan50 agar plates were first inoculated into 10 ml BG11 + Kan50 medium and grown for 2–3 days to OD_{750} ∼1.0. To initiate the assay, the seed cultures were diluted to a starting OD_{750} of 0.2 in 24-well plates (Costar Corning Incorporated) containing fresh BG11 + Kan50 medium to a final volume of 2 ml. Cultures were grown for three days under culturing conditions and high humidity (95%) to avoid evaporation. eYFP and mTagBFP fluorescence were measured by flow cytometry for individual cells (minimum 10,000 cells per culture) with an LSRFortessa SORP with HTS Flow Cytometer (Becton Dickinson). Cells were

gated using forward and side scatter. Median eYFP and mTagBFP fluorescence levels were calculated from excitation/emission wavelengths 488 nm/515–545 nm and 407 nm/ 425–475 nm, respectively. As above, a base line control was included for each species.

Calculations for Termination Efficiency

TE was calculated as a percentage from the ratio of the mTagBFP fluorescence signal downstream of the terminator to the eYFP fluorescence signal upstream relative to a control containing no terminator between fluorescent reporters:

$$\Delta Term_0 = \frac{BFP_0}{YFP_0} \tag{1}$$

Where BFP_0 and YFP_0 are the mTagBFP and eYFP fluorescence signals, respectively, of the strain containing either pCA1.376 or pCA1.377.

$$TE = 100 - \left(\frac{BFP_{Term}}{YFP_{Term}} \times \frac{1}{\Delta Term_0} \times 100 \right) \tag{2}$$

Where BFP_{Term} and YFP_{Term} are the mTagBFP and eYFP fluorescence signals, respectively, of a strain carrying a given level 1 terminator vector (**Supplementary Table S2**).

Statistical Analysis

Significant differences between sample groups were assessed by one-way ANOVA followed by Tukey's honest significant difference (HSD) *post-hoc* test using GraphPad Prism (version. 8.4.2).

Estimation of Gibbs Free Energy

Estimated Gibbs free energy values were generated using mFold v3.0[6] (Zuker, 2003). Free energy values were calculated without adjustment of the standard parameters, which included a fixed temperature of 37°C.

RESULTS

Generating a Screening System for Level 0 Terminator Parts

The RSF1010-based level T acceptor vector pPMQAK1-T from the CyanoGate toolkit was modified to generate the new level 1 acceptor vector pDUOTK1-L1 for terminator screening (**Figure 1A** and **Supplementary Information S1**) (Vasudevan et al., 2019). pDUOTK1-L1 comprises a dual fluorescent reporter system with eYFP and mTagBFP, similar to that in Liu and Pakrasi (2018). Terminators can be assembled as level 0 parts into pDUOTK1-L1 using Golden Gate assembly (**Figure 1B**), while the RSF1010 origin of replication allows for screening in a wide range of species (Mermet-Bouvier et al., 1993).

We compiled a library of 34 level 0 vectors containing intrinsic transcription terminators (**Table 1** and **Figure 1C**), and then

[6]http://unafold.rna.albany.edu/?q=mfold

assembled these into pDUOTK1-L1 (**Supplementary Table S2**). In order to maximize potential orthogonality with terminators in cyanobacterial genomes, we primarily targeted heterologous terminator sequences. The library included 22 native terminators from *E. coli* and eight synthetic terminators based on *E. coli* sequences that have been previously characterized in *E. coli* (Chen et al., 2013). We also included T_{rrnB} (i.e., T_{rrnB} from *E. coli* and the T7 viral terminator in tandem (Vasudevan et al., 2019)) and the pSB1AK3 terminator ($T_{pSB1AK3}$) that was derived from the *E. coli* ribosomal RNA rrnC operon and is used in several BioBricks vectors, including pPMQAK1, to flank the cloning site (Huang et al., 2010). From PCC 6803, the terminator of the highly expressed D1 subunit of photosystem II was included (T_{psbA2}), as we expected it to have a high efficiency of termination. In contrast, T_{psaB} was included as a potentially low efficiency terminator based on previous work (Liu and Pakrasi, 2018). Two "no terminator" control vectors, pC1.376 and pC1.377, were assembled based on sequences used in previous *E. coli* studies (Cambray et al., 2013; Chen et al., 2013). In pC1.376, eYFP and mTagBFP were separated only by an RBS-associated sequence, while pCA1.377 included a spacer sequence reported to be inert (i.e., free from promoter or terminator activity in *E. coli*) (**Supplementary Information S1**).

Validation of the Dual Reporter Testing System in *E. coli*

We first assessed the dual fluorescent reporter system in *E. coli* by generating TE values for each terminator and compared these to the data reported by Chen et al. (2013) (**Figure 2A**). Terminator strength (TS) values reported by Chen et al. (2013) were converted to a more commonly reported TE (**Supplementary Table S3**; Hess and Graham, 1990; Yager and von Hippel, 1991; Cambray et al., 2013; Mairhofer et al., 2015).

E. coli cultures measured at early exponential growth phase had similar levels of eYFP fluorescence across different strains with an average value of 7034 ± 134 arbitrary units (a.u.) (**Supplementary Figure S1**). In contrast, the strains showed a wide range of mTagBFP fluorescence values from 1.3 ± 3.4 a.u. to 9094 ± 446 a.u. Both eYFP and mTagBFP fluorescence values showed a unimodal and narrow distribution (**Supplementary Figure S2**). As expected, the two "no terminator" controls pC1.376 and pC1.377 produced the highest mTagBFP fluorescence values. Previous reports have indicated that translation efficiency is dependent on the length of the transcript (Lim et al., 2011), so we checked if eYFP levels might be decreased in the "no terminator" controls compared to plasmid with terminators. However, we observed no significant differences in eYFP levels between different plasmids, indicating that efficiency of eYFP translation was not reduced in either "no terminator" controls (**Supplementary Figure S1B**). The mTagBFP:eYFP ratio (i.e., Equation 1) for pC1.376 was 22% higher than for pC1.377, which indicated that pC1.376 produced more transcripts containing both mTagBFP and eYFP. Thus, we decided to use pC1.376 for all TE calculations in this study.

Sixteen terminators had TE values of >95% in *E. coli* (**Figure 2A** and **Supplementary Table S3**), with $T_{L3S2P21}$ and
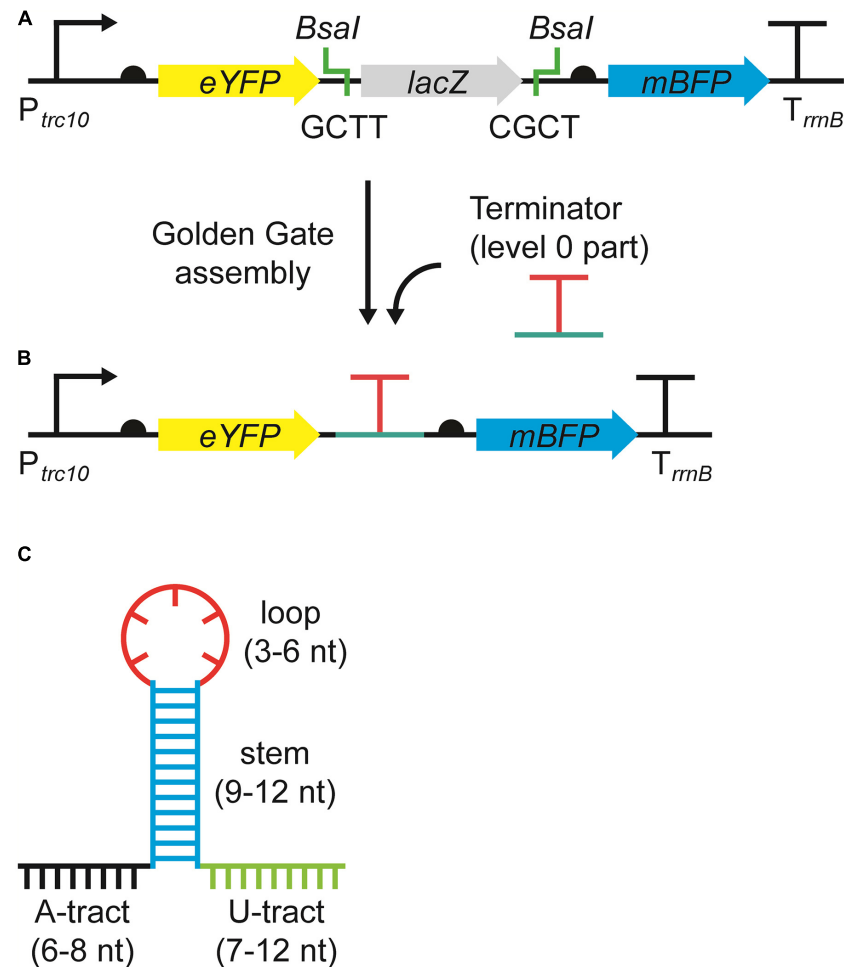
FIGURE 1 | The dual fluorescence reporter system for screening terminators. **(A)** The acceptor vector pDUOTK1-L1 contains two *BsaI* sites that generate 4 nucleotide (nt) overhangs (i.e., GCTT and CGCT following restriction, which are compatible with standard level 0 terminator parts (Engler et al., 2014). **(B)** Following a level 1 Golden Gate assembly reaction (Vasudevan et al., 2019), the level 0 terminator part is inserted between eYFP and mTagBFP and the dual fluorescent reporter system is formed, which can then be used to evaluate termination efficiency (TE). The reporter system is driven by the strong promoter P_{trc10} and is terminated by the terminator T_{rrnB}. Ribosome binding sites (half circles) are indicated (see **Supplementary Information S1** for sequence details). **(C)** Example of an intrinsic terminator structure and nt sequence, comprised of an adenine rich region (A-tract) (black), followed by a G-C rich stem (blue), a hairpin loop (red), and a uracil rich region (U-tract) (green).

T_{Bba_B0011} producing the highest (99.9%) and lowest values (40.8%), respectively. TE values for both PCC 6803 terminators were relatively low in *E. coli* (*ca.* 60%). Overall, the terminator library demonstrated a corresponding 10-fold change reduction in normalized downstream reporter expression (**Figure 2B**). We then compared the TE values for 30 native *E. coli* and synthetic terminators with those also reported in Chen et al. (2013) and observed a reasonable correlation (coefficient of determination (R^2) = 0.78), with 19 of the observed TE values differing by less than 5% (**Figure 2C**). The latter included 14 of the 16 strongest terminators with TE values of >95%. Similarly, the three weakest terminators (T_{Bba_B0011}, $T_{ECK120010842}$, and $T_{ECK120010820}$) were the same in both data sets. Six terminators showed a greater difference in TE values (i.e., 12–26%), which comprised four native *E. coli* terminators ($T_{ECK120030798}$, $T_{ECK120010820}$, T_{Bba_B0011}, and T_{Bba_B0061}) and two synthetic

terminators ($T_{L3S1P22}$ and $T_{L3S1P13}$). These variations may have been due to differences in experimental setup (e.g., the vector, origin of replication (ori) and reporter genes) and the different strain of *E. coli* used, as significant differences in the behavior of some terminators has been reported between different *E. coli* strains (Kelly et al., 2019).
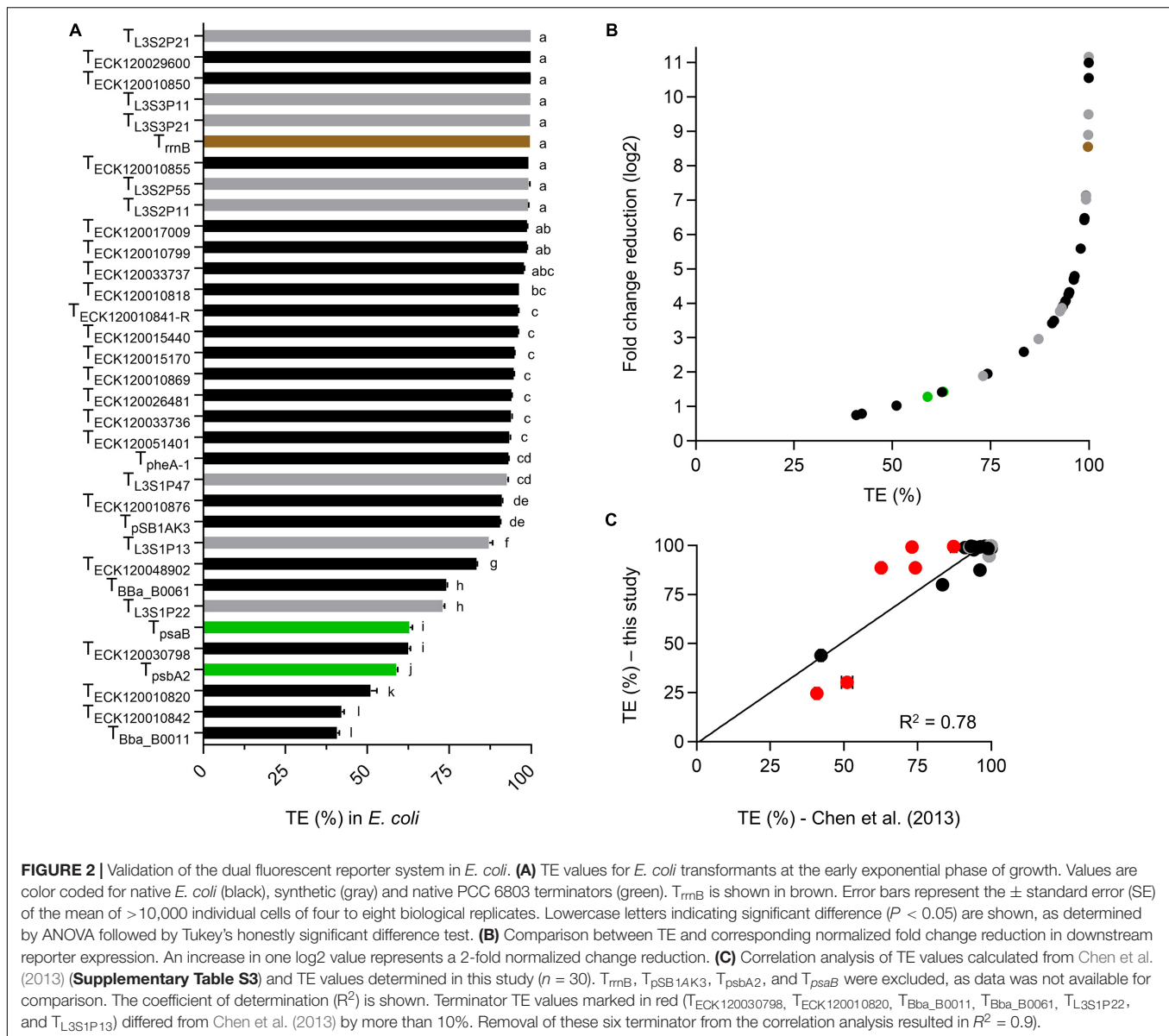
Performance of the Terminator Library in *Synechocystis* sp. PCC 6803

We next evaluated the terminator library in PCC 6803. Due to the slower growth rates of PCC 6803 compared to *E. coli* (**Supplementary Figure S3A**), we measured fluorescence levels at 24, 48, and 72 h (**Supplementary Figure S3B**). The cyanobacterial strains grew at comparable rates and the majority expressed eYFP at similar levels between strains at each time

TABLE 1 | List of terminators used in this study.

Vector ID	Part name	ΔG_A (kcal/mol)	Length (bp)	Terminator sequence	Origin	Reference
pC0.291	$T_{L3S2P21}$	−11.0	61	CTCGGTACCAAATTCCAGAAAAGAGAGGCCTCCCGAAAGGGGGGGCCTTTTTTCGTTTTGGTCC		
pC0.292	$T_{L3S2P11}$	−11.0	57	CTCGGTACCAAATTCCAGAAAAGAGACGCTTTCGAGCGTCTTTTTTCGTTTTGGTCC		
pC0.293	$T_{L3S2P55}$	−2.8	57	CTCGGTACCAAAGACGAACAATAAGACGCTGAAAAGCGTCTTTTTTCGTTTTGGTCC		
pC0.294	$T_{L3S3P21}$	−4.1	53	CCAATTATTGAAGGCCTCCCTAACGGGGGGCCTTTTTTTGTTTCTGGTCTCCC	Synthetic	Chen et al., 2013
pC0.295	$T_{L3S1P13}$	−2.8	51	GACGAACAATAAGGCCTCCCTAACGGGGGGCCTTTTTTATTGATAACAAAA		
pC0.296	$T_{L3S3P11}$	−4.2	47	CCAATTATTGAACACCCTTCGGGGTGTTTTTTTGTTTCTGGTCTCCC		
pC0.306	$T_{L3S1P22}$	−2.8	48	GACGAACAATAAGGCCGCAAATCGCGGCCTTTTTTATTGATAACAAAA		
pC0.307	$T_{L3S1P47}$	−8.4	52	TTTTCGAAAAAAGGCCTCCCAAATCGGGGGGCCTTTTTTTATAGCAACAAAA		
pC0.066	T_{pheA-1}	−2.8	52	GACGAACAATAAGGCCTCCCAAATCGGGGGGCCTTTTTTATTGATAACAAAA		
pC0.068	$T_{ECK120010850}$	−4.4	45	AGTTAACCAAAAAGGGGGGGATTTTATCTCCCCTTTAATTTTTCCT		
pC0.069	$T_{ECK120026481}$	−6.3	54	TACCACCGTCAAAAAAAACGGCGCTTTTTAGCGCCGTTTTTATTTTTCAACCTT		
pC0.072	$T_{ECK120010842}$	−2.5	47	CCGACGTAAAAAGACGGTAAGTATCGCTTTCAGTCTTATGAATATCG		
pC0.074	$T_{ECK120048902}$	−7.9	36	GCGTAAAAAAGCACCTTTTTAGGTGCTTTTTTGTGG	E. coli	Chen et al., 2013;
pC0.062	T_{Bba_B0011}	−5.5	46	AGAGAATATAAAAGCCAGATTATTAATCCGGCTTTTTTATTATTT		Vasudevan et al., 2019
pC0.064	$T_{ECK120010820}$	−5.3	33	CTAAGCGTTGTCCCCAGTGGGGATGTGACGAAG		
pC0.070	T_{Bba_B0061}	−13.1	31	AAGTCAAAAGCCTCCGGTCGGAGGCTTTTGACTTT		
pC0.071	$T_{ECK120030798}$	−5.9	42	AGAATAAATTCAACCGCCCGTCAGGGCGGTTGTCATATGGAG		
pC0.073	$T_{ECK120010869}$	−5.6	35	TAACGTAAAAACCCGCTTCGGCGGGTTTTTTATG		
pC0.077	$T_{ECK120010841-R}$	−3.0	41	AAAAACAAAAACCCCGGACTCTCATCCAGGGTTCTCTGCTT		
pC0.308	$T_{ECK120033737}$	−8.0	57	GGAAACACAGAAAAAAGCCCGCACCTGACAGTGCGGGCTTTTTTTTTCGACCAAAGG		
pC0.309	$T_{ECK120033736}$	−8.7	53	AACGCATGAGAAAGCCCCCGGAAGATCACCTTCCGGGGGCTTTTTTATTGCGC		
pC0.310	$T_{ECK120010818}$	−10.8	54	CACCTGTTTTACGTAAAAACCCGCTTCGGCGGGTTTTTACTTTTGG		
pC0.311	$T_{ECK120015440}$	−6.4	49	TCCGGCAATTAAAAAAGCGGCTAACCACGCCGCTTTTTTTACGTCTGCA		
pC0.312	$T_{ECK120029600}$	−4.8	90	TTCAGCCAAAAAACTTAAGACCGCCGGTCTTGTCCACTACCTTGCAGTAATGCGGTGGACAGGATCGGCGGTTTTCTTTTCTCTTCTCAA		
pC0.313	$T_{ECK120010799}$	−10.6	60	TCAGGAAAAAAGGCGACAGAGTAATCTGTCGCCTTTTTTCTTTGC	E. coli	Chen et al., 2013
pC0.314	$T_{ECK120010876}$	−5.6	55	GAAAAATAAAAACGGCGCTAAAAAGCGCCGTTTTTTTTGACGGT		
pC0.315	$T_{ECK120015170}$	−8.5	47	TTTCGAAAAAACCCGCTTCGGCGGGTTTTTTTATAGC		
pC0.316	$T_{ECK120017009}$	−5.5	44	GATCTAACTAAAAAGGCCGCTCTGCGGCCTTTTTTCTTTTCACT		
pC0.317	$T_{ECK120051401}$	−7.4	47	ATAGCAAAAAAGCGCCTTTAGGGCGCTTTTTTACATTG		
pC0.318	$T_{ECK120010855}$	−5.7	42	AACAACGGAAACCGGCCATTGCGCCGGTTTTTTTTGGCC		
pC0.082	T_{rrnB}	−12.2	123	CAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCTACTAGAGTCACACTGGCTCACCTTCGGGTGGGCCTTTCTGCG	E. coli/Viral	
pC0.063	$T_{pSB1AK3}$	−11.8	44	ATTTCAGATAAAAAAAAATCCTTAGCTTTCGCTAAGGATGATTTC	E. coli	Vasudevan et al., 2019
pC0.079	T_{psbA2}	−1.9	83	CCAACTGAATAATCTGCAAATTGCACTCTCCTTCAATGGGGGGTGCTTTTTGCTTGACTGAGTAATCTTCTGATTGCTGATCT		
pC0.081	T_{psaB}	−10.6	53	TTAAGCTTGTCCCCTGCCCTCGTTGGTGGGGGATTTGCTTTAATTGGCTGATC	PCC 6803	

*The sequences have been annotated with features common to intrinsic terminators, including the A-tract (black underlined), stem (blue), loop (red), and U-tract (green underlined) (see **Figure 1C**) as reported by Chen et al. (2013). The features for the additional terminators were predicted using ARNold (http://rna.igmors.u-psud.fr/toolbox/arnold), Kinefold (http://kinefold.curie.fr/), or FindTerm (http://www.softberry.com/) (Xayaphoummine et al., 2005; Naville et al., 2011). Gibbs free energy values for the extended hairpin formation (i.e., the A- and U-tract) (ΔG_A) were calculated according to the equation $\Delta G_A = \Delta G_{HA} - \Delta G_H$, where ΔG_{HA} is the free energy of the hairpin loop with the inclusion of eight nucleotides upstream and downstream, and ΔG_H is the free energy of the hairpin loop alone (for generation of ΔG_{HA} and ΔG_H values, see section "Materials and Methods"). Further free energy values are shown in **Supplementary Table S4**.*

FIGURE 2 | Validation of the dual fluorescent reporter system in *E. coli*. **(A)** TE values for *E. coli* transformants at the early exponential phase of growth. Values are color coded for native *E. coli* (black), synthetic (gray) and native PCC 6803 terminators (green). T_{rrnB} is shown in brown. Error bars represent the ± standard error (SE) of the mean of >10,000 individual cells of four to eight biological replicates. Lowercase letters indicating significant difference ($P < 0.05$) are shown, as determined by ANOVA followed by Tukey's honestly significant difference test. **(B)** Comparison between TE and corresponding normalized fold change reduction in downstream reporter expression. An increase in one log2 value represents a 2-fold normalized change reduction. **(C)** Correlation analysis of TE values calculated from Chen et al. (2013) (**Supplementary Table S3**) and TE values determined in this study ($n = 30$). T_{rrnB}, $T_{pSB1AK3}$, T_{psbA2}, and T_{psaB} were excluded, as data was not available for comparison. The coefficient of determination (R^2) is shown. Terminator TE values marked in red ($T_{ECK120030798}$, $T_{ECK120010820}$, T_{Bba_B0011}, T_{Bba_B0061}, $T_{L3S1P22}$, and $T_{L3S1P13}$) differed from Chen et al. (2013) by more than 10%. Removal of these six terminator from the correlation analysis resulted in $R^2 = 0.9$).

point. The single exception was $T_{L3S2P21}$, which produced eYFP values consistently 2.5-fold higher than other strains. We are unsure why eYFP values were higher for $T_{L3S2P21}$, but we did re-confirm the terminator sequence in this strain by Sanger sequencing. In *E. coli* and bacteriophages, some intrinsic terminators can enhance upstream gene expression by enhancing the stability of the mRNA transcript via the hairpin loop (Abe and Aiba, 1996; Cisneros et al., 1996). Enhancement of mRNA stability by several putative intrinsic terminators has also been demonstrated for the marine species *Synechococcus* sp. PCC 7002, where transcripts with a canonical intrinsic terminator downstream were found to have a longer a half-life compared to transcripts without a downstream terminator (Gordon et al., 2020). However, $T_{L3S2P21}$ shares the same U-tract as both $T_{L3S2P11}$ and $T_{L3S2P55}$ but no increased eYFP expression was observed in the latter strains. mRNA transcript stability is a

subject of ongoing research, but some examples of causative factors in heterotrophic bacteria include starvation in *E. coli* and *Lactococcus lactis* (Redon et al., 2005; Morin et al., 2020), and temperature induced stress in *Staphylococcus aureus* and *Mycobacterium tuberculosis* (Anderson et al., 2006; Rustad et al., 2013). mRNA concentration can influence mRNA stability, with increasing transcript concentration leading to decreased stability and mRNA turnover in *E. coli* and *L. lactis* (Nouaille et al., 2017). Similar examples have not been reported yet for PCC 6803.

Similarly to *E. coli*, PCC 6803 strains produced a wide range of mTagBFP fluorescence values at each time point (**Supplementary Figure S3B**), while the mTagBFP:eYFP ratio for the "no terminator" control pCA1.376 was also consistently higher by $21 \pm 2\%$ compared to pCA1.377. A strong correlation was shown between TE values measured at different time points with R^2 values ranging from 0.982 to 0.988 (**Supplementary Figure 3C**).
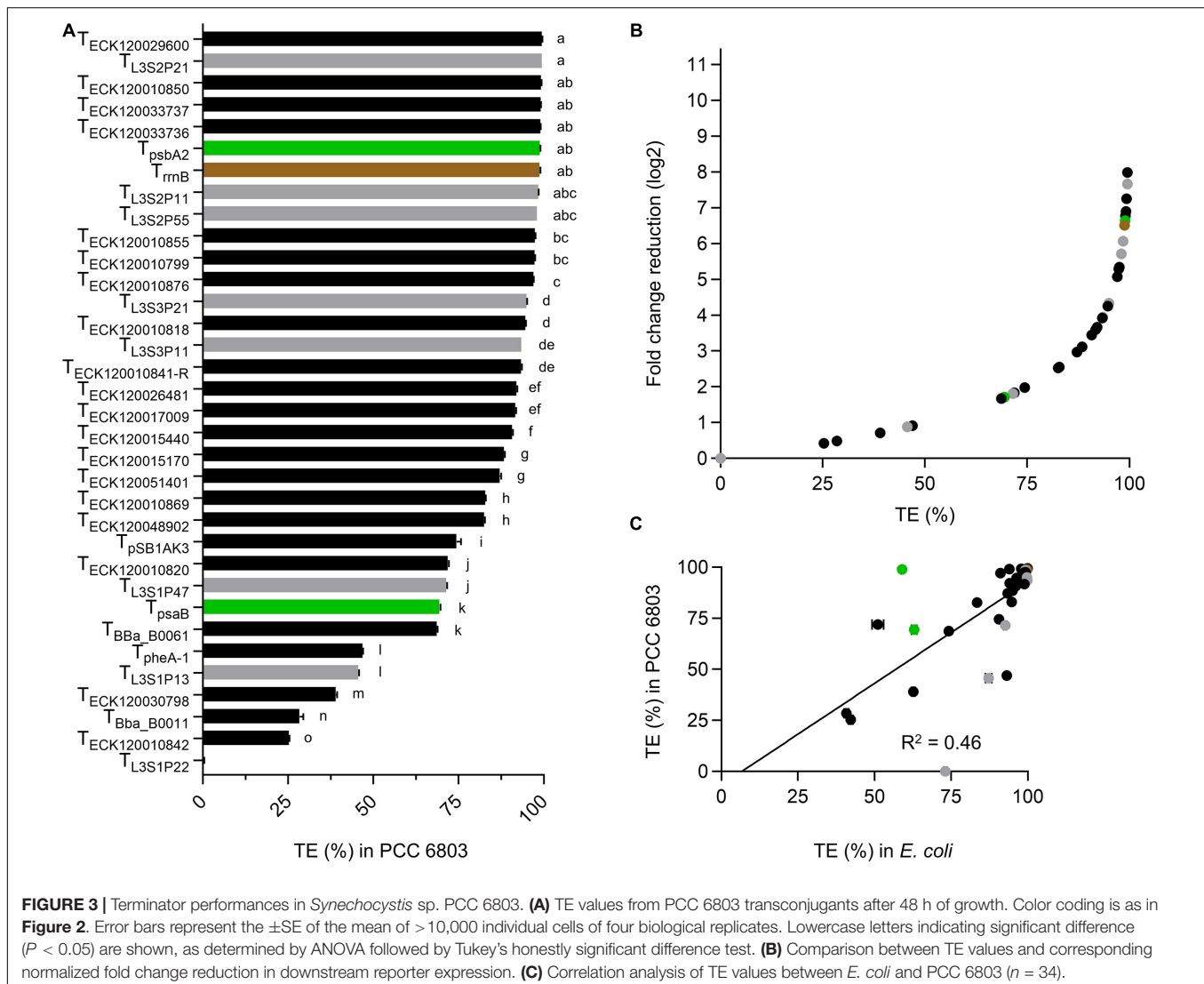
Comparison of TE values over the three time points were consistent for strong terminators (**Supplementary Table S3**). In contrast, weaker terminators tended to show a small decline in TE over time, although there was no significant change in the rankings observed. Overall, terminator behavior in PCC 6803 was consistent between on OD_{750} of 0.4 and 5.9 (**Supplementary Table S3**). Thus, we focused on reporting TE values at a single time point (48 h) below.

Thirteen terminators had TE values of >95% in PCC 6803 (**Figure 3A** and **Supplementary Table S3**), with $T_{L3S2P21}$ and $T_{ECK120029600}$ producing the highest value (99.5%) and $T_{ECK120010842}$ producing the lowest value (25.3%). Ten of the 13 strongest terminators in PCC 6803 also produced TE of >95% in *E. coli* (**Figure 2A**). Similarly, the two weakest terminators in PCC 6803 ($T_{ECK120010842}$ and T_{Bba_B0011}) were also the weakest in *E. coli*. Notably, $T_{L3S1P22}$ showed no detectable terminator activity in PCC 6803, but had a TE value of 73% in *E. coli*. Overall, the terminator library demonstrated a corresponding 8-fold change reduction in normalized downstream reporter

expression in PCC 6803 (**Figure 3B**). The TE values of 10 terminators differed more widely from those in *E. coli* (i.e., by 12–46%). Thus, the correlation of TE values between *E. coli* and PCC 6803 was modest ($R^2 = 0.46$) (**Figure 3C**). Removal of $T_{L3S1P22}$ led to only a marginal improvement ($R^2 = 0.53$).

Performance of the Terminator Library in *Synechococcus elongatus* UTEX 2973 and Comparison Between Species

Lastly, we evaluated our terminator library in the high-light tolerant strain UTEX 2973. UTEX 2973 generally grew faster than PCC 6803, but showed more variability in growth rates (**Supplementary Figure S4A**). This was likely due to a greater relative difference in light distribution within the growth incubator under the higher light levels used for culturing UTEX 2973, as strains in the same plate showed more similar rates of growth compared to those located at different positions within the incubator. As for PCC 6803, we measured fluorescence

FIGURE 3 | Terminator performances in *Synechocystis* sp. PCC 6803. **(A)** TE values from PCC 6803 transconjugants after 48 h of growth. Color coding is as in **Figure 2**. Error bars represent the ±SE of the mean of >10,000 individual cells of four biological replicates. Lowercase letters indicating significant difference ($P < 0.05$) are shown, as determined by ANOVA followed by Tukey's honestly significant difference test. **(B)** Comparison between TE values and corresponding normalized fold change reduction in downstream reporter expression. **(C)** Correlation analysis of TE values between *E. coli* and PCC 6803 (*n* = 34).

levels for UTEX 2973 at 24, 48, and 72 h (**Supplementary Figure S4B**). Consistent with the observed differences in growth, the expression levels of eYFP were variable between strains at 24 hr. However, this variation decreased over time.

As for PCC 6803, mTagBFP fluorescence values for the UTEX 2973 strains showed a wide spread at each time point, while the mTagBFP:eYFP ratio for pCA1.376 was consistently higher by $20 \pm 5\%$ compared to pCA1.377. Furthermore, the expression levels of mTagBFP and eYFP for pCA1.337 were more variable over time in UTEX 2973, with large increases in both eYFP and mTagBFP fluorescence values observed at 48 h (**Supplementary Figure S4B**). The TE values over the three time points were similar for most strains, with R^2 values ranging from 0.964 to 0.978 (**Supplementary Figure 4C**), indicating that terminator behavior in UTEX 2973 was consistent between an OD_{750} of 0.4–11 (**Supplementary Table S3**). Thus, as for PCC 6803 we also focused on reporting TE values at 48 h below.

Eleven terminators had TE values of >95% in UTEX 2973 (**Figure 4A** and **Supplementary Table S3**), with $T_{ECK120029600}$ producing a very high value of 99.9% and T_{Bba_B0061} producing the lowest value (29.7%). Six of the 10 strongest terminators in UTEX 2973 produced TE values of >95% in *E. coli* (**Figure 2A**), while seven of these terminators also produced TE values of >95% in PCC 6803 (**Figure 3A**). The three weakest terminators in UTEX 2973 (T_{Bba_B0061}, $T_{ECK120030798}$, and $T_{ECK120010820}$) were among the bottom ten ranked terminators in PCC 6803 and *E. coli*. $T_{ECK120010820}$ achieved the same ranking (i.e., 3rd weakest terminator) in both UTEX 2973 and *E. coli*. Overall, the terminator library demonstrated a corresponding 10-fold change reduction of normalized downstream reporter expression in UTEX 2973 (**Figure 4B**). Similarly to PCC 6803, the correlation of TE values between UTEX 2973 and *E. coli* was low ($R^2 = 0.35$) (**Figure 4C**). More surprisingly, the correlation of TE values between UTEX 2973 and PCC 6803 was even lower ($R^2 = 0.12$) (**Figure 4D**).

We next compared the TE values for *E. coli*, PCC 6803 and UTEX 2973 to identify terminators that were consistently strong between different species (**Supplementary Table S3**). The overall strongest terminator was $T_{ECK120029600}$, which had TE values of >99.5% across all three species. A further four terminators ($T_{L3S2P21}$, $T_{ECK120010850}$, $T_{L3S2P11}$, and T_{rrnB}) also had consistent cross-species TE values of >96%. For the two cyanobacterial species alone, $T_{ECK120033736}$ and T_{psbA2} had TE values of >95.8%. The TE values for these seven strong terminators was also very consistent over time for PCC 6803 and UTEX 2973.

The Performance of the Seven Strongest Terminators Was Consistent Under Suboptimal Growth Conditions

To examine if terminator performance might be affected by the growth environment, we measured the TE values for the seven strongest terminators in PCC 6803 and UTEX 2973 grown under suboptimal conditions. Both species were cultured at 30°C in 300 μM photons m^{-2} s^{-1}, which is considered high light for PCC 6803 (typically grown at 100 μM photons m^{-2} s^{-1}) and a low temperature for UTEX 2973 (typically grown at 40°C) (Vasudevan et al., 2019).

Both PCC 6803 and UTEX 2973 grew at similar rates and reached an OD_{750} of 5.9 and 5.7 after 72 h, respectively (**Supplementary Figure S5A**). In higher light PCC 6803 grew faster than under typical conditions, while growth rates were reduced in UTEX 2973 due to the lower temperature. Fluorescence measurements for eYFP and mTagBFP in PCC 6803 were comparable to those under typical growth conditions (**Supplementary Figure S5B**). In contrast, fluorescence values were generally reduced at all time points in UTEX 2973 (**Supplementary Figure S5C**). TE values for each day were calculated as before (**Supplementary Table S3**), and the mean values for the three time points were compared (**Table 2**). Overall, all seven terminators retained TE values of >95.8% for both species under the suboptimal growth conditions, and $T_{ECK120029600}$ remained the strongest terminator. Overall, our results indicated that the performance of these terminators was generally consistent and robust between the two growth conditions.

DISCUSSION

Here, we adapted a dual reporter tool for the CyanoGate MoClo Assembly system that provides a normalized quantification of terminator efficiency within and between species. The pDUOTK1-L1 vector is compatible with several available libraries and thus facilitates easy adoption and sharing of parts with the community (Andreou and Nakayama, 2018; Lai et al., 2018; Valenzuela-Ortega and French, 2019; Vasudevan et al., 2019), and is accessible to any lab currently using Golden Gate cloning. The robustness of our system was validated by comparing results in *E. coli* against previously published data (Chen et al., 2013).

The pDUOTK1-L1 vector contains the broad host range replicative origin RSF1010, which has been shown to be functional in a wide diversity of prokaryotic species, including cyanobacteria from all five subsections (Mermet-Bouvier et al., 1993; Stucken et al., 2012; Bishé et al., 2019). Thus, pDUOTK1-L1 could help to make terminator characterization more accessible, as promising new strains are discovered (Włodarczyk et al., 2019; Jaiswal et al., 2020; Nies et al., 2020). To the best of our knowledge, this is the first study to compare the efficiencies of terminators between two different cyanobacterial species. We identified five strong terminators with consistent TE values in *E. coli*, PCC 6803 and UTEX 2973. These findings should help to inform future strategies for building gene expression systems or more advanced gene circuit designs.

Besides the double terminator T_{rrnB}, no unique features could be identified for any of the five strong terminators that behaved consistently between all three species (i.e., the hairpin loop length and GC content, and adenine and uracil content for the A-tract and U-tract, respectively). Overall, our results showed that terminator performances was highly reproducible at different growth points for the same strain but generally differed between the three species examined, and significant differences were observed between PCC 6803 and UTEX 2973 even though both are subsection I species (Castenholz et al., 2001). We also demonstrated that the performance of the seven strongest
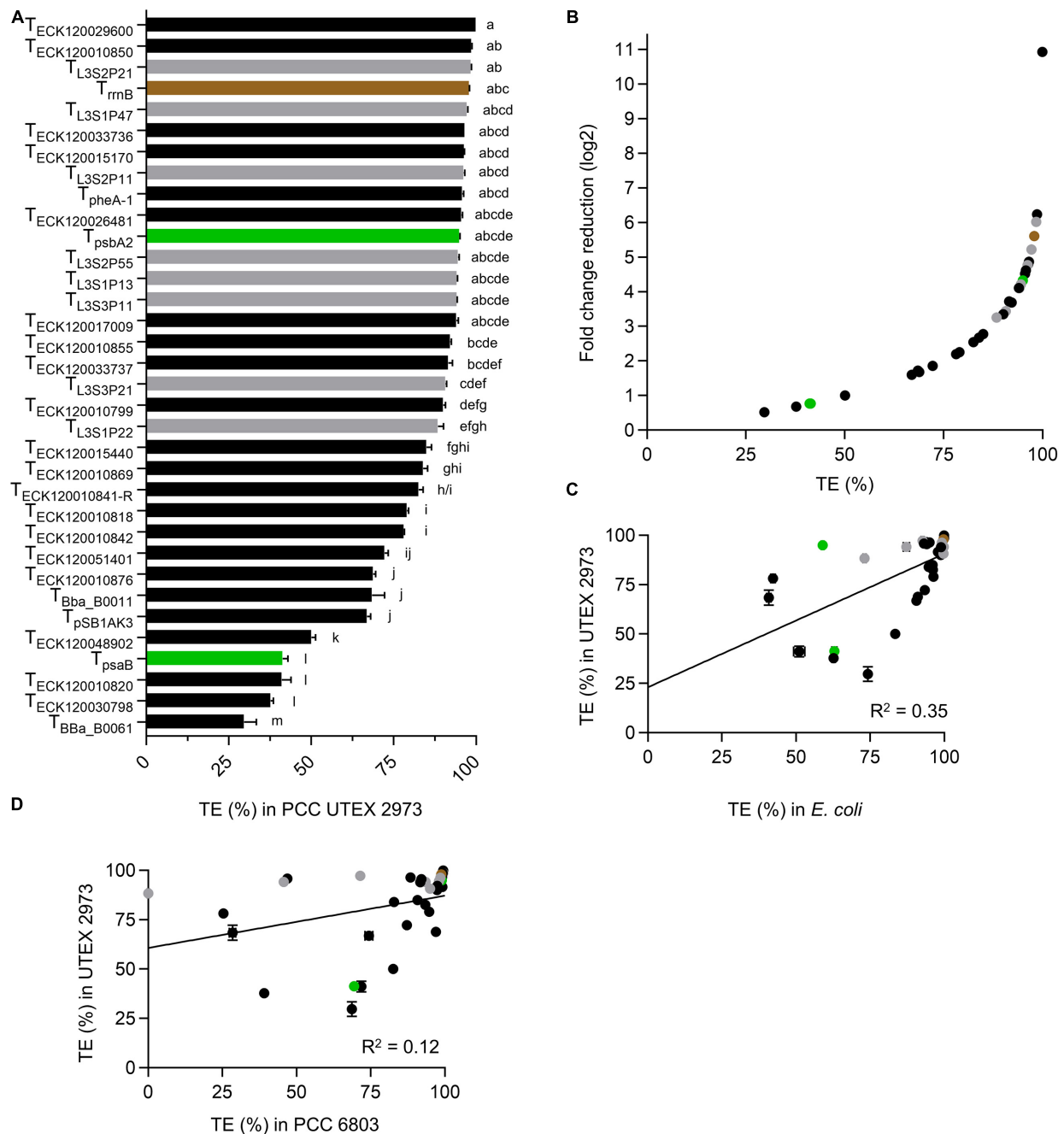
FIGURE 4 | Terminators performances in *Synechococcus elongatus* UTEX 2973. **(A)** TE value from UTEX 2973 after 48 h growth. Color coding is as in **Figure 2**. Error bars represent the ± SE of the mean of >10,000 individual cells of four biological replicates. Lowercase letters indicating significant difference ($P < 0.05$) are shown, as determined by ANOVA followed by Tukey's honestly significant difference test. **(B)** Comparison between TE values and corresponding normalized fold change reduction in downstream reporter expression. **(C)** Correlation analysis of TE values between *E. coli* and UTEX 2973 ($n = 34$). **(D)** Correlation analysis of TE values between PCC 6803 and UTEX 2973 ($n = 34$).

terminators was consistent in different growth conditions for PCC 6803 and UTEX 2793. Cyanobacterial RNAPs do differ in structure compared to other bacterial RNAPs [for a recent review see Stensjö et al. (2018)]. In addition, RNAP subunits also differ between cyanobacterial species [for a recent review see Srivastava et al. (2020)]. For example, the primary vegetative sigma factor (sigA) in PCC 6803 (srl0653) and UTEX 2973 (WP_071818124.1)

have a shared identity and similarity of only 70.5 and 74.1%, respectively (**Supplementary Figure S7**). Furthermore, cyanobacteria lack transcription elongation factors commonly found in heterotrophic bacteria to restart elongation and for proofreading of transcripts. To compensate, cyanobacterial RNAPs have evolved additional proof-reading and elongation functionalities (Riaz-Bradley et al., 2020). These differences may

TABLE 2 | Terminator performances in *Synechocystis* sp. PCC 6803 and Synechococcus elongatus UTEX 2973 under suboptimal growth conditions.

	TE (%) in PCC 6803		TE (%) in UTEX 2973	
	30°C, 100 μM photons m^{-2} s^{-1}	30°C, 300 μM photons m^{-2} s^{-1}	40°C, 300 μM photons m^{-2} s^{-1}	30°C, 300 μM photons m^{-2} s^{-1}
T$_{L3S2P21}$	99.5 ± 0.1	99.6 ± 0.1	97.1 ± 2.2	98.1 ± 0.9
T$_{L3S2P11}$	98.3 ± 0.2	98.3 ± 0.6	95.9 ± 0.5	95.9 ± 1.0
T$_{ECK120010850}$	99.2 ± 0.2	99.4 ± 0.2	98.1 ± 0.6	98.5 ± 0.3
T$_{ECK120033736}$	99.0 ± 0.1	99.3 ± 0.3	97.0 ± 0.9	98.1 ± 1.3
T$_{ECK120029600}$	99.6 ± 0.3	99.7 ± 0.2	99.9 ± 0.1	99.9 ± 0.1
T$_{rrnB}$	98.4 ± 0.4	98.4 ± 0.9	98.5 ± 0.5	98.0 ± 0.7
T$_{psbA2}$	98.7 ± 0.2	98.8 ± 0.4	95.8 ± 0.9	96.9 ± 1.2

*The mean TE values for PCC 6803 and UTEX 2973 at 24, 48, and 72 h are shown for typical and suboptimal growth conditions (see **Supplementary Table S3** for daily values). The standard deviation represents four biological replicates from each time point (n = 12).*

account for the observed disparity in terminator performance between *E. coli* and cyanobacteria. However, the differences between PCC 6803 and UTEX 2973 were intriguing, and could suggest that RNAP activities differ between cyanobacterial species and/or that other unknown factors are involved.

Several methods and prediction tools exist for the identification and mapping of intrinsic terminators in different species (Carafa et al., 1990; de Hoon et al., 2005; Gardner et al., 2011; Naville et al., 2011; Fritsch et al., 2015; Millman et al., 2017). Traditionally, these approaches have relied on identifying sequence features associated with intrinsic terminators (e.g., the hairpin loop). Previous studies have suggested a relationship between terminator performance and the estimated Gibbs free energy of the extended hairpin (ΔG_A), the U-tract (ΔG_U) and to a lesser extent the hairpin loop (ΔG_H) (Cambray et al., 2013; Chen et al., 2013). In our study, we did not find a strong correlation between TE values and ΔG_A, ΔG_H or the estimated Gibbs free energy of the complete terminator sequence (**Supplementary Figure S6**). Although our terminator library was relatively small, the differences in terminator behavior within and between species indicated that there may be more factors involved in determining intrinsic termination than can be attributed to the properties of individual structural components. For example, the U-tract appears dispensable for intrinsic termination in mycobacteria (Ahmad et al., 2020). Cutting edge approaches utilizing RNA-seq methods have also been applied for the identification of previously unknown terminators in the *E. coli* genome, which go beyond that which has been achieved with previous structural identification models (Ju et al., 2019). In addition, recent work has shown that terminator sequences can be designed as tunable control elements that can be "turned on" to attenuate gene transcription at low temperatures (Roßmanith et al., 2018). With the growing evidence that the structural components of terminators may be malleable depending on species, future work should focus on understanding the combined contributions of terminator components, including those beyond transcriptional control (e.g., modulation of protein expression) for metabolic engineering (Curran et al., 2013; Ito et al., 2020). This may lead to better designs for strong synthetic terminators with consistent cross-species performance. As terminator research and cyanobacterial synthetic biology progresses, tools such as pDUOTK1-L1 will be useful for reliable and convenient determination of terminator efficiency across a broad-host range.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

GG and AM: conceptualization and writing–original draft preparation. GG: performing the experiments. BW and AM: supervision. All authors: experimental design and writing–review and editing.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.624011/full#supplementary-material

REFERENCES

Abe, H., and Aiba, H. (1996). Differential contributions of two elements of rho-independent terminator to transcription termination and mRNA stabilization. *Biochimie* 78, 1035–1042. doi: 10.1016/S0300-9084(97)86727-2

Ahmad, E., Hegde, S. R., and Nagaraja, V. (2020). Revisiting intrinsic transcription termination in mycobacteria: u-tract downstream of secondary structure is dispensable for termination. *Biochem. Biophys. Res. Commun.* 522, 226–232. doi: 10.1016/j.bbrc.2019.11.062

Anderson, K. L., Roberts, C., Disz, T., Vonstein, V., Hwang, K., Overbeek, R., et al. (2006). Characterization of the Staphylococcus aureus heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. *J. Bacteriol.* 188, 6739–6756. doi: 10.1128/JB.00609-06

Andreou, A. I., and Nakayama, N. (2018). Mobius assembly: a versatile golden-gate framework towards universal DNA assembly. *PLoS One* 13:e0189892. doi: 10.1371/journal.pone.0189892

Behle, A., Saake, P., Germann, A. T., Dienst, D., and Axmann, I. M. (2020). Comparative dose–response analysis of inducible promoters in cyanobacteria. *ACS. Synth. Biol.* 9, 843–855. doi: 10.1021/acssynbio.9b00505

Bishé, B., Taton, A., and Golden, J. W. (2019). Modification of RSF1010-based broad-host-range plasmids for improved conjugation and cyanobacterial bioprospecting. *IScience* 20, 216–228. doi: 10.1016/J.ISCI.2019.09.002

Cambray, G., Guimaraes, J. C., Mutalik, V. K., Lam, C., Mai, Q. A., Thimmaiah, T., et al. (2013). Measurement and modeling of intrinsic transcription terminators. *Nucleic Acids Res.* 41, 5139–5148. doi: 10.1093/nar/gkt163

Camsund, D., Heidorn, T., and Lindblad, P. (2014). Design and analysis of LacI-repressed promoters and DNA-looping in a cyanobacterium. *J. Biol. Eng.* 8:4. doi: 10.1186/1754-1611-8-4

Carafa, Y. D. A., Brody, E., and Thermes, C. (1990). Prediction of rho-independent Escherichia coli transcription terminators. a statistical analysis of their RNA stem-loop structures. *J. Mol. Biol.* 216, 835–858. doi: 10.1016/S0022-2836(99)80005-9

Castenholz, R. W., Wilmotte, A., Herdman, M., Rippka, R., Waterbury, J. B., Iteman, I., et al. (2001). "Phylum BX. Cyanobacteria," in *Bergey's Manual® of Systematic Bacteriology: Volume One?: the Archaea and the Deeply Branching and Phototrophic Bacteria*, eds D. R. Boone, R. W. Castenholz, and G. M. Garrity (New York, NY: Springer), 473–599.

Chen, Y. J., Liu, P., Nielsen, A. A. K., Brophy, J. A. N., Clancy, K., Peterson, T., et al. (2013). Characterization of 582 natural and synthetic terminators and quantification of their design constraints. *Nat. Methods* 10, 659–664. doi: 10.1038/nmeth.2515

Cisneros, B., Court, D., Sanchez, A., and Montañez, C. (1996). Point mutations in a transcription terminator, λtI, that affect both transcription termination and RNA stability. *Gene* 181, 127–133. doi: 10.1016/S0378-1119(96)00492-1

Curran, K. A., Karim, A. S., Gupta, A., and Alper, H. S. (2013). Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metab. Eng.* 19, 88–97. doi: 10.1016/j.ymben.2013.07.001

de Hoon, M. J. L., Makita, Y., Nakai, K., and Miyano, S. (2005). Prediction of transcriptional terminators in bacillus subtilis and related species. *PLoS Comput. Biol.* 1:e25. doi: 10.1371/journal.pcbi.0010025

Engler, C., Youles, M., Gruetzner, R., Ehnert, T. M., Werner, S., Jones, J. D. G., et al. (2014). A golden gate modular cloning toolbox for plants. *ACS Synth. Biol.* 3, 839–843. doi: 10.1021/sb4001504

Englund, E., Liang, F., and Lindberg, P. (2016). Evaluation of promoters and ribosome binding sites for biotechnological applications in the unicellular cyanobacterium *Synechocystis*. PCC 6803. *Sci. Rep.* 6:36640. doi: 10.1038/srep36640

Eungrasamee, K., Miao, R., Incharoensakdi, A., Lindblad, P., and Jantaro, S. (2019). Improved lipid production via fatty acid biosynthesis and free fatty acid recycling in engineered *Synechocystis*. PCC 6803. *Biotechnol. Biofuels* 12:8. doi: 10.1186/s13068-018-1349-8

Ferreira, E. A., Pacheco, C. C., Pinto, F., Pereira, J., Lamosa, P., Oliveira, P., et al. (2018). Expanding the toolbox for *Synechocystis*. PCC 6803: validation of replicative vectors and characterization of a novel set of promoters. *Synth. Biol.* 3:ysy014. doi: 10.1093/synbio/ysy014

Fritsch, T. E., Siqueira, F. M., and Schrank, I. S. (2015). Intrinsic terminators in *Mycoplasma hyopneumoniae* transcription. *BMC Genomics* 16:273. doi: 10.1186/s12864-015-1468-6

Gale, G. A. R., Schiavon Osorio, A. A., Puzorjov, A., Wang, B., and McCormick, A. J. (2019). Genetic modification of cyanobacteria by conjugation using the cyanogate modular cloning toolkit. *J. Vis. Exp.* 152:e60451. doi: 10.3791/60451

Gardner, P. P., Barquist, L., Bateman, A., Nawrocki, E. P., and Weinberg, Z. (2011). RNIE: genome-wide prediction of bacterial intrinsic terminators. *Nucleic Acids Res.* 39, 5845–5852. doi: 10.1093/nar/gkr168

Gordon, G. C., Cameron, J. C., Gupta, S. T. P., Engstrom, M. D., Reed, J. L., and Pfleger, B. F. (2020). Genome-wide analysis of RNA decay in the cyanobacterium *Synechococcus sp*. Strain PCC 7002. *MSystems* 5:e00224–20 doi: 10.1128/msystems.00224-20

Herbert, K. M., Greenleaf, W. J., and Block, S. M. (2008). Single-molecule studies of RNA polymerase: motoring along. *Annu. Rev. Biochem.* 77, 149–176.

Hess, G. F., and Graham, R. S. (1990). Efficiency of transcriptional terminators in *Bacillus subtilis*. *Gene* 95, 137–141.

Huang, H. H., Lindblad, P. (2013). Wide-dynamic-range promoters engineered for cyanobacteria. *J. Biol. Eng.* 7:10. doi: 10.1186/1754-1611-7-10

Huang, H. H., Camsund, D., Lindblad, P., and Heidorn, T. (2010). Design and characterization of molecular tools for a synthetic biology approach towards developing cyanobacterial biotechnology. *Nucleic Acids Res.* 38, 2577–2593. doi: 10.1093/nar/gkq164

Huang, I.-S., and Zimba, P. V. (2019). Cyanobacterial bioactive metabolites—a review of their chemistry and biology. *Harmful Algae* 83, 42–94. doi: 10.1016/J.HAL.2018.11.008

Immethun, C. M., DeLorenzo, D. M., Focht, C. M., Gupta, D., Johnson, C. B., and Moon, T. S. (2017). Physical, chemical, and metabolic state sensors expand the synthetic biology toolbox for *Synechocystis sp*. PCC 6803. *Biotechnol. Bioeng.* 114, 1561–1569. doi: 10.1002/bit.26275

Ito, Y., Terai, G., Ishigami, M., Hashiba, N., Nakamura, Y., Bamba, T., et al. (2020). Exchange of endogenous and heterogeneous yeast terminators in *Pichia pastoris* to tune mRNA stability and gene expression. *Nucleic Acids Res.* 48, 13000–13012. doi: 10.1093/nar/gkaa1066

Jaiswal, D., Sengupta, A., Sengupta, S., Madhu, S., Pakrasi, H. B., and Wangikar, P. P. (2020). A novel cyanobacterium *Synechococcus elongatus* PCC 11802 has distinct genomic and metabolomic characteristics compared to its neighbor PCC 11801. *Sci. Rep.* 10:191. doi: 10.1038/s41598-019-57051-0

Ju, X., Li, D., and Liu, S. (2019). Full-length RNA profiling reveals pervasive bidirectional transcription terminators in bacteria. *Nat. Microbiol.* 4, 1907–1918. doi: 10.1038/s41564-019-0500-z

Kelly, C. L., Taylor, G. M., Hitchcock, A., Torres-Méndez, A., and Heap, J. T. (2018). A rhamnose-inducible system for precise and temporal control of gene expression in cyanobacteria. *ACS Synth. Biol.* 7, 1056–1066. doi: 10.1021/acssynbio.7b00435

Kelly, C. L., Taylor, G. M., Satkute, A., Dekker, L., and Heap, J. T. (2019). Transcriptional terminators allow leak-free chromosomal integration of genetic constructs in cyanobacteria. *BioRxiv [Preprint]* doi: 10.1101/689281

Kim, W. J., Lee, S.-M., Um, Y., Sim, S. J., and Woo, H. M. (2017). Development of synebrick vectors as a synthetic biology platform for gene expression in *Synechococcus elongatus* PCC 7942. *Front. Plant Sci.* 8:293. doi: 10.3389/fpls.2017.00293

Knoot, C. J., Ungerer, J., Wangikar, P. P., and Pakrasi, H. B. (2018). Cyanobacteria: promising biocatalysts for sustainable chemical production. *J. Biol. Chem.* 293, 5044–5052. doi: 10.1074/JBC.R117.815886

Lai, H. E., Moore, S., Polizzi, K., and Freemont, P. (2018). EcoFlex: a multifunctional moclo kit for E. coli synthetic biology. *Methods Mol. Biol.* 1772, 429–444. doi: 10.1007/978-1-4939-7795-6_25

Lea-Smith, D. J., Vasudevan, R., and Howe, C. J. (2016). Generation of marked and markerless mutants in model cyanobacterial species. *J. Vis. Exp.* 2016:54001. doi: 10.3791/54001

Lim, H. N., Lee, Y., and Hussein, R. (2011). Fundamental relationship between operon organization and gene expression. *Proc. Natl. Acad. Sci. USA* 108, 10626–10631. doi: 10.1073/pnas.1105692108

Lin, P. C., and Pakrasi, H. B. (2019). Engineering cyanobacteria for production of terpenoids. *Planta* 249, 145–154. doi: 10.1007/s00425-018-3047-y

Lin, P. C., Saha, R., Zhang, F., and Pakrasi, H. B. (2017). Metabolic engineering of the pentose phosphate pathway for enhanced limonene production in the cyanobacterium *Synechocysti s sp*. PCC. *Sci. Rep.* 7:17503. doi: 10.1038/s41598-017-17831-y

Liu, D., and Pakrasi, H. B. (2018). Exploring native genetic elements as plug-in tools for synthetic biology in the cyanobacterium *Synechocystis* sp. PCC 6803. *Microb. Cell. Fact.* 17, 1–8. doi: 10.1186/s12934-018-0897-8

Ma, A. T., Schmidt, C. M., and Golden, J. W. (2014). Regulation of gene expression in diverse cyanobacterial species by using theophylline-responsive riboswitches. *Appl. Environ. Microbiol.* 80, 6704–6713. doi: 10.1128/AEM.01697-14

Mairhofer, J., Wittwer, A., Cserjan-Puschmann, M., and Striedner, G. (2015). Preventing T7 RNA polymerase read-through transcription—a synthetic termination signal capable of improving bioprocess stability. *ACS Synth. Biol.* 4, 265–273. doi: 10.1021/sb5000115

Mermet-Bouvier, P., Cassier-Chauvat, C., Marraccini, P., and Chauvat, F. (1993). Transfer and replication of RSF1010-derived plasmids in several cyanobacteria of the genera *Synechocystis* and *Synechococcus*. *Curr. Microbiol.* 27, 323–327. doi: 10.1007/BF01568955

Millman, A., Dar, D., Shamir, M., and Sorek, R. (2017). Computational prediction of regulatory, premature transcription termination in bacteria. *Nucleic Acids Res.* 45, 886–893. doi: 10.1093/nar/gkw749

Morin, M., Enjalbert, B., Ropers, D., Girbal, L., and Cocaign-Bousquet, M. (2020). Genomewide stabilization of mRNA during a "Feast-To-Famine" growth transition in ⁢span class="named-content genus-species" id="named-content-1">*Escherichia coli*⁢/span&gt. *MSphere* 5, e276–e220. doi: 10.1128/mSphere.00276-20

Moser, F., Espah Borujeni, A., Ghodasara, A. N., Cameron, E., Park, Y., and Voigt, C. A. (2018). Dynamic control of endogenous metabolism with combinatorial logic circuits. *Mol. Syst. Biol.* 14:e8605. doi: 10.15252/msb.20188605

Naville, M., Ghuillot-Gaudeffroy, A., Marchais, A., and Gautheret, D. (2011). ARNold: a web tool for the prediction of Rho-independent transcription terminators. *RNA Biol.* 8, 11–13. doi: 10.4161/rna.8.1.13346

Nies, F., Mielke, M., Pochert, J., and Lamparter, T. (2020). Natural transformation of the filamentous cyanobacterium *Phormidium lacuna*. *PLoS One* 15:e023440. doi: 10.1371/journal.pone.0234440

Nouaille, S., Mondeil, S., Finoux, A. L., Moulis, C., Girbal, L., and Cocaign-Bousquet, M. (2017). The stability of an mRNA is influenced by its concentration: a potential physical mechanism to regulate gene expression. *Nucleic Acids Res.* 45, 11711–11724. doi: 10.1093/nar/gkx781

Peters, J. M., Mooney, R. A., Kuan, P. F., Rowland, J. L., Keleş, S., and Landick, R. (2009). Rho directs widespread termination of intragenic and stable RNA transcription. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15406–15411. doi: 10.1073/pnas.0903846106

Peters, J. M., Vangeloff, A. D., and Landick, R. (2011). Bacterial transcription terminators: the RNA 3′-end chronicles. *J. Mol. Biol.* 412, 793–813. doi: 10.1016/j.jmb.2011.03.036

Redon, E., Loubiere, P., and Cocaign-Bousquet, M. (2005). Transcriptome analysis of the progressive adaptation of *Lactococcus lactis* to carbon starvation. *J. Bacteriol.* 187, 3589–3592. doi: 10.1128/JB.187.10.3589-3592.2005

Riaz-Bradley, A., James, K., and Yuzenkova, Y. (2020). High intrinsic hydrolytic activity of cyanobacterial RNA polymerase compensates for the absence of transcription proofreading factors. *Nucleic Acids Res.* 48, 1341–1352. doi: 10.1093/nar/gkz1130

Roßmanith, J., Weskamp, M., and Narberhaus, F. (2018). Design of a temperature-responsive transcription terminator. *ACS Synth. Biol.* 7, 613–621. doi: 10.1021/acssynbio.7b00356

Rustad, T. R., Minch, K. J., Brabant, W., Winkler, J. K., Reiss, D. J., Baliga, N. S., et al. (2013). Global analysis of mRNA stability in *Mycobacterium tuberculosis*. *Nucleic Acids Res.* 41, 509–517. doi: 10.1093/nar/gks1019

Shen, P., and Huang, H. V. (1986). Homologous recombination in *Escherichia coli*: dependence on substrate length and homology. *Genetics* 112, 441–457.

Sleight, S. C., Bartley, B. A., Lieviant, J. A., and Sauro, H. M. (2010). Designing and engineering evolutionary robust genetic circuits. *J. Biol. Eng.* 4:12. doi: 10.1186/1754-1611-4-12

Srivastava, A., Varshney, R. K., and Shukla, P. (2020). Sigma factor modulation for cyanobacterial metabolic engineering. *Trends Microbiol.* doi: 10.1016/j.tim.2020.10.012 Online ahead of print.

Stensjö, K., Vavitsas, K., and Tyystjärvi, T. (2018). Harnessing transcription for bioproduction in cyanobacteria. *Physiol. Plant.* 162, 148–155. doi: 10.1111/ppl.12606

Stucken, K., Ilhan, J., Roettger, M., Dagan, T., and Martin, W. F. (2012). Transformation and conjugal transfer of foreign genes into the filamentous multicellular cyanobacteria (subsection V) *Fischerella* and *Chlorogloeopsis*. *Curr. Microbiol.* 65, 552–560. doi: 10.1007/s00284-012-0193-5

Sun, T., Li, S., Song, X., Pei, G., Diao, J., Cui, J., et al. (2018). Re-direction of carbon flux to key precursor malonyl-CoA via artificial small RNAs in photosynthetic *Synechocystis sp*. PCC 6803. *Biotechnol. Biofuels* 11:26. doi: 10.1186/s13068-018-1032-0

Tsinoremas, N. F., Kutach, A. K., Strayer, C. A., and Golden, S. S. (1994). Efficient gene transfer in *Synechococcus sp*. strains PCC 7942 and PCC 6301 by interspecies conjugation and chromosomal recombination. *J. Bacteriol.* 176, 6764–6768. doi: 10.1128/jb.176.21.6764-6768.1994

Valenzuela-Ortega, M., and French, C. (2019). Joint universal modular plasmids (JUMP): a flexible and comprehensive platform for synthetic biology. *BioRxiv [Preprint]* doi: 10.1101/799585

Vasudevan, R., Gale, G. A. R., Schiavon, A. A., Puzorjov, A., Malin, J., Gillespie, M. D., et al. (2019). Cyanogate: a modular cloning suite for engineering cyanobacteria based on the plant moclo syntax. *Plant Physiol.* 180, 39–55. doi: 10.1104/pp.18.01401

Vijayan, V., Jain, I. H., and O'Shea, E. K. (2011). A high resolution map of a cyanobacterial transcriptome. *Genome Biol.* 12:R47. doi: 10.1186/gb-2011-12-5-r47

Williams, J. G. K. (1988). Methods in synechocystis 6803. *Methods Enzymol.* 167, 766–778.

Wilson, K. S., and Von Hippel, P. H. (1995). Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proc. Natl. Acad. Sci. USA* 92, 8793–8797. doi: 10.1073/pnas.92.19.8793

Włodarczyk, A., Selão, T. T., Norling, B., and Nixon, P. J. (2019). Unprecedented biomass and fatty acid production by the newly discovered cyanobacterium *Synechococcus sp*. PCC 11901. *BioRxiv [Preprint]* doi: 10.1101/684944

Xayaphoummine, A., Bucher, T., and Isambert, H. (2005). Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots. *Nucleic Acids Res.* 33(Suppl. 2), 605–610. doi: 10.1093/nar/gki447

Yager, T. D., and von Hippel, P. H. (1991). A thermodynamic analysis of RNA transcript elongation and termination in *Escherichia coli*. *Biochemistry* 30, 1097–1118. doi: 10.1021/bi00218a032

Yao, L., Shabestary, K., Björk, S. M., Asplund-Samuelsson, J., Joensson, H. N., Jahn, M., et al. (2020). Pooled CRISPRi screening of the cyanobacterium *Synechocystis sp* PCC 6803 for enhanced industrial phenotypes. *Nat. Commun.* 11:1666. doi: 10.1038/s41467-020-15491-7

Yu, J., Liberton, M., Cliften, P. F., Head, R. D., Jacobs, J. M., Smith, R. D., et al. (2015). Synechococcus elongatus UTEX 2973, a fast growing cyanobacterial chassis for biosynthesis using light and CO_2. *Sci. Rep.* 5:8132. doi: 10.1038/srep08132

Zavřel, T., Očenášová, P., and Červený, J. (2017). Phenotypic characterization of *Synechocystis sp*. PCC 6803 substrains reveals differences in sensitivity to abiotic stress. *PLoS One* 12:e0189130. doi: 10.1371/journal.pone.0189130

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406–3415. doi: 10.1093/nar/gkg595

Check for
updates

Slow Protein Turnover Explains Limited Protein-Level Response to Diurnal Transcriptional Oscillations in Cyanobacteria

Jan Karlsen[1,2], Johannes Asplund-Samuelsson[1,2†], Michael Jahn[1,2†], Dóra Vitay[1,2,3†] and Elton P. Hudson[1,2*†]

[1] Department of Protein Science, KTH Royal Institute of Technology, Stockholm, Sweden, [2] Science for Life Laboratory, Stockholm, Sweden, [3] Biosyntia ApS, Copenhagen, Denmark

Metabolically engineered cyanobacteria have the potential to mitigate anthropogenic CO_2 emissions by converting CO_2 into renewable fuels and chemicals. Yet, better understanding of metabolic regulation in cyanobacteria is required to develop more productive strains that can make industrial scale-up economically feasible. The aim of this study was to find the cause for the previously reported inconsistency between oscillating transcription and constant protein levels under day-night growth conditions. To determine whether translational regulation counteracts transcriptional changes, *Synechocystis* sp. PCC 6803 was cultivated in an artificial day-night setting and the level of transcription, translation and protein was measured across the genome at different time points using mRNA sequencing, ribosome profiling and quantitative proteomics. Furthermore, the effect of protein turnover on the amplitude of protein oscillations was investigated through *in silico* simulations using a protein mass balance model. Our experimental analysis revealed that protein oscillations were not dampened by translational regulation, as evidenced by high correlation between translational and transcriptional oscillations ($r = 0.88$) and unchanged protein levels. Instead, model simulations showed that these observations can be attributed to a slow protein turnover, which reduces the effect of protein synthesis oscillations on the protein level. In conclusion, these results suggest that cyanobacteria have evolved to govern diurnal metabolic shifts through allosteric regulatory mechanisms in order to avoid the energy burden of replacing the proteome on a daily basis. Identification and manipulation of such mechanisms could be part of a metabolic engineering strategy for overproduction of chemicals.

Keywords: cyanobacteria, diurnal gene expression, protein turnover, post-transcriptional regulation, metabolic regulation, RNA sequencing, ribosome profiling, proteomics

INTRODUCTION

Knowledge of cyanobacterial metabolism and its regulation can guide metabolic engineering efforts to create more efficient strains for renewable fuel and chemical production. As their energy source is limited to the light hours of the day, cyanobacteria have evolved to shift between photosynthetic and respiratory metabolism between day and night, respectively. During the day, CO_2 is fixed in the

Calvin cycle and converted into biomass, including storage compounds such as glycogen. During the night, CO_2 fixation and most biosynthetic pathways are inactive while glycogen is degraded to support cellular maintenance and a small subset of pathways that prepare the cell for the next light period (Saha et al., 2016; Reimers et al., 2017; Welkie et al., 2019; Werner et al., 2019). Metabolic shifts that occur at specific time points over the day-night cycle are governed by regulating the flux through key enzymes and pathways. The flux through an enzyme is regulated by changing its abundance, product/substrate concentration, or through post-translational effects that alter its apparent kinetic parameters. Several studies have investigated abundance-controlled regulation by tracking changes in the cyanobacterial transcriptome and proteome across the day-night cycle. Transcriptomic data collected from a range of cyanobacteria showed that a large fraction of cyanobacterial transcripts oscillates diurnally (30–87%), with peak expression mostly during the transitions between day and night (Stöckel et al., 2008; Ito et al., 2009; Waldbauer et al., 2012; Saha et al., 2016). Additionally, many transcripts tend to peak just before the time when the gene product's function is expected to be needed by the cell. For example, transcripts of Calvin cycle and pentose phosphate pathway genes peaked in the beginning of the light and dark period, respectively (Waldbauer et al., 2012). Yet surprisingly, a few proteomics studies have shown that abundance of most proteins remains nearly constant (Stöckel et al., 2011; Waldbauer et al., 2012; Guerreiro et al., 2014; Angermayr et al., 2016). This makes the regulatory purpose of time-dependent transcription seem insignificant for regulating enzyme activity and diurnal metabolic shifts.

The underlying cause for a broad discrepancy between transcript and protein dynamics is still not clear, but it could be attributed to post-transcriptional regulation or low daily *de novo* protein synthesis relative to the protein abundance. One possibility is that translational regulation counteracts changes in mRNA abundance, resulting in reduced variation in protein synthesis rate of genes despite their altered transcript levels. Protein synthesis rates can be measured genome-wide through ribosome profiling (Ribo-Seq), which quantifies the total number of ribosomes translating a gene's transcripts (Brar and Weissman, 2015; Liu et al., 2017). A translationally-regulated gene would show a change in ribosome abundance that is not equal to the change in transcript abundance, or vice versa. Translational regulation was shown to occur in 7% of the genome of *Synechocystis* sp. PCC 6803 (*Synechocystis*) in response to CO_2 starvation (Karlsen et al., 2018). A second possibility is that protein levels are held relatively constant by active protein degradation. However, rapid degradation of newly synthesized proteins would waste energy and cellular resources and reduce fitness. Lastly, relatively low variation in protein levels could also occur without any post-transcriptional regulation, if the daily variation in protein synthesis rate is low compared to the protein abundance, i.e., if the turnover rate of the proteome is low.

Here, we apply a systems biology approach to take a closer look at the discrepancies between transcription and protein abundances during day-night cycles in cyanobacteria. The model cyanobacterium *Synechocystis* was grown in controlled turbidostat cultures under artificial day-night cycles. To assess the impact of translational regulation on the protein level, the transcriptome, translatome, and proteome was measured at different time points using mRNA sequencing, ribosome profiling, and quantitative proteomics. We found that protein synthesis rates tracked with transcriptional oscillations, while protein abundances remained relatively constant, indicating that translational regulation does not significantly impact the protein-level behavior. We further investigated the effect of protein turnover on protein dynamics *in silico*. Simulation of protein oscillations using biologically relevant parameter settings, resulted in a protein amplitude similar to experimental observations. The data and model simulations demonstrate that post-translational regulation is not necessary for the proteome to remain stable, even under significant transcriptional oscillations.

MATERIALS AND METHODS

Cultivation and Sampling

Synechocystis sp. PCC 6803 was cultivated in 1.6 L BG-11 (pH = 7.8) at 30°C in a cylindrical photobioreactor (D = 10 cm, V = 2 L, baffled). The culture was illuminated with an LED light jacket covering the sides of the cylinder (90% red light, 10% blue light). CO_2-enriched air was sparged into the culture (7% CO_2, 330 mL min^{-1}) and the impeller stirring rate was set to 150 rpm. Cells were grown in turbidostat mode (OD_{730} set point: 0.65–0.80) under an artificial day-night light regime (Day: sinusoidal, max 500 µmol photons m^{-2} s^{-1}; Night: dark) for seven days, at which point the diurnal pattern of dissolved oxygen (growth rate proxy) became stable over subsequent days. Five time points (1 h before/after sunrise, midday, 1 h before/after sunset) were then sampled at -1, 1, 6, 11, 13, 30, 35, 37, 47, and 49 h relative to the first subjected sunrise. Two replicate cultivations were conducted. In the first cultivation, two replicate samples were collected at all five diurnal time points for mRNA sequencing and ribosome profiling. In the second cultivation, one and two replicate samples were collected for ribosome profiling and quantitative proteomics, respectively. Eleven out of fifteen collected ribosome profiling samples were analyzed, which resulted in two replicate measurements at all time points, except 1 h after subjected sunrise which had three. The correlation between ribosome profiling replicates within the same cultivation was similar to the correlation between replicates of different cultivations ($r \approx$ 0.99 and $r \approx$ 0.97, respectively), which indicated that results were reproducible across cultivations (**Supplementary Figure 1**). Cultivation data is shown in **Supplementary Figure 2**.

Determination of Diurnal Growth Rate

Analysis was performed using R v.3.6 scripts[1]. Specific growth rates (µ) were determined over the time course of the first cultivation at 30-second intervals according to a mass balance-derived equation of biomass in the culture:

$$\mu = \left(\frac{dOD_{730}}{dt} + D \cdot OD_{730} \right) / OD_{730}.$$

---

[1]https://github.com/J-Karlsen/diurnal-gene-expression

Where D is the dilution rate (h^{-1}) and OD_{730} is the optical density (cell mass concentration proxy) of the culture. Dilution rates were calculated by dividing time-specific medium feed rates (L h^{-1}, automatically regulated) with the volume of the culture (1.6 L). To remove noise, each time point was assigned with the 40% truncated average of the closest 480 time points (40% most extreme values removed in each window). The noise filtering step was repeated once. The optical density was measured automatically at 880 nm and converted to OD_{730} using a conversion factor based on offline OD_{730} measurements during the time course. Noise was removed from OD_{730} values in two subsequent filtration steps. In the first step, each time point was assigned with the 40% truncated average of the 960 closest time points. In the second step, each time point was assigned with the average of the closest 360 time points. The change in OD_{730} over time ($\frac{dOD_{730}}{dt}$) was calculated at each time point as the slope in noise-removed OD_{730} (1 h centered time intervals). Noise was finally removed from calculated growth rates by assigning each time point with the 40% truncated average of the 720 closest time points, and negative values were replaced by zero. The average diurnal growth rate was based on growth rates determined across the two days of sampling (0–48 h). Growth rates determined in the second cultivation experiment were prone to error and therefore not reported (higher noise levels and longer data acquisition intervals). Data is shown in **Supplementary Figure 2**.

Sample Preparation for mRNA Sequencing

Culture medium was removed by centrifugation and cell pellets snap-frozen in liquid nitrogen (stored at –80°C). Cell lysis was performed using lysozyme treatment and vortexing with glass beads. Total RNA was extracted from the cleared lysate with hot phenol/chloroform and isopropanol precipitation and remaining DNA was removed using DNase I. The amount of rRNA was subsequently reduced using the *RiboMinus Kit, Bacteria* (*ThermoFisher, K155004*) according to the manufacturer's instructions. Sequencing libraries were prepared using the *NEBNext Ultra Directional RNA Library Prep Kit* (*New England Biolabs, E7420*). Libraries were sequenced on an Illumina NextSeq500 platform (75 bp read length, single end). For details, see Karlsen et al. (2018). Raw sequencing data are available at the European Nucleotide Archive under accession number PRJEB42778.

Sample Preparation for Ribosome Profiling

Cells were rapidly harvested by vacuum filtration and snap-frozen in liquid nitrogen (stored at –80°C). Frozen cells were lysed using cryogenic grinding. The frozen lysate was thawed and cell debris was removed by centrifugation. Polysomes in the lysate were immediately digested with micrococcal nuclease, and generated monosomes were extracted by sucrose gradient ultracentrifugation and fractionation. Total RNA was extracted from monosomes with hot phenol/chloroform and isopropanol precipitation. Ribosome protected mRNA fragments were then extracted by size selection on a denaturing polyacrylamide gel (20–40 nt) and subsequently converted into a sequencing library using the *NEBNext small RNA library prep set (New England Biolabs; E7300)*. Libraries were sequenced on an Illumina NextSeq500 platform (75 bp read length, single end). For details, see Karlsen et al. (2018). Raw sequencing data are available at the European Nucleotide Archive under accession number PRJEB42778.

Sample Preparation for LC-MS-MS

Culture medium was removed by centrifugation and cell pellets snap-frozen in liquid nitrogen (stored at –80°C). Thawed cell pellets were suspended in 125 μL solubilization buffer (200 mM TEAB, 8 M Urea, protease inhibitor). 100 μL glass beads (100 μm diameter) were added to the cell suspension and cells were lysed by bead beating in a Qiagen TissueLyzer II (5 min, f = 30/s, precooled cassettes). Cell debris was removed by centrifugation at 14,000 × g, 30 min, 4°C, and supernatant was transferred to a new tube. Protein concentration was determined using the Bradford assay (Bio-Rad). For reduction and alkylation of proteins, 2.5 μL 200 mM DTT (5 mM final) and 5 μL 200 mM CAA (10 mM final) were added, respectively, and samples incubated for 60 min at RT in the dark. Samples were diluted 8-fold with 700 μL 200 μM TEAB. For digestion, Lys-C was added in a ratio of 1:75 w/w to protein concentration, and samples were incubated at 37°C and 600 RPM for 12 h. Trypsin was added (1:75 w/w) and samples incubated for 24 h at the same conditions. Samples were acidified with 100 μL 10% formic acid (FA) and insoluble compounds were removed by centrifugation (14,000 × g, 15 min, RT). Peptide samples were then cleaned up using a solid phase extraction protocol (Sep-Pak 1cc 50 mg A C18 cartridges, Waters) according to the manufacturer's recommendations. Briefly, Sep-Pak columns were equilibrated with 1 mL acetonitrile (ACN) and 2 × 1 mL 0.6% acetic acid. Samples were loaded on columns and washed twice with 1 mL 0.6% acetic acid. Peptides were eluted from the column in 500 μL elution buffer (0.6% acetic acid, 80% ACN) and dried in a speedvac for 2 h, 37°C. Dried peptides were frozen at –80°C and dissolved in 10% FA to a final concentration of 1 μg/μL before MS measurement.

LC-MS-MS Analysis of Lysates

Lysates were analyzed using a Thermo Fisher Q Exactive HF mass spectrometer (MS) coupled to a Dionex UltiMate 3000 UHPLC system (Thermo Fisher). The UHPLC was equipped with a trap column (Acclaim PepMap 100, 75 μm × 2 cm, C18, P/N 164535, Thermo Fisher Scientific) and a 50 cm analytical column (Acclaim PepMap 100, 75 μm × 50 cm, C18, P/N ES803, Thermo Fisher Scientific). The injection volume was 2 μL out of 18 μL in which the samples were dissolved in the autosampler. Chromatography was performed using solvent A (3% ACN, 0.1% FA) and solvent B (95% ACN, 0.1% FA) as the mobile phases. The peptides were eluted from the UHPLC system over a 90 min gradient at a flow rate of 250 nL/min with the following mobile phase gradient: 2% solvent B for 4 min, 2–4% solvent B for 1 min, 4–45% solvent B for 90 min, 45–99% solvent B for 3 min, 99% solvent B for 10 min and 99–2% solvent B for 1 min following re-equilibration of the column at 2% solvent B for 6 min. The MS was operated in a data-dependent acquisition mode with a

Top 8 method. The MS was configured to perform a survey scan from 300 to 2,000 m/z with resolution of 120,000, AGC target of 1×10^6, maximum IT of 250 ms and 8 subsequent MS/MS scans at 30,000 resolution with isolation window of 2.0 m/z, AGC target of 2×10^5, maximum IT 150 ms and dynamic exclusion set to 20 s. LC-MS shotgun proteomics data are available at the PRIDE Archive[2] under accession number PXD023812.

Relative Quantification of Cellular Protein Content

The protein content was quantified in the cell extracts used for LC-MS-MS (Bradford assay). Measured concentrations were normalized to the sample's cell mass concentration (based on external OD_{730} measurements).

Sequencing Data Processing and Quantification of mRNA and Ribosomes

Analysis of sequencing data was conducted using python v.2.7 scripts adapted from Becker et al. (2013), R v.3.4 scripts, and bash commands parallelized using GNU Parallel v.20161222 (Tange, 2011)[3]. FastQC was used to assess the quality and general features of sequencing datasets (Andrews, 2010). Adapter sequences were trimmed off using Cutadapt v1.18 (Martin, 2011). Base calls with a Sanger quality score lower than 20 were trimmed off the ends of mRNA sequencing reads using Sickle (Joshi and Fass, 2011). Ribosome profiling reads with an average Sanger quality score lower than 25 were removed using Seqmagick v0.6.2[4]. Reads shorter than 6 nt were discarded. Reads that mapped to tRNA and rRNA genome sequences were subsequently removed using Bowtie v.1.2.2 (Langmead et al., 2009). Bowtie was used to map remaining reads to the genome, including plasmids (NC_000911.1 + NC_005229.1 + NC_005230.1 + NC_005231.1 + NC_005232.1 + NC_020289.1 + NC_020290.1 + NC_020298.1). A maximum of two alignment mismatches were allowed. If a read mapped to several locations, only the one best alignment was kept. The read was discarded if it could not be mapped to a unique location in this way. The total number of mapped non-tRNA/rRNA reads was ∼2 million and 34–78 million in mRNA sequencing and ribosome profiling samples, respectively. For each mapped mRNA sequencing read, a read count equal to 1 was distributed evenly over all its aligned genome positions. In contrast, the read count of each mapped ribosome profiling read was assigned to a single genome position, 12 nt upstream of the aligned 3' end. This assigns the read count to the genome position covered by the A-site of the ribosome (Karlsen et al., 2018). As only ribosome profiling reads longer than 24 nt were counted, the total number of counted mapped reads per sample was between 17 and 68 million. The mRNA/ribosome abundance of a gene (RPKM) was quantified by dividing the read count on the gene's coding sequence with the length of the coding sequence (in 1,000 base pairs) and the total number of counted reads on all coding sequences (in million). Coding sequences were defined

according to GenBank files for the NCBI reference sequences NC_000911.1, NC_005229.1, NC_005230.1, NC_005231.1, NC_005232.1, NC_020289.1, NC_020290.1, and NC_020298.1.

Protein Identification and Quantification

Thermo raw spectra files were converted to the mzML standard using Proteowizard's MSConvert tool (Adusumilli and Mallick, 2017). Peptide identification and label-free quantification were performed using OpenMS 2.4.0 in KNIME (Röst et al., 2016). The KNIME pipeline for MS data processing was deposited on https://github.com/m-jahn/openMS-workflows (labelfree_MSGFplus_Percolator_FFI.knwf). MS/MS spectra were subjected to sequence database searching using the OpenMS implementation of MS-GF+ (Granholm et al., 2014) with the *Synechocystis* sp. PCC 6803 reference proteome as database (as of 04 April 2019). Carbamidomethylation was considered as a fixed modification on cysteine and oxidation as a variable modification on methionine. The precursor ion mass window tolerance was set to 10 ppm. The *PeptideIndexer* module was used to annotate peptide hits with their corresponding target or decoy status, *PSMFeatureExtractor* was used to annotate additional characteristics to features, *PercolatorAdapter* was used to estimate the false discovery rate (FDR), and *IDFilter* was used to keep only peptides with q-values lower than 0.01 (1% FDR). The quantification pipeline is based on the *FeatureFinderIdentification* workflow allowing feature propagation between different runs (Weisser and Choudhary, 2017). MzML files were retention time corrected using *MapRTTransformer*, and identifications (idXML files) were combined using the *IDMerger* module. *FeatureFinderIdentification* was then used to generate featureXML files based on all identifications combined from different runs. Individual feature maps were combined to a consensus feature map using *FeatureLinkerUnlabelledKD*, and global intensity was normalized using *ConsensusMapNormalizer* (by median). Protein quantity was determined by summing up the intensities of all unique peptides per protein using *ProteinQuantifier*.

Integrated Analysis of Diurnal mRNA, Ribosome and Protein Oscillations

Analysis was performed using R v.3.6 scripts[1]. Genes with less than 30 and 60 reads were initially removed from the mRNA sequencing dataset and ribosome profiling dataset, respectively. In the proteomics dataset, inaccurately measured proteins were removed by discarding those with a $\log_2$ intensity standard deviation greater than 1 at any time point. Only genes with at least two replicate measurements across all time points and across all three datasets were analyzed. The total number of genes ($n = 1,126$) was mainly limited by the proteomics dataset. Abundance values for each gene were $\log_2$ transformed and then centered around the gene's daily average $\log_2$ abundance to reflect relative fold changes. To identify genes with diurnal changes in mRNA abundance (considered "cyclic"), differential abundance between time points was analyzed with one-way ANOVA. A gene's mRNA abundance was considered to change

---

[2]http://www.ebi.ac.uk/pride

[3]https://github.com/Asplund-Samuelsson/ribopipe

[4]https://github.com/fhcrc/seqmagick

significantly over the day-night cycle if (1) the Benjamini-Hochberg adjusted q-value was less than 0.1, and if (2) the absolute $\log_2$ fold change was greater than 1, between any two time points. Cyclic genes were then clustered into four groups (G1–G4) according to their diurnal mRNA abundance pattern using hierarchical clustering (R function: *hcluster*; distance measure: Pearson correlation; linkage method: Ward). Maximum cluster separation was obtained when choosing a cluster number of 2 and 4 (average silhouette width of 0.63 and 0.49, respectively). Four clusters were chosen as relatively unique diurnal patterns were visible in each, despite lower average silhouette width. Non-cyclic genes were assigned to a fifth group (G0). The peak-to-peak relative amplitude in mRNA, ribosome and protein abundance was calculated as the maximum $\log_2$ fold change between sample mean values across the day-night cycle. The median relative amplitude of cyclic genes in each dataset was used to summarize and compare the overall relative amplitude observed at transcriptional, translational and protein level. Differential $\log_2$ protein abundance across time points was assessed for cyclic and non-cyclic genes using one-way ANOVA. A Benjamini-Hochberg adjusted q-value less than 0.1 was considered significant.

Modeling of Protein Oscillations

The change over time of an arbitrary gene's (J) protein concentration (P_J) was expressed according to the cellular mass balance of that protein:

$$\frac{dP_J}{dt} = F_{S,\,J} \cdot S_{TOT} - F_{P,\,J} \cdot P_{TOT} \cdot \mu - F_{P,\,J} \cdot P_{TOT} \cdot k_{D,\,J} \quad (1)$$

Where $F_{S,J}$ is the fraction of total bulk protein synthesis (S_{TOT}) dedicated to protein J, $F_{P,J}$ is the fraction of the total cellular protein concentration (P_{TOT}) made up by protein J, μ is the growth rate, and $k_{D,J}$ is the gene-specific degradation rate of protein J.

In a similar manner, the rate change of P_{TOT} was expressed according to the cellular mass balance of total protein:

$$\frac{dP_{TOT}}{dt} = S_{TOT} - P_{TOT} \cdot \mu - P_{TOT} \cdot k_{D,\,MEAN} \quad (2)$$

Where $k_{D,MEAN}$ is the bulk protein degradation rate.

Under the assumption of a constant P_{TOT}, the term S_{TOT} is constrained to be proportional to the sum of μ and $k_{D,MEAN}$:

$$S_{TOT} = P_{TOT} \cdot (\mu + k_{D,\,MEAN}) \quad (3)$$

Substitution of Eq. 3 into Eq. 1 yields:

$$\frac{dP_J}{dt} = P_{TOT} \cdot \left(F_{S,\,J} \cdot (\mu + k_{D,\,MEAN}) - F_{P,\,J} \cdot (\mu + k_{D,\,J}) \right) \quad (4)$$

For the model to better reflect the abundance fraction of protein J measured by protein mass spectrometry, the expression $P_J = F_{P,J} \cdot P_{TOT}$ as substituted into Eq. 4:

$$\frac{dF_{P,\,J}}{dt} = F_{S,\,J} \cdot (\mu + k_{D,\,MEAN}) - F_{P,\,J} \cdot (\mu + k_{D,\,J}) \quad (5)$$

Considering an "average" protein J for which $k_{D,J} = k_{D,MEAN}$, Eq. 5 can be further simplified to:

$$\frac{dF_{P,\,J}}{dt} = (\mu + k_{D,\,MEAN}) \cdot (F_{P,\,J} - F_{S,\,J}) \quad (6)$$

Simulations of diurnal protein abundance oscillations were performed using R v.3.6 scripts[1]. The model equation (Eq. 5) was solved numerically using an ordinary differential equation solver (R function: *ode*). $F_{S,J}$ was set as a function of time with an amplitude fold change corresponding to the experimentally observed median value:

$$F_{S,\,J} = (3.05 - 1) / (3.05 + 1) \cdot sin\,(2 \cdot \pi / 24 \cdot t + 0.5\pi\,) + 1.$$

The behavior of $F_{P,J}$ over time in response to different protein turnover scenarios was analyzed by altering the settings of the remaining input parameters μ, $k_{D,MEAN}$ and $k_{D,J}$. To simulate anticorrelated degradation vs. synthesis, $k_{D,J}$ was expressed as a sine function with a 12 h phase shift relative to $F_{S,J}$:

$$k_{D,\,J} = 0.1 \cdot sin\,(2 \cdot \pi / 24 \cdot t + 1.5\pi\,) + 0.1.$$

A time depended growth rate was modeled by expressing μ as a sine function during day time:

$$\mu = 0.05 \cdot sin\,(2 \cdot \pi / 24 \cdot t)$$

and as zero during night time.

RESULTS

Diurnal Transcriptional Oscillations Are Not Dampened by Translation but Protein Levels Are Largely Constant

To investigate whether translational regulation causes protein levels to remain constant during day-night cycles in cyanobacteria, we performed genome-wide measurements of the transcriptome, translatome, and proteome in *Synechocystis* using mRNA sequencing, ribosome profiling, and quantitative shotgun proteomics, respectively. Cells were adapted to an artificial day-night regime for seven days in a controlled turbidostat culture and samples were collected over separate days at five time points: 1 h before and after artificial sunrise, midday, and 1 h before and after sunset. The maximum and average growth rate was 0.05 and 0.018 h^{-1}, respectively, and correlated with the light intensity curve (**Supplementary Figure 2**).

A total of 1126 genes were analyzed, which had at least two replicate measurements across all time points in all three datasets. Of these, 43% showed cyclic diurnal mRNA expression ($|\log_2 FC| > 1$, FDR < 0.1, **Supplementary Table 1**) which is in agreement with microarray-based transcriptomics studies from *Synechocystis* and *Synechococcus elongatus* PCC 7942 grown in diurnal light conditions (Guerreiro et al., 2014; Saha et al., 2016). These genes were designated "cyclic genes" and clustered according to their diurnal mRNA abundance pattern into four groups: "G1-G4." Non-cyclic genes were assigned to a fifth group: "G0" (**Figure 1A**). Protein synthesis rates, inferred from the
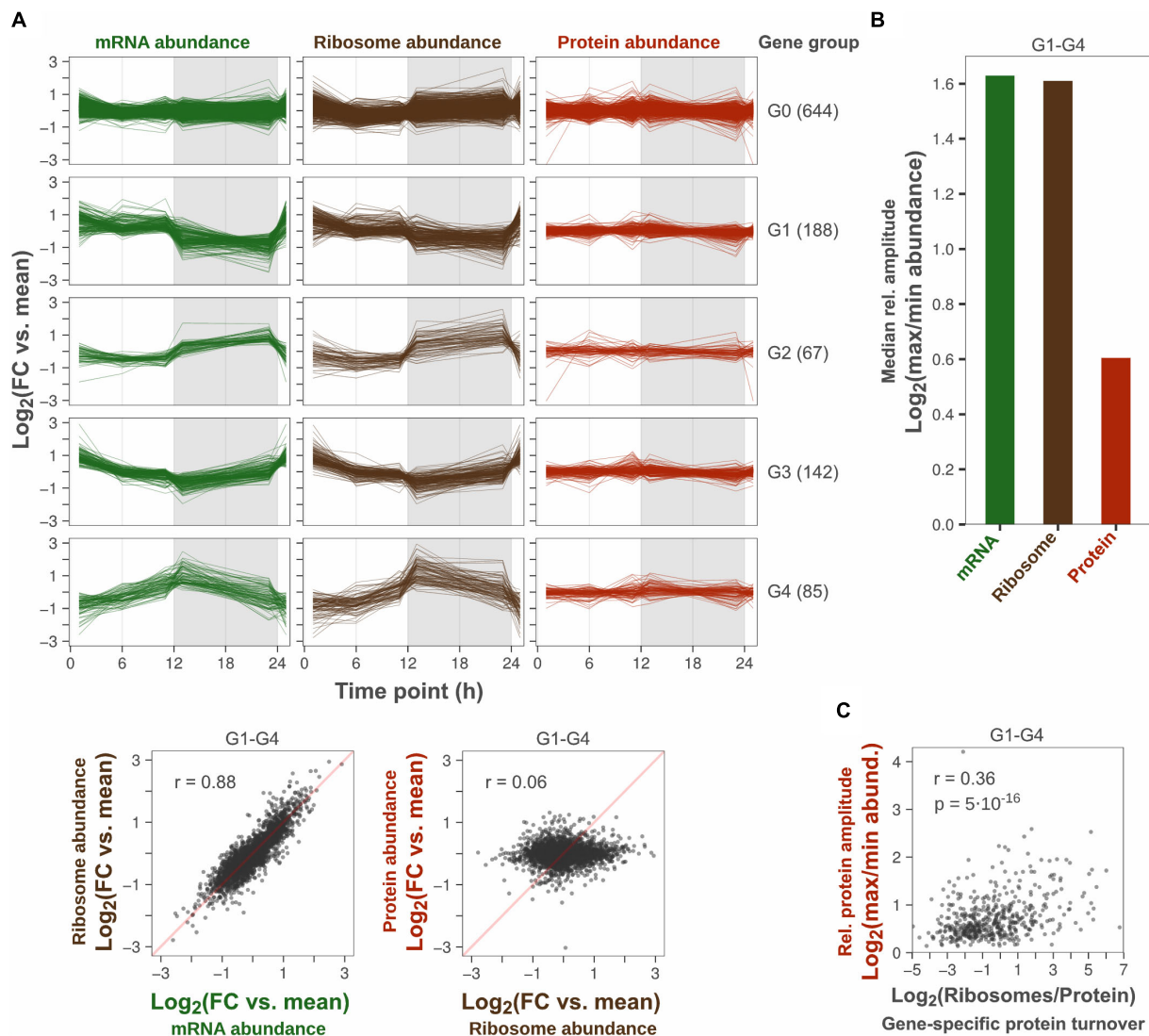
FIGURE 1 | Protein amplitudes over a diurnal cycle are small relative to amplitudes in transcription and translation. **(A)** Comparison of diurnal expression patterns at the level of transcription, translation and protein. The log$_2$-transformed fold change (FC) vs. the daily average gene-specific abundance of mRNA, ribosome and protein was plotted for each gene detected across all three levels (top). Time point 1 was plotted twice to visualize changes at sunrise. Genes were grouped according to their daily mRNA abundance pattern: No significant change (G0), Cyclic (G1–G4). The number of genes in each group is shown in parentheses. Scatter plots show the correlation of diurnal abundance patterns between levels among cyclic genes (bottom). Insets show the Pearson correlation coefficient. Duplicate measurements were performed at all time points, except 1 h after sunrise where ribosome abundance was measured in triplicates. **(B)** Comparison of the median peak-to-peak relative amplitude among cyclic genes. **(C)** Correlation of relative protein amplitude vs. protein synthesis to protein abundance ratio (inferred gene-specific protein turnover) among cyclic genes. The protein synthesis to protein abundance ratio was based on the daily average abundance of ribosomes and protein. Insets show the Pearson correlation coefficient (*r*) and the statistical significance of the trend (*p*-value).

number of translating ribosomes, correlated well with cyclic transcription patterns (*r* = 0.88, **Figure 1A**). This implies a low degree of translational regulation and confirms that protein synthesis rates oscillate significantly over the day night cycle, in concert with transcript levels. In contrast, oscillation patterns could not be distinguished at the protein level which remained relatively constant. Since the variance between time points was low relative to the variance between replicates, no significant change in protein abundance was found across any time point (ANOVA, FDR < 0.1, **Supplementary Table 1**). This further

explains the low correlation (*r* = 0.06) between protein synthesis and protein abundance patterns in this dataset (**Figure 1A**).

To analyze the peak-to-peak oscillation amplitude at the level of transcription, translation and protein, the log$_2$ fold change between the minimum and maximum time point mean abundance of mRNA, translating ribosome, and protein was calculated for each cyclic gene (termed "relative amplitude"). The median relative amplitude across genes was compared to quantify the overall amplitude reduction from protein synthesis to protein abundance (**Figure 1B**). The median amplitude was

two times lower at level of protein (1.5-fold) compared to the level of transcription and translation (3.0-fold), which is similar to the 2.3-fold median ratio between transcript and protein oscillations reported by Waldbauer et al. (2012). However, the median protein amplitude was probably overestimated here since the error of time point means was high relative to the variation between time point means. For example, if the error of time point means are large and the true protein amplitude is small for a gene, the measured variation across time points will be mostly noise. Consequently, the calculated relative amplitude is likely to be mostly noise, as it will be determined from the maximum fold change across five error-prone time point means. The variance between replicates was larger than the variance between time points for 41% of cyclic genes (ANOVA, SSB/SSW > 1). Thus, the determined median relative protein amplitude is likely to provide a certain over-estimated relative amplitude. Even though gene-specific protein amplitudes could not be determined with precision, there was a trend that proteins with high turnover rate (daily mean protein synthesis rate/daily mean protein abundance) had stronger oscillations (**Figure 1C**).

In conclusion, our multi-omics analysis shows that the decrease in oscillation amplitude between the mRNA level and the protein level is not caused by translational regulation of protein synthesis. The ratio between the median mRNA oscillation amplitude and the median protein oscillation amplitude was estimated to be more than twofold.

A Slow Protein Turnover Reduces the Amplitude of Protein Abundance Oscillations

Protein concentrations in cyanobacteria remain largely constant over diurnal cycles, despite significant fluctuations in transcription and protein synthesis. A possible and intuitive explanation is that diurnal peaks in synthesis are counteracted by increased protein degradation. At the same time, this seems unlikely in an evolutionary context, as it implies an ineffective use of cellular resources which would result in decreased fitness. Therefore, we sought to determine whether this observation could solely be the result of a slow protein turnover. For this purpose, we applied a mass balance-based model that describes the change in an arbitrary protein's concentration in response to diurnal synthesis oscillations (see section "Materials and Methods"). The model takes into account the synthesis, degradation, and growth dilution of the modeled protein (protein J) as well as the synthesis, degradation, and growth dilution of the bulk proteome. The model assumes that the total cellular protein concentration is constant, which has been shown experimentally in cyanobacteria over a range of growth rates and genetic perturbations (Touloupakis et al., 2015; Zheng and O'Shea, 2017). This assumption constrains bulk protein synthesis to be proportional to protein depletion, which is the sum of two processes, bulk protein degradation (described by $k_{D,MEAN}$) and dilution by cell growth (described by μ). Bulk protein turnover, defined as bulk protein synthesis rate divided by bulk protein abundance, is thus also proportional to protein depletion (see section "Materials and Methods," Eq. 3). We measured the total

protein concentration in cell extracts and found it to be constant across time points ($p \geq 0.4$, **Supplementary Figure 3**). In any case, deviations from this assumption do not have a significant impact on the amplitude of protein oscillations as long as the change is restored within the time span of the day-night cycle (see section "Discussion").

The effect of protein turnover rate on diurnal protein oscillations was explored using the mathematical model described above. Model parameters were selected to simulate biologically relevant cellular scenarios. The synthesis rate of protein J was set to oscillate with an amplitude equal to the median amplitude determined by ribosome profiling (**Figure 1B**) and μ was set to the observed daily average (**Supplementary Figure 2**). The bulk protein degradation rate was set to the median degradation rate reported for microalgae and plants (0.01 h^{-1}), as it has not been determined experimentally in cyanobacteria (**Table 1**). In a first simulation, the gene-specific degradation rate of protein J ($k_{D,J}$) was set equal to the bulk degradation rate so as to mimic the response of an "average gene." This resulted in a relative protein amplitude of 1.11-fold, which is similar to the relative protein amplitude determined from the experimental data (**Figure 2A**). Increasing the bulk turnover rate ($\mu + k_{D,MEAN}$) from 0.028 to 0.118 h^{-1} (by increasing $k_{D,MEAN}$ from 0.01 to 0.1 h^{-1}) resulted in increased relative amplitude (1.11–1.53-fold) and decreased lag time of a protein J's oscillations. The change in relative amplitude was in this case caused by changes in the absolute protein abundance difference between peak and trough (termed "absolute amplitude," see legend **Figure 2**). A positive correlation between protein amplitude and bulk protein turnover was also predicted implicitly in model Eq. 6, where a higher bulk turnover increases the protein response dP/dt, which leads to a faster change and increased amplitude (see section "Materials and Methods"). Equation 6 further shows that the direction of the protein change is determined by the difference in protein synthesis fraction and protein abundance fraction ($F_{S,J}$–$F_{P,J}$). The abundance fraction will therefore become equal to the synthesis fraction over time, if the synthesis rate of a gene J is constant (e.g., at steady state growth) and $k_{D,J} = k_{D,MEAN}$. More importantly, this implies that a change in synthesis rate from one steady state to a new one, will result in an abundance change that is at most equal to the synthesis change, if given enough time to reach the new steady state ($\sim$5 protein half-lives). Thus, under a diurnally changing synthesis rate, the protein amplitude is bound to be less than (or at most equal to) the protein synthesis amplitude, unless the protein half-life is much shorter than the time period of the day-night cycle (i.e., relatively high protein turnover).

A positive effect on protein J's relative amplitude was also observed when its gene-specific protein turnover rate was high relative to the bulk protein turnover rate (by setting $k_{D,J} > k_{D,MEAN}$, **Figure 2B**). In this case, the increased relative amplitude was a result of reduced daily mean abundance of the modeled protein (and unchanged absolute amplitude), instead of increased absolute amplitude as in **Figure 2A**. Similarly, a gene-specific turnover rate lower than the bulk

TABLE 1 | Reported median protein degradation rates and growth rates in different organisms.

Organism	Median protein deg. rate ($k_{D,MEAN}$, h^{-1})	Growth rate (μ, h^{-1})
Chlamydomonas reinhardtii (Algae)	0.015[a]	0.011[a]
Arabidopsis thaliana (Plant)	0.010[b]	0.0097[b]
	0.0092[c]	0.0013–0.0063[c]
Lactococcus lactis (Heterotrophic bacteria)	0.12–0.91[d]	0.1–0.5[d]
Saccharomyces Cerevisiae (Budding yeast)	0.97[e]	0.46[e]
Human (HeLa cells)	0.02[f*]	0.03[h*]
	0.034[g*]	

[a]*Mastrobuoni et al. (2012)*; [b]*Li et al. (2012)*; [c]*Li et al. (2017)*; [d]*Lahtvee et al. (2014)*; [e]*Belle et al. (2006)*; [f]*Cambridge et al. (2011)*; [g]*Doherty et al. (2009)*; [h]*Kono et al. (2015)*; *Similar growth conditions (Culture medium: DMEM + 10% FPS).
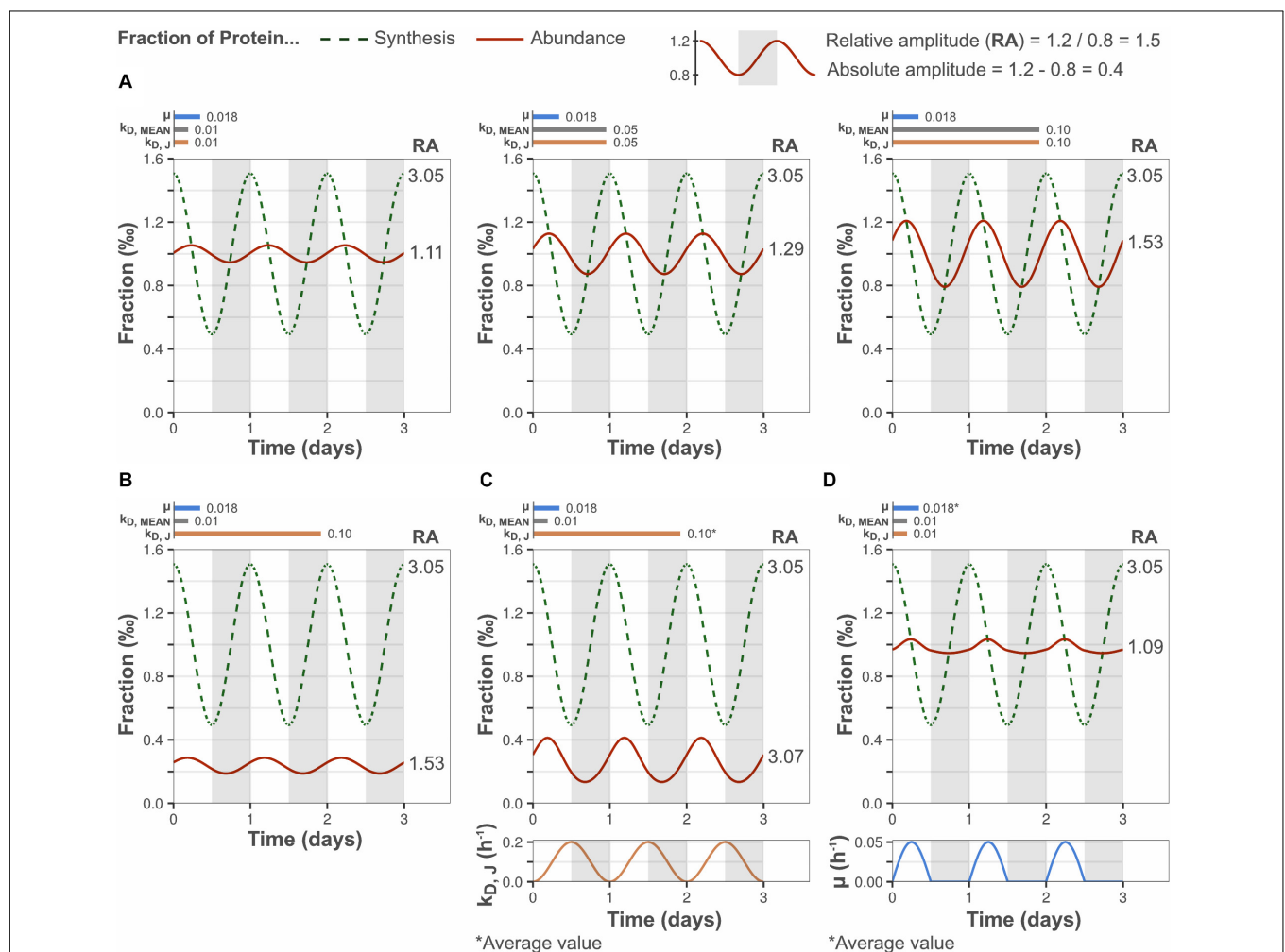
FIGURE 2 | Low protein turnover reduces the relative amplitude of diurnal protein oscillations. Y-axes display the fraction of bulk protein synthesis (dashed lines) and the fraction of bulk protein abundance (solid lines) taken up by a modeled gene J. Bar plots show the selected growth rate (μ), bulk protein degradation rate ($k_{D,MEAN}$) and gene-specific protein degradation rate ($k_{D,J}$). **(A)** Higher bulk protein turnover ($\mu + k_{D,MEAN}$) increased the relative amplitude (RA) of diurnal protein oscillations, by increasing the absolute amplitude. $k_{D,J}$ was set equal to $k_{D,MEAN}$ to simulate the behavior of an "average" gene. **(B)** Increasing the gene-specific turnover, by setting the gene-specific protein degradation rate greater than bulk protein degradation rate ($k_{D,J} > k_{D,MEAN}$; gene-specific turnover > bulk turnover), increased the relative protein amplitude by decreasing the daily average protein abundance. However, the absolute protein amplitude was not affected (compare **A**, left panel). **(C)** A high and fluctuating gene-specific protein degradation that is anticorrelated to the protein synthesis rate increased the relative protein amplitude by increased absolute protein amplitude, and by reduced daily average protein abundance. **(D)** A diurnally fluctuating growth rate had no significant effect on the protein amplitude, although the diurnal pattern was altered. The daily average growth rate was equal to the set value in **Figure 1A**. The growth rate curve was based on the experimentally determined pattern.

degradation rate ($k_{D,J} < k_{D,MEAN}$) results in increased daily mean abundance and reduced relative amplitude (data not shown). A positive correlation between the daily mean protein synthesis to abundance ratio and the gene-specific protein turnover, as determined by $k_{D,J}$ in the model, is consistent with the general definition of protein turnover (turnover rate = synthesis rate/protein abundance, at constant protein abundance). Thus, the model predicts a positive trend between the relative protein amplitude and the gene-specific protein turnover rate, which possibly explains the experimentally observed positive trend between relative protein amplitude and gene-specific turnover (**Figure 1C**). The model further suggests that high-amplitude proteins are likely to possess a high gene-specific degradation rate relative to the bulk protein degradation rate. Increased absolute amplitude was also observed when the gene-specific degradation rate was actively regulated (time-dependent) and anticorrelated to the synthesis rate (**Figure 2C**). Thus, particularly high relative amplitudes are possible for a subset of genes even at slow bulk protein turnover, if regulated degradation is fast enough to also reduce the daily mean abundance. Allowing the growth rate to fluctuate according to our experimental data altered the pattern of protein oscillations but did not have a significant effect on the protein amplitude (**Figure 2D**).

These results demonstrate that the relative amplitude of a protein depends on the bulk protein turnover and the protein's specific turnover. The bulk protein turnover acts on the protein's absolute amplitude (and all other proteins), while the latter acts on the daily mean abundance of the protein. The model further shows that the observed reduction in oscillation amplitude of a given protein can be attributed solely to a low bulk protein turnover, corresponding to the experimentally determined growth rate of 0.018 h^{-1} and a bulk protein degradation rate of 0.01–0.05 h^{-1}.

DISCUSSION

Post-transcriptional regulation is an intuitive explanation for the discrepancy between cyclic diurnal transcription and relatively constant protein levels in cyanobacteria. Our transcriptomic, translatomic and proteomic data confirmed this discrepancy and showed that it is not caused by translational regulation. In addition, modeling of the protein response to transcriptional oscillations under biologically relevant parameter settings demonstrated that the experimentally observed decrease in protein oscillation amplitude can be attributed to a slow bulk protein turnover, without the requirement of regulated protein degradation that counteracts transcriptional oscillations. Modeling results further suggested that the bulk protein degradation rate was similar to the daily average growth rate.

The strong correlation between ribosome and mRNA abundance fold changes indicates that protein synthesis oscillates significantly over the day-night cycle and that translation is not regulated between time points (**Figure 1A**). Synthesis rates were solely based on the ribosome abundance and did not account for within-gene changes in ribosome elongation rate. However, elongation rates were not expected to change significantly on global level between time points, since elongation rates primarily depend on gene-specific properties of the mRNA structure (Riba et al., 2019). Furthermore, variation in elongation rate would more likely result in reduced correlation with mRNA abundance.

In contrast, diurnal protein abundance patterns generally did not show a clear cyclic behavior and did not correlate with protein synthesis oscillations (**Figure 1A**). Small cyclic patterns were most likely present, but concealed by technical variation and therefore not detectable. As measurement errors were high relative to diurnal changes in protein abundance, the determined median relative protein amplitude of 1.5 was probably overestimated (**Figure 1B**). The proteome-wide 2.0-fold reduction in amplitude from synthesis to abundance, was comparable to the 2.3-fold reduction determined previously with higher statistical power (Waldbauer et al., 2012). With this approximate ratio taken into account, our model suggests that the bulk degradation rate was in the range of 0.01–0.05 h^{-1}, i.e., similar to the daily average growth rate, and in line with degradation rates measured in microalgae and plants (**Figure 2A**). This was further supported by bulk degradation rates measured in other organisms which are typically in the same magnitude as the growth rate (**Table 1**). The positive correlation between growth rate and bulk protein degradation has been attributed to a high energy burden of protein turnover when nutrients are limited (Lahtvee et al., 2014).

Our modeling analysis showed that the bulk protein turnover rate (proportional to $\mu + k_{D,MEAN}$) determines the proteome-wide reduction in amplitude between the synthesis level and the abundance level (mean synthesis:protein amplitude ratio). The model further suggested that gene-specific deviations from the mean synthesis:protein amplitude ratio are determined by deviations in individual protein degradation rates relative to the bulk degradation rate (**Figure 2B**). Waldbauer et al. (2012) reported variation in the synthesis:protein amplitude ratio (synthesis = mRNA level) across the genome of *Prochlorococcus* MED4 during diurnal growth. While the vast majority of genes in this study also exhibited low amplitude or no oscillations at the protein level, approximately 30 of the 548 analyzed proteins showed an amplitude fold change greater than 2. However, the relatively high amplitude of these proteins was not caused by particularly strong oscillations in protein synthesis relative to other genes. Instead, protein synthesis oscillations of these genes appeared to be less dampened at the level of protein relative to other genes, as indicated by a lower synthesis:protein amplitude ratio. Our model suggests that such outlier proteins are subjected to a high gene-specific degradation rate (i.e., gene-specific protein turnover), which increases the relative amplitude of oscillations by reducing the protein's daily mean abundance without affecting the absolute amplitude. This was further indicated in our experimental data (**Figure 1C**), where a positive trend between the relative protein amplitude and gene-specific protein turnover (daily mean synthesis rate/daily mean

abundance) was detected. A degradation rate for a given protein that is 10-fold higher than the bulk degradation rate is not unrealistic, as gene-specific degradation rates were shown to span two to three orders of magnitude in *Lactococcus lactis* (Lahtvee et al., 2014). Furthermore, artificially increasing degradation rate, by fusing a ssrA degradation peptide, increased the relative amplitude and decreased the phase shift of a diurnally expressed yellow fluorescent protein in *Synechococcus elongatus* PCC 7942 (Chabot et al., 2007).

The protein oscillation model assumes a constant cellular protein concentration. This assumption was largely satisfied over the day-night cycle, according to measurements of total protein content in cell extracts. The assumption of a constant cellular protein concentration constrains bulk protein synthesis to be proportional to the sum of bulk protein degradation and growth dilution. Consequently, a decreasing protein concentration during night time will lead to an overestimated bulk protein synthesis rate by the model. This will in turn result in an overestimated rate change of each protein's (J) concentration during night time. However, as the cellular protein concentration increases to its original level during sunrise, the opposite effect will occur. That means bulk protein synthesis will be underestimated and the rate change of each protein's concentration will be underestimated, which compensates for the overestimated rate change during the night. Thus, small changes in cellular protein concentration will not change the simulated protein amplitude significantly, but rather alter the diurnal pattern of protein abundance. This is analogous to the effect of setting a constant growth rate vs. setting a fluctuating growth rate (**Figure 2D**).

Cyclic transcription has been shown to peak near time points of the day-night cycle when the corresponding function is expected to be needed by the cell (Waldbauer et al., 2012; Beck et al., 2014; Saha et al., 2016; Strenkert et al., 2019). However, the regulatory purpose of a diurnally shifting transcriptome appears less meaningful, since the impact on the functional protein level is significantly diminished. It is nonetheless possible that well-timed, yet small, changes in protein abundance results in a growth benefit that increases survival fitness in a natural environment. Furthermore, our model shows that these changes would become increasingly relevant in a condition that permits higher growth rates, such as an eutrophicated lake exposed to intense sunlight (**Figure 2A**, right). Indeed, *Synechocystis* can grow with a growth rate as high as 0.16 h^{-1} (van Alphen et al., 2018). This growth rate would correspond to a daily average protein turnover (μ + k$_{D,MEAN}$) of approximately 0.12 h^{-1}, considering a diurnal growth pattern and that the bulk degradation rate is typically similar and dependent on the growth rate (**Table 1**). Protein levels in cyanobacteria do change significantly in response to changes in light intensity, if allowed to adjust to a steady state (Jahn et al., 2018). Yet, during diurnal growth, the co-occurrence of a largely constant proteome and considerable metabolic shifts suggests that allosteric interactions play an important regulatory role. For example, CO$_2$ fixation is inactivated during the night through an allosteric mechanism where the regulatory protein CP-12 binds and inactivates the

Calvin cycle enzymes phosphoribulokinase and glyceraldehyde-3-phosphate dehydrogenase (Tamoi et al., 2005). Glycogen degradation is another potential target of allosteric regulation since it mostly occurs during the night, even though the abundance of glycogen phosphorylase does not change over the day-night cycle (**Supplementary Table 1**).

Our results also have implications for synthetic biology in cyanobacteria. There have been many efforts to control the abundance of heterologous proteins in *Synechocystis,* at both the level of translation, through alteration of RBS sequence (Thiel et al., 2018), and at the level of degradation, through a synthetic ssrA peptide with a calculated homology to the native sequence (Landry et al., 2013). The perceived ribosome binding site affinity is not an accurate predictor of protein levels, even when comparing ribosome binding sites with the same heterologous protein (Thiel et al., 2018). It is possible that ribosome profiling, which provides a measure of ribosome occupancy across the entire transcript, could provide insight as to how genetic context affects translation of heterologous proteins. The findings in this study suggest that faster changes in a heterologously expressed protein's abundance can be achieved, if its synthesis rate and degradation rate is high, i.e., if its gene-specific protein turnover is high. In case transcription of the heterologous gene is from a promoter that has an inherent oscillation, then an increased degradation rate, through e.g., a strong degradation tag, could increase oscillations in the protein level. At the same time, a slow bulk protein turnover will extend the time needed for that protein to reach its steady-state abundance, since the cellular protein space is limited. This appears to be the case in cyanobacteria cultures grown at constant light. In a study on the induction kinetics of YFP from various promoters in *Synechocystis*, the protein accumulated for five days after induction with rhamnose before reaching a steady state (Behle et al., 2020). In day/night cultivations, the change in the target's protein abundance will be slower still, as total transcription and/or translation is globally downregulated at night, by inactivation of RNA polymerases and/or ribosomes (Hood et al., 2016). Therefore, comparisons of gene expression constructs, such as promoters or ribosome binding sites, should occur only after steady-state has been reached.

In conclusion, we show that the relatively constant proteome during diurnal growth can be explained by low protein turnover. A relatively high bulk protein turnover is required to obtain significant diurnal changes at the global proteome level. To minimize protein turnover energy costs and improve fitness under growth limited conditions, cyanobacteria may instead have evolved allosteric mechanisms to regulate metabolic shifts. Such adaptation may be particularly relevant for photosynthetic organisms as their energy supply is limited to times of the day with sunlight exposure. Identifying potential allosteric regulation of key enzymes in cyanobacteria could assist future metabolic engineering attempts to accelerate carbon fixation or divert metabolic flux, as these enzymes could become targets for protein engineering. Incorporating allosteric regulation into metabolic models would also improve their prediction capability when simulating genetic knockouts that result in altered

metabolic flux patterns. Furthermore, our results suggest that changes in transcription or translation are not necessarily a good predictor of diurnal changes in enzyme concentration, or metabolic flux.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: ProteomeXchange dataset PXD023812, https://www.ncbi.nlm.nih.gov/bioproject/PRJEB42778.

AUTHOR CONTRIBUTIONS

EH and JK conceived and designed the experiments. DV, JK, and MJ performed the experiments. JA-S, JK, and MJ conducted the computational data analysis. All authors contributed to the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.657379/full#supplementary-material

REFERENCES

Adusumilli, R., and Mallick, P. (2017). Data conversion with proteoWizard msConvert. *Methods Mol. Biol.* 1550, 339–368. doi: 10.1007/978-1-4939-6747-6_23

Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data.* Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [Accessed January 21, 2021]

Angermayr, S. A., van Alphen, P., Hasdemir, D., Kramer, G., Iqbal, M., van Grondelle, W., et al. (2016). Culturing of Synechocystis sp. PCC6803 with N2 /CO2 in a diel regime shows multi-phase glycogen dynamics and low maintenance costs. *Appl. Environ. Microbiol.* 2016, 256–216. doi: 10.1128/AEM.00256-16

Beck, C., Hertel, S., Rediger, A., Lehmann, R., Wiegard, A., Kölsch, A., et al. (2014). Daily expression pattern of protein-encoding genes and small noncoding RNAs in synechocystis sp. strain PCC 6803. *Appl. Environ. Microbiol.* 80, 5195–5206. doi: 10.1128/AEM.01086-14

Becker, A. H., Oh, E., Weissman, J. S., Kramer, G., and Bukau, B. (2013). Selective ribosome profiling as a tool for studying the interaction of chaperones and targeting factors with nascent polypeptide chains and ribosomes. *Nat. Protoc.* 8, 2212–2239. doi: 10.1038/nprot.2013.133

Behle, A., Saake, P., Germann, A. T., Dienst, D., and Axmann, I. M. (2020). Comparative Dose–Response Analysis of Inducible Promoters in Cyanobacteria. *ACS Synthetic Biol.* 9, 843–855. doi: 10.1021/acssynbio.9b00505

Belle, A., Tanay, A., Bitincka, L., Shamir, R., and O'Shea, E. K. (2006). Quantification of protein half-lives in the budding yeast proteome. *Proc. Natl. Acad. Sci. U. S. A.* 103, 13004–13009. doi: 10.1073/pnas.0605420103

Brar, G. A., and Weissman, J. S. (2015). Ribosome profiling reveals the what, when, where and how of protein synthesis. *Nat. Rev. Mol. Cell Biol.* 16, 651–664. doi: 10.1038/nrm4069

Cambridge, S. B., Gnad, F., Nguyen, C., Bermejo, J. L., Krüger, M., and Mann, M. (2011). Systems-wide proteomic analysis in mammalian cells reveals conserved, functional protein turnover. *J. Proteome Res.* 10, 5275–5284. doi: 10.1021/pr101183k

Chabot, J. R., Pedraza, J. M., Luitel, P., and Van Oudenaarden, A. (2007). Stochastic gene expression out-of-steady-state in the cyanobacterial circadian clock. *Nature* 450, 1249–1252. doi: 10.1038/nature06395

Doherty, M. K., Hammond, D. E., Clague, M. J., Gaskell, S. J., and Beynon, R. J. (2009). Turnover of the human proteome: Determination of protein intracellular stability by dynamic SILAC. *J. Proteome Res.* 8, 104–112. doi: 10.1021/pr800641v

Granholm, V., Kim, S., Navarro, J. C. F., Sjölund, E., Smith, R. D., and Käll, L. (2014). Fast and accurate database searches with MS-GF+percolator. *J. Proteome Res.* 13, 890–897. doi: 10.1021/pr400937n

Guerreiro, A. C. L., Benevento, M., Lehmann, R., van Breukelen, B., Post, H., Giansanti, P., et al. (2014). Daily Rhythms in the Cyanobacterium Synechococcus elongatus Probed by High-resolution Mass Spectrometry–based Proteomics Reveals a Small Defined Set of Cyclic Proteins. *Mol. Cell. Proteomics* 13, 2042–2055. doi: 10.1074/mcp.M113.035840

Hood, R. D., Higgins, S. A., Flamholz, A., Nichols, R. J., and Savage, D. F. (2016). The stringent response regulates adaptation to darkness in the cyanobacterium *Synechococcus elongatus. Proc. Natl. Acad. Sci. U. S. A.* 113, 4867–4876. doi: 10.1073/pnas.1524915113

Ito, H., Mutsuda, M., Murayama, Y., Tomita, J., Hosokawa, N., and Terauchi, K. (2009). Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in Synechococcus elongatus. *Proc. Natl. Acad. Sci. U. S. A.* 106, 14168–14173. doi: 10.1073/pnas.0902587106

Jahn, M., Vialas, V., Karlsen, J., Maddalo, G., Edfors, F., Forsström, B., et al. (2018). Growth of Cyanobacteria Is Constrained by the Abundance of Light and Carbon Assimilation Proteins. *Cell Rep.* 25, 478.e–486.e. doi: 10.1016/j.celrep.2018.09.040

Joshi, N., and Fass, J. (2011). *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33).* Available online at: https://github.com/najoshi/sickle [Accessed January 21, 2021].

Karlsen, J., Asplund-samuelsson, J., Thomas, Q., Jahn, M., Elton, P., Plant, C., et al. (2018). Ribosome profiling of Synechocystis reveals altered ribosome allocation at carbon starvation. *mSystems* 3, e126–e118. doi: 10.1128/mSystems.00126-18

Kono, K., Takada, N., Yasuda, S., Sawada, R., Niimi, S., Matsuyama, A., et al. (2015). Characterization of the cell growth analysis for detection of immortal cellular impurities in human mesenchymal stem cells. *Biologicals* 43, 146–149. doi: 10.1016/j.biologicals.2014.11.007

Lahtvee, P. J., Seiman, A., Arike, L., Adamberg, K., and Vilu, R. (2014). Protein turnover forms one of the highest maintenance costs in Lactococcus lactis. *Microbiology* 160, 1501–1512. doi: 10.1099/mic.0.078089-0

Landry, B. P., Stöckel, J., and Pakrasi, H. B. (2013). Use of Degradation Tags To Control Protein Levels in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803. *Appl. Environ. Microbiol.* 79, 2833–2835. doi: 10.1128/AEM.03741-12

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25–R25. doi: 10.1186/gb-2009-10-3-r25

Li, L., Nelson, C. J., Solheim, C., Whelan, J., and Millar, A. H. (2012). Determining degradation and synthesis rates of arabidopsis proteins using the kinetics of

progressive 15N labeling of two-dimensional gel-separated protein spots. *Mol. Cell. Proteomics* 11, 1–16. doi: 10.1074/mcp.M111.010025

Li, L., Nelson, C. J., Trösch, J., Castleden, I., Huang, S., and Millar, A. H. (2017). Protein degradation rate in Arabidopsis thaliana leaf growth and development. *Plant Cell* 29, 207–228. doi: 10.1105/tpc.16.00768

Liu, T. Y., Huang, H. H., Wheeler, D., Xu, Y., Wells, J. A., Song, Y. S., et al. (2017). Time-Resolved Proteomics Extends Ribosome Profiling-Based Measurements of Protein Synthesis Dynamics. *Cell Syst.* 4, 636.e–644.e. doi: 10.1016/j.cels.2017.05.001

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12. doi: 10.14806/ej.17.1.200

Mastrobuoni, G., Irgang, S., Pietzke, M., Aßmus, H. E., Wenzel, M., Schulze, W. X., et al. (2012). Proteome dynamics and early salt stress response of the photosynthetic organism Chlamydomonas reinhardtii. *BMC Genomics* 13:1–13. doi: 10.1186/1471-2164-13-215

Reimers, A. M., Knoop, H., Bockmayr, A., and Steuer, R. (2017). Cellular trade-offs and optimal resource allocation during cyanobacterial diurnal growth. *Proc. Natl. Acad. Sci. U. S. A.* 114, E6457–E6465. doi: 10.1073/pnas.1617508114

Riba, A., Di Nanni, N., Mittal, N., Arhné, E., Schmidt, A., and Zavolan, M. (2019). Protein synthesis rates and ribosome occupancies reveal determinants of translation elongation rates. *Proc. Natl. Acad. Sci.* 116, 15023–15032. doi: 10.1073/pnas.1817299116

Röst, H. L., Sachsenberg, T., Aiche, S., Bielow, C., Weisser, H., Aicheler, F., et al. (2016). OpenMS: A flexible open-source software platform for mass spectrometry data analysis. *Nat. Methods* 13, 741–748. doi: 10.1038/nmeth.3959

Saha, R., Liu, D., Hoynes-O'Connor, A., Liberton, M., Yu, J., Bhattacharyya-Pakrasi, M., et al. (2016). Diurnal regulation of cellular processes in the cyanobacterium Synechocystis sp. strain PCC 6803: Insights from transcriptomic, fluxomic, and physiological analyses. *MBio* 7, 1–14. doi: 10.1128/mBio.00464-16

Stöckel, J., Jacobs, J. M., Elvitigala, T. R., Liberton, M., Welsh, E. A., Polpitiya, A. D., et al. (2011). Diurnal rhythms result in significant changes in the cellular protein complement in the cyanobacterium Cyanothece 51142. *PLoS One* 6:0016680. doi: 10.1371/journal.pone.0016680

Stöckel, J., Welsh, E. A., Liberton, M., Kunnvakkam, R., Aurora, R., and Pakrasi, H. B. (2008). Global transcriptomic analysis of Cyanothece 51142 reveals robust diurnal oscillation of central metabolic processes. *Proc. Natl. Acad. Sci. U. S. A.* 105, 6156–6161. doi: 10.1073/pnas.0711068105

Strenkert, D., Schmollinger, S., Gallaher, S. D., Salomé, P. A., Purvine, S. O., Nicora, C. D., et al. (2019). Multiomics resolution of molecular events during a day in the life of Chlamydomonas. *Proc. Natl. Acad. Sci. U. S. A.* 116, 2374–2383. doi: 10.1073/pnas.1815238116

Tamoi, M., Miyazaki, T., Fukamizo, T., and Shigeoka, S. (2005). The Calvin cycle in cyanobacteria is regulated by CP12 via the NAD(H)/NADP(H) ratio under light/dark conditions. *Plant J.* 42, 504–513. doi: 10.1111/j.1365-313X.2005.02391.x

Tange, O. (2011). GNU Parallel: the command-line power tool.;login. *USENIX Mag.* 36, 42–47. doi: 10.5281/zenodo.16303

Thiel, K., Mulaku, E., Dandapani, H., Nagy, C., Aro, E. M., and Kallio, P. (2018). Translation efficiency of heterologous proteins is significantly affected by the genetic context of RBS sequences in engineered cyanobacterium Synechocystis sp. PCC 6803. *Microb. Cell Fact.* 17:34. doi: 10.1186/s12934-018-0882-2

Touloupakis, E., Cicchi, B., and Torzillo, G. (2015). A bioenergetic assessment of photosynthetic growth of Synechocystis sp. PCC 6803 in continuous cultures. *Biotechnol. Biofuels* 8, 1–11. doi: 10.1186/s13068-015-0319-7

van Alphen, P., Abedini Najafabadi, H., Branco, dos Santos, F., and Hellingwerf, K. J. (2018). Increasing the Photoautotrophic Growth Rate of Synechocystis sp. PCC 6803 by Identifying the Limitations of Its Cultivation. *Biotechnol. J.* 13, 1–8. doi: 10.1002/biot.201700764

Waldbauer, J. R., Rodrigue, S., Coleman, M. L., and Chisholm, S. W. (2012). Transcriptome and Proteome Dynamics of a Light-Dark Synchronized Bacterial Cell Cycle. *PLoS One* 7:0043432. doi: 10.1371/journal.pone.0043432

Weisser, H., and Choudhary, J. S. (2017). Targeted Feature Detection for Data-Dependent Shotgun Proteomics. *J. Proteome Res.* 16, 2964–2974. doi: 10.1021/acs.jproteome.7b00248

Welkie, D. G., Rubin, B. E., Diamond, S., Hood, R. D., Savage, D. F., and Golden, S. S. (2019). A Hard Day's Night: Cyanobacteria in Diel Cycles. *Trends Microbiol.* 27, 231–242. doi: 10.1016/j.tim.2018.11.002

Werner, A., Broeckling, C. D., Prasad, A., and Peebles, C. A. M. (2019). A comprehensive time-course metabolite profiling of the model cyanobacterium Synechocystis sp. PCC 6803 under diurnal light:dark cycles. *Plant J.* 99, 379–388. doi: 10.1111/tpj.14320

Zheng, X. Y. U., and O'Shea, E. K. (2017). Cyanobacteria Maintain Constant Protein Concentration despite Genome Copy-Number Variation. *Cell Rep.* 19, 497–504. doi: 10.1016/j.celrep.2017.03.067

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Multi-Omic Analyses Reveal Habitat Adaptation of Marine Cyanobacterium *Synechocystis* sp. PCC 7338

Yujin Jeong[1], Seong-Joo Hong[2,3], Sang-Hyeok Cho[1], Seonghoon Yoon[2,3], Hookeun Lee[4], Hyung-Kyoon Choi[5], Dong-Myung Kim[6], Choul-Gyun Lee[2,3], Suhyung Cho[1,7]* and Byung-Kwan Cho[1,7,8]*

[1] Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon, South Korea, [2] Department of Biological Engineering, Inha University, Incheon, South Korea, [3] Department of Biological Sciences and Bioengineering, Inha University, Incheon, South Korea, [4] Institute of Pharmaceutical Research, College of Pharmacy, Gachon University, Incheon, South Korea, [5] College of Pharmacy, Chung-Ang University, Seoul, South Korea, [6] Department of Chemical Engineering and Applied Chemistry, Chungnam National University, Daejeon, South Korea, [7] Innovative Biomaterials Center, KI for the BioCentury, Korea Advanced Institute of Science and Technology, Daejeon, South Korea, [8] Intelligent Synthetic Biology Center, Daejeon, South Korea

Cyanobacteria are considered as promising microbial cell factories producing a wide array of bio-products. Among them, *Synechocystis* sp. PCC 7338 has the advantage of growing in seawater, rather than requiring arable land or freshwater. Nonetheless, how this marine cyanobacterium grows under the high salt stress condition remains unknown. Here, we determined its complete genome sequence with the embedded regulatory elements and analyzed the transcriptional changes in response to a high-salt environment. Complete genome sequencing revealed a 3.70 mega base pair genome and three plasmids with a total of 3,589 genes annotated. Differential RNA-seq and Term-seq data aligned to the complete genome provided genome-wide information on genetic regulatory elements, including promoters, ribosome-binding sites, 5′- and 3′-untranslated regions, and terminators. Comparison with freshwater *Synechocystis* species revealed *Synechocystis* sp. PCC 7338 genome encodes additional genes, whose functions are related to ion channels to facilitate the adaptation to high salt and high osmotic pressure. Furthermore, a ferric uptake regulator binding motif was found in regulatory regions of various genes including SigF and the genes involved in energy metabolism, suggesting the iron-regulatory network is connected to not only the iron acquisition, but also response to high salt stress and photosynthesis. In addition, the transcriptomics analysis demonstrated a cyclic electron transport through photosystem I was actively used by the strain to satisfy the demand for ATP under high-salt environment. Our comprehensive analyses provide pivotal information to elucidate the genomic functions and regulations in *Synechocystis* sp. PCC 7338.

Keywords: cyanobacteria, marine *Synechocystis* sp., photosynthesis, genome, transcriptome

INTRODUCTION

Cyanobacteria are the bacteria that generate energy through oxygenic photosynthesis (Camsund and Lindblad, 2014; Hitchcock et al., 2020). In addition to the photosynthetic capability, cyanobacteria, owing to their rapid growth and applicability in genetic engineering, are considered as potential industrial hosts for the production of value-added biochemicals (Lindberg et al., 2010; Ducat et al., 2011; Lan and Liao, 2011; Oliver et al., 2013; Savakis et al., 2013; Singh et al., 2016). Among those cyanobacterial hosts, freshwater cyanobacteria *Synechocystis* sp. PCC 6803 has been intensively studied to understand its genetic traits and metabolic networks, thus providing insights for effective engineering of the strain (Mitschke et al., 2011; Kopf et al., 2014; Hernandez-Prieto et al., 2016; Jablonsky et al., 2016; Vavitsas et al., 2017). Compared to this, *Synechocystis* sp. PCC 7338 is a marine cyanobacterium that grows under stressful environmental conditions characterized by high salinity and high osmotic pressure. In a recent metabolic and lipidomic profiling study, *Synechocystis* sp. PCC 7338 was reported to contain more unsaturated fatty acids, phosphatidylglycerol, and amino acids than *Synechocystis* sp. PCC 6803, which was suggested to be the result of adapting to the high-salt environment (Noh et al., 2020). Considering that seawater accounts for approximately 96.5% of the water on earth, and is a promising alternative for large-scale cultivation of cyanobacteria (Cui et al., 2020), these features make *Synechocystis* sp. PCC 7338 an attractive host for the mass production of commercially valuable bioproducts. Although several marine *Synechococcus* species, such as *Synechococcus* sp. PCC 7002 and *Synechococcus* sp. PCC 11901, have been used as bioproduction hosts, no marine *Synechocystis* species has been used. *Synechococcus* is an obligate photoautotroph, whereas *Synechocystis* is a facultative photoautotroph and has the advantage of increasing biomass and bioproduction in the presence of glucose (Varman et al., 2013). Moreover, it has been shown that Synechocystis is exceptional at growth and nutrient removal in waste treatment owing to the mixotrophic nature (Trentin et al., 2019). However, essential information on utilizing *Synechocystis* sp. PCC 7338 as an engineering host, including its genome sequence and regulatory elements is yet to be unraveled.

With the development of advanced next-generation sequencing techniques, recent studies have provided genome-wide information on the bacteria. In addition to the RNA-seq method that allows a systematic measurement of the transcriptome changes within the cells, techniques for cataloging the genome architecture have also been developed. In particular, the differential RNA-seq (dRNA-seq) method enabled the identification of the transcription start sites (TSSs), thereby revealing diverse regulatory elements, such as the promoters, 5′-untranslated regions (5′-UTRs), and small RNAs (sRNAs) (Qiu et al., 2010; Sharma et al., 2010; Kim et al., 2012; Jeong et al., 2016). Term-seq method has also been proven to be an effective tool for identifying genome-wide transcript 3′-end positions (TEPs) and 3′-untranslated regions (3′-UTRs) in various organisms (Dar et al., 2016a,b; Lee et al., 2019). In this study, the genome architecture and regulatory features of

Synechocystis sp. PCC 7338 were identified with the detection of genome-wide TSSs and TEPs. We also describe the differences in the gene sets and the unique energy generation strategy in *Synechocystis* sp. PCC 7338 that facilitates its habitat adaptation compared to freshwater *Synechocystis* sp. PCC 6803.

MATERIALS AND METHODS

Cell Growth

Synechocystis sp. PCC 7338 and *Synechocystis* sp. PCC 6803 cells were cultured under continuous illumination at 30 μmol photons m^{-2} s^{-1} at 30°C, and aerated with 2% CO_2 balanced air (flow rate of 0.1 V–V^{-1} min^{-1}). *Synechocystis* sp. PCC 7338 cells were cultured in ASN-III medium composed of 3.50 g MgSO$_4$·7H$_2$O, 2.00 g MgCl$_2$·6H$_2$O, 0.50 g CaCl$_2$·2H$_2$O, 0.75 g NaNO$_3$, 0.015 g K$_2$HPO$_4$, 0.04 g Na$_2$CO$_3$, 0.003 g citric acid, 0.003 g (NH$_4$)$_5$Fe citrate, 0.50 g KCl, 25.00 g NaCl, 0.0005 g EDTA K$_2$Mg·2H$_2$O, and 1 mL of trace metal in 1 L distilled water. The composition of trace metal was as follow: 2.86 g H$_3$BO$_3$, 1.81 g MnCl$_2$·4H$_2$O, 0.22 g ZnSO$_4$·7H$_2$O, 0.39 g Na$_2$MoO$_4$·2H$_2$O, 0.079 g CuSO$_4$·5H$_2$O, and 0.049 g Co(NO$_3$)$_2$·6H$_2$O in 1 L distilled water. The pH of the ASN-III medium was adjusted to 7.4. Cells in the mid-exponential phase of growth (fresh cell weight of 0.91 $\pm$ 0.051 gl^{-1}) were harvested for library construction. *Synechocystis* sp. PCC 6803 cells were cultured in BG-11 medium, and cells in the mid-exponential phase of growth (optical density at 730 nm = 0.8) were harvested for library construction.

Genomic DNA and RNA Extraction

For genomic DNA extraction, the cells were collected by centrifugation at 4°C for 10 min at 3,000 × g, and the cell pellet was resuspended in 1 mL of lysis buffer composed of 10 mM Tris-HCl (pH 7.6), 5 mM MgCl$_2$, and 40 mM NaCl. The resuspended cells were then dripped into a mortar filled with liquid nitrogen and grounded with a pestle. The powdered cells were thawed, and the cell debris was removed by centrifugation at 4°C for 5 min at 3,000 × g. The supernatant was further clarified and collected by centrifugation at 4°C for 10 min at 16,000 × g. The collected lysate was used for construction of the genome sequencing library. Genomic DNA was prepared using a genomic DNA extraction kit (Promega) according to the manufacturer's protocol. For RNA-seq, dRNA-seq, and Term-seq library construction, the cells were resuspended in 1 mL solution composed of 25 mM Tris-HCl (pH 8.0), 10 mM EDTA, 50 mM glucose, and 2 mg/mL lysozyme (Sigma-Aldrich) and incubated at 30°C for 10 min. The cell pellet was collected by centrifugation at 4°C for 10 min at 16,000 × g, and then resuspended in 1 mL ice-cold solution composed of 50 mM sodium acetate (pH 5.3) and 10 mM EDTA. The resuspended sample was mixed with 100 μL of 10% sodium dodecyl sulfate (SDS) and 1.2 mL of phenol-chloroform mixture in the ratio of 5:1. The sample was incubated at 65°C for 5 min with periodic vortexing every 1 min. After centrifugation at 4°C for 20 min at 16,000 × g, 700 μL supernatant was mixed with 700 μL phenol-chloroform (5:1) solution. The sample was centrifuged at 4°C

for 20 min at 16,000 × g, and the supernatant was used for total RNA extraction by isopropanol precipitation. To remove genomic DNA, the isolated RNA was incubated at 37°C for 1 h with 2 U of DNase I (New England Biolabs) and 5 µL of 10 × DNase I buffer (New England Biolabs). The RNA devoid of any DNA was purified by phenol-chloroform extraction and ethanol precipitation. Ribosomal RNA (rRNA) was removed using the Ribo-Zero rRNA Removal Kit (Epicenter) according to the manufacturer's protocol.

Genome Sequencing Library Preparation and Next-Generation Sequencing

The genome sequencing library for long-read sequencing was constructed by following the PacBio 20-kb library preparation protocol (Pacific Biosciences), and the library for short-read sequencing was constructed using the TruSeq DNA Sample Prep Kit (Illumina), according to the manufacturer's protocol. Sequencing was performed using PacBio RS II with P6-C4 chemistry for the long-read sequencing library, and with the Illumina MiSeq v2 instrument with 1 × 50 bp read length for the short-read sequencing library. *De novo* assembly was conducted using the hierarchical genome assembly process workflow (HGAP v2.3), generating four contigs including one complete genome and three plasmids. The draft assemblies were improved by error correction using the Pilon software (Walker et al., 2014). The genes were annotated with the latest version of the National Center for Biotechnology Information (NCBI) Prokaryotic Genome Annotation Pipeline. For phylogenetic analysis, genome sequences of cyanobacteria were downloaded from the NCBI genome portal. The phylogenetic tree was generated by calculating the distance with the Up-to-date Bacterial Core Gene (UBCG) analysis pipeline and represented using the Randomized Axelerated Maximum Likelihood (RAxML) program. For the pan-genome analysis with other *Synechocystis* sp. strains, PGAP v1.12 program was used with the Gene Family (GF) method (Zhao et al., 2012). To find orthologs in *Synechocystis* sp. PCC 6803, total CDSs were compared using BLASTP search (E-value $< 1.00 \times 10^{-6}$). BLAST hits covering over 80% length of the CDS remained and were considered to be orthologs.

RNA-Seq Library Preparation and Next-Generation Sequencing

RNA-seq libraries were constructed using TruSeq Stranded mRNA LT Sample Prep Kit (Illumina) according to the manufacturer's protocol. The amplified library was purified using AgencourtAMPure XP beads (Beckman Coulter) and quantified using a Qubit 2.0 fluorometer (Invitrogen). The quantified libraries were sequenced using Illumina HiSeq 2000 platform with 1 × 50 bp read length.

Differential RNA-Seq (dRNA-Seq) Library Preparation and Next-Generation Sequencing

The rRNA-depleted RNAs were split into two samples to construct two different libraries: a library of whole transcriptome (TAP +) and a library without primary transcripts (TAP −). To construct the TAP + library, 20 U of RNA 5′-polyphosphatase (Epicenter) was treated with 2 µL of 10 × RNA 5′-polyphosphatase Reaction buffer (Epicenter) and 20 U of SUPERase.In (Invitrogen) at 37°C for 1 h. RNA 5′-polyphosphatase converts the triphosphates at the 5′-end of the primary transcript to monophosphate and sequentially enables the ligation of 5′-RNA adaptor. For the TAP − library, RNA 5′-polyphosphatase was excluded from the reaction. After the RNA was purified by ethanol precipitation 0.5 µL of 10 µM 5′-RNA adaptor (5′-ACACUCUUUCCCUACACGACGCUCUUCCGAUCU   -3′) was added to the purified mRNA with 10 U of T4 RNA Ligase (Thermo), 2 µL 10 × ligation buffer for T4 RNA Ligase (Thermo), and 0.1 mgmL^{-1} bovine serum albumin (BSA). The ligation reaction was incubated at 37°C for 90 min, and then the adaptor-ligated RNA was purified using AgencourtAMPure XP beads. The purified RNA was incubated with random 3′ overhanging primer (5′- GTGACTGGAGTTCAGAC GTGTGCTCTTCCGATCTNNNNNNNNN -3′) and 1 µL of 10 mM dNTPs (Invitrogen) at 65°C for 10 min. After the sample was chilled on ice, 2 µL of 10 × reverse transcription (RT) buffer, 2 µL of 100 mM dithiothreitol (DTT), 4 µL of 25 mM MgCl$_2$, 20 U of SUPERase.In, and 200 U of SuperScript III Reverse Transcriptase (Invitrogen) were added followed by incubation at 25°C for 10 min, 50°C for 1 h, 85°C for 5 min, and 4°C for chilling, sequentially. To remove residual RNAs, 2 U of RNase H (Invitrogen) was added to the reaction, and the mixture was incubated at 37°C for 20 min. The synthesized complementary DNA (cDNA) was purified using AgencourtAMPure XP beads, and amplified by polymerase chain reaction (PCR) with the indexed primer for Illumina sequencing. The amplification step was monitored on a CFX96 Real-Time PCR Detection System (Bio-Rad) before the PCR reaction was fully saturated. Finally, the amplified library was purified using AgencourtAMPure XP beads, and the concentration of the library was measured with Qubit 2.0 fluorometer. The size distribution of the library was checked by 2% agarose gel electrophoresis. The constructed dRNA-seq library was sequenced by using Illumina HiSeq 2000 platform with 1 × 50 bp read length.

Term-Seq Library Preparation and Next-Generation Sequencing

Term-seq library was prepared as described in a previous study with some modifications (Dar et al., 2016b). Briefly, modified DNA adaptor (5′- NNAGATCGGAAGAGCGTCGTGT -3′) was ligated to the 3′ end of the mRNA. The adaptor-ligated mRNA was purified using AgencourtAMPure XP beads followed by fragmentation with 10 × fragmentation buffer (Ambion). The fragments were purified with AgencourtAMPure XP beads, and then reverse transcribed using SuperScript III Reverse Transcriptase. The synthesized cDNA was then purified with AgencourtAMPure XP beads, cDNA 3′-adaptor (5′- NNAGATCGGAAGAGCACACGTCTGAACTCCAGTCAC -3′) was ligated to the 3′ end of the cDNA. The adaptor-ligated cDNA was purified using AgencourtAMPure XP beads. For amplification, the indexed primer for Illumina sequencing was

used, and the amplification step was monitored on a CFX96 Real-Time PCR Detection System. The amplified library was removed from the PCR machine at the beginning of the saturation point, and purified using AgencourtAMPure XP beads. The concentration and size distribution of the final library were checked with a Qubit 2.0 fluorometer and an Agilent 2200 TapeStation System (Agilent), respectively. The Term-seq library was sequenced by using Illumina HiSeq 2000 platform with 1×50 bp read length.

Quantitative Real-Time PCR

cDNAs were synthesized from rRNA-depleted RNAs by using SuperScript III First-Strand Synthesis System (Invitrogen) according to the manufacturer's instructions. KAPA SYBR® FAST qPCR Master Mix (KAPA Biosystems) and StepOnePlus Real-Time PCR System (ThermoFisher) was used to measure the different expression level of these genes. The expression fold changes were calculated from the Ct values. The primers used for amplification are indicated in **Supplementary Table 1**.

Data Processing

The adaptor sequences were trimmed from the reads of the dRNA-seq and Term-seq libraries. Random (N) sequences from dRNA-seq and Term-seq reads were also trimmed. After trimming, the reads shorter than 15 bp were discarded and the remaining reads were aligned to the genome of *Synechocystis* sp. PCC 7338 using CLC genomics workbench with the following parameters: mismatch cost = 2, deletion cost = 3, insertion cost = 3, length fraction = 0.9, and similarity fraction = 0.9. Only uniquely mapped reads were retained. The 8.8–11.4 million sequence reads were mapped to the reference genomes of *Synechocystis* sp. PCC 6803 and PCC 7338 with at least $99 \times$ and $109 \times$ coverage, respectively. To visualize the data with SignalMap (Roche Nimblegen), the read depth of each genomic position of each library was normalized by multiplying a size factor (one million divided by the total number of reads in each library). To compare the mRNA expression of orthologs between the two strains, a size factor was applied to the number of reads mapped to the orthologs, and thereafter the gene expression was normalized with DESeq2 package in R (Love et al., 2014). The reproducibility between biological triplicates was confirmed by drawing a heatmap and a principal component analysis (PCA) plot (**Supplementary Figures 4B,C**). Compared to the mRNA expression of orthologs in *Synechocystis* sp. PCC 6803, genes with over 2-fold changes in mRNA levels and adjusted $p < 0.01$, were selected as DEGs.

TSS Identification and Data Analysis

The 5′ end positions of the reads of TAP-treated (TAP +) libraries were considered as potential TSSs. The depths of the reads in each library were normalized by multiplying it with a size factor. Next, the potential TSSs were clustered if the distance between them were <100 nt. To subdivide the clusters, the standard deviation in the peak positions within a cluster was calculated. A standard deviation of <10 for two or more nearby located peaks indicated a region with densely located peaks. Thus, the peaks were sub-clustered together and the one with highest depth

in the sub-cluster remained. The remaining peaks were compared with data from the respective non-TAP-treated (TAP −) libraries. The peaks that only existed in the TAP + library or had over 2-fold depth than the peak of the TAP − library remained. If the remaining peak was absent in one of the biological duplicate data, the peak was removed. Further manual curation was performed to finalize the TSS list by comparing the peaks and dRNA-seq profiles with RNA-seq profiles. If a peak was removed due 1–2 nt difference between the biological duplicate data, the peak was restored and selected as a TSS. In addition, if a peak was present in one of the biological duplicate data and had a clear RNA-seq profile, it was selected as a TSS. In contrast, if the expression level of a gene was extremely high, it was difficult to remove the peaks of processed transcripts by the above curation steps; thus, we removed the peaks by manual curation. Among the TSSs located within 500 nt upstream to 100 nt downstream of the start codon of a gene, the one with the highest read depth was categorized as primary (P) TSS, while the others were categorized as secondary (S) TSSs. The TSSs located inside or the antisense orientation of a gene were categorized as internal (I) TSSs or antisense (A) TSSs, respectively. The TSSs that did not belong to any of these categories were considered intergenic (N) TSSs. To calculate the 5′-UTR sequence difference from the 5′-UTR of *Synechocystis* sp. PCC 6803, the 5′-UTR sequences were aligned by MUSCLE, and the distance between each sequence was calculated using the p-distance method with pairwise deletion option in MEGA X software (Kumar et al., 2018). The 5′-UTR lengths were compared for 397 genes with 5′-UTR identified in both the strains; Those having length differences under 2 nt were considered as conserved 5′-UTRs, and length differences over 10 nt as degenerated 5′-UTRs. For comparison of the promoter region, the 50 nt upstream sequences (promoter region) of 754 TSSs associated with the CDSs of *Synechocystis* sp. PCC 7338 were searched against the 550 nt upstream sequence of ortholog from *Synechocystis* sp. PCC 6803 by BLASTN. Conversely, the 50 nt upstream sequences of 1,384 TSSs associated with the CDSs of *Synechocystis* sp. PCC 6803 were searched against the 550 nt upstream sequence of ortholog from *Synechocystis* sp. PCC 7338 by BLASTN. If there was a matched promoter region with associated TSS from both search results, the promoter region was classified as 'conserved'. If there was a matched promoter region in the upstream sequence of orthologs from each other with no TSS or mismatched TSS position, the promoter region was classified as 'mismatched'. If there was no matched promoter region in the upstream sequence of ortholog from each other, the promoter region was classified as 'orphan'. Finally, if the promoter region is associated with the genes that have no ortholog, it was classified as 'specific'.

TEP Identification and Data Analysis

The 5′end positions of the Term-seq library reads were extracted and the strand information of the 5′end positions were reversed to be considered as potential TEPs. The TEPs were determined by a combined method of previous studies with a machine-learning algorithm (Dar et al., 2016b; Lalanne et al., 2018). First, to select the positive control learning set, the clustering method used for TSS identification was used. The selected peaks were

curated by calculating the Z-score of their read depth (Z-score >6), which indicated the enrichment of the peak compared to adjacent peaks. Among the remaining peaks, the positive control learning set was manually selected by considering the decreasing RNA expression profiles near the peaks. The negative control learning set was selected within the peaks located at ± 10 nt position of the positive control learning sets. Using the positive and negative control sets, the TEPs were identified with an in-house Python script based on the scikit-learn package. Briefly, for a reversed 5′ end position, which is a potential TEP, the read depth in ± 10 nt positions were submitted to the K-nearest neighbor (KNN) machine classifier. The Python script is available at http://cholab.or.kr. Finally, only the TEPs present in at least two or more of the triplicate data (± 1 nt) remained and were manually curated by considering the RNA expression profiles near them. Among the TEPs located within 500 nt downstream of the stop codon, the one with the highest depth was classified as primary P-TEP, and the others were classified as secondary S-TEPs. Based on the criteria, 329 P-TEPs and 25 S-TEPs were assigned. If a TEP was located between the start codon and the TSS assigned to the gene, it was classified as 5′-UTR TEP. There were 28 genes with 5′-UTR TEPs at their 5′-UTRs. A TEP located inside a gene was classified as internal I-TEP, and one located in the antisense strand of a gene was classified as antisense A-TEP. Finally, the remaining TEPs were classified as intergenic N-TEPs. The TEPs were classified as L-shaped TEPs if there were ≥4 uridines near TEP, and I-shaped TEPs if there were <4 uridines. For comparison of the terminator regions, we downloaded and analyzed the Term-seq data of *Synechocystis* sp. PCC 6803 deposited in NCBI Sequence Read Archive (Cho and Jeong, 2020). The same strategy for the comparison of the promoter region was applied with modifications to compare and classify the terminator regions; the terminator regions were considered as 40 nt upstream to 20 nt downstream sequences of the TEPs, and they were BLASTN searched against the 520 nt downstream of the orthologs.

Motif Discovery and $\Delta G_{folding}$ Calculation

The conserved sequences were examined using the MEME suite (Bailey et al., 2009). The 20 nt upstream sequences of each TSS and the sequences between 40 nt and 20 nt upstream of each TSS were used for searching the − 10 promoter motif and the − 35 promoter motif, respectively. RBS was searched in 20 nt to 1 nt upstream sequences of start codons of genes that have 5′-UTR lengths >10 nt. The 40 nt upstream to 20 nt downstream sequences of each TEP were used for searching the terminator sequences. We extracted the conserved sequences and obtained the motif sequences with WebLogo (Crooks et al., 2004). The $\Delta G_{folding}$ values from the upstream sequences of TEPs were predicted using RNAfold (Lorenz et al., 2011).

Accession Numbers

The datasets, genome-seq, RNA-seq, dRNA-seq, and Term-seq, generated for this study have been deposited in National Center for Biotechnology Information (NCBI) as BioProject PRJNA629670[1].

RESULTS

Genome Completion and Annotation

To obtain a high-quality genome sequence of *Synechocystis* sp. PCC 7338, we utilized two sequencing platforms, the PacBio and Illumina providing long read and short read outputs, respectively. The sequencing reads were assembled to obtain the complete genome sequence without any sequence gaps, resulting a 3.70 mega base pair (Mbp) circular chromosome and three plasmids of sizes 81.33 kilo base pair (kbp), 45.22 kbp, and 1.56 kbp (**Supplementary Table 2**). The benchmarking universal single-copy orthologs (BUSCO) analysis showed 98.39% of complete BUSCO, 1.61% of fragmented BUSCO, and no missing BUSCO in *Synechocystis* sp. PCC 7338 genome, validating that the completed genome sequence was of high-quality (Simao et al., 2015). Compared to the genome sequence of model cyanobacteria, *Synechocystis* sp. PCC 6803, the calculated average nucleotide identity (ANI) value was 86.96, which indicates that they are closely related species with moderately divergent genomes (Jain et al., 2018). The phylogenetic analysis revealed a close relationship between *Synechocystis* sp. PCC 7338 and other freshwater *Synechocystis* sp. strains, in addition to similar sizes and GC contents of the genomes (**Figure 1A** and **Supplementary Table 2**).

Next, of the 3,589 genes annotated in the complete genome sequence, 3,385 were coding sequences (CDSs) with 152 pseudogenes, 42 transfer RNAs (tRNAs), six ribosomal RNAs (rRNAs), and four RNAs (**Supplementary Table 2**). The annotated CDSs were categorized by the functions based on Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology (KO). Among the 3,385 CDSs, 2,573 KO IDs were annotated, and a large number of genes were involved in carbohydrate metabolism (17.27%), energy metabolism (14.01%), amino acid metabolism (12.84%), and metabolism of cofactors and vitamins (11.97%) (**Supplementary Figure 1A**). Particularly, in *Synechocystis* sp. PCC 7338, the energy metabolism genes were highly enriched owing to the presence of photosynthesis-associated genes, similar to other photosynthetic organisms (Nakayama et al., 2014).

Additional Genes in *Synechocystis sp.* PCC 7338 Compared to Freshwater *Synechocystis sp.* Strains

We compared genes in *Synechocystis* sp. PCC 7338 with those in other freshwater *Synechocystis* sp. strains to investigate the differences in gene sets that enables them to survive in different environmental conditions. The genome information of six *Synechocystis* sp. strains were subject to pan-genome analysis, revealing 396 specific genes in *Synechocystis* sp. PCC 7338, which can be responsible for the distinctive

---
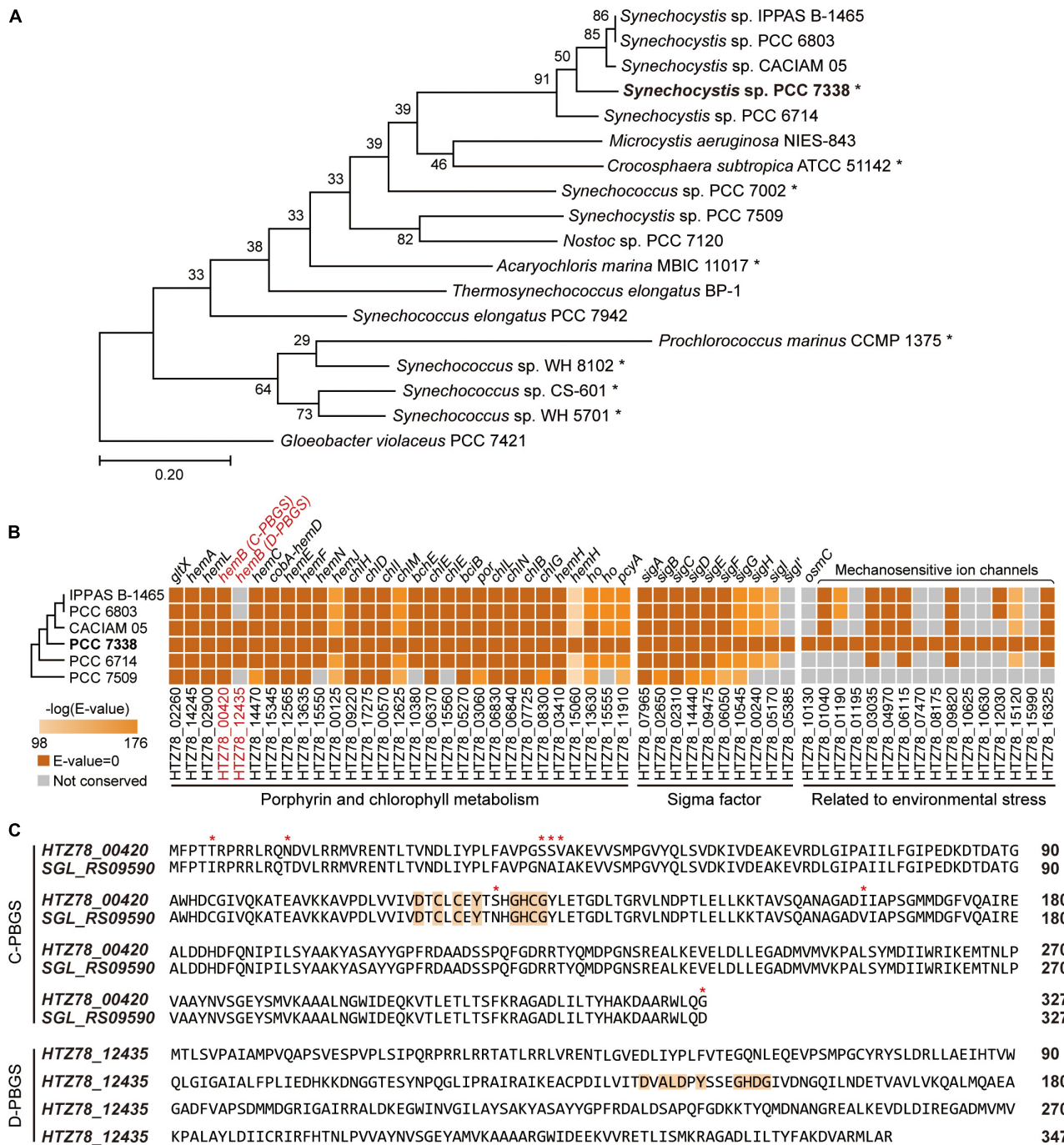
FIGURE 1 | Genome completion of *Synechocystis* sp. PCC 7338 and comparison with other cyanobacteria. **(A)** Phylogenetic analysis of *Synechocystis* sp. PCC 7338 and 12 genome sequenced cyanobacteria. *Gloeobacter violaceus* PCC 7421 was selected as the outgroup. The evolutionary distances were calculated by Up-to-date Bacterial Core Gene analysis pipeline (UBCG) and represented by Randomized Axelerated Maximum Likelihood (RAxML). Asterisks indicate marine cyanobacteria. **(B)** Comparison of gene sequences related to porphyrin and chlorophyll metabolism, sigma factors, and environmental stress in six *Synechocystis* sp. strains by BLASTP search. The names of strains are indicated with omitted species name (*Synechocystis* sp.). The extra sigma factor detected only in *Synechocystis* sp. PCC 7338 was denoted as SigI'. The tree is not to scale. Keys: C-PBGS, cysteine-rich porphobilinogen synthase; D-PBGS, aspartate-rich porphobilinogen synthase. **(C)** The amino acid sequences of C-PBGS in *Synechocystis* sp. PCC 7338 (HTZ78_00420) and *Synechocystis* sp. PCC 6803 (SGL_RS09590) and D-PBGS in *Synechocystis* sp. PCC 7338 (HTZ78_12435). The amino acid sequences of the other *Synechocystis* species are compared in **Supplementary Figure 1B**. Orange boxes indicate the active sites and red asterisks indicate the varying sequences.

characteristics of the strain (**Supplementary Table 3**). Interestingly, there were core or dispensable genes that were strain-specifically duplicated, which can also describe evolutionary characteristics. For example, while three species possessed two copies of *hemB,* which encode porphobilinogen synthase (PBGS) involved in the porphyrin and chlorophyll metabolism pathways, the rest had a single copy of *hemB* (**Figure 1B**). The PBGS exhibits phylogenetic variation according to the active site sequence, which is suggested to evolve from being cysteine-rich (C-) to aspartate-rich (D-) PBGS (Jaffe, 2003). Interestingly, between the two *hemB* genes in *Synechocystis* sp. PCC 7338 genome, one encoded C-PBGS and the other D-PBGS (**Figure 1C** and **Supplementary Figure 1B**). To date, only *Nostoc* sp. PCC 7120 has been reported to express both types of PBGS (Jaffe, 2003). It is noteworthy that although cyanobacteria are the ancestors of plant plastids, the plants express only D-PBGS (Jaffe, 2003). Thus, cyanobacteria expressing both C- and D-PBGSs may be involved in the evolution of C-PBGS to D-PBGS and thus closely resembling the plant plastids. When the transcription levels of the two genes were compared by using RNA-seq data (see "Materials and Methods"), gene expression of C-PBGS was 13.5-fold higher than that of D-PBGS (**Supplementary Figure 1C**). Despite the existence of both type of PBGSs, the different transcription levels infer that *Synechocystis* sp. PCC 7338 mainly uses C-PBGS, and presumably has D-PBGS as a trace of evolution.

Synechocystis sp. PCC 7338 is presumed to possess specific genes that confer the ability to withstand high salinity and high osmotic pressure. Those are osmotically inducible protein C (OsmC) family of proteins (*HTZ78_10130*) and six additional mechanosensitive ion channels (*HTZ78_01195,* *HTZ78_07470, HTZ78_08175, HTZ78_10625, HTZ78_10630,* and *HTZ78_15990*) (**Figure 1B** and **Supplementary Table 4**). The OsmC accumulates under osmotic stress, and the mechanosensitive ion channels regulate the turgor in bacteria in response to changes in osmotic pressure (Perozo and Rees, 2003). Thus, *Synechocystis* sp. PCC 7338 has more mechanosensitive ion channels than freshwater *Synechocystis* sp. strains to respond to the high osmotic pressure. Also, we observed an additional sigma factor and an anti-sigma factor present exclusively in *Synechocystis* sp. PCC 7338 (**Figure 1B** and **Supplementary Table 4**). Among the annotated sigma factors in *Synechocystis* sp. PCC 7338 genome, seven (SigA, SigB, SigC, SigD, SigE, SigF, and SigG) were conserved in five other *Synechocystis* sp. strains, and two (SigH and SigI) were conserved in four strains (**Figure 1B**). Interestingly, one sigma factor, encoded by *HTZ78_05385*, was not found in any *Synechocystis* sp. strains. Among the annotated sigma factors, SigI encoded by *HTZ78_05170* had the highest similarity to *HTZ78_05385* with a 50.24% identity (BLASTP $e = 2.00 \times 10^{-52}$) and similar amino acid lengths (185 and 187, respectively) (**Supplementary Figure 1D**). The evolution of cyanobacterial SigI was believed to be strain-specific, which responded rapidly under specific stress conditions, contributing to cell survival (Imamura and Asayama, 2009). Thus, it suggests that *HTZ78_05385*

encodes a strain-specific sigma factor in *Synechocystis* sp. PCC 7338, which responds to specific environmental stress conditions.

Identification of TSSs, Promoters, and 5′-UTRs

The existence of an exclusive strain-specific sigma factor suggests the presence of specific regulatory networks in *Synechocystis* sp. PCC 7338, which are involved in the stress conditions. To examine the regulatory traits, we identified the TSS positions in *Synechocystis* sp. PCC 7338 genome using dRNA-seq. A total of 897 TSSs were identified and categorized based on their positions relative to the annotated genes (see Materials and Methods for details) (**Figure 2A** and **Supplementary Data 1**). The dinucleotide preference, which is high purine usage at the TSS position (84.62%) and high pyrimidine usage at the − 1 position (67.56%), was observed, similar to other previously reported bacteria (**Figure 2B**; Jeong et al., 2016; Hwang et al., 2019; Lee et al., 2019). Given the significance of the promoter motifs in regulating the transcription initiation step, we next investigated the conserved promoter motifs from the upstream sequences of TSSs. Among the 897 TSSs, − 10 promoter and − 35 promoter motif sequences were detected as 5′-TANAAT and 5′-TTGCCAA in the upstream regions of 746 TSSs and 372 TSSs, respectively (**Figure 2C**). Similar motif sequences were detected in *Synechocystis* sp. PCC 6803 by searching the upstream sequence of previously published TSSs (**Supplementary Figure 2A**) (Mitschke et al., 2011; Kopf et al., 2014).

Interestingly, we found an AT-rich ferric uptake regulator (Fur) binding motif in the upstream sequences of 77 genes, including 43 specific genes in *Synechocystis* sp. PCC 7338, by MEME and FIMO search (**Figure 2D** and **Supplementary Table 5**) (Bailey et al., 2009; Grant et al., 2011). Fur is involved in iron homeostasis by regulating several iron responsive genes (Dang et al., 2012; Kaushik et al., 2016). The Fur binding motif was found at the TSS upstream regions of the genes encoding bacterioferritin (HTZ78_10445), which is a ferritin-type storage complex, and PerR (HTZ78_00325), which is a Fur family transcriptional regulator/peroxide stress response regulator (**Supplementary Table 5**; Shcolnick et al., 2009). Among the genes having the Fur motif sequence and KEGG orthology identifier, five were included in the energy metabolism category (29.41%), containing the photosynthesis and oxidative phosphorylation categories, consistent with the high iron requirement for photosynthesis in cyanobacteria (**Figure 2E**; Xie et al., 2011; Morrissey and Bowler, 2012). In addition, four genes were included in the signal transduction category, two of which were regulators involved in pilus function: *pilH* (HTZ78_05555) encoding a twitching motility two-component system response regulator and *chpA* (HTZ78_02660) encoding a chemosensory pili system protein (sensor histidine kinase/response regulator) (**Figure 2E** and **Supplementary Table 5**). Interestingly, SigF, a sigma factor known to regulate pili gene expression in cyanobacteria, also has the Fur binding motif (Imamura and Asayama, 2009). These results agree with the previous reports
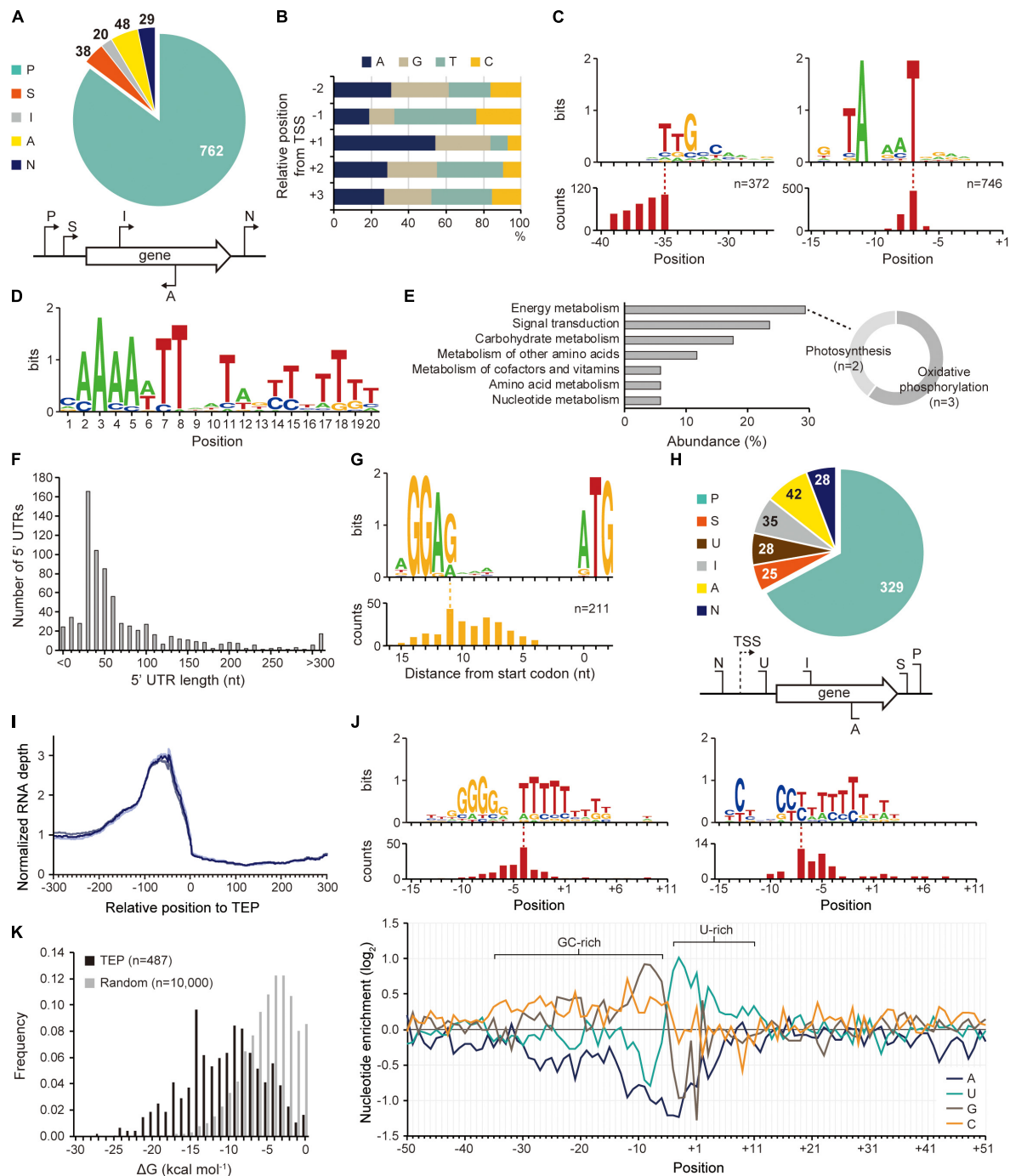
FIGURE 2 | Identification of transcription start site (TSS), transcript 3′-end position (TEP), and regulatory elements involved in transcription regulation. **(A)** TSS categorization by their relative positions to adjacent genes. **(B)** The nucleotide frequency calculated near the TSSs shows purine preference at TSS (+ 1 position) and pyrimidine preference at − 1 position. **(C)** Conserved − 10 and − 35 promoter motifs. The relative position of the motif to the TSS is represented at the bottom. **(D)** The Fur binding motif found at the upstream regions of *Synechocystis* sp. PCC 7338-specific genes. **(E)** The KEGG pathway analysis of the genes having the Fur motif. The sub-categories of the category with the highest abundance (energy metabolism category) were indicated as a doughnut chart. **(F)** The length distribution of 5′-untranslated region (5′-UTR). **(G)** The ribosome binding site (RBS) sequence detected at 5′-UTR. The relative position of the RBS to the start codon is represented at the bottom. **(H)** TEP categorization by their relative positions to adjacent genes. Keys: P, primary TEP; S, secondary TEP; U, 5′-UTR TEP; I, internal TEP; A, antisense TEP; N, intergenic TEP. **(I)** RNA expression profiles near the identified TEPs. Each line shows each result from triplicate RNA-seq data. **(J)** The top panel shows conserved sequences detected near the TEPs. The relative positions of the sequences to the TEPs are represented at the bottom of each sequence. The bottom panel shows nucleotide enrichment calculated in ± 50 nt sequences from TEPs. The ratio of each nucleotide at each position was normalized with those of randomly selected intergenic positions (*n* = 10,000). **(K)** The folding energy was calculated at upstream sequences of TEPs or at randomly selected intergenic positions.

suggesting that pili are involved in the reduction of iron oxides in bacteria, including *Synechocystis* sp. PCC 6803 (Lamb et al., 2014). SigF is also known to be involved in the long-term application of high salt stress in *Synechocystis* sp. PCC 6803 (Huckauf et al., 2000). In addition, an extracytoplasmic function (ECF) sigma factor SigG was found to have the Fur binding motif in its TSS upstream sequence. SigG showed high similarities to two ECF sigma factors of *Escherichia coli*, RpoE and FecI, which are involved in strong heat shock response and iron uptake regulation, respectively (Huckauf et al., 2000). Thus, the Fur binding motif indicates the possibility that the sigma factor is regulated by iron concentration, and orchestrates the regulatory network involved in iron homeostasis in *Synechocystis* sp. PCC 7338.

The SigG and other ECF sigma factors (SigH, SigI, and HTZ78_05385) were further analyzed to investigate their functions in stress responses. The amino acid sequences of region 2 and region 4, which are the conserved domains of ECF sigma factors, were aligned with previously known bacterial ECF sigma factors (Todor et al., 2020). In results, it was revealed that both region 2 and region 4 sequences of SigG, SigH, and SigI of *Synechocystis* sp. PCC 7338 were close to those of *Synechocystis* sp. PCC 6803 and *Synechocystis* sp. PCC 6714 (**Supplementary Figure 2B**) (Todor et al., 2020). On the other hand, the region 2 sequence of HTZ78_05385 was close to an ECF sigma factor of another marine cyanobacteria, *Nodularia spumigena*, inferring the function of HTZ78_05385 is related to the adaptation to marine environment. Considering the promoter motifs predicted for clusters of ECF sigma factors, which had been clustered by region 2 and region 4 sequences in the previous study, we found 5′-GTC of − 10 motif and 5′-GGAAC of − 35 motif for SigG, and the extended − 10 and − 35 motifs for SigH; SigI and HTZ78_05385 were predicted to have 5′-CGTA of − 10 motif and 5′-CATCC of − 35 motif, which are distinctive from those of SigG and SigH (**Supplementary Figure 2C**) (Todor et al., 2020).

The identified TSS positions provide information on the 5′-UTR, an important genetic element involved in the regulation of gene expression at transcriptional and post-transcriptional levels. Among the 737 5′-UTRs, 55.77% had a length varying between 20–59 nucleotides (nt), and the median length was 42 nt (**Figure 2F**). The conserved ribosome binding site (RBS) was identified within the 5′-UTRs as 5′-AGGAG, and major portion of the RBS (74.41%) was at a distance of 5–10 nt from the start codon (**Figure 2G**). The RBS sequence was not detected at the upstream sequence of the leaderless transcripts with a 5′-UTR length <9 nt (**Supplementary Figure 2D**). The 5′-AGGAG located at a distance of 5–10 nt from the start codon has been identified as the third most frequently used RBS sequence in bacteria (Omotajo et al., 2015).

Identification of TEPs and 3′-UTRs

The 3′-UTR is also a crucial genetic element that affects gene expression, mRNA stability, transcription termination rate, and interaction with the RNA-binding proteins (Dar et al., 2016a,b). Using Term-seq coupled with machine-learning approach, a total of 487 TEPs were identified and categorized based on their relative positions to the adjacent genes (see Materials

and Methods for details) (**Figure 2H** and **Supplementary Data 2**). To confirm the identified TEPs, we used RNA-seq data to investigate the RNA expression profiles near TEPs (**Figure 2I**). We observed that the RNA expression levels decreased steeply from 50 nt upstream of TEPs, supporting the reliability of the determined TEPs.

Next, the sequences near TEPs were analyzed to search for conserved sequences involved in the regulation of transcription termination. One of the searched motifs had a G-rich region followed by an uracil (U)-rich region, and the other one had a C-rich region followed by a U-rich region (**Figure 2J**). These motifs resemble the shape of a bacterial rho-independent terminator, which has a GC-stem and U-tract (Mitra et al., 2009). We observed that a similar motif was also found in *Synechocystis* sp. PCC 6803 (**Supplementary Figure 2E**). Since no homologs of the rho factor were identified in cyanobacteria, most transcription termination were suggested to occur via a rho-independent mechanism (Ramey et al., 2015). The patterns were also observed during the calculation of nucleotide enrichment near TEP, which is similar to other bacteria such as *E. coli* and *Streptomyces lividans* (**Figure 2J**; Dar and Sorek, 2018; Lee et al., 2019). The formation of a secondary structure at the upstream sequences of TEPs was investigated by calculating the free energy of folding ($\Delta G_{folding}$). The $\Delta G_{folding}$ distribution was relatively lower than that of random intergenic regions ($n = 10,000$), indicating that secondary structure formation upstream of TEP is more stable (**Figure 2K**).

Based on the shape of the rho-independent terminators in bacteria, we classified the identified TEPs into L- or I-shaped TEPs (Unniraman et al., 2002). The nucleotide enrichment near the L-shaped TEPs was similar to that of the total TEPs, whereas the I-shaped TEPs showed an omitted U-rich region and stretched GC-rich region (**Supplementary Figure 2F**). However, despite the broader distribution of the GC-rich region in the I-shaped TEPs, the $\Delta G_{folding}$ values were lower in the L-shaped TEPs, indicating that the secondary structures of the L-shaped TEPs were more stable (**Supplementary Figure 2G**). The RNA expression profile decreased from 100 nt upstream regions of L-shaped TEPs, and a sharper decrease was observed from 50 nt upstream regions of the I-shaped TEPs, suggesting both types of TEPs are the 3′-termini of the transcripts (**Supplementary Figure 2H**). Previously, the read-through effect of RNA polymerase at I-shaped TEP was reported in *S. lividans*, however, the read-through effect at the I-shaped TEP appears uncommon in *Synechocystis* sp. PCC 7338 (Lee et al., 2019).

Regulatory Regions in *Synechocystis* sp. PCC 7338 Compared to *Synechocystis* sp. PCC 6803

We compared the regulatory regions of the two strains to investigate the differences in gene regulation (Mitschke et al., 2011; Kopf et al., 2014; Cho and Jeong, 2020). From a total of 397 5′-UTR and 93 3′-UTR from orthologs in the two strains, 63.98% of the 5′-UTR pairs have conserved length, while 25.81% of the 3′-UTR pairs have conserved length (**Figure 3A**). We observed that > 27% of the orthologs with conserved

5′-UTR length was associated with carbohydrate metabolism, while > 30% of the orthologs with degenerated 5′-UTR length was involved in energy metabolism, especially photosynthesis (**Supplementary Figure 3A**).

Next, the promoter regions (50 nt upstream sequences of TSSs) of the orthologs in the two strains were compared (see Materials and Methods for details). The results were classified into four categories; the ones with conserved promoter regions and associated TSSs were classified as 'conserved', the ones with conserved promoter regions but mismatched TSSs as 'mismatched', the ones whose promoter regions were not conserved as 'orphan', and the ones associated with no orthologs as 'specific' (**Figure 3B** and **Supplementary Data 3**). For example, among the two TSSs associated with *HTZ78_12890* encoding photosystem I (PSI) core protein (PsaA), one of them has the promoter region conserved with its ortholog, *SGL_RS06305* (**Figure 3C**). On the other hand, the promoter region from the *HTZ78_15610* encoding photosystem II (PSII) protein D1 (PsbA) was not conserved with its ortholog, *SGL_RS10440*, and the TSS position also differed, thus making it an orphan promoter region (**Figure 3C** and **Supplementary Figure 3B**). Considering the previous result that 17.70% of the promoter regions were conserved between *E. coli* and *Klebsiella pneumoniae*, the lower ratio of the conserved promoter regions between the two *Synechocystis* species (12.47%) is noticeable (Kim et al., 2012). It can be inferred that the regulatory mechanism of the two species differentiated in a large degree to adapt to the different environmental conditions. In addition, among the genes related to photosynthesis, the genes involved in cytochrome b6-f complex, PSI, and carbon fixation tend to have more conserved promoter regions than PSII genes, suggesting that the regulatory regions of the PSII genes had diverged more between the two environments (**Figure 3D**).

For comparison of the terminator regions, the proximate sequences of the TEPs of the orthologs in the two strains were compared (see Materials and Methods for details). The overall characteristics such as $\Delta G_{folding}$ distribution and the nucleotide enrichment were similar between the two strains, however, only 2.40% of the compared terminator region were found to be conserved (**Figure 3B** and **Supplementary Figure 3C**). When considering that the 3′-UTR length is also less conserved than the 5′-UTR lengths, the terminator region seems to operate by the formation of the stem-loop structure rather than the precise sequence-mediated mechanism.

Comparison of Gene Expression Patterns of *Synechocystis sp.* PCC 7338 With *Synechocystis sp.* PCC 6803

Regulatory elements, such as promoters, 5′-UTRs, RBSs, 3′-UTRs, and terminators, are involved in gene expression regulation of bacteria confronting diverse environmental conditions. Although the two strains have 2,790 orthologs, their transcription levels vary given the differential growing environmental conditions. Thus, for comparison, we additionally performed RNA-seq in *Synechocystis* sp. PCC 6803 (**Supplementary Figures 4A-C**). The reads aligned to

2,790 orthologs in each strain were normalized and compared among each other, obtaining a list of 1,504 differentially expressed genes (DEGs) (**Supplementary Data 4**). We observed a significant upregulation of the key enzymes involved in the glucosylglycerol biosynthesis pathway, such as 19.6-fold increase in glycerol kinase (*HTZ78_01960*, DESeq2 $p = 1.02 \times 10^{-72}$), 35.6-fold increase in glucosylglycerol-phosphate synthase (*HTZ78_01970*, DESeq2 $p = 2.56 \times 10^{-84}$), and a 4.6-fold increase in glucosylglycerol 3-phosphatase (*HTZ78_06805*, DESeq2 $p = 4.29 \times 10^{-25}$) (**Figure 4A**). A previous metabolic profiling study also detected a significantly higher level (23.1-fold increase with Mann-Whitney test, $p = 0.002$) of glucosylglycerol in *Synechocystis* sp. PCC 7338 compared to *Synechocystis* sp. PCC 6803, which uses glucosylglycerol as the main compatible solute. Thus, our results support that *Synechocystis* sp. PCC 7338 uses glucosylglycerol as a compatible solute to respond to the osmotic stress. In addition, the ABC transporters with specific functions, such as transport of amino acids, branched-chain amino acids, and carbohydrates were highly expressed in *Synechocystis* sp. PCC 7338 (**Supplementary Figure 4D**). The highly expressed genes in *Synechocystis* sp. PCC 7338 also include aquaporin (*HTZ78_10655*) and sodium/proton antiporters (*HTZ78_01555*, *HTZ78_09010*, and *HTZ78_13585*), which are known to be involved in response to osmolarity oscillations and salt tolerance in cyanobacteria, respectively (**Supplementary Data 4**) (Waditee et al., 2002; Akai et al., 2012). The bacterial secretion system genes and motility genes encoding membrane proteins that utilize ATP to operate their functions were highly expressed in *Synechocystis* sp. PCC 7338, justifying their higher demand for ATP. Interestingly, comparing the regulatory regions of the highly expressed ABC transporters, aquaporin, antiporters, secretion system genes, and motility genes, most of them (15 out of 17) had degenerated (orphan) promoter regions (**Supplementary Data 3**).

Among the photosynthesis genes, those related to PSI, ATP synthase, and antenna proteins were upregulated in *Synechocystis* sp. PCC 7338 (**Figure 4B**). In addition, 50.82% of the genes related to porphyrin and chlorophyll metabolism (31 out of 61) were upregulated while 16.39% were downregulated (**Figure 4C**). On the contrary, genes related to PSII tend to be downregulated, especially the genes encoding D1 proteins which are responsible for the PSII repair. The low D1 protein synthesis rate can cause vulnerability to even low illumination and generating reactive oxygen species (ROS), which induces oxidative stress within the cell (Nishiyama et al., 2001; Allakhverdiev and Murata, 2004; Nishiyama et al., 2006). However, the expression of genes responding to oxidative stress decreased, indicating that the light energy transferred to PSII was insufficient to cause oxidative stress in cells (**Supplementary Data 4**). An antenna protein phycobilisome (PBS) mainly binds to PSII to transfer the light energy (Chang et al., 2015; Luimstra et al., 2019). However, the energy is transferred directly from PBS to PSI through state transition or PBS-PSI complex formation under certain circumstances, such as plastoquinone pool reduction and high ATP demand (Allen et al., 1981; Mullineaux and Allen, 1990). For example, in *Anabaena* sp., the cyclic electron transport through PSI provides ATP for
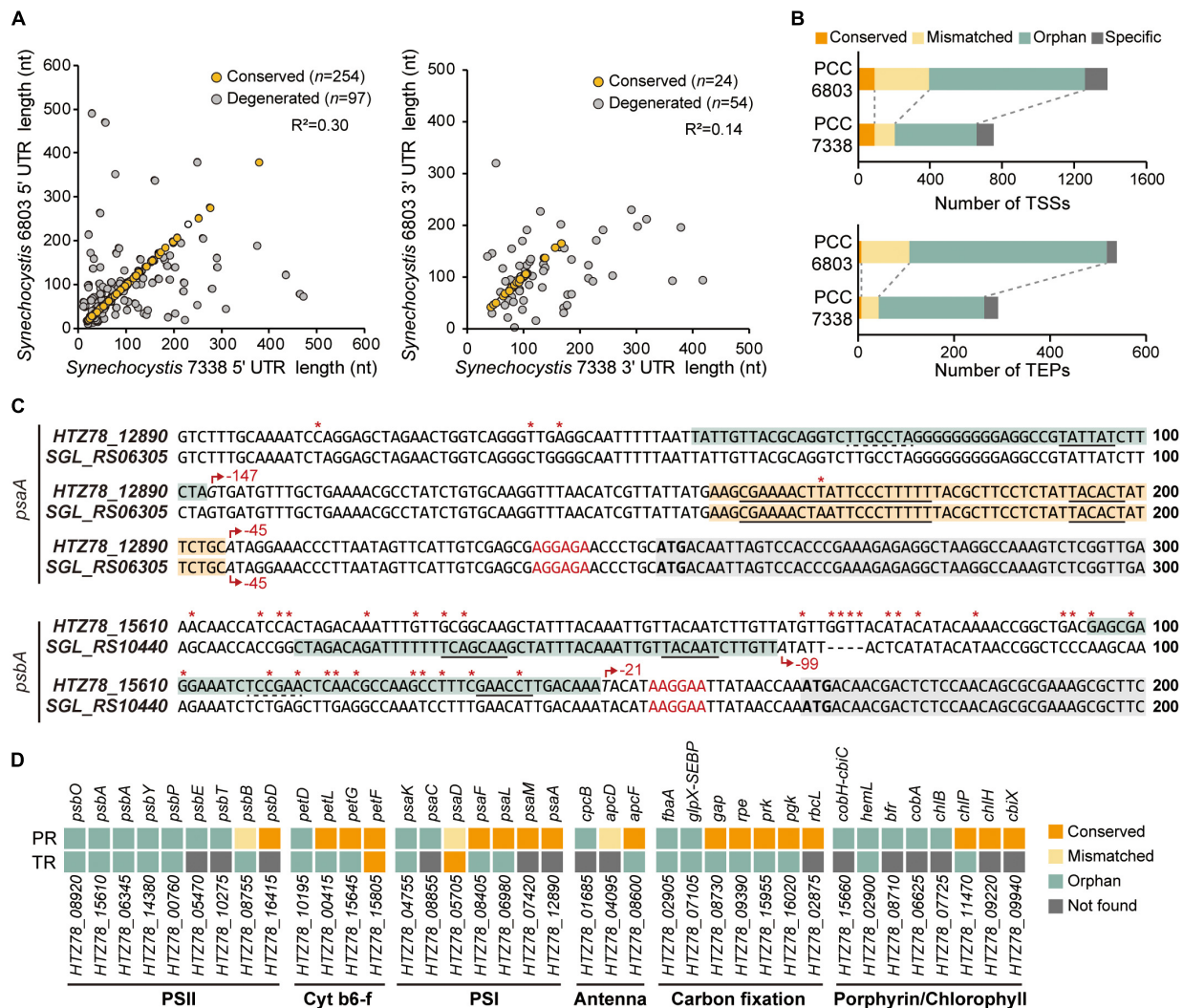
FIGURE 3 | Comparison of the regulatory regions between *Synechocystis* sp. PCC 7338 and *Synechocystis* sp. PCC 6803. **(A)** The 5′-UTR and the 3′-UTR lengths of orthologs were compared with *Synechocystis* sp. PCC 6803. The UTR lengths were classified as conserved or degenerated according to the length difference between the two strains. **(B)** The promoter regions and the terminator regions were compared between the two species and classified as four categories. **(C)** Examples of the compared promoter regions. *psaA* has a conserved promoter region (orange boxes) and an orphan promoter region (cyan box), and *psbA* has orphan promoter regions (cyan boxes). Grey boxes indicate the coding sequences. Red asterisks indicate the varying sequences. Red arrows indicate the TSSs, and the relative positions of the TSSs to the start codons are designated near the red arrows. The promoter sequences detected from MEME search are underlined with solid lines, and the predicted promoter sequences are underlined with dotted lines. The RBS sequences are indicated as red characters, and the start codons are indicated as bold characters. **(D)** The classification of the regulatory regions associated with the genes related to photosynthesis. Keys: PR, promoter region; TR, terminator region, PSII, photosystem II; Cyt b6-f, cytochrome b6-f; PSI, photosystem I.

nitrogen fixation under the light-limited conditions by forming a PBS-CpcL-tetrameric PSI supercomplex (Watanabe et al., 2014). CpcL is a variant of the rod-core linker CpcG, which possesses a hydrophobic region at its C-terminus. Similar to *Anabaena* sp., increased expression of PSI, ferredoxin-NADP (+) reductase (FNR), and ATP synthase genes were observed in *Synechocystis* sp. PCC 7338, indicating the activation of cyclic electron transport chain. Furthermore, genes encoding NAD(P)H-quinone oxidoreductase complex (NDH-1), which is involved in NDH-1-mediated cyclic electron transport, were also upregulated (**Supplementary Table 6**) (Laughlin et al.,

2020). In addition, among the two *cpcG* in *Synechocystis* sp. PCC 7338, the one with a hydrophobic region in the C-terminus (*HTZ78_06190*; CpcL) was highly upregulated, suggesting the formation of a PBS-CpcL-PSI supercomplex in *Synechocystis* sp. PCC 7338 (**Figure 4D**). Independent validation of the RNA-seq data was performed through quantitative PCR by selection of genes that have a broad range of fold changes between *Synechocystis* sp. PCC 7338 and *Synechocystis* sp. PCC 6803. The comparison of the RNA-seq data and quantitative PCR results showed high R^2 value (0.97), indicating the reliability of the quantification based on RNA-seq data
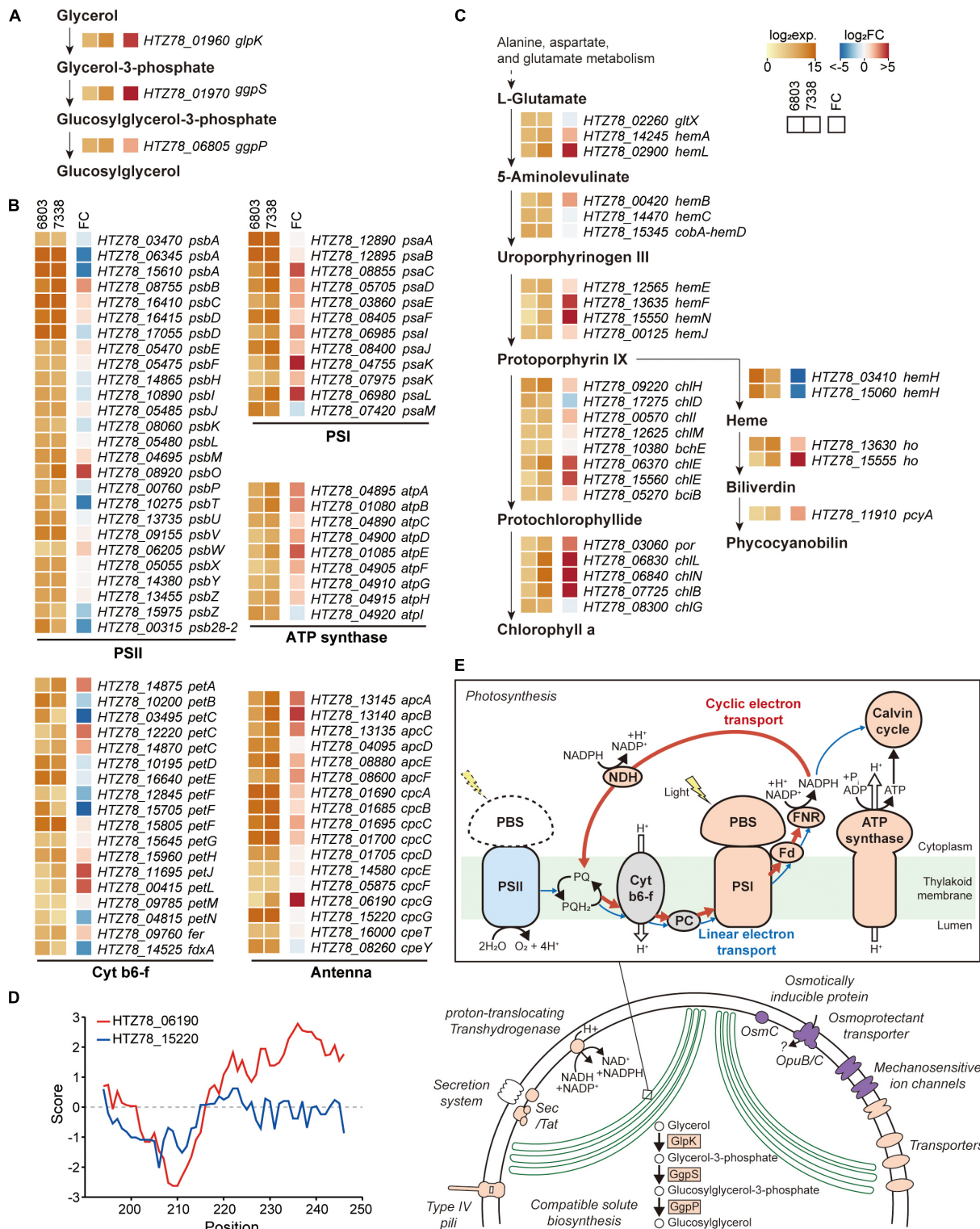
FIGURE 4 | Messenger RNA (mRNA) expressions in *Synechocystis* sp. PCC 7338. **(A–C)** Log$_2$ mRNA expression levels and log$_2$ fold changes of orthologs in *Synechocystis* sp. PCC 7338 compared to those of *Synechocystis* sp. PCC 6803. The orthologs related to glucosylglycerol biosynthesis **(A)**, photosynthesis including photosystem II (PSII), cytochrome b6-f (cyt b6-f), photosystem I (PSI), ATP synthase, and antenna proteins **(B)**, and porphyrin and chlorophyll metabolism pathway **(C)** were described. **(D)** Hydrophobicity calculated at the C-terminal of 2 *cpcG* in *Synechocystis* sp. PCC 7338. **(E)** Schematic illustration of proposed response in *Synechocystis* sp. PCC 7338. Orange color indicates the upregulation of the component compared to *Synechocystis* sp. PCC 6803, and blue color indicates the downregulation. Purple color indicates the specific component that is exclusively present in *Synechocystis* sp. PCC 7338.

(**Supplementary Figure 4E**). In summary, *Synechocystis* sp. PCC 7338 may have evolved by increased distribution of energy to PSI than PSII, thus allowing it to adapt to situations with high ATP demand and insufficient light (**Figure 4E**). This interpretation agrees with the previous observations of increased PSI/PSII ratio under high salt stress, and consistent with the suggestion that optimal photosynthesis can be achieved with short-term state transition and long-term PSI-specific complex formation in cyanobacteria (Hagemann, 2011; Watanabe et al., 2014).

DISCUSSION

We obtained a high-quality genome sequence of marine cyanobacteria *Synechocystis* sp. PCC 7338 and aligned the same with multi-omics data for systematic analysis of its survival under harsh environmental conditions of high salinity and high osmotic pressure. Compared to freshwater *Synechocystis* sp. PCC 6803, *Synechocystis* sp. PCC 7338 is expected to actively use the cyclic electron transport through PSI to replenish the insufficient ATP availability under the conditions (**Figure 4E**). It has been suggested that ATP demand increases under salt stress, which might reduce the carbon fixation rate (van Thor et al., 2000). However, in *Synechocystis* sp. PCC 7338, the gene expression related to carbon fixation increased slightly, indicating no significant reduction in carbon fixation rate (**Figure 4E** and **Supplementary Figure 5**).

A significant downregulation of the D1 protein gene was observed in *Synechocystis* sp. PCC 7338, which agrees with the previous results that the transcription and subsequent production D1 protein is inhibited by salt stress (**Figure 4B**; Allakhverdiev and Murata, 2008; Yang et al., 2020). As a result, the repair rate of damaged PSII would slow down, and oxidative stress would occur when excessive energy is transferred to PSII. Thus, adaptation toward distributing electrons to PSI rather than PSII may be beneficial for survival of marine cyanobacteria. Interestingly, the comparison of 5′-UTR length of orthologs between the two strains showed that the downregulated D1 encoding genes have a much shorter 5′-UTR in *Synechocystis* sp. PCC 7338 (**Figure 3C** and **Supplementary Figure 3B**). Considering the regulation by cis- or trans-regulatory elements in the longer 5′-UTR, it is possible that the reduction in the 5′-UTR length of D1 encoding genes was a strategic change to avoid negative regulation from salt stress or oxidative stress. When the regulatory regions of the orthologs were compared, the promoter regions of PSII genes were more divergent than those of PSI

genes, suggesting that the different regulatory mechanism of PSII to respond more sensitively to stress such as excessive light or oxidative stress in the two strains (**Figure 3D**). In addition to the photosynthesis-related genes, the degenerated promoter regions of the genes involved in response to stress conditions, such as aquaporin and sodium/proton antiporters, inferred the diverged regulatory mechanisms in the two strains. Taken together, our analysis suggests that *Synechocystis* sp. PCC 7338 adapted to a different homeostatic balance compared to the freshwater *Synechocystis* sp. PCC 6803, by evolving its photosynthetic mechanism and different regulatory network through changes in regulatory sequences to avoid various stresses, such as high salinity, high osmotic pressure, and oxidative stress in the marine environment.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA629670.

AUTHOR CONTRIBUTIONS

SC and B-KC designed and supervised the project. YJ, S-JH, and S-HC performed experiments. YJ, S-JH, S-HC, SC, and B-KC analyzed the data. YJ, HL, H-KC, D-MK, C-GL, SC, and B-KC wrote the manuscript. All authors have read and approved the final manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.667450/full#supplementary-material

REFERENCES

Akai, M., Onai, K., Morishita, M., Mino, H., Shijuku, T., Maruyama, H., et al. (2012). Aquaporin AqpZ is involved in cell volume regulation and sensitivity to osmotic stress in *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* 194, 6828–6836. doi: 10.1128/JB.01665-12

Allakhverdiev, S. I., and Murata, N. (2004). Environmental stress inhibits the synthesis de novo of proteins involved in the photodamage-repair cycle of

Photosystem II in *Synechocystis* sp. PCC 6803. *Biochim. Biophys. Acta* 1657, 23–32. doi: 10.1016/j.bbabio.2004.03.003

Allakhverdiev, S. I., and Murata, N. (2008). Salt stress inhibits photosystems II and I in cyanobacteria. *Photosynth. Res.* 98, 529–539.

Allen, J. F., Bennett, J., Steinback, K. E., and Arntzen, C. J. J. N. (1981). Chloroplast protein phosphorylation couples plastoquinone redox state to distribution of excitation energy between photosystems. *Nature* 291, 25–29. doi: 10.1023/A: 1020414106993

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids. Res.* 37, W202–W208.

Camsund, D., and Lindblad, P. (2014). Engineered transcriptional systems for cyanobacterial biotechnology. *Front. Bioeng. Biotechnol.* 2:40. doi: 10.3389/fbioe.2014.00040

Chang, L., Liu, X., Li, Y., Liu, C. C., Yang, F., Zhao, J., et al. (2015). Structural organization of an intact phycobilisome and its association with photosystem II. *Cell Res.* 25, 726–737. doi: 10.1038/cr.2015.59

Cho, S.-H., and Jeong, Y. (2020). *Data from Synechocystis sp. PCC 6803 NGS data.* (Sequence Read Archive). Available online at: https://www.ncbi.nlm.nih.gov/sra/SRR12763770 and https://www.ncbi.nlm.nih.gov/sra/SRR12763771 (accessed October 01, 2020).

Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190.

Cui, J., Sun, T., Chen, L., and Zhang, W. (2020). Engineering salt tolerance of photosynthetic cyanobacteria for seawater utilization. *Biotechnol. Adv.* 43:107578.

Dang, T. C., Fujii, M., Rose, A. L., Bligh, M., and Waite, T. D. (2012). Characteristics of the freshwater cyanobacterium *Microcystis aeruginosa* grown in iron-limited continuous culture. *Appl. Environ. Microbiol.* 78, 1574–1583. doi: 10.1128/AEM.06908-11

Dar, D., Prasse, D., Schmitz, R. A., and Sorek, R. (2016a). Widespread formation of alternative 3' UTR isoforms via transcription termination in archaea. *Nat. Microbiol.* 1:16143. doi: 10.1038/nmicrobiol.2016.143

Dar, D., Shamir, M., Mellin, J. R., Koutero, M., Stern-Ginossar, N., Cossart, P., et al. (2016b). Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science* 352:aad9822. doi: 10.1126/science.aad9822

Dar, D., and Sorek, R. (2018). High-resolution RNA 3'-ends mapping of bacterial Rho-dependent transcripts. *Nucleic Acids Res.* 46, 6797–6805.

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

Grant, C. E., Bailey, T. L., and Noble, W. S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018. doi: 10.1093/bioinformatics/btr064

Hagemann, M. (2011). Molecular biology of cyanobacterial salt acclimation. *FEMS Microbiol. Rev.* 35, 87–123.

Hernandez-Prieto, M. A., Semeniuk, T. A., Giner-Lamia, J., and Futschik, M. E. (2016). The transcriptional landscape of the photosynthetic model cyanobacterium *Synechocystis* sp. PCC6803. *Sci. Rep.* 6:22168. doi: 10.1038/srep22168

Hitchcock, A., Hunter, C. N., and Canniffe, D. P. (2020). Progress and challenges in engineering cyanobacteria as chassis for light-driven biotechnology. *Microb. Biotechnol.* 13, 363–367. doi: 10.1111/1751-7915.13526

Huckauf, J., Nomura, C., Forchhammer, K., and Hagemann, M. (2000). Stress responses of *Synechocystis* sp. strain PCC 6803 mutants impaired in genes encoding putative alternative sigma factors. *Microbiology* 146(Pt 11), 2877–2889. doi: 10.1099/00221287-146-11-2877

Hwang, S., Lee, N., Jeong, Y., Lee, Y., Kim, W., Cho, S., et al. (2019). Primary transcriptome and translatome analysis determines transcriptional and translational regulatory elements encoded in the *Streptomyces clavuligerus* genome. *Nucleic Acids Res.* 47, 6114–6129. doi: 10.1093/nar/gkz471

Imamura, S., and Asayama, M. (2009). Sigma factors for cyanobacterial transcription. *Gene Regul. Syst. Biol.* 3, 65–87.

Jablonsky, J., Papacek, S., and Hagemann, M. (2016). Different strategies of metabolic regulation in cyanobacteria: from transcriptional to biochemical control. *Sci. Rep.* 6:33024. doi: 10.1038/srep33024

Jaffe, E. K. (2003). An unusual phylogenetic variation in the metal ion binding sites of porphobilinogen synthase. *Chem. Biol.* 10, 25–34. doi: 10.1016/s1074-5521(02)00296-x

Jain, C., Rodriguez, R. L., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114.

Jeong, Y., Kim, J. N., Kim, M. W., Bucca, G., Cho, S., Yoon, Y. J., et al. (2016). The dynamic transcriptional and translational landscape of the model antibiotic producer *Streptomyces coelicolor* A3(2). *Nat. Commun.* 7:11605. doi: 10.1038/ncomms11605

Kaushik, M. S., Singh, P., Tiwari, B., and Mishra, A. K. (2016). Ferric uptake regulator (FUR) protein: properties and implications in cyanobacteria. *Ann. Microbiol.* 66, 61–75.

Kim, D., Hong, J. S., Qiu, Y., Nagarajan, H., Seo, J. H., Cho, B. K., et al. (2012). Comparative analysis of regulatory elements between *Escherichia coli* and *Klebsiella pneumoniae* by genome-wide transcription start site profiling. *PLoS Genet.* 8:e1002867. doi: 10.1371/journal.pgen.1002867

Kopf, M., Klahn, S., Scholz, I., Matthiessen, J. K., Hess, W. R., and Voss, B. (2014). Comparative analysis of the primary transcriptome of *Synechocystis* sp. PCC 6803. *DNA Res.* 21, 527–539. doi: 10.1093/dnares/dsu018

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Lalanne, J. B., Taggart, J. C., Guo, M. S., Herzel, L., Schieler, A., and Li, G. W. (2018). Evolutionary convergence of pathway-specific enzyme expression stoichiometry. *Cell* 173, 749.e38–761.e38. doi: 10.1016/j.cell.2018.03.007

Lamb, J. J., Hill, R. E., Eaton-Rye, J. J., and Hohmann-Marriott, M. F. (2014). Functional role of PilA in iron acquisition in the cyanobacterium *Synechocystis* sp. PCC 6803. *PLoS One* 9:e105761. doi: 10.1371/journal.pone.0105761

Lan, E. I., and Liao, J. C. (2011). Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. *Metab. Eng.* 13, 353–363.

Laughlin, T. G., Savage, D. F., and Davies, K. M. (2020). Recent advances on the structure and function of NDH-1: the complex I of oxygenic photosynthesis. *Biochim. Biophys. Acta Bioenerg.* 1861:148254. doi: 10.1016/j.bbabio.2020.148254

Lee, Y., Lee, N., Jeong, Y., Hwang, S., Kim, W., Cho, S., et al. (2019). The transcription unit architecture of *Streptomyces lividans* TK24. *Front. Microbiol.* 10:2074. doi: 10.3389/fmicb.2019.02074

Lindberg, P., Park, S., and Melis, A. (2010). Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metab. Eng.* 12, 70–79. doi: 10.1016/j.ymben.2009.10.001

Lorenz, R., Bernhart, S. H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., et al. (2011). ViennaRNA Package 2.0. *Algorithms Mol. Biol.* 6:26. doi: 10.1186/1748-7188-6-26

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Luimstra, V. M., Schuurmans, J. M., De Carvalho, C. F. M., Matthijs, H. C. P., Hellingwerf, K. J., and Huisman, J. (2019). Exploring the low photosynthetic efficiency of cyanobacteria in blue light using a mutant lacking phycobilisomes. *Photosynth. Res.* 141, 291–301. doi: 10.1007/s11120-019-00630-z

Mitra, A., Angamuthu, K., Jayashree, H. V., and Nagaraja, V. (2009). Occurrence, divergence and evolution of intrinsic terminators across eubacteria. *Genomics* 94, 110–116. doi: 10.1016/j.ygeno.2009.04.004

Mitschke, J., Georg, J., Scholz, I., Sharma, C. M., Dienst, D., Bantscheff, J., et al. (2011). An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2124–2129. doi: 10.1073/pnas.1015154108

Morrissey, J., and Bowler, C. (2012). Iron utilization in marine cyanobacteria and eukaryotic algae. *Front. Microbiol.* 3:43. doi: 10.3389/fmicb.2012.00043

Mullineaux, C. W., and Allen, J. F. (1990). State 1-State 2 transitions in the cyanobacterium *Synechococcus* 6301 are controlled by the redox state of electron carriers between Photosystems I and II. *Photosynth. Res.* 23, 297–311. doi: 10.1007/BF00034860

Nakayama, T., Kamikawa, R., Tanifuji, G., Kashiyama, Y., Ohkouchi, N., Archibald, J. M., et al. (2014). Complete genome of a nonphotosynthetic cyanobacterium in a diatom reveals recent adaptations to an intracellular lifestyle. *Proc. Natl. Acad. Sci. U.S.A.* 111, 11407–11412. doi: 10.1073/pnas.1405222111

Nishiyama, Y., Allakhverdiev, S. I., and Murata, N. (2006). A new paradigm for the action of reactive oxygen species in the photoinhibition of photosystem II. *Biochim. Biophys. Acta* 1757, 742–749. doi: 10.1016/j.bbabio.2006.05.013

Nishiyama, Y., Yamamoto, H., Allakhverdiev, S. I., Inaba, M., Yokota, A., and Murata, N. (2001). Oxidative stress inhibits the repair of photodamage to the photosynthetic machinery. *EMBO J.* 20, 5587–5594.

Noh, Y., Lee, H., Hong, S.-J., Lee, H., Cho, B.-K., Lee, C.-G., et al. (2020). Comparative primary metabolic and lipidomic profiling of freshwater and marine *Synechocystis* strains using by GC-MS and NanoESI-MS analyses. *Biotechnol. Bioproc. Eng.* 25, 308–319.

Oliver, J. W., Machado, I. M., Yoneda, H., and Atsumi, S. (2013). Cyanobacterial conversion of carbon dioxide to 2,3-butanediol. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1249–1254.

Omotajo, D., Tate, T., Cho, H., and Choudhary, M. (2015). Distribution and diversity of ribosome binding sites in prokaryotic genomes. *BMC Genomics* 16:604. doi: 10.1186/s12864-015-1808-6

Perozo, E., and Rees, D. C. (2003). Structure and mechanism in prokaryotic mechanosensitive channels. *Curr. Opin. Struct. Biol.* 13, 432–442.

Qiu, Y., Cho, B. K., Park, Y. S., Lovley, D., Palsson, B. O., and Zengler, K. (2010). Structural and operational complexity of the *Geobacter sulfurreducens* genome. *Genome Res.* 20, 1304–1311. doi: 10.1101/gr.107540.110

Ramey, C. J., Baron-Sola, A., Aucoin, H. R., and Boyle, N. R. (2015). Genome engineering in cyanobacteria: where we are and where we need to go. *ACS Synth. Biol.* 4, 1186–1196. doi: 10.1021/acssynbio.5b00043

Savakis, P. E., Angermayr, S. A., and Hellingwerf, K. J. (2013). Synthesis of 2,3-butanediol by *Synechocystis* sp. PCC6803 via heterologous expression of a catabolic pathway from lactic acid- and enterobacteria. *Metab. Eng.* 20, 121–130. doi: 10.1016/j.ymben.2013.09.008

Sharma, C. M., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., et al. (2010). The primary transcriptome of the major human pathogen *Helicobacter pylori*. *Nature* 464, 250–255. doi: 10.1038/nature08756

Shcolnick, S., Summerfield, T. C., Reytman, L., Sherman, L. A., and Keren, N. (2009). The mechanism of iron homeostasis in the unicellular cyanobacterium *Synechocystis* sp. PCC 6803 and its relationship to oxidative stress. *Plant Physiol.* 150, 2045–2056. doi: 10.1104/pp.109.141853

Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212.

Singh, J. S., Kumar, A., Rai, A. N., and Singh, D. P. (2016). Cyanobacteria: a precious bio-resource in agriculture, ecosystem, and environmental sustainability. *Front. Microbiol.* 7:529. doi: 10.3389/fmicb.2016.00529

Todor, H., Osadnik, H., Campbell, E. A., Myers, K. S., Li, H., Donohue, T. J., et al. (2020). Rewiring the specificity of extracytoplasmic function sigma factors. *Proc. Natl. Acad. Sci. U.S.A.* 117, 33496–33506. doi: 10.1073/pnas.2020204117

Trentin, G., Bertucco, A., and Sforza, E. (2019). Mixotrophy in *Synechocystis* sp. for the treatment of wastewater with high nutrient content: effect of CO2 and light. *Bioprocess Biosyst. Eng.* 42, 1661–1669. doi: 10.1007/s00449-019-02162-1

Unniraman, S., Prakash, R., and Nagaraja, V. (2002). Conserved economics of transcription termination in eubacteria. *Nucleic Acids Res.* 30, 675–684.

van Thor, J. J., Jeanjean, R., Havaux, M., Sjollema, K. A., Joset, F., Hellingwerf, K. J., et al. (2000).). Salt shock-inducible photosystem I cyclic electron transfer in Synechocystis PCC6803 relies on binding of ferredoxin: NADP+ reductase to the thylakoid membranes via its CpcD phycobilisome-linker homologous N-terminal domain. *Biochim. Biophys. Acta Bioenerget.* 1457, 129–144. doi: 10.1016/s0005-2728(00)00072-4

Varman, A. M., Xiao, Y., Pakrasi, H. B., and Tang, Y. J. (2013). Metabolic engineering of *Synechocystis* sp. strain PCC 6803 for isobutanol production. *Appl. Environ. Microbiol.* 79, 908–914.

Vavitsas, K., Rue, E. O., Stefansdottir, L. K., Gnanasekaran, T., Blennow, A., Crocoll, C., et al. (2017). Responses of *Synechocystis* sp. PCC 6803 to heterologous biosynthetic pathways. *Microb. Cell Fact.* 16:140. doi: 10.1186/s12934-017-0757-y

Waditee, R., Hibino, T., Nakamura, T., Incharoensakdi, A., and Takabe, T. (2002). Overexpression of a Na^+/H^+ antiporter confers salt tolerance on a freshwater cyanobacterium, making it capable of growth in sea water. *Proc. Natl. Acad. Sci. U.S.A.* 99, 4109–4114. doi: 10.1073/pnas.052576899

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. doi: 10.1371/journal.pone.0112963

Watanabe, M., Semchonok, D. A., Webber-Birungi, M. T., Ehira, S., Kondo, K., Narikawa, R., et al. (2014). Attachment of phycobilisomes in an antenna-photosystem I supercomplex of cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2512–2517. doi: 10.1073/pnas.1320599111

Xie, C., Mao, X., Huang, J., Ding, Y., Wu, J., Dong, S., et al. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* 39, W316–W322. doi: 10.1093/nar/gkr483

Yang, W., Wang, F., Liu, L. N., and Sui, N. (2020). Responses of membranes and the photosynthetic apparatus to salt stress in cyanobacteria. *Front. Plant Sci.* 11:713. doi: 10.3389/fpls.2020.00713

Zhao, Y., Wu, J., Yang, J., Sun, S., Xiao, J., and Yu, J. (2012). PGAP: pan-genomes analysis pipeline. *Bioinformatics* 28, 416–418.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership