Biosynthesis system of Synechan, a sulfated exopolysaccharide, in the model cyanobacterium

***Synechocystis* sp. PCC 6803**

Kaisei Maeda^{1,2}, Yukiko Okuda¹, Gen Enomoto¹, Satoru Watanabe², Masahiko Ikeuchi^{1,3}

¹Department of Life Sciences (Biology), Graduate School of Arts and Sciences, University of Tokyo,

Komaba, Meguro, Tokyo 153-8902, Japan

²(present address) Department of Bioscience, Tokyo University of Agriculture, Sakuragaoka, Setagaya-

ku, Tokyo 156-8502, Japan

³Faculty of Education and Integrated Arts and Sciences, Waseda University, Tokyo 162-8480, Japan

Abstract

Extracellular polysaccharides of bacteria contribute to biofilm formation, stress tolerance, and infectivity. Cyanobacteria, the oxygenic photoautotrophic bacteria, uniquely and widely have sulfated extracellular polysaccharides and they may utilize the polysaccharides for survival in nature. In addition, sulfated polysaccharides of cyanobacteria and other organisms have been focused as beneficial biomaterial. However, very little is known about their biosynthesis machinery and function in cyanobacteria. Here we found that the model cyanobacterium, *Synechocystis* sp. PCC 6803, formed bloom-like cell aggregates using sulfated extracellular polysaccharides (designated as synechan) and identified whole set of genes responsible for synechan biosynthesis and its transcriptional regulation, thereby suggesting a model for the synechan biosynthesis apparatus. Because similar genes are found in many cyanobacterial genomes with wide variation, our findings may lead elucidation of various sulfated polysaccharides, their functions, and their potential application in biotechnology.

23 **Introduction**

24 Bacterial extracellular polysaccharides establish biofilms for nutrient supply and stress

25 avoidance, and they sometimes support cellular activities such as motility and infectivity (*Woodward*

26 *and Naismith, 2016*). Generally, the polysaccharide chains consist of a few types of sugars (with or

27 without chemical modifications) and are anchored on cells (capsular polysaccharides, CPS) or exist as

28 nonanchored exopolysaccharides (EPS). Nonetheless, their molecular structures vary greatly, e.g.,

29 branching schemes, sugar constituents, and modifications, and thus their physical properties also vary.

30 Bacterial extracellular polysaccharides and lipopolysaccharides are produced and exported via three

31 distinct pathways: the Wzx/Wzy-dependent pathway, ABC-dependent pathway, and synthase-dependent

32 pathway (*Schmid et al., 2015*). Every bacterium can produce several extracellular polysaccharides, and

33 production often depends on environmental conditions. Some extracellular polysaccharides have been

34 appropriated for use as biopolymers for food, cosmetics, medicine (*Freitas et al., 2014, Lapasin and*

35 *Pricl, 1995*).

Cyanobacteria, the oxygenic photoautotrophic bacteria that inhabit almost every ecosystem on Earth, contribute to the global photosynthetic production (*Flombaum et al., 2013, Mangan et al., 2016*). Cyanobacteria produce various extracellular polysaccharides to form colonies, which are planktonic or attached on solid surfaces, likely to stay in a phototrophic niche in nature (*De Philippis and Vincenzini, 1998*). A notable example is the water bloom, a dense population of cyanobacterial cells that floats on the water surface and often produces cyanotoxins and extracellular polysaccharides (*Huisman et al., 2018*). The extracellular polysaccharides are also important for photosynthetic production of cyanobacteria and their application (*Kumar et al., 2018*). However, very little is known about their biosynthesis except for extracellular cellulose. A thermophilic cyanobacterium (*Thermosynechococcus vulcanus*) accumulates cellulose to form cell aggregation (*Kawano et al., 2011*). This cellulose is produced by cellulose synthase with unique tripartite system and regulations (*Enomoto et al., 2015, Maeda et al., 2018*). In the cyanobacterial genomes, there are still many putative genes for extracellular polysaccharide biosynthesis.

Uniquely, many cyanobacterial extracellular polysaccharides are sulfated, i.e., as a sugar modification (*Pereira et al., 2009*). Sulfated polysaccharides are also produced by animals (as glycosaminoglycan in the extracellular matrix such as heparan sulfate) and algae (as cell-wall components such as carrageenan) but are scarcely known in other bacteria or plants (*Ghosh et al., 2009*). Major examples of cyanobacterial sulfated polysaccharides are spirulan from *Arthrospira platensis* (vernacular name, “Spirulina”), sacran from *Aphanothece sacrum* (vernacular name, “Suizenji-Nori”) and cyanoflan from *Cyanothece* sp. CCY 0110 (*Mota et al., 2020, Mouhim et al., 1993, Okajima et al., 2008*). These sulfated polysaccharides are used for formation of colony and biofilm and may be functionally relevant to the ecology of cyanobacteria (*Fujishiro et al., 2004*). In addition, the bioactivities (antiviral, antitumor, and anti-inflammatory) of sulfated polysaccharides from cyanobacteria were reported, too (*Flores et al., 2019a, Hayashi et al., 1996, Ngatu et al., 2012*). However, very little is known about their biosynthesis machinery and physiological functions. On the other hand, biosynthesis and modification of animal sulfated polysaccharides have been extensively

studied because of their importance to tissue protection, tissue development, and immunity (*Karamanos et al., 2018, Sasisekharan et al., 2006*) and potential applications in healthy foods, biomaterials, and medicines (*Jiao et al., 2011, Wardrop and Keeling, 2008*).

The poor understanding about cyanobacterial sulfated polysaccharide biosynthesis is probably due to low or no accumulation of sulfated polysaccharides in typical model species (*Pereira et al., 2009*).

More than three decades ago, Panoff et al. reported sulfated polysaccharides in two related model cyanobacteria, *Synechocystis* sp. PCC 6803 and *Synechocystis* sp. PCC 6714 (hereafter *Synechocystis* 6803 and *Synechocystis* 6714) (*Panoff et al., 1988*). Recently, Flores et al. confirmed sulfated polysaccharides and reported their enhanced accumulation in a sigma factor *sigF* mutant for global cell surface regulation in *Synechocystis* 6803 (*Flores et al., 2019b*). In parallel, several papers have studied genes that could be involved in extracellular polysaccharides biosynthesis in *Synechocystis* 6803, but no clear results were obtained about the sulfated polysaccharide (*Fisher et al., 2013, Foster et al., 2009, Jittawuttipoka et al., 2013, Pereira et al., 2019*). Here we found that a motile substrain of *Synechocystis*

6803 showed bloom-like cell aggregation and sulfated EPS production, but a non-motile substrain (a standard substrain for photosynthesis study) did not. By gene disruption and overexpression, we first identified a whole set of genes responsible for sulfated EPS biosynthesis and its regulatory system, opening the way to engineering of their production.

Results

Bloom formation and EPS accumulation in Synechocystis sp. PCC 6803

We fortuitously found that a motile substrain of *Synechocystis* 6803 produces EPS and forms floating cell aggregates resembling a typical cyanobacterial bloom. We established a two-step culture regime (2-day bubbling culture and subsequent standing culture without bubbling under continuous light) for reproducible formation of bloom-like aggregates (Fig. 1A, B). The first (bubbling) step allows for cell propagation and EPS production, whereas the second (standing) step allows for heavy-cell aggregation and flotation, even though *Synechocystis* 6803 does not possess genes for intracellular gas

vesicles (*Harke et al., 2016*). In *Synechocystis* 6803, cell flotation accompanying the generation of extracellular gas bubbles was suppressed by inactivation of photosynthesis (Fig. 1C), suggesting that gas derived from photosynthesis drives the upward movement of cells embedded in viscous EPS. The non-motile, glucose-tolerant substrain—commonly used for photosynthesis research—did not aggregate or float.

We first isolated crude EPS from the bloomed culture by membrane filtration (Fig. S1A). The crude EPS consisted of polysaccharide but very little protein or nucleic acid, and its abundance remained unchanged during the second culture step (Fig. S1B). As a common feature of diverse EPS biosynthesis systems in bacteria, membrane-bound glycosyltransferases are particularly important (*Schmid et al., 2015*). So, we screened such glycosyltransferase genes by disruption and revealed that *slr5054* is essential for bloom formation (Fig. 1D, E and Fig. S2). The EPS preparation was improved by removing cells before filtration to avoid cell-associated polysaccharides such as CPS (Fig. 1F). $\Delta slr5054$ lacked most of the EPS present in the wild type (WT), whereas the CPS and free polysaccharide fractions were

Figure 1 Bloom formation and EPS isolation.

A, Time course of bloom formation by WT *Synechocystis* 6803 during the second step of culture. Extracellular gas bubbles are formed and trapped in viscous EPS (~1 h). Green vertical columns with bubbles become apparent at 4 h. Those trapped gas bubbles slowly rise together with the viscous columns. **B**, Enlarged images showing gas bubbles trapped in EPS. Vertically aligned bubbles are indicated by red arrows. **C**, Lack of bloom formation in the non-motile substrain (NM) or WT with or without light and the photosynthesis inhibitor DCMU at 48 h of the second step of culture. **D**, Domain architecture of Slr5054. GT, glycosyltransferase domain; TM, transmembrane region. **E**, Lack of bloom formation in $\Delta slr5054$ after standing culture for 48 h. **F**, Isolation of EPS from the first step of culture. Cells and CPS were removed from the culture by centrifugation, and EPS in the supernatant was separated from “free” polysaccharide (PS) by membrane filtration followed by a second centrifugation to remove residual cells. CPS was collected from the cell pellet after vortexing and centrifugation. **G**, Sugar content of fractions from WT and $\Delta slr5054$. Error bars represent SD (n = 3, **P < 0.005).

Figure S1. Isolation of crude EPS and sugar analysis.

A, Protocol for isolating crude EPS from the bloomed culture. The bloom, including EPS and cells at the end of the second (standing) step of culture, was trapped by membrane filtration, and crude EPS was recovered from the bloom by vortexing and centrifugation. PS, polysaccharide; ppt, precipitate; sup, supernatant. **B**, Time course of sugar accumulation in the crude EPS during the standing culture. **C**, Sugar content of the crude EPS from WT, and a non-motile glucose-tolerant substrain (NM). Error bars represent SD (n = 3, **P < 0.005).

Figure S2. Bloom formation by several glycosyltransferase mutants.

A, Domain architecture of membrane-bound glycosyltransferases. Red box, glycosyltransferase domain (GT); blue box, transmembrane region; orange box, glycogen phosphorylase domain. **B**, Bloom formation by mutants after the second step of culture.

similar in the WT and $\Delta slr5054$ (Fig. 1G). Then we performed Alcian blue staining to examine the acidity of the EPS (Fig. S3). Generally, sulfated polysaccharides are stained at pH 0.5 condition, while acidic polysaccharides, which contain sulfate groups and/or carboxylate groups (such as uronic acids and carboxylate modification) are stained at pH 2.5 condition (Bellezza *et al.*, 2006). The EPS from WT was clearly stained under both pH conditions, strongly suggestive of the sulfate modification.

Gene cluster for the biosynthesis of viscous polysaccharides

slr5054 resides on a megaplasmid, pSYSM, in a large gene cluster (*slr5042–60*), which we named *xss* (extracellular sulfated polysaccharide biosynthesis) (Fig. 2A, *xssA–xssS*). This cluster includes two genes for sulfotransferases (*xssA*, *xssE*), eight genes for glycosyltransferases (*xssB*, *xssC*, *xssG*, *xssI*, *xssM*, *xssN*, *xssO*, *xssP*), three genes for the polysaccharide polymerization system (Wzx/flippase; *xssH*, Wzy/polymerase; *xssF*, and polysaccharide co-polymerase [PCP]; *xssK*), one gene for a putative transcriptional regulator (*xssQ*), a pair of genes for the bacterial two-component

Figure S3 Alcian blue staining of EPS from WT.

The microscopy images of isolated EPS from WT culture (Fig. 1g) stained with alcian blue at pH 2.5 (left) and pH 0.5 (right).

Figure 2 The *xss* gene cluster and phenotype of *xss* mutants.

A, The *xss* gene cluster. Red, polysaccharide biosynthesis genes; green, regulatory genes; black, genes of unknown function. **B**, Bloom formation by the mutants carrying disruptions in the polysaccharide biosynthesis *xss* genes. **C**, Total sugar content (μg glucose per 1×10^8 cells) of the EPS fraction from mutants in b. Red bars, bloom-forming mutants; blue bars, non-bloom-forming mutants. Error bars represent SD (WT grown at 20°C, $n = 6$; others, $n = 3$). Statistical significance was determined using Welch's t test (* $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$). **D**, Bloom formation by regulatory mutants, WT grown at 20°C, and OPX mutant ($\Delta sll1581$). **E**, Total sugar content of the EPS fraction from mutants in d. Red bars, bloom-forming mutants; green bars, excess-bloom-forming mutants. **F**, A sheet of OX-xssR cells was stripped off from the agar plate by tweezers. The culture temperature was 31°C unless otherwise stated.

160 phosphorelay system (*xssR*, *xssS*), and genes encoding several small proteins of unknown function

161 (Table S1, Fig. 2A). All genes except those of unknown function were disrupted individually with a

162 read-through cassette, and segregation was confirmed by colony PCR (Fig. S4). Bloom formation and

163 sugar content of the EPS fraction were reduced in many mutants (Fig. 2B, C). In particular, bloom

164 formation was completely abolished in $\Delta xssA$, $\Delta xssB$, $\Delta xssF$, $\Delta xssH$, $\Delta xssK$, $\Delta xssM$, $\Delta xssN$, and $\Delta xssP$,

165 in which EPS accumulation was also suppressed. Certain glycosyltransferase mutants ($\Delta xssC$, $\Delta xssG$,

166 $\Delta xssI$, $\Delta xssO$) formed blooms but accumulated little EPS, and neither bloom formation nor EPS

167 accumulation was substantially altered in one sulfotransferase mutant ($\Delta xssE$). In general, the Wzx/Wzy

168 system in bacteria produces various EPS, lipopolysaccharides, and CPS through four steps: (i)

169 biosynthesis of a heterooligosaccharide repeat unit on a lipid linker at the cytoplasmic side of the plasma

170 membrane by a series of glycosyltransferases and modification enzymes, (ii) flip-out of the unit to the

171 periplasmic side by Wzx, (iii) polymerization by transfer of the nascent polysaccharide chain to the

172 repeat unit by Wzy, and (iv) export of the EPS chain through the periplasm and outer membrane via PCP

Table S1. Summary of *Synechocystis* 6803 Xss proteins.

Category	Name	Locus tag	Function	Importance	Regulation
Polysaccharide chain biosynthesis	XssP	<i>sll5057</i>	Priming GT	+++	+
	XssB	<i>sll5043</i>	GT	+++	+
	XssM	<i>slr5054</i>	GT	++	—
	XssN	<i>slr5055</i>	GT	++	+
	XssC	<i>sll5044</i>	GT	+	+
	XssG	<i>sll5048</i>	GT	+	—
	XssI	<i>sll5050</i>	GT	+	—
	XssO	<i>slr5056</i>	GT	+	+
Modification	XssA	<i>sll5043</i>	ST	+++	+
	XssE	<i>sll5046</i>	ST	±	+
Translocation	XssH	<i>sll5049</i>	Wzx/Flippase	++	—
Polymerization and exportation	XssF	<i>sll5047</i>	Wzy/Polymerase	++	—
	XssK	<i>sll5052</i>	PCP-2a/Exporter	+++	—
	XssT	<i>sll1581</i>	OPX/Exporter	++	—
Regulation	XssQ	<i>slr5058</i>	Transcriptional regulator	+++	n.d.
	XssR	<i>sll5059</i>	Response regulator	+++	n.d.
	XssS	<i>sll5060</i>	Sensor histidine kinase	+++	n.d.

Function: GT, glycosyltransferase; ST, sulfotransferase; PCP, polysaccharide co-polymerase; OPX, outer-membrane polysaccharide export protein.

Importance: contribution in EPS accumulation.

Regulation: transcriptional regulation of *xss* genes by XssQ/R/S. n.d., not determined.

Figure S4. Agarose gel electrophoresis to assess segregation of mutants based on PCR data.

M indicates marker lane, and the numbers above each lane indicate the different clones. The band positions of wild type (WT) and mutants (MT) are shown at the right, with theoretical lengths in nt. A solid arrow indicates the existence of the band; a dotted arrow indicates the absence of the band. The bold text in the strain names indicates the region assessed by PCR.

and the outer-membrane polysaccharide export protein (OPX) (*Islam and Lam, 2014, Schmid et al., 2015*). It is very likely that the *xss* cluster harbors a whole set of genes for the Wzx/Wzy-dependent pathway except a gene for OPX.

Regulation of the sulfated EPS biosynthesis

The sensory histidine kinase mutant $\Delta xssS$ accumulated a much larger amount of EPS than WT, whereas mutants of the cognate response regulator *xssR* and transcriptional regulator *xssQ* had a null phenotype with regard to both bloom formation and EPS accumulation (Fig. 2D, E, Table S2). The double mutant $\Delta xssS/\Delta xssR$ had a phenotype similar to that of $\Delta xssR$. Overexpression of *xssR* or *xssQ* (OX-*xssR*, OX-*xssQ*) resulted in strong bloom formation as well as marked accumulation of viscous EPS, similar to that seen for $\Delta xssS$. The combination of *xssQ* disruption and *xssR* overexpression ($\Delta xssQ$ +OX-*xssR*) abrogated bloom formation and EPS accumulation, whereas the combination of *xssQ* overexpression and *xssR* disruption ($\Delta xssR$ +OX-*xssQ*) resulted in a pronounced phenotype of bloom

formation and EPS accumulation. These results suggested that the sensor histidine kinase XssS suppresses the response regulator XssR, leading to activation of the transcriptional activator XssQ. Notably, the OX-*xssR* and OX-*xssQ* strains formed sticky, non-motile, biofilm-like colonies on agar plates that could be picked by tweezers (Fig. 2F).

XssQ is a new type of the signal transduction ATPase with numerous domains (STAND) protein, because it harbors an N-terminal helix-turn-helix transcriptional DNA-binding domain (Fig. S5). Typical STAND proteins possess a three-domain module with ATPase activity and are involved in processes such as apoptosis and immunity in animals, plants, and some bacteria (Danot *et al.*, 2009). Using real-time quantitative PCR (qPCR), we compared gene expression in the *xss* cluster for WT, $\Delta xssS$, and $\Delta xssQ$ (Fig. 3A). Expression of five genes (*xssA*, *xssB*, *xssE*, *xssN*, *xssP*) was very low in $\Delta xssQ$ compared with WT, whereas that of *xssF*, *xssH*, and *xssK* was not substantially affected. These results suggested that XssQ transcriptionally activates genes encoding sulfotransferases and certain glycosyltransferases but not genes for polymerization and export via the Wzx/Wzy system. qPCR

Table S2. EPS accumulation and bloom formation by WT *Synechocystis* 6803 and *xss* mutants.

Strain name	Classification	Glc (μ g) per 1.0×10^8 cells	Ratio (% WT)	Bloom
WT(31°C)		0.526 ± 0.062	100.0	+
WT(20°C)		1.629 ± 0.472	309.9	+
$\Delta xssA$	ST3	0.054 ± 0.042	10.4	-
$\Delta xssB$	GT2	0.040 ± 0.029	7.7	-
$\Delta xssC$	GT25	0.148 ± 0.020	28.1	+
$\Delta xssE$	ST3	0.402 ± 0.127	76.6*	+
$\Delta xssF$	Wzy / polymerase	0.107 ± 0.008	20.3	-
$\Delta xssG$	GT4	0.105 ± 0.018	20.1	+
$\Delta xssH$	Wzx / Flippase	0.112 ± 0.059	23.0	-
$\Delta xssI$	GT4	0.113 ± 0.044	21.6	+
$\Delta xssJ$	-	0.224 ± 0.025	42.6	+
$\Delta xssK$	PCP2a	0.037 ± 0.012	7.0	-
$\Delta xssM$	GT2	0.049 ± 0.025	14.6	-
$\Delta xssN$	GT26	0.115 ± 0.025	21.8	-
$\Delta xssO$	GT2	0.161 ± 0.025	30.6	+
$\Delta xssP$	Priming GT	0.039 ± 0.010	7.4	-
$\Delta xssQ$	Transcriptional regulator	0.036 ± 0.010	6.8	-
$\Delta xssR$	Response regulator	0.059 ± 0.009	11.2	-
$\Delta xssS$	Histidine kinase	9.770 ± 0.524	1860	++
$\Delta xssR-S$		0.062 ± 0.008	11.8	-
$\Delta xssK \sim M + \Delta xssS$		0.044 ± 0.013	8.4	-
OX- $xssR$		10.426 ± 0.225	1984	++
OX- $xssQ$		7.359 ± 2.36	1400	++
$\Delta xssQ + OX-xssR$		0.077 ± 0.027	14.6	-
$\Delta xssR + OX-xssQ$		4.837 ± 2.169	920	+
$\Delta sll1581$ ($xssT$)	OPX / exporter	0.103 ± 0.012	19.6	-

Classification: ST, sulfotransferase; GT, glycosyltransferase; PCP, polysaccharide co-polymerase; OPX, outer-membrane polysaccharide export protein. Total sugar content of the EPS fraction is expressed as μ g glucose / 1×10^8 cells. The errors are based on SD (n = 3). *Close to the detection limit. EPS accumulation ratio, percent of WT grown at 31°C. Bloom formation is summarized from Fig. 2.

Figure S5. Domain architecture and proposed signal transduction pathway for XssS/XssR/XssQ.

Hybrid-type histidine kinase XssS presumably senses a cold signal and transduces it to the response regulator XssR, which in turn activates the transcription of *xss* genes via the STAND protein XssQ. Generally, STAND proteins consist of the three-domain module, sensor region, and effector region¹³. In XssQ, the three-domain module consisting of the NACHT (NAIP [neuronal apoptosis inhibitor protein], CIIA [MHC class II transcription activator], HET-E, and TP1 [mammalian telomerase-associated proteins]) domain (PF05729), HD1 (helical domain 1), and WHD (winged helix domain) responds to a signal from XssR and oligomerizes, leading to the activation of the N-terminal effector domain (helix-turn-helix XRE family domain [SM00530]).

analysis of gene expression in $\Delta xssS$ revealed a tendency for slight upregulation of *xssB*, *xssE*, *xssN*, and *xssP*. We performed RNA-seq analysis of WT, $\Delta xssS$, and $\Delta xssQ$ to see the transcriptome (Fig. S6). The genes down-regulated in $\Delta xssQ$ and up-regulated in $\Delta xssS$ were mostly *xss* genes. In detail, the regulated genes were *xssA-E* and *xssL-P*, which were roughly consistent with the qPCR analysis. We conclude that *xssA-E* and *xssL-P* were specifically regulated by XssS/XssR/XssQ. In a previous report, *xssA-xssE* and *xssL-xssP* were up-regulated at low temperature in another substrain of *Synechocystis* 6803 (Kopf *et al.*, 2014b). To test this in our substrain, we measured the sulfated EPS accumulation of WT culture at 20°C, and it was 3.1-fold greater than that at normal growth temperature (31°C; Fig. 2D, E and Table S2). This result suggests that XssS/XssR/XssQ is a unique temperature sensor for *xss* gene expression.

We aligned nucleotide sequences near the transcription start sites of the regulated genes (*xssA*, *xssE*, *xssL*, *xssN*, and *xssP*) to find the consensus sequences for XssQ binding (Fig. 3B), according to the differential RNA-seq-type transcriptomic analysis of *Synechocystis* 6803 (Kopf *et al.*, 2014b). There are single or tandem consensus sequence, AAGTTXXAC. To confirm the binding of XssQ to this region,

Figure 3 Transcriptional regulation of *xss* genes.

A, Transcript levels for *xss* genes in WT, $\Delta xssS$, and $\Delta xssQ$ measured by qPCR. The internal standard was *rnpB*. Relative expression levels were obtained by normalization to the transcript levels of each gene in WT. Error bars represent SD (n = 3, biological triplicates). **B**, upper; Sequence comparison of upstream regions of the five regulated genes. Putative consensus regions are shown in red and fully conserved nucleotides are shown in bold letter. Underlines represent transcribed regions based on the report (Kopf et al., 2014b) and initiation codons of regulated genes are shown in black/bold letters. **B**, lower: Sequences of DNA probe and competitors for *xssE* (native and mutant) used for EMSA of C. Consensus regions are shown in red, and mutated region are shown in blue. Total DNA size is 251 bp, where identical sequences are mostly not shown except 20 bp at both ends. **C**, The autoradiogram image of EMSA of XssQ protein and the DNA probe of *xssE* with some competitors.

Figure S6. Scatter plot of transcriptome comparison among WT, $\Delta xssS$, and $\Delta xssQ$.

Transcriptional levels were based on RPKM (Reads Per Kilobase of exon per Million mapped reads) value shown in supplementally data 1.

we performed electrophoretic mobility shift assay (EMSA) using purified recombinant XssQ protein and a PCR-amplified DNA fragment of *xssE* upstream (Fig. 3C). The band position of the radiolabeled probe DNA shifted reflecting the concentration of XssQ. This shift was largely eliminated by excess addition of the unlabeled native competitor, but not by addition of the mutant competitor with mutations in the consensus region. These results suggest that XssQ recognizes the consensus sequence of *xssE* and other target genes for their transcriptional activation.

The chemical composition of the sulfated EPS

EPS of WT and the 19-fold overproduction mutant ($\Delta xssS$) were subjected to chemical composition analysis (Table 1, Fig. S7). EPS from WT included various sugars and some sulfate groups, whereas EPS from $\Delta xssS$ consisted of only four types of sugars and sulfate groups with the near stoichiometric molar ratio of rhamnose:mannose:galactose:glucose:sulfate of 1:1:1:5:2. This finding roughly fits with the gene number, i.e., eight glycosyltransferase genes and two sulfotransferase genes.

We speculated that the overaccumulation of EPS in $\Delta xssS$ reflects the true product of the *xss* cluster. The EPS from WT may contain a considerable amount of unrelated polysaccharides, which were erroneously recovered together with the *xss* product. Here, the sulfated EPS produced by the *xss* cluster in *Synechocystis* 6803 was designated “Synechan”.

The OPX protein for synechan biosynthesis

There is no candidate gene in the *xss*-carrying plasmid for the OPX protein of the Wzx/Wzy system, whereas *sll1581*, an OPX homolog, was found on the main chromosome. Disruption of *sll1581* ($\Delta sll1581$) abolished bloom formation and EPS accumulation (Fig. 2D, E). Thus, the chromosomal OPX protein Sll1581 (XssT) appears to serve as the outer-membrane exporter for synechan. Interestingly, *Synechocystis* 6803 possesses *xssT* (OPX gene) and *sll0923* (a second PCP-2a gene) on the main chromosome and *xssK* (PCP-2a gene) on the plasmid pSYSM, whereas its close relative *Synechocystis* 6714 harbors only homologs of *sll0923* and *xssT* but lacks the entire plasmid carrying the *xss* cluster.

Table 1. Chemical composition of the EPS from WT *Synechocystis* 6803 and $\Delta xssS$ mutant.

	<i>Sugars</i>									<i>Sulfate residues</i>		
	Neutral sugars (mol/mol %)								Uronic acids (mol/mol %)		Substitution degree (mol/mol %)	
	Rhamnose	Ribose	Mannose	Fucose	Galactose	Xylose	Glucose	Total	Galacturonic acid	Glucuronic acid		Total
WT	16.6	N.D.	25.7	16.2	4.7	10.6	23.1	96.9	N.D.	3.1	3.1	10.4
$\Delta xssS$	13.1	N.D.	14.2	1.2	12.5	1.0	57.9	99.9	N.D.	0.1	0.1	26.6

Figure S7. Chromatograms of HPLC and anion exchange column chromatography of EPS.

A and B, HPLC profiles for neutral sugars in wild-type EPS (A) and $\Delta xssS$ EPS (B). **c and d,** HPLC profiles for uronic acids in wild-type EPS (C) and $\Delta xssS$ EPS (D). The corresponding monosaccharide and retention time are noted at each peak. **E and F,** HPLC profiles for SO_4^{2-} after hydrolysis of EPS samples of wild-type (E) and $\Delta xssS$ (F).

This suggested that XssT serves as an OPX for dual function for both XssK and Sll0923. It is likely that *Synechocystis* 6803 acquired pSYSM and borrowed the chromosomal OPX gene *xssT* to produce synechan or, alternatively, *Synechocystis* 6714 may have lost the plasmid.

Discussion

Summarizing these data, we propose models for synechan biosynthesis apparatus including OPX and temperature-responsive regulation (Fig. 4A, B and Fig. S5). The model of the Xss apparatus fits well with the known Wzx/Wzy-dependent apparatus represented by xanthan biosynthesis in *Xanthomonas campestris* (Katzen *et al.*, 1998). The eight glycosyltransferases including XssP (the priming glycosyltransferase) produce oligosaccharide repeat unit of eight sugars, which is consistent with the sugar composition of synechan. These findings suggest that the *xss* cluster on the pSYSM plasmid harbors a whole set of genes for synechan biosynthesis except the OPX gene (*xssT* on the main chromosome). Notably, the cluster harbors two sulfotransferase genes, which have not been found to

Figure 4 Proposed models for the synechan biosynthesis apparatus, transcriptional regulation, and bloom formation.

A, Model for the synechan biosynthesis apparatus with sugar polymerization and modification. Red boxes represent biosynthesis components, and red boxes with blue stripes represent transcriptionally regulated components. A putative lipid linker is shown as a black rod. Each monosaccharide is shown as a small hexagon, and each sulfate group is shown as a red spot. **B**, Signaling and transcriptional regulation model. Green arrows and ellipses represent regulatory genes and proteins, respectively. Genes for synechan biosynthesis are shown in red, and transcriptionally regulated genes are depicted with blue stripes. Arrows with double lines represent transcriptional activation. **C**, Two flotation models for bloom formation in cyanobacteria. Left, flotation of cells (green circles) with intracellular gas-filled vesicles (white circles). Right, flotation of EPS (red shading)-entrapped cells (green circles) and extracellular gas bubbles (white circles), which are generated by photosynthesis.

our knowledge in other bacterial gene clusters for extracellular polysaccharide biosynthesis.

Sulfotransferases, XssA and XssE, belong to distinct subfamilies of bacterial sulfotransferases. We found many sulfotransferase genes in various cyanobacterial genomes by Pfam search (PF00685, PF03567, PF13469). They are mostly found in gene clusters for putative extracellular polysaccharide biosynthesis (Wzx/Wzy-type and ABC-type) (Fig. S8). It should be noted that they are more or less partial as a cluster for extracellular polysaccharide biosynthesis system, whereas the *xss* cluster appears to be complete except the OPX gene in *Synechocystis* 6803. It is well established that the polysaccharide moiety of membrane-anchored lipopolysaccharides and CPS of bacteria are produced and exported by the Wzx/Wzy-dependent or ABC transporter-dependent pathways, whereas free EPS, i.e., xanthan and cellulose, are produced by the Wzx/Wzy-dependent and synthase-dependent pathways but not by the ABC transporter-dependent pathway (Schmid *et al.*, 2015, Willis and Whitfield, 2013). In the literature, a sulfated CPS was reported in *Arthrospira platensis* (formerly *Spirulina platensis*) (Mouhim *et al.*, 1993). This sulfated CPS may be produced by an ABC transporter-type gene cluster in Fig. S8. Gene

Cyanothece sp. PCC 7822

Cyan7822_0906~0913

A horizontal gene cluster for Cyanothece sp. PCC 7822. It consists of eight genes represented by colored arrows pointing right. From left to right: a yellow arrow, a red arrow, a red arrow, a red arrow, a red arrow, a yellow arrow, a magenta arrow, and a red arrow.

Arthrospira platensis C1

SPLC1_S542070~S542140

A horizontal gene cluster for Arthrospira platensis C1. It consists of eight genes represented by colored arrows pointing right. From left to right: a yellow arrow, a yellow arrow, a blue arrow, a yellow arrow, a yellow arrow, a magenta arrow, a yellow arrow, and a yellow arrow.

Acaryochloris marina MBIC11017

AM1_5713~5721

A horizontal gene cluster for Acaryochloris marina MBIC11017. It consists of eight genes represented by colored arrows pointing right. From left to right: a yellow arrow, a yellow arrow, a yellow arrow, a magenta arrow, a yellow arrow, a magenta arrow, a blue arrow, and a yellow arrow.

Figure S8. Typical examples of putative gene clusters for biosynthesis of sulfated polysaccharides in cyanobacteria.

A part of each cluster harboring genes for sulfotransferases, glycosyltransferases, and polysaccharide biosynthesis/export systems (Wzx/Wzy type and ABC transporter type) in cyanobacterial genomes.

disruption will confirm such predictions deduced from the gene cluster analyses, although targeted disruption is not so easy in many cyanobacteria due to poor transformation efficiency except *Synechocystis* 6803. In contrast, no sulfated polysaccharide has been reported in the other bacteria, though many sulfotransferases are also registered in Pfam database. Some of them are known to transfer a sulfuryl group to lipo-oligosaccharides in rhizobia (Nod factor) and mycobacteria (Mougous *et al.*, 2002).

XssQ, a STAND protein with a DNA binding domain is indeed the transcriptional activator for *xssE* and other induced genes. XssQ homologs are found widely throughout the cyanobacteria but the set of XssS/XssR/XssQ is found near the gene cluster for sulfated EPS biosynthesis with sulfotransferases in many cyanobacteria (Fig. S9, Table S3). Consensus sequences are also found in upstream of some genes in the cluster, suggesting that the XssS/XssR/XssQ system may operate universally for induction of sulfated EPS production under certain environmental conditions such as cold temperature.

Acidic polysaccharides containing uronic acids and other carboxylic groups are common in bacteria, but sulfated polysaccharides are produced exclusively by cyanobacteria (*Pereira et al., 2009*). To speculate on the physiological significance of sulfated polysaccharides, we summarized the distribution of sulfotransferase genes in cyanobacteria of various habitats (Table S3). Species living in high salinity environments such as a salt lake or seawater mostly produce sulfated polysaccharides. The freshwater species *Synechocystis* sp. PCC 6714 lacks the *xss* genes including sulfotransferases and salt-resistance genes that are present in the more salt-resistant *Synechocystis* 6803 (*Kopf et al., 2014a*). A cyanobacterial sulfated polysaccharide, sacran, shows much higher capacity for saline absorption than does hyaluronic acid, a uronic acid-containing polysaccharide, whereas both absorb pure water efficiently (*Okajima et al., 2008*). The wide distribution of sulfated polysaccharides among many cyanobacteria may reflect their inherent compatibility with survival in some saline environments.

There are many putative genes that may be involved in biosynthesis and export of extracellular polysaccharides and lipopolysaccharides in the genome of *Synechocystis* 6803

Figure S9. Homologous gene clusters of *xssS*/*xssR*/*xssQ* in some cyanobacteria.

The colored blocks indicate genes with direction. The homologs of each gene are represented by their corresponding colors (red, *xssQ*; blue, *xssR*; green, *xssS*).

Table S3. Number of sulfotransferase genes (STs) in the genome of cyanobacteria collected from various habitats.

Habitat	Species	STs	Comment
Marine	<i>Trichodesmium erythraeum</i> IMS101	30	
Marine	<i>Crocospaera subtropica</i> ATCC 51142	14	Formerly <i>Cyanothece</i> sp. ATCC 51142
Marine	<i>Acaryochloris marina</i> MBIC 11017	13	
Marine	<i>Synechococcus</i> sp. WH8102	12	
Marine	<i>Pseudanabaena</i> sp. PCC 7367	8	
Marine	<i>Rivularia</i> sp. PCC 7116	5	
Marine	<i>Synechococcus</i> sp. PCC 7002	4	
Marine	<i>Prochlorococcus marinus</i> MIT 9313	2	
Marine	<i>Nodularia spumigena</i> UHCC 0039	0	
Salt lake	<i>Arthrospira platensis</i> NIES-39	16	
Salt lake	<i>Acaryochloris</i> sp. CCMEE 5410	6	
Terrestrial (rice field)	<i>Gloeotheca verrucosa</i> PCC 7822	17	Formerly <i>Cyanothece</i> sp. PCC 7822
Terrestrial (rice field)	<i>Rippakea orientalis</i> PCC 8801	8	Formerly <i>Cyanothece</i> sp. PCC 8801
Terrestrial (rice field)	<i>Cyanothece</i> sp. PCC 7425	6	
Terrestrial	<i>Gloeobacter violaceus</i> PCC 7421	10	
Terrestrial	<i>Oscillatoria nigro-viridis</i> PCC 7112	10	
Terrestrial	<i>Leptolyngbya</i> sp. PCC 7376	8	
Terrestrial	<i>Oscillatoria acuminata</i> PCC 6304	6	
Terrestrial	<i>Cylindrospermum stagnale</i> PCC 7417	3	
Terrestrial	<i>Microcoleus</i> sp. PCC 7113	2	
Terrestrial	<i>Crinalium epipsammum</i> PCC 9333	2	
Terrestrial	<i>Nostoc punctiforme</i> NIES-2108	1	
Terrestrial	<i>Nostoc</i> sp. HK-01	1	
Terrestrial	<i>Nostoc punctiforme</i> PCC 73102	0	
Terrestrial	<i>Chroococcidiopsis thermalis</i> PCC 7203	0	
Hot spring	<i>Pleurocapsa</i> sp. PCC 7327	10	
Hot spring	<i>Thermoleptolyngbya</i> sp. O-77	7	
Hot spring	<i>Thermosynechococcus elongatus</i> BP-1	0	
Hot spring	<i>Thermosynechococcus vulcanus</i> NIES-2134	0	
Hot spring	<i>Synechococcus</i> sp. JA-3-3Ab	0	
Fresh water	<i>Aphanothece sacrum</i>	11	Sacran (<i>Okajima et al.</i> , 2008)
Fresh water	<i>Microcystis aeruginosa</i> NIES-843	9	
Fresh water	<i>Synechocystis</i> sp. PCC 6803	2	Salt tolerant (<i>Kopf et al.</i> , 2014a)
Fresh water	<i>Synechocystis</i> sp. PCC 6714	0	Salt sensitive (<i>Kopf et al.</i> , 2014a)
Fresh water	<i>Anabaena cylindrica</i> PCC 7122	4	
Fresh water	<i>Calothrix</i> sp. PCC 7507	3	
Fresh water	<i>Synechococcus elongatus</i> PCC 7942	0	Salt sensitive
Fresh water	<i>Nostoc (Anabaena)</i> sp. PCC 7120	0	
Fresh water	<i>Leptolyngbya boryana</i> IAM M-101	0	
Fresh water	<i>Anabaena variabilis</i> ATCC 29413	0	
Fresh water	<i>Calothrix</i> sp. PCC 6303	0	
Fresh water	<i>Calothrix</i> sp. 336/3	0	

377 (Fisher *et al.*, 2013, Pereira *et al.*, 2015). Many of them have been disrupted for characterization. With
378 regard to the Wzx/Wzy-dependent system, *sll5052* (*xssK*) disruption mutant showed no clear phenotypes
379 for bloom formation or EPS production (Jittawuttipoka *et al.*, 2013), probably because the parent strain
380 did not produce discernable amount of EPS like our nonmotile strain. Similarly, the deletion mutant of
381 *sll5049* (*xssH*) did not show any defect in EPS or CPS accumulation, though related mutants (Δ *sll0923*
382 for a second PCP-2a) were shown to be depleted slightly of both CPS and EPS (Pereira *et al.*, 2019).
383 These results contrast with our null phenotype of Δ *sll5052* (*xssK*) and Δ *sll5049* (*xssH*), probably because
384 of the difference in the parent strains. In addition, we found that our Δ *sll0923* did not show any defect
385 in the bloom formation. On the other hand, disruption of *sigF* (*slr1564*) for a sigma factor of global cell
386 surface regulation increased three to four fold accumulation of sulfated EPS (Flores *et al.*, 2019a, Flores
387 *et al.*, 2019b). The proteome analysis of Δ *sigF* revealed many (more than 160) proteins except for any
388 Xss proteins were up-regulated, leaving the sulfated EPS biosynthesis pathway elusive. The sugar
389 composition of the sulfated EPS of their WT is similar to our WT, although the composition of EPS of

$\Delta sigF$ was different for WT or synechan from our $\Delta xssS$.

To get insights into the difference in bloom formation between the motile and nonmotile substrains, we compared transcription data (Table S4). It is evident that many *xss* genes on the plasmid are expressed several times higher in the motile substrain than the nonmotile one except for *xssT* on the main chromosome, despite that the nucleotide sequence of the *xss* gene cluster was completely conserved between them. This fact suggests a possibility that another mechanism besides the XssS/XssR/XssQ contributes to the difference between the substrains. For example, the plasmid copy number of pSYSM may be higher in the motile substrain than the nonmotile substrain. The plasmid function is often affected depending on variations in the main chromosome (Vial and Hommais, 2020). Anyway, it is very important to select the parent strain depending on the research purpose.

The cyanobacterial bloom rapidly accumulates in populations of cyanobacterial cells floating on the water surface, which often produce potent cyanotoxins (hepatotoxins, neurotoxins, etc.) (Merel et al., 2013). Blooms are thought to be supported mainly by cellular buoyancy due to intracellular

Table S4. Transcriptional levels of *xss* genes in the motile (WT) and nonmotile (NM) substrains.

locas tag	gene name	WT (RPKM)	NM (RPKM)	WT / NM
<i>sll5042</i>	<i>xssA</i>	46.7	28.5	1.6
<i>sll5043</i>	<i>xssB</i>	43.2	9.8	4.4
<i>sll5044</i>	<i>xssC</i>	53.3	4.5	12.0
<i>sll5045</i>	<i>xssD</i>	77.6	6.2	12.4
<i>sll5046</i>	<i>xssE</i>	291.8	24.0	12.1
<i>sll5047</i>	<i>xssF</i>	17.9	8.7	2.1
<i>sll5048</i>	<i>xssG</i>	71.7	28.9	2.5
<i>sll5049</i>	<i>xssH</i>	41.8	21.3	2.0
<i>sll5050</i>	<i>xssI</i>	26.8	33.6	0.8
<i>slr5051</i>	<i>xssJ</i>	231.6	112.0	2.1
<i>sll5052</i>	<i>xssK</i>	29.0	23.7	1.2
<i>slr5053</i>	<i>xssL</i>	43.9	6.5	6.8
<i>slr5054</i>	<i>xssM</i>	33.2	22.5	1.5
<i>slr5055</i>	<i>xssN</i>	455.5	14.4	31.6
<i>slr5056</i>	<i>xssO</i>	120.5	23.3	5.2
<i>sll5057</i>	<i>xssP</i>	514.1	12.6	40.9
<i>slr5058</i>	<i>xssQ</i>	54.3	27.5	2.0
<i>sll5059</i>	<i>xssR</i>	40.5	45.9	0.9
<i>sll5060</i>	<i>xssS</i>	12.8	11.7	1.1
<i>sll1581</i>	<i>xssT</i>	133.3	181.8	0.7

Transcriptional level is given as RPKM (Reads Per Kilobase of exon per Million mapped reads).

proteinaceous gas vesicles constructed by gas vesicle proteins (*Beard et al., 2002, Walsby, 1994*).

Moreover, recent studies suggested that extracellular polysaccharides are also important for the bloom

formation (*Chen et al., 2019*). Some papers reported that the cells without gas vesicle can form blooms

by EPS-dependent manner after artificial addition of divalent cations (Ca^{2+} or Mg^{2+}) (*Dervaux et al.,*

2015, Wei et al., 2019). On the other hand, our study demonstrated that the gas-trapped EPS is sufficient

for bloom formation of *Synechocystis* 6803, which does not produce gas vesicles, without addition of

any divalent cations. This result is consistent with reports that cyanobacteria without gas vesicles form

booms in natural environments, including freshwater lakes (*Casero et al., 2019, du Plooy et al., 2015,*

Steffen et al., 2012).

Finally, sulfated polysaccharides are expected to be healthy foods, industrial materials and

medicines (*Jiao et al., 2011, Wardrop and Keeling, 2008*). Some sulfated EPS from *Synechocystis* 6803

showed antitumor activity (*Flores et al., 2019a*), though this EPS may not be identical to synechan. The

Xss-dependent biosynthesis of synechan in *Synechocystis* 6803 should be a good model for studies of

420 other cyanobacterial sulfated polysaccharides. Combinatorial expression of sulfotransferases and

421 glycosyltransferases from other cyanobacteria in *Synechocystis* cells will provide clues to their functions.

422 Heterologous expression of *Synechocystis xss* genes in other organisms will also open a possibility of

423 large scale production of modified synechan species. Further molecular studies of *xss* genes and related

424 genes from the database should accelerate screening and potential applications of cyanobacterial

425 sulfated polysaccharides.

426 **Materials and Methods**

427 *Cyanobacterial strains and cultures*

428 The motile substrain PCC-P of the unicellular cyanobacterium *Synechocystis* sp. PCC 6803,
429 which exhibits phototaxis (*Yoshihara et al., 2000*) and forms bloom-like aggregates, was used as the
430 WT in this work. A non-motile glucose-tolerant substrain, which has been widely used for studies of
431 photosynthesis, was used for comparison (*Chin et al., 2018*). Cells were maintained in BG11 liquid
432 medium (*Stanier et al., 1971*) under continuous illumination with bubbling of 1% CO₂ in air at 31°C, or
433 on 1.5% agar plates. White light of 30 μmol photons m⁻²s⁻¹ was generated by fluorescent lamps. Cell
434 density was monitored at 730 nm.

435

436 *Construction of plasmids and mutants*

437 Primers used are listed in Table S5. Plasmids and mutants were constructed as described
438 (*Chin et al., 2018*). In brief, the DNA fragments, antibiotic-resistance cassettes, the *trc* promoter, and

plasmid vectors were amplified by PCR using PrimeSTAR MAX DNA polymerase (Takara, Shiga, Japan) and combined using the In-Fusion System (Takara). The resulting plasmid constructs were confirmed by DNA sequencing.

Gene disruption was performed in two different ways. One method was replacement of a large portion of a targeted gene(s) with an antibiotic resistance cassette. The other method was replacement of the translation initiation codon with a stop codon. In both cases, the screening cassette without the terminator was inserted in the direction of the targeted gene(s) to allow transcriptional readthrough of the downstream gene(s). For overexpression, gene expression was constitutively driven by the strong *trc* promoter in two ways: integration of a target gene with the strong *trc* promoter into a neutral site near *slr0846* or IS203c, or replacement of the target-gene promoter with the *trc* promoter.

Natural transformation and subsequent homologous recombination were performed as described (Chin *et al.*, 2018). The antibiotic concentration for the selection of transformants was 20 $\mu\text{g}\cdot\text{mL}^{-1}$ chloramphenicol, 20 $\mu\text{g}\cdot\text{mL}^{-1}$ kanamycin, and/or 20 $\mu\text{g}\cdot\text{mL}^{-1}$ spectinomycin. Complete segregation of

452 **Table S5. Primer sequences.**

Product (purpose)	Primer name	Sequence (5' => 3')
Antibiotic resistance cassette	Sp-1F	ATGCGCGAAGCCGTGATTGCC
	Sp-2R	TAGTAAAGCCCTCGCTAG
	Km-1Fup	AATTCCTGCTCGCGCTGCCGCAAGCACTCAGGG
	Km-2Rcom	AATATTGCGAAGCGGCCAACCTTTTCATAGAAGGCG
	Cm-1Fup	AATTCCTGCTCGCGCAGCGCTGATGTCCGGCGGTGCT
	Cm-2Rcom	AATATTGCGAAGCGGCTAACCGTTTTATCAGGC
Glycosyltransferase disruption	slr1943-1FpS	AACGACGGCCAGTGAAGACCGCATGTGTGGCAGT
	slr1943-2Rup	GCGAGCAGGGGAATTGGCTTGGTGAATTTGTAACC
	slr1943-3Fcom	CCGCTTCGCAATATTGGCTGGCTTTGAACAACTTC
	slr1943-4RpS	AAACAGCTATGACCAATCGGCAATGTAGGTGAAGG
	slr1943-5F	AAGCCCGTCTGGAGTTTG
	slr1943-6R	AATCCGGGTCGATGGTAC
	slr1063-1FpS	AACGACGGCCAGTGAATGTCCAAAAATGACAAGC
	slr1063-2Rup	GCGAGCAGGGGAATTACACGATTGGCGGAAATTTTC
	slr1063-3Fcom	CCGCTTCGCAATATTCTGGTATATCAGATTACTCG
	slr1063-4RpS	AAACAGCTATGACCAGCTTGGAGTAAGCTGCATAA
	slr1063-5F	ATGCGTCGTTTCGGTCAC
	slr1063-6R	TGCAAGACATTCCTGGAC
	sll0501-1FpS	AACGACGGCCAGTGAGGGATTAGGCAGGGGTAATA
	sll0501-2Rup	GCGAGCAGGGGAATTCCTGTCAATTTAGTCAGTGC
	sll0501-3Fcom	CCGCTTCGCAATATTGACCATTTGAAGTGTCTATTGTG
	sll0501-4RpS	AAACAGCTATGACCAGGTGAGAATCGATACCATGG
	sll0501-5F	TCAAAGGGAACGTGCTGG
	sll0501-6R	GTTGCCTTGGGCGTAATC
	slr1118-1FpS	AACGACGGCCAGTGATCCTGCCGCTCAAAGTATTC
	slr1118-2Rup	GCGAGCAGGGGAATTCGATCGCCGCTTTAGCTAAA
	slr1118-3Fcom	CCGCTTCGCAATATTTCAAGCTCCCGTTGGTTAA
	slr1118-4RpS	AAACAGCTATGACCAGCAGGGGAAATCGTTTGCAT
	slr1118-5F	CGCACCATTTATCAGCAG
	slr1118-6R	ATCGCTTCGGTTCCATGC
	sll5052_54-1FpS	AACGACGGCCAGTGATCGCTGTTTGTGGTGAGG
	sll5052_54-2Rup	GCGAGCAGGGGAATTCATGCTGGCTTGGCCAAATTA
	sll5052_54-3Fcom	CCGCTTCGCAATATTGCTGCTGGATTACAGGAATTG
	sll5052_54-4RpS	AAACAGCTATGACCACAGGCGGGAGATTGCTAAAA
	sll5052_54-5F	GGTACTTCGCTGTAGTTC
	sll5052_54-6R	GGGCAAACTTCGTCTTC
xss gene disruption and overexpression	xssA-1FpS	AACGACGGCCAGTGACCCAGCTTGTAGTGGATTG
	xssA-2Rup	GCGAGCAGGGGAATTTTCATGCCGCCGATAATGAT
	xssA-3Fcom	CCGCTTCGCAATATTAGAAATGTGCCACCACTACC
	xssA-4RpS	AAACAGCTATGACCAATCTAGGTTGACCCCTACC
	xssA-5Ftrc	GAGGAATAAACCATGAAATGTGCCACCACTACC
	xssA-6F	GATCGCCAATGGCAAGAT
	xssA-7R	ATTGGGTTCTTCAGCGA
	xssB-1FpS	AACGACGGCCAGTGAGAGCGTGTGTTGAAAAGCCCC
	xssB-2Rup	GCGAGCAGGGGAATTTGGCCATAATCTTGCAAATACTC
	xssB-3Fcom	CCGCTTCGCAATATTAGGCCATAATTGCTTCAAACAATAG
	xssB-4RpS	AAACAGCTATGACCAGAAATCTACTCCGGACGGT
	xssB-5Ftrc	GAGGAATAAACCATGGCCATAATTGCTTCAAACAA
	xssB-6F	GCCAAACAAATCCCCAGA
	xssB-7R	AATCCACTACAAGCTGGG
	xssC-1FpS	AACGACGGCCAGTGAAGGCTACTGGGTTTGAGGTT
	xssC-2Rup	GCGAGCAGGGGAATTATCATTTCCCGACGGCGATC
	xssC-3Fcom	CCGCTTCGCAATATTAGATCCAAGAGTTTCAGCGGTT
	xssC-4RpS	AAACAGCTATGACCACCGCACAATGGTATCGCTAT
	xssC-5F	CATTACTTTTTCCCCGCC
	xssC-6R	CCGCTGTAAAAAATGCAGC

453

454

455 **Table S5.** Continuing.

Product (purpose)	Primer name	Sequence (5' => 3')
xss gene disruption and overexpression	xssE-1FpS	AACGACGGCCAGTGACCGTCGAGCGAGTTAAGG
	xssE-2Rup	GCGAGCAGGGGAATTAAGTGTTTCATGGGGCGTGA
	xssE-3Fcom	CCGCTTCGCAATATTTAGAACAACCTCTTTCTCCTCGAAT
	xssE-4RpS	AAACAGCTATGACCAGCTCTTAGGGCGGGGAAAA
	xssE-5Ftrc	AAGGAGGAATAAACCATGAACACTTCTTTCTCCTCGAAT
	xssE-6F	TAGTGAGCAGGGGAAACT
	xssE-7R	CACTTACCTGATCGATGAC
	xssF-1FpS	AACGACGGCCAGTGACTGCAAAAACGAGGCCATGA
	xssF-2Rup	GCGAGCAGGGGAATTTGTCATGGATCGTTTGCTGTTG
	xssF-3Fcom	CCGCTTCGCAATATTAGACCCCCAAGCCCAACTG
	xssF-4RpS	AAACAGCTATGACCAATAGACCCGATTGCGTTTCC
	xssF-5F	CCCATGATGTTGTGGTGT
	xssF-6R	TTGCCTGACCCTGACTTT
	xssG-1FpS	AACGACGGCCAGTGACTCCACGGCATTATTTACCC
	xssG-2Rup	GCGAGCAGGGGAATTCATACTCCTAACTCCATAACGC
	xssG-3Fcom	CCGCTTCGCAATATTAGAAAATCACCTTTGTTTTACCTA
	xssG-4RpS	AAACAGCTATGACCATCATGGATCGTTTGCTGTTGC
	xssG-5F	ATCTAGCAGTGGCTACG
	xssG-6R	CCGAAGCCATTGGCTAAT
	xssH-1FpS	AACGACGGCCAGTGAGCTGAAAACCCTCGGCTTAA
	xssH-2Rup	GCGAGCAGGGGAATTCATAGTTACGGCGCCGATG
	xssH-3Fcom	CCGCTTCGCAATATTAGGCTTCCCTCAATAAACTGG
	xssH-4RpS	AAACAGCTATGACCATTGGCAAAGTTACCGATCGC
	xssH-5F	GTAGAAGCAATGCAGTCTG
	xssH-6R	GGTATTTACCAACGCCATC
	xssI-1FpS	AACGACGGCCAGTGACCAATTTGCAACATCCCGAC
	xssI-2Rup	GCGAGCAGGGGAATTTAGCCCCGAATCAATTATCC
	xssI-3Fcom	CCGCTTCGCAATATTAGGTGAGATCGCCCCAATTCC
	xssI-4RpS	AAACAGCTATGACCAAAATGTTGCTGCACCATGGC
	xssI-5F	AATTCGGTCTAGCTTGCC
	xssI-6R	GCTCTTCCAAAGCCAGAA
	xssK-1FpS	AACGACGGCCAGTGATCTGCCACTTGGTAAACCTC
	xssK-2Rup	GCGAGCAGGGGAATTTGGAATATTCGGGCGCTTTGA
	xssK-3Fcom	CCGCTTCGCAATATTTGGTGCGGGTTTATCTCAAC
	xssK-4RpS	AAACAGCTATGACCAGCAGATTGTTCCGTTTTTGAA
	xssK-5F	CGCTCCCATAAAGATCTC
	xssK-6R	TTCTGGAATCTGGGGGAGAA
	xssM-1FpS	AACGACGGCCAGTGATCGCTGTTTGTGTTGGTGAGG
	xssM-2Rup	GCGAGCAGGGGAATTTCAATGCTGGCTTGCCAAAT
	xssM-3Fcom	CCGCTTCGCAATATTCTTCCATGGTGGTTAGTGGA
	xssM-4RpS	AAACAGCTATGACCACAGGCGGGAGATTGCTAAAA
	xssM-5F	CTGATAGGCTGTTGCTAACC
	xssM-6R	GCAGTTGTTTATGGACCCTC
	xssN-1FpS	AACGACGGCCAGTGAGGAATTGTCGTTATTGGGCG
	xssN-2Rup	GCGAGCAGGGGAATTTGGAATTCATTGCTTAGCCC
	xssN-3Fcom	CCGCTTCGCAATATTAGAATTCATACGCTTCTCAA
	xssN-4RpS	AAACAGCTATGACCATATAGGACAGGGCGATCTTC
	xssN-5Ftrc	GAGGAATAAACCATGAATTCCATCACGCTTCTCAA
	xssN-6F	TCCGCCTCACCAACAAA
	xssN-7R	GCCACTTGGTAAACCTCT
	xssO-1FpS	AACGACGGCCAGTGAGTTAGCCTGGCTCTATGCCA
	xssO-2Rup	GCGAGCAGGGGAATTAACACCTAACGGTCCGAGA
	xssO-3Fcom	CCGCTTCGCAATATTAGTTTCGGGGTTGCTAGCCAA
	xssO-4RpS	AAACAGCTATGACCAAAATAGCGGCTTTCAGCAACC
	xssO-5F	ATGAGCAAAGTCGGGTTG
	xssO-6R	GGAGAATGAAACGAGACC
	xssP-1FpS	AACGACGGCCAGTGACCTCTGCTTCAGATAAAGCTG
	xssP-2Rup	GCGAGCAGGGGAATTTGCCATAGGTAATCACCAAGGAG
	xssP-3Fcom	CCGCTTCGCAATATTAGGCTAGCACTATTCAACTGATTG
	xssP-4RpS	AAACAGCTATGACCATTCTCACCCACTCGGGTGTA
	xssP-5Ftrc	GAGGAATAAACCATGGCTAGCACTATTCAACTGA
	xssP-6F	CTGGTGAAGCTTTAAGGG
	xssP-7R	AAAATCGAACACCAGCCG

457 **Table S5.** Continuing.

Product (purpose)	Primer name	Sequence (5' => 3')
<i>xss</i> gene disruption and overexpression	xssQ-1Rup	GCGAGCAGGGGAATTGCCATGTTTACAATGATAGAGA
	xssQ-2Fcom	CCGCTTCGCAATATTAGGCAAAACGTTCCCTTAAAGCT
	xssQ-3Ftrc	GAGGAATAAACCATGGCAAAACGTTCCCTTAAAG
	xssQ-4Rrrn	CAGACCGCTTCTGCGTTAGCTAACCTTTTCGCC
	xssR-1FpS	AACGACGGCCAGTGAGGAGGAAGTAGATGGGTTAC
	xssR-2Rup	GCGAGCAGGGGAATTGTCGGAGAAGTCTCAGTCAT
	xssR-3Fcom	CCGCTTCGCAATATTGAGCAATTGCGACAGTTTCC
	xssR-4RpS	AAACAGCTATGACCAACGTTAACAGCGCCGTTGTT
	xssR-5F	CAGACGGAGCAACCTTTT
	xssR-6R	GGAAACTGTCGCAATTGC
	xssR-7Ftrc	GAGGAATAAACCATGACTGAGACTTCTCCG
	xssR-8Rrrn	CAGACCGCTTCTGCGCTACCCCATGACCAGGGC
	xssS-1Rup	GCGAGCAGGGGAATTACCTGCAAAAAGGTTGCTCC
	xssS-2Fcom	CCGCTTCGCAATATTGCGGGGTTGATGGACAATAT
<i>sll1581</i> disruption	sll1581-1FpS	AACGACGGCCAGTGATGGCTTTTCAGGCAATTCCG
	sll1581-2Rup	GCGAGCAGGGGAATTTTCCAAGTTCAAGGCCACCA
	sll1581-3Fcom	CCGCTTCGCAATATTGAGAGTGAACCTATAACCCA
	sll1581-4RpS	AAACAGCTATGACCATCTGCCTGTAATCTCTGCAC
	sll1581-5F	ATCTGGTGGCCTTGAAC
	sll1581-6R	TTCAGTTGCTTGCAGGCA
RT-qPCR	rnxB-1F	GAGAGTTAGGGAGGGAGTTGC
	rnxB-2R	AAGGGCGGTATTTTCTGTG
	xssA-8F	TTGCCACCTACCCTGACAC
	xssA-9R	AGGCGATCGATGGGATGACG
	xssB-8F	CGGCCAGGGGCAATTCAAAC
	xssB-9R	ATTTGGGGAGTGCCATCGGG
	xssE-8F	CAACGCGGCTATGAAGCCAG
	xssE-9R	GCCCCGTAAACGCTGTAAACG
	xssF-7F	GCTGGAAACGCAATCGGGTC
	xssF-8R	GTTTCAGCAGGCCAAAGGCTC
	xssH-7F	CCGCAATCGCTTTACCTGGG
	xssH-8R	GGCGAGAAAGGTCATGGCTG
	xssK-7F	ATTTCCCACCATTTGCGGCTC
	xssK-8R	CAAGGGACATCCTGTTGGGC
	xssM-7F	GCCATGTTCTGTGCTGGGTTT
	xssM-8R	GCCCAGAACCATACTGCGG
	xssN-8F	GCAAAGTCGGGTTGGAGTGG
	xssN-9R	CCAAGGCAAAGAACGGTAGGC
	xssP-8F	CGGCTTAACCGGAGAATGGC
	xssP-9R	CCAAGCTCCAGCGTTTCTG
EMSA	xssEup-1F	AGGAAAGAAGTGTTTCATGGG
	xssEup-2R	ATCTTTAAAGCCCCAGTCCC
	xssQ-5F28a	TATTTTCAGAGCCATATGGCAAAACGTTCCCTTA
	xssQ-6R28a	CTCGAATTCGGATCCTTAGCTAACCTTTTCGCC
	pET28a-1R	CATATGGCTCTGAAAATACAG
	pET28a-2F	GGATCCGAATTCGAGCTC

459 **Table S6.** List of mutants.

Strain name	Design
<i>Δslr1943</i>	Complete deletion by Sp ^R
<i>Δslr1063</i>	Complete deletion by Sp ^R
<i>Δslr0501</i>	Complete deletion by Sp ^R
<i>Δslr1118</i>	Complete deletion by Sp ^R
<i>Δslr5054</i>	Partial deletion by Sp ^R (= Δ xssK~M)
<i>ΔxssA</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssB</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssC</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssE</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssF</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssG</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssH</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssI</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssJ</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssK</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssM</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssN</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssO</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssP</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssQ</i>	Replacement of the the start codon by Sp ^R and stop codon
<i>ΔxssR</i> [*]	Partial deletion by Sp ^R
<i>ΔxssS</i> [*]	Partial deletion by Sp ^R
<i>ΔxssR-S</i> [*]	Partial deletion by Sp ^R
<i>ΔxssK~M</i> + <i>ΔxssS</i> [*]	<i>xssK~M::Sp^R + xssS::Cm^R</i>
<i>OX-xssR</i> [#]	<i>slr0846_platform-Km^R-P_{trc}-XssS-T_{rmB}</i>
<i>OX-xssQ</i> [#]	<i>IS203_platform-Cm^R-P_{trc}-XssQ-T_{rmB}</i>
<i>ΔxssQ</i> + <i>OX-xssR</i>	Simple combination
<i>ΔxssR</i> [*] + <i>OX-xssQ</i> [#]	Simple combination
<i>OX-xssP</i> [#]	Replacement of the the native promoter to Cm ^R -P _{trc}
<i>ΔxssT</i>	Complete deletion by Km ^R

461 *Almost complete segregation. #Partial segregation.

462

the transformed DNA in the multicopy genome was confirmed by PCR using primers listed in Table S5,

and the transformants are listed in Table S6.

Bloom formation

The bloom was reproducibly formed using the two-step culture regime we developed in this work. Before the bloom formation experiment, cells were precultured once in liquid after transfer from plates. In the first step, cells inoculated at $OD_{730} = 0.2$ were grown with vigorous aeration under continuous light at 31°C or 20°C for 48 h. Typically the cell density reached $OD_{730} \sim 2$. In the second step, the culture was shifted to the standing condition without bubbling under the same continuous light for another 48 h (or longer) for cells to rise to the surface. Regarding the mutants of transmembrane glycosyltransferases, bloom formation was examined after 168 h of the second-step culture. The final concentration of the photosynthesis inhibitor DCMU (3-(3,4-dichlorophenyl)-1,1-dimethylurea) was 100 μ M.

476

477 *EPS fractionation*

478 The fractionation method to isolate the crude EPS is shown in Fig. S1A. The viscous
479 materials including cells after the second step of culture were collected by filtration using a 1.0- μ m pore
480 PTFE membrane (Millipore). The trapped materials were gently and carefully recovered from the
481 membrane using MilliQ water with the aid of flat-tip tweezers. The collected sample was vortexed and
482 then centrifuged at $20,000 \times g$ for 10 min to remove cells. The supernatant constituted the crude EPS
483 that contained viscous EPS and possibly CPS.

484 The refined fractionation method to isolate EPS is shown in Fig. 1F. The entire culture at the
485 end of the first step, which did not contain gas bubbles, was first centrifuged at $10,000 \times g$ for 10 min
486 to remove cells and CPS and then filtered through a 1.0- μ m pore PTFE membrane. The trapped EPS
487 was carefully recovered as described above. The flowthrough of the filtration was regarded as free
488 polysaccharides, which were recovered by ethanol precipitation. CPS was released from the cell pellet

by vigorous vortexing with MilliQ water and recovered by centrifugation to remove cells ($20,000 \times g$ for 10 min).

Sugar quantification

Total sugar was quantified using the phenol-sulfate method (*DuBois et al., 1956*). A 100- μ L aliquot of 5% (w/w) phenol was added to 100 μ L of a sample in a glass tube and vortexed three times for 10 s. Then, 500 μ L of concentrated sulfuric acid was added, and the tube was immediately vortexed three times for 10 s and then kept at 30°C for 30 min in a water bath. Sugar content was measured by absorption at 487 nm using a UV-2600PC spectrophotometer (Shimadzu, Japan, Tokyo). Any contamination of the BG11 medium was evident by slight background coloration. This background was subtracted on the basis of the extrapolation of absorption at 430 nm, where the coloration due to sugars was minimal. Glucose was used as the standard. Some EPS samples were highly viscous, so we vortexed and sonicated them before measurement. Statistical significance was determined using Welch's *t* test.

502

503 *Sugar composition analysis*

504 The collected EPS samples were dialyzed with MilliQ water and then freeze-dried for 3 days.

505 Sugar composition was analysed by Toray Research Center, Inc. (Tokyo, Japan). A part of the fluffy

506 sample (WT, 0.298 mg; $\Delta xssS$, 0.203 mg) was dissolved in 200 μ L of 2 M trifluoroacetic acid and

507 hydrolysed at 100°C for 6 h. The treated sample was vacuum-dried with a centrifugal evaporator,

508 redissolved in 400 μ L MilliQ water, and filtered through a 0.22- μ m pore filter. This sample was used for

509 the analysis.

510 Monosaccharide composition was determined by HPLC with the LC-20A system (Shimadzu).

511 For neutral sugars, the column was TSK-gel Sugar AXG (TOSOH, Japan) and the temperature was 70°C.

512 The mobile phase was 0.5 M potassium borate (pH 8.7) at 0.4 mL/min. Post-column labelling was

513 performed using 1% (w/v) arginine and 3% (w/v) boric acid at 0.5 mL/min, 150°C. For uronic acids, the

514 column was a Shimpack ISA-07 (Shimadzu) and the temperature was 70°C. The mobile phase was 1.0

M potassium borate (pH 8.7) at 0.8 mL/min. Post-column labelling was performed using 1% (w/v) arginine and 3% (w/v) boric acid at 0.8 mL/min, 150°C. The detector was a RF-10A_{XL} (Shimadzu), with excitation at 320 nm and emission at 430 nm. The standard curves were prepared for each monosaccharide with standard samples.

The SO₄²⁻ content was determined by anion exchange column chromatography using the ISC-2100 system (Thermo Fisher Scientific, USA, Massachusetts). The column was eluted via a gradient of 0–1.0 M KOH. The separation column was IonPac ASI I-HC-4 µm (Thermo Fisher Scientific). Electric conductivity was used for detection.

Alcian blue staining

The polysaccharides were stained with 1% Alcian blue 8GX (Merck) for 10 min in 3 % acetic acid (pH 2.5) or 0.5 N HCl (pH 0.5) as previously described (*Di Pippo et al., 2013*).

528 *qPCR*

529 The qPCR was performed as described in our previous work (*Maeda et al., 2018*). Cells were
530 harvested by centrifugation at $5000 \times g$ for 10 min at 4°C. Cell disruption and RNA extraction were
531 done using an RNeasy Mini kit for bacteria (Qiagen, Venlo, Netherlands). In addition, cells were
532 disrupted five times by mechanical homogenization with zirconia beads (0.1-mm diameter) in a
533 microhomogenizing system (Micro Smash MS-100, TOMY SEIKO, Tokyo, Japan) at 5,000 rpm for 40
534 s. For cDNA preparation, RNA was reverse-transcribed using random primers (PrimeScript RT reagent
535 kit with gDNA eraser, Takara). Real-time PCR was performed using the THUNDERBIRD SYBR qPCR
536 Mix (Toyobo) and the Thermal Cycler Dice Real Time System II (Takara). The transcript level in each
537 strain was normalized to the internal control (*rnpB*). The primers used are listed in Table S5.

538

539 *EMSA* (electrophoretic mobility shift assay)

540 The expression and purification of recombinant His-tagged proteins and EMSA were

performed as described in our previous works (*Hirose et al., 2010, Maeda et al., 2014*). In brief, His-tagged XssQ was expressed using pET28a vector system and *E. coli* C41(DE3) strain. The protein was purified by Histrap HP column (Cytiva, Tokyo, Japan) and AKTA prime system (Cytiva). For probe and native competitor, the upstream region of *xssE* was amplified with the primer set xssEup-1F/2R (total 251 bp). As a mutant competitor, the same region of the chemically synthesized DNA fragment containing mutations in the two consensus sequences was used for amplification with the same primer set as mention above. Labelling of the DNA probe, electrophoresis, and autoradiography were performed as described (*Midorikawa et al., 2009*). We incubated the aliquots of the XssQ protein (0, 500, 1000, or 3000 ng/lane) with the radiolabeled probe for 30 min at room temperature. For competition, 3000 ng of XssQ was incubated with the probe and 20 pmol of unlabeled competitors (native or mutant).

RNA-seq analysis

RNA-seq analysis was performed as described in our previous work (*Ohbayashi et al., 2016*).

Total RNA of *Synechocystis* 6803 was extracted as described in the qPCR protocol. The contaminated genome DNA was removed by TURBO DNA-free™ Kit (Thermo Fisher Scientific).

Bioinformatics analysis

The sequences of the proteins were obtained from NCBI (<http://www.ncbi.nlm.nih.gov/>) and Pfam (<http://pfam.xfam.org/>) (Finn *et al.*, 2016). The domain architecture was searched using the Simple Modular Architecture Research Tool, SMART (Letunic *et al.*, 2015). Glycosyltransferase classifications were based on the CAZy database (<http://www.cazy.org/>) (Henrissat, 1991, Lombard *et al.*, 2013). Amino acid sequence similarity was evaluated by NCBI BLAST search.

Acknowledgements

This study was supported by Grants-in-Aid for JSPS Fellows 15J07605 and 19J01251 (to K.M.), the Japan Society for the Promotion of Science for Scientific Research (16H06558 to M.I.), and

567 the JST for CREST program (JPMJCR1653 to M.I.).

568

References

- Beard SJ, Hayes PK, Pfeifer F, Walsby AE. 2002. The sequence of the major gas vesicle protein, GvpA, influences the width and strength of halobacterial gas vesicles. *FEMS microbiology letters* **213**:149-157
- Bellezza S, Albertano P, de Philippis R, Paradossi G. 2006. Exopolysaccharides of two cyanobacterial strains from Roman hypogea. *Geomicrobiology Journal* **23**:301-310
- Casero MC, Velázquez D, Medina-Cobo M, Quesada A, Cirés S. 2019. Unmasking the identity of toxigenic cyanobacteria driving a multi-toxin bloom by high-throughput sequencing of cyanotoxins genes and 16S rRNA metabarcoding. *Science of The Total Environment* **665**:367-378
- Chen M, Tian L-L, Ren C-Y, Xu C-Y, Wang Y-Y, Li L. 2019. Extracellular polysaccharide synthesis in a bloom-forming strain of *Microcystis aeruginosa*: implications for colonization and buoyancy. *Scientific reports* **9**:1-11
- Chin T, Okuda Y, Ikeuchi M. 2018. Sorbitol production and optimization of photosynthetic supply in the cyanobacterium *Synechocystis* PCC 6803. *Journal of biotechnology* **276**:25-33
- Danot O, Marquenet E, Vidal-Ingigliardi D, Richet E. 2009. Wheel of Life, Wheel of Death: A Mechanistic Insight into Signaling by STAND Proteins. *Structure* **17**:172-182. DOI: 10.1016/j.str.2009.01.001
- De Philippis R, Vincenzini M. 1998. Exocellular polysaccharides from cyanobacteria and their possible applications. *FEMS Microbiology Reviews* **22**:151-175
- Dervaux J, Mejean A, Brunet P. 2015. Irreversible collective migration of cyanobacteria in eutrophic conditions. *PLoS One* **10**:e0120906
- Di Pippo F, Ellwood NT, Gismondi A, Bruno L, Rossi F, Magni P, De Philippis R. 2013. Characterization of exopolysaccharides produced by seven biofilm-forming cyanobacterial strains for biotechnological applications. *Journal of Applied Phycology* **25**:1697-1708

- du Plooy S, Perissinotto R, Smit A, Muir D. 2015. Role of salinity, nitrogen fixation and nutrient assimilation in prolonged bloom persistence of *Cyanothece* sp. in Lake St Lucia, South Africa. *Aquatic Microbial Ecology* **74**:73-83
- DuBois M, Gilles KA, Hamilton JK, Rebers Pt, Smith F. 1956. Colorimetric method for determination of sugars and related substances. *Analytical chemistry* **28**:350-356
- Enomoto G, Narikawa R, Ikeuchi M. 2015. Three cyanobacteriochromes work together to form a light color-sensitive input system for c-di-GMP signaling of cell aggregation. *Proceedings of the National Academy of Sciences* **112**:8082-8087
- Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic acids research* **44**:D279-D285
- Fisher ML, Allen R, Luo Y, Curtiss R, 3rd. 2013. Export of extracellular polysaccharides modulates adherence of the Cyanobacterium *synechocystis*. *PLoS One* **8**:e74514. DOI: 10.1371/journal.pone.0074514, PMID: PMC3769361
- Flombaum P, Gallegos JL, Gordillo RA, Rincón J, Zabala LL, Jiao N, Karl DM, Li WK, Lomas MW, Veneziano D. 2013. Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proceedings of the National Academy of Sciences* **110**:9824-9829
- Flores C, Lima RT, Adessi A, Sousa A, Pereira SB, Granja PL, De Philippis R, Soares P, Tamagnini P. 2019a. Characterization and antitumor activity of the extracellular carbohydrate polymer from the cyanobacterium *Synechocystis* DeltasigF mutant. *Int J Biol Macromol* **136**:1219-1227. DOI: 10.1016/j.ijbiomac.2019.06.152
- Flores C, Santos M, Pereira SB, Mota R, Rossi F, De Philippis R, Couto N, Karunakaran E, Wright PC, Oliveira P, Tamagnini P. 2019b. The alternative sigma factor SigF is a key player in the control of secretion mechanisms in *Synechocystis* sp. PCC 6803. *Environ Microbiol* **21**:343-359. DOI: 10.1111/1462-2920.14465

- 620 Foster JS, Havemann SA, Singh AK, Sherman LA. 2009. Role of mrgA in peroxide and light stress
621 in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEMS Microbiol Lett* **293**:298-304.
622 DOI: 10.1111/j.1574-6968.2009.01548.x
- 623 Freitas F, Alves V, Reis M. 2014. Bacterial polysaccharides: production and applications in cosmetic
624 industry Springer International Publishing Cham.
- 625 Fujishiro T, Ogawa T, Matsuoka M, Nagahama K, Takeshima Y, Hagiwara H. 2004. Establishment
626 of a pure culture of the hitherto uncultured unicellular cyanobacterium *Aphanothece*
627 *sacrum*, and phylogenetic position of the organism. *Applied and environmental*
628 *microbiology* **70**:3338-3345. DOI: 10.1128/AEM.70.6.3338-3345.2004, PMID: PMC427778
- 629 Ghosh T, Chattopadhyay K, Marschall M, Karmakar P, Mandal P, Ray B. 2009. Focus on antivirally
630 active sulfated polysaccharides: from structure-activity analysis to clinical evaluation.
631 *Glycobiology* **19**:2-15. DOI: 10.1093/glycob/cwn092
- 632 Harke MJ, Steffen MM, Gobler CJ, Otten TG, Wilhelm SW, Wood SA, Paerl HW. 2016. A review of
633 the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp.
634 *Harmful Algae* **54**:4-20
- 635 Hayashi T, Hayashi K, Maeda M, Kojima I. 1996. Calcium spirulan, an inhibitor of enveloped virus
636 replication, from a blue-green alga *Spirulina platensis*. *Journal of natural products* **59**:83-
637 87
- 638 Henrissat B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities.
639 *Biochemical Journal* **280**:309-316
- 640 Hirose Y, Narikawa R, Katayama M, Ikeuchi M. 2010. Cyanobacteriochrome CcaS regulates
641 phycoerythrin accumulation in *Nostoc punctiforme*, a group II chromatic adapter.
642 *Proceedings of the National Academy of Sciences* **107**:8854-8859
- 643 Huisman J, Codd GA, Paerl HW, Ibelings BW, Verspagen JM, Visser PM. 2018. Cyanobacterial
644 blooms. *Nature Reviews Microbiology* **16**:471-483
- 645 Islam ST, Lam JS. 2014. Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent

pathway. *Canadian journal of microbiology* **60**:697-716

Jiao G, Yu G, Zhang J, Ewart HS. 2011. Chemical structures and bioactivities of sulfated polysaccharides from marine algae. *Marine Drugs* **9**:196-223. DOI: 10.3390/md9020196, PMID: PMC3093253

Jittawuttipoka T, Planchon M, Spalla O, Benzerara K, Guyot F, Cassier-Chauvat C, Chauvat F. 2013. Multidisciplinary evidences that *Synechocystis* PCC6803 exopolysaccharides operate in cell sedimentation and protection against salt and metal stresses. *PLoS One* **8**:e55564

Karamanos NK, Piperigkou Z, Theocharis AD, Watanabe H, Franchi M, Baud S, Brezillon S, Gottel M, Passi A, Vigetti D. 2018. Proteoglycan chemical diversity drives multifunctional cell regulation and therapeutics. *Chemical reviews* **118**:9152-9232

Katzen F, Ferreira DU, Oddo CG, Ielmini MV, Becker A, Pühler A, Ielpi L. 1998. *Xanthomonas campestris* pv. *campestris* gum Mutants: Effects on Xanthan Biosynthesis and Plant Virulence. *Journal of bacteriology* **180**:1607-1617

Kawano Y, Saotome T, Ochiai Y, Katayama M, Narikawa R, Ikeuchi M. 2011. Cellulose accumulation and a cellulose synthase gene are responsible for cell aggregation in the cyanobacterium *Thermosynechococcus vulcanus* RKN. *Plant and Cell Physiology* **52**:957-966

Kopf M, Klähn S, Pade N, Weingärtner C, Hagemann M, Voß B, Hess WR. 2014a. Comparative genome analysis of the closely related *Synechocystis* strains PCC 6714 and PCC 6803. *DNA research* **21**:255-266

Kopf M, Klähn S, Scholz I, Matthiessen JK, Hess WR, Voß B. 2014b. Comparative analysis of the primary transcriptome of *Synechocystis* sp. PCC 6803. *DNA research* **21**:527-539

Kumar D, Kastanek P, Adhikary SP. 2018. Exopolysaccharides from cyanobacteria and microalgae and their commercial application. *Current Science* **115**:234-241

Lapasin R, Pricl S, 1995. Industrial applications of polysaccharides *Rheology of Industrial Polysaccharides: Theory and Applications*. Springer. p134-161.

672 Letunic I, Doerks T, Bork P. 2015. SMART: recent updates, new developments and status in 2015.
673 *Nucleic acids research* **43**:D257-D260

674 Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2013. The carbohydrate-
675 active enzymes database (CAZy) in 2013. *Nucleic acids research* **42**:D490-D495

676 Maeda K, Narikawa R, Ikeuchi M. 2014. CugP is a novel ubiquitous non-GalU-type bacterial UDP-
677 glucose pyrophosphorylase found in cyanobacteria. *Journal of bacteriology* **196**:2348-2354

678 Maeda K, Tamura J, Okuda Y, Narikawa R, Midorikawa T, Ikeuchi M. 2018. Genetic identification
679 of factors for extracellular cellulose accumulation in the thermophilic cyanobacterium
680 *Thermosynechococcus vulcanus*: proposal of a novel tripartite secretion system. *Molecular*
681 *Microbiology* **109**:121-134

682 Mangan NM, Flamholz A, Hood RD, Milo R, Savage DF. 2016. pH determines the energetic
683 efficiency of the cyanobacterial CO₂ concentrating mechanism. *Proceedings of the National*
684 *Academy of Sciences* **113**:E5354-E5362

685 Merel S, Walker D, Chicana R, Snyder S, Baurès E, Thomas O. 2013. State of knowledge and
686 concerns on cyanobacterial blooms and cyanotoxins. *Environment international* **59**:303-327

687 Midorikawa T, Matsumoto K, Narikawa R, Ikeuchi M. 2009. An Rrf2-type transcriptional regulator
688 is required for expression of psaAB genes in the cyanobacterium *Synechocystis* sp. PCC
689 6803. *Plant physiology* **151**:882-892

690 Mota R, Vidal R, Pandeirada C, Flores C, Adessi A, De Philippis R, Nunes C, Coimbra MA,
691 Tamagnini P. 2020. Cyanoflan: A cyanobacterial sulfated carbohydrate polymer with
692 emulsifying properties. *Carbohydrate Polymers* **229**:115525. DOI:
693 10.1016/j.carbpol.2019.115525

694 Mougous JD, Green RE, Williams SJ, Brenner SE, Bertozzi CR. 2002. Sulfotransferases and
695 sulfatases in mycobacteria. *Chemistry & biology* **9**:767-776

696 Mouhim RF, Cornet J-F, Fontane T, Fournet B, Dubertret G. 1993. Production, isolation and
697 preliminary characterization of the exopolysaccharide of the cyanobacterium *Spirulina*

platensis. *Biotechnology letters* **15**:567-572

Ngatu NR, Okajima MK, Yokogawa M, Hirota R, Eitoku M, Muzembo BA, Dumavibhat N, Takaishi M, Sano S, Kaneko T. 2012. Anti-inflammatory effects of sacran, a novel polysaccharide from *Aphanothece sacrum*, on 2, 4, 6-trinitrochlorobenzene-induced allergic dermatitis in vivo. *Annals of Allergy, Asthma & Immunology* **108**:117-122. e112

Ohbayashi R, Watanabe S, Ehira S, Kanesaki Y, Chibazakura T, Yoshikawa H. 2016. Diversification of DnaA dependency for DNA replication in cyanobacterial evolution. *The ISME journal* **10**:1113-1121

Okajima MK, Bamba T, Kaneko Y, Hirata K, Fukusaki E, Kajiyama Si, Kaneko T. 2008. Supergiant ampholytic sugar chains with imbalanced charge ratio form saline ultra-absorbent hydrogels. *Macromolecules* **41**:4061-4064

Panoff J-M, Priem B, Morvan H, Joset F. 1988. Sulphated exopolysaccharides produced by two unicellular strains of cyanobacteria, *Synechocystis* PCC 6803 and 6714. *Archives of microbiology* **150**:558-563

Pereira S, Zille A, Micheletti E, Moradas-Ferreira P, De Philippis R, Tamagnini P. 2009. Complexity of cyanobacterial exopolysaccharides: composition, structures, inducing factors and putative genes involved in their biosynthesis and assembly. *FEMS Microbiology Reviews* **33**:917-941

Pereira SB, Mota R, Vieira CP, Vieira J, Tamagnini P. 2015. Phylum-wide analysis of genes/proteins related to the last steps of assembly and export of extracellular polymeric substances (EPS) in cyanobacteria. *Scientific reports* **5**

Pereira SB, Santos M, Leite JP, Flores C, Eisfeld C, Büttel Z, Mota R, Rossi F, De Philippis R, Gales L. 2019. The role of the tyrosine kinase Wzc (Slr0923) and the phosphatase Wzb (Slr0328) in the production of extracellular polymeric substances (EPS) by *Synechocystis* PCC 6803. *MicrobiologyOpen* **8**:e00753

Sasisekharan R, Raman R, Prabhakar V. 2006. Glycomics approach to structure-function

relationships of glycosaminoglycans. *Annu Rev Biomed Eng* **8**:181-231

Schmid J, Sieber V, Rehm B. 2015. Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. *Frontiers in microbiology* **6**:496. DOI: 10.3389/fmicb.2015.00496, PMID: PMC4443731

Stanier R, Kunisawa R, Mandel M, Cohen-Bazire G. 1971. Purification and properties of unicellular blue-green algae (order *Chroococcales*). *Bacteriological reviews* **35**:171-205

Steffen MM, Li Z, Effler TC, Hauser LJ, Boyer GL, Wilhelm SW. 2012. Comparative metagenomics of toxic freshwater cyanobacteria bloom communities on two continents. *PLoS One* **7**:e44002

Vial L, Hommais F. 2020. Plasmid - chromosome cross - talks. *Environmental Microbiology* **22**:540-556

Walsby A. 1994. Gas vesicles. *Microbiological reviews* **58**:94-144

Wardrop D, Keeling D. 2008. The story of the discovery of heparin and warfarin. *British journal of haematology* **141**:757-763

Wei K, Jung S, Amano Y, Machida M. 2019. Control of the buoyancy of *Microcystis aeruginosa* via colony formation induced by regulating extracellular polysaccharides and cationic ions. *SN Applied Sciences* **1**:1573

Willis LM, Whitfield C. 2013. Structure, biosynthesis, and function of bacterial capsular polysaccharides synthesized by ABC transporter-dependent pathways. *Carbohydrate Research* **378**:35-44

Woodward L, Naismith JH. 2016. Bacterial polysaccharide synthesis and export. *Current Opinion in Structural Biology* **40**:81-88

Yoshihara S, Suzuki F, Fujita H, Geng XX, Ikeuchi M. 2000. Novel putative photoreceptor and regulatory genes required for the positive phototactic movement of the unicellular motile cyanobacterium *Synechocystis* sp. PCC 6803. *Plant and Cell Physiology* **41**:1299-1304