Effects of rising CO_2 on the harmful cyanobacterium Microcystis

Sandrini, G.

Publication date
2016
Document Version
Final published version

Chapter 6

Diel variation in gene expression of the CO_2-concentrating mechanism during a harmful cyanobacterial bloom

Diel variation in gene expression of the CO_2-concentrating mechanism during a harmful cyanobacterial bloom

Giovanni Sandrini[1], Robert P. Tann[1], J. Merijn Schuurmans[1,2], Sebastiaan A. M. van Beusekom[1], Hans C. P. Matthijs[1], Jef Huisman[1]

[1]*Department of Aquatic Microbiology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands*
[2]*Department of Aquatic Ecology, Netherlands Institute of Ecology, Wageningen, The Netherlands*

Keywords: bicarbonate; climate change; CO_2-concentrating mechanism; gene expression/regulation; harmful algal blooms; lakes; *Microcystis aeruginosa*

This chapter is submitted for publication.

Abstract

Dense phytoplankton blooms in eutrophic waters often experience large daily fluctuations in environmental conditions. We investigated how this diel variation affects *in situ* gene expression of the CO_2-concentrating mechanism (CCM) and other selected genes of the harmful cyanobacterium *Microcystis aeruginosa*. Photosynthetic activity of the cyanobacterial bloom depleted the dissolved CO_2 concentration, raised pH to 10, and caused large diel fluctuations in the bicarbonate and O_2 concentration. The *Microcystis* population consisted of three C_i uptake genotypes that differed in the presence of the low-affinity and high-affinity bicarbonate uptake genes *bicA* and *sbtA*. Expression of the bicarbonate uptake genes *bicA*, *sbtA* and *cmpA* (encoding a subunit of the high-affinity bicarbonate uptake system BCT1), the CCM transcriptional regulator gene *ccmR* and the photoprotection gene *flv4* increased at first daylight and was negatively correlated with the bicarbonate concentration. In contrast, genes of the two CO_2 uptake systems were constitutively expressed, whereas expression of the RuBisCO chaperone gene *rbcX*, the carboxysome gene *ccmM*, and the photoprotection gene *isiA* was highest at night and down-regulated during daytime. In total, our results show that the harmful cyanobacterium *Microcystis* is very responsive to the large diel variations in carbon and light availability often encountered in dense cyanobacterial blooms.

Introduction

Harmful cyanobacterial blooms are a recurring problem in many eutrophic lakes worldwide (Chorus and Bartram, 1999; Huisman *et al.*, 2005; Michalak *et al.*, 2013). Cyanobacteria can produce a variety of hepatotoxins, gastrointestinal toxins and neurotoxins causing liver, digestive and neurological disease in birds and mammals, including humans (Carmichael, 2001; Codd *et al.*, 2005). This may lead to the closure of water bodies used for recreational use, drinking or irrigation water, and aquaculture (Verspagen *et al.*, 2006; Qin *et al.*, 2010; Steffen *et al.*, 2014). Cyanobacterial blooms are often favored by excessive nutrients and high temperatures (Jöhnk *et al.*, 2008; Kosten *et al.*, 2012), and it is foreseen that rising CO_2 concentrations and global warming will further promote harmful cyanobacterial blooms (Paerl and Huisman, 2009; O'Neil *et al.*, 2012; Verspagen *et al.*, 2014).

Yet, effects of daily fluctuations in environmental conditions on harmful cyanobacterial blooms remain poorly investigated. The day-night cycle leads to large changes in the activity of phototrophic organisms. The high photosynthetic activity of cyanobacterial blooms during daytime can deplete the dissolved CO_2 (CO_2(aq)) concentration, causing a concomitant increase in pH (Ibelings and Maberly, 1998; Balmer and Downing, 2011; Gu *et al.*, 2011). Conversely, at night, cyanobacteria obtain their energy from respiration of the

stored carbon compounds, releasing CO_2 back in the water column. Cyanobacteria adjust to this daily variation by changes in gene expression and the built-up or break-down of specific enzymes and cellular components (Kucho *et al.*, 2005; Straub *et al.*, 2011; Brauer *et al.*, 2013). In particular, the daily excursions in the availability of both light and inorganic carbon (C_i) imply that cyanobacteria may not only acclimate their light reactions of photosynthesis, but may also adjust their C_i uptake machinery to the diel cycle.

Cyanobacteria have evolved a CO_2-concentrating mechanism (CCM) to efficiently fix CO_2 at a wide range of C_i conditions (Kaplan and Reinhold, 1999; Price *et al.*, 2008; Raven *et al.*, 2012; Burnap *et al.*, 2015). The cyanobacterial CCM is based on the uptake of dissolved CO_2 and bicarbonate by different uptake systems and subsequent concentration of C_i in specialized compartments, called carboxysomes, that also contain the CO_2 fixing enzyme RuBisCO. Five different C_i uptake systems are known in cyanobacteria, three for bicarbonate uptake and two for CO_2 uptake (Price, 2011). The bicarbonate transporter BCT1 and the two CO_2 uptake systems are present in most freshwater cyanobacteria (Badger *et al.*, 2006; Rae *et al.*, 2011; Sandrini *et al.*, 2014). BCT1 is directly ATP-dependent, and combines a high affinity with a low flux rate (Omata *et al.*, 1999). The two NADPH-dependent CO_2 uptake systems, NDH-I_3 and NDH-I_4, have contrasting properties: NDH-I_3 has a high affinity and low flux rate, while NDH-I_4 has a low affinity and high flux rate (Maeda *et al.*, 2002; Price *et al.*, 2002). The other two bicarbonate uptake systems BicA and SbtA are less widespread among freshwater cyanobacteria (Badger *et al.*, 2006; Rae *et al.*, 2011; Sandrini *et al.*, 2014). Both BicA and SbtA are sodium-dependent symporters, where SbtA usually has a high affinity but low flux rate, and BicA has a low affinity but high flux rate (Price *et al.*, 2004; Du *et al.*, 2014). We recently compared CCM gene sequences of 20 strains of the harmful cyanobacterium *Microcystis aeruginosa* (Kützing) (Sandrini *et al.*, 2014). Interestingly, some strains lacked the high-flux bicarbonate uptake gene *bicA*, whereas others lacked the high-affinity bicarbonate uptake gene *sbtA*. Hence, in *Microcystis*, three different C_i uptake genotypes can be distinguished: *bicA* strains (with *bicA* but no *sbtA*), *sbtA* strains (with *sbtA* but no or incomplete *bicA*), and *bicA*+*sbtA* strains.

In laboratory studies, several cyanobacterial CCM genes were shown to be responsive to changes in both C_i availability and light intensity (Hihara *et al.*, 2001; Wang *et al.*, 2004; Schwarz *et al.*, 2011; Straub *et al.*, 2011; Burnap *et al.*, 2013; Sandrini *et al.*, 2015a). However, whereas lab studies usually focus on a single strain, cyanobacterial blooms often consist of multiple strains (Saker *et al.*, 2005; Kardinaal *et al.*, 2007). Recent work shows that different strains may respond differently to changes in CO_2 levels (Van de Waal *et al.*, 2011; Sandrini *et al.*, 2014, 2015c), which may complicate how daily fluctuations in C_i availability affect gene expression and cellular physiology of cyanobacterial communities in lakes. Two lake studies

used metatranscriptomics to study diel variation in cyanobacterial gene expression (Penn *et al.*, 2014) or to compare gene expression between different locations (Steffen *et al.*, 2015), yet detailed information on C_i concentrations and differences in expression of, for example, the C_i uptake genes were missing. Another interesting field study used RT-qPCR to investigate diel variation in the expression of CCM and other selected genes of a *Synechococcus* strain in the hot spring microbial mats from Yellowstone National Park (Jensen *et al.*, 2011). However, detailed data about the C_i conditions in the mats were not presented.

In this study, we investigate diel variation in gene expression patterns of *Microcystis* in a dense cyanobacterial bloom of a eutrophic lake. *Microcystis* is a ubiquitous harmful cyanobacterium that often threatens water quality by the production of dense blooms and hepatotoxins called microcystins (Carmichael, 2001; Huisman *et al.*, 2005; Qin *et al.*, 2010; Dittmann *et al.*, 2013). We quantified expression of the CCM genes as well as a few other selected genes of *Microcystis* using RT-qPCR. Furthermore, we monitored diel changes in environmental conditions, to link observed gene expression patterns to fluctuations in C_i and light availability. Our study provides a better understanding of the daily variations in the CCM during dense cyanobacterial blooms.

Results

Community composition

Changes in gene expression and environmental conditions were monitored in Lake Kennemermeer (**Figure 6.1**) for 24 h, from 13:00 to 13:00 the next day on July 17 and 18, 2013. At this time, the lake was dominated by cyanobacteria comprising >97% of the total phytoplankton biovolume. The four major cyanobacterial groups included Pseudanabaenaceae (380 mm^3 L^{-1}), *Anabaenopsis hungarica* (67 mm^3 L^{-1}), small Chroococcales (40 mm^3 L^{-1}) and *Microcystis* spp. (1 mm^3 L^{-1}; 70×10^6 cells L^{-1}). All three C_i uptake genotypes were present in the *Microcystis* population, which consisted of 23±1% *bicA* strains, 15±1% *sbtA* strains and 62±2% *bicA*+*sbtA* strains. Based on quantification of the microcystin synthetase gene *mcyB*, 82±5% of the *Microcystis* population was potentially toxic.

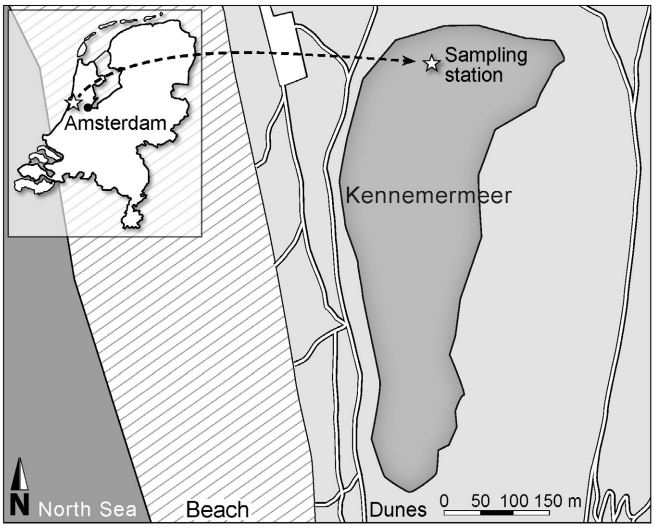

Figure 6.1. Map of Lake Kennemermeer with the sampling point.

Diel variation in environmental conditions

During the study period it was sunny and dry at daytime, and the photosynthetically active radiation (PAR) increased above 2,000 µmol photons m^{-2} s^{-1} in the early afternoon (**Figure 6.2A**). Sunset was at 21:51, the night was dry with clear skies, with sunrise at 5:42. Water temperature peaked at 22.9°C during early evening, and dropped to 20.6°C in the early morning of the next day (**Figure 6.2A**).

The dissolved CO_2 (CO_2(aq)) concentration was depleted to 0.05 µmol L^{-1} (equivalent to a partial pressure of 1 ppm) during daytime, but increased during the night to 0.32 µmol L^{-1} (8 ppm) at the early morning of the next day (**Figure 6.2B**). In equilibrium with an atmospheric partial pressure of 390 ppm, one would expect a CO_2(aq) concentration of ~15 µmol L^{-1}. Hence, the lake was strongly CO_2 undersaturated, even at night. Similar to CO_2(aq), the bicarbonate concentration showed large diel variation, from ~400 µmol L^{-1} in the early evening to ~850 µmol L^{-1} in the early morning of the next day (Pearson correlation of CO_2(aq) vs bicarbonate: ρ = 0.94, n = 8, p < 0.01; Fig. 2B).

CO_2 depletion by the phytoplankton bloom led to an increase in pH with values up to 10.3 in the afternoon (Pearson correlation of pH vs log CO_2(aq): ρ = -0.99, n = 8, p < 0.001; **Figure 6.2B,C**). CO_2 depletion was correlated with oxygen production by the bloom (Pearson correlation of CO_2(aq) vs dissolved oxygen: ρ = -0.95, n = 8, p < 0.001 **Figure 6.2B,C**), with dissolved oxygen peaking at ~200% during the early evening. During most of the day the lake was supersaturated with oxygen, except for the early morning when it was slightly undersaturated. The sodium concentration in the lake was 12.7±0.4 mmol L^{-1}.

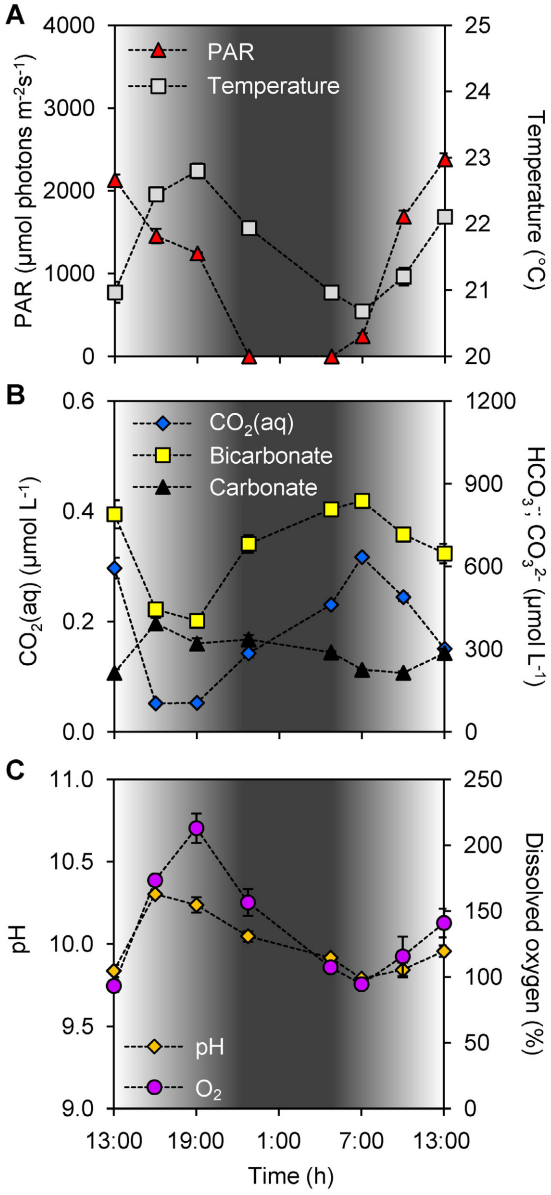

Figure 6.2. Diel changes in environmental conditions. (A) Photosynthetically active radiation (PAR) and water temperature. **(B)** Dissolved CO_2 (CO_2(aq)), bicarbonate and carbonate concentrations. **(C)** pH and dissolved oxygen. Shading indicates the light intensity measured at the water surface, with complete darkness between 22:30 and 5:00. Each data point shows the mean ±s.d. of three independent measurements.

Diel variation in gene expression

Most of the investigated genes of *Microcystis* showed significant diel variation in gene expression (**Figure 6.3; Table S6.1, Supplementary Information**). Hierarchical cluster analysis revealed that the expression patterns of the studied genes could be grouped into four distinct clusters (**Figure 6.4**).

Expression of the genes in cluster 1 increased during daytime and decreased at night (**Figure 6.4**). This cluster comprised the three bicarbonate uptake genes *cmpA*, *bicA* and *sbtA*, the transcriptional regulator gene *ccmR* and the flavodiiron protein gene *flv4*. Expression of *cmpA*, encoding for one of the subunits of the bicarbonate uptake system BCT1, showed the largest diel variation in this cluster. Its maximum expression during daytime was almost 360-fold higher (log2 difference of 8.5) than its minimum expression at night (**Figure 6.3A; Table S6.1, Supplementary Information**). Expression of *flv4* and *ccmR*, which is most likely a transcriptional regulator of the *cmpABCD* operon of *Microcystis* (Sandrini *et al.*, 2015c), also showed large diel variation (**Figure 6.3C,F; Table S6.1, Supplementary Information**). Expression of *bicA* and *sbtA*, encoding for the two sodium-dependent bicarbonate uptake systems, varied in tandem but at a lower amplitude than the other genes in this cluster (**Figure 6.3A; Table S6.1, Supplementary Information**).

Cluster 2 consisted only of *ccmR2* (**Figure 6.4**), a transcriptional regulator of the *bicA-sbtA* operon of *Microcystis* (Sandrini *et al.*, 2014). Similar to the genes in cluster 1, the expression of *ccmR2* increased during daytime and decreased at night, although at a lower amplitude than *ccmR* (**Figure 6.3C; Table S6.1, Supplementary Information**). Its expression pattern differed from the genes in cluster 1 by a dip during the morning of the second day.

Cluster 3 comprised the genes *chpX* and *chpY* (encoding the hydration subunits of the low-affinity and high-affinity CO_2 uptake system, respectively), *ccaA* (encoding the carboxysomal carbonic anhydrase) and *mcyB* (encoding a microcystin synthetase) (**Figure 6.4**). The cluster analysis suggests that these genes displayed a higher expression during the first daytime than during the second daytime period. However, diel variation of these genes was only minor and in most cases not significant, except for *mcyB* which decreased at night and increased briefly but significantly at first daylight of the second day (**Figure 6.3; Table S6.1, Supplementary Information**).

Expression of the genes in cluster 4 was lowest during daytime and highest at night (**Figures 6.3 and 6.4; Table S6.1, Supplementary Information**). This cluster comprised the CCM genes *rbcX* located in the RuBisCO operon and *ccmM* encoding for a carboxysomal shell protein, the gas vesicle protein gene *gvpC* and the gene *isiA* encoding for the iron-starvation induced protein.

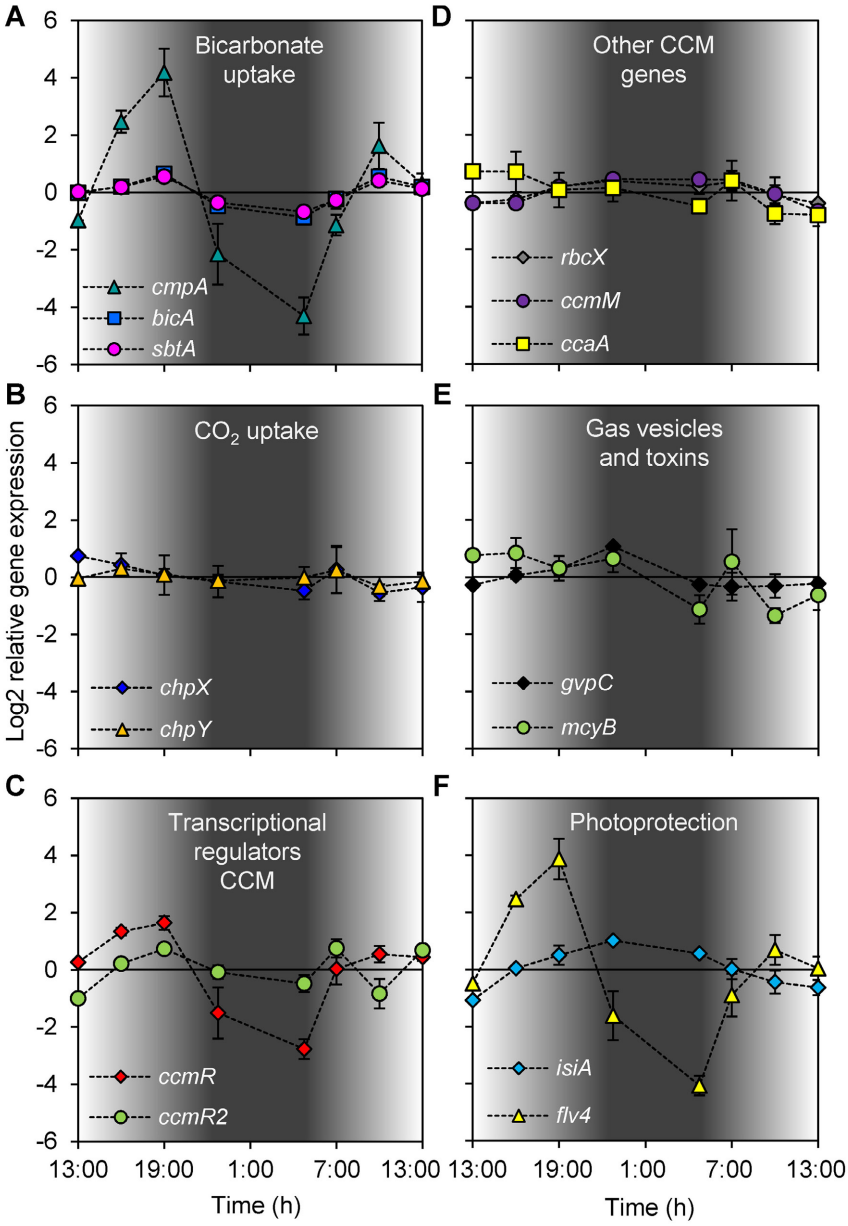

Figure 6.3. Diel changes in gene expression . (A) Bicarbonate uptake genes *cmpA*, *bicA* and *sbtA*. **(B)** CO_2 uptake genes *chpX* and *chpY*. **(C)** CCM transcriptional regulator genes *ccmR* and *ccmR2*. **(D)** RuBisCO chaperone gene *rbcX*, carboxysome gene *ccmM* and carbonic anhydrase gene *ccaA*. **(E)** Gas vesicle gene *gvpC* and microcystin gene *mcyB* **(F)** Photoprotection genes *isiA* and *flv4*. Gene expression was quantified as log2 ratios of the expression at the given time point relative to the mean expression over the 24-hour period. Shading indicates the light intensity measured at the water surface, with complete darkness between 22:30 and 5:00. Each data point shows the mean ±s.d. of three independent lake samples.

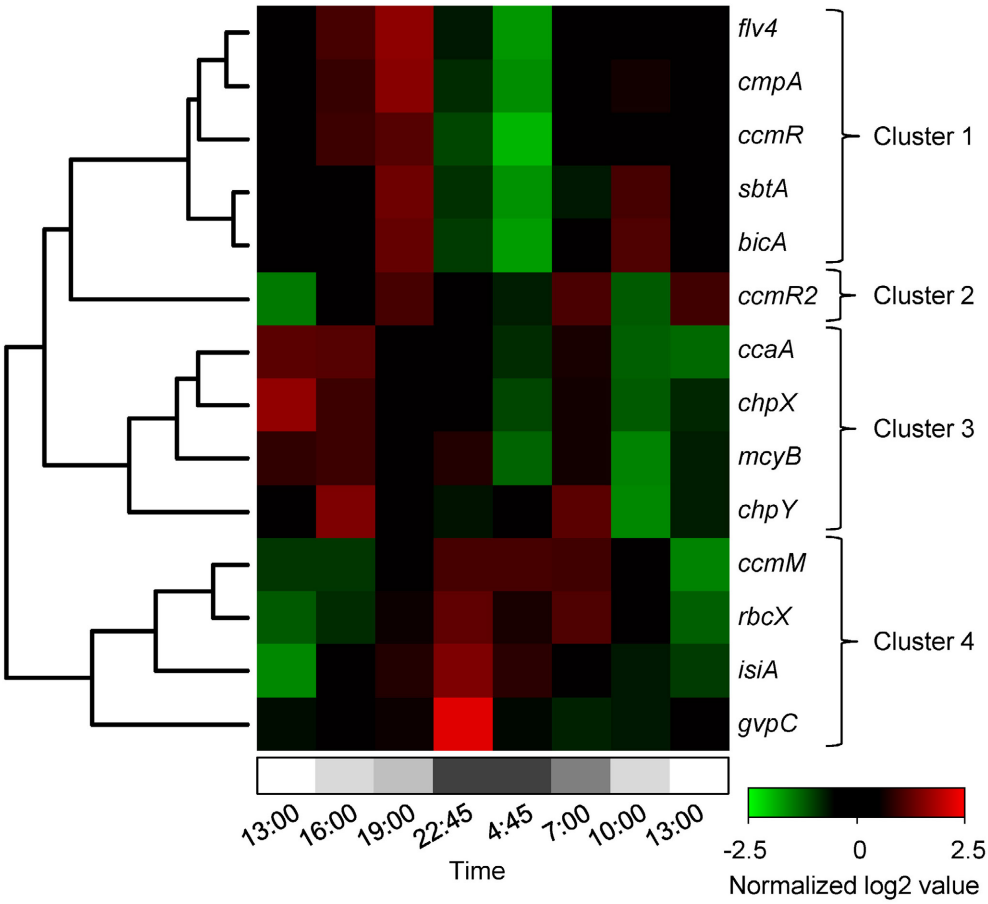

Figure 6.4. Hierarchical clustering of the diel changes in gene expression. The heatmap shows normalized expression data. The log2 expression values of each gene were normalized with respect to the mean and standard deviation for that gene over the 24-hour period. Upregulated genes are in red and downregulated genes in green. Grey shading at the bottom indicates the light intensity at the water surface, with complete darkness between 22:30 and 5:00. Hierarchical clustering resulted in four distinct gene clusters.

Comparison of gene expression patterns with environmental conditions

Several of the gene expression patterns were associated with diel changes in environmental conditions (**Table 6.1**). Correlations with the bicarbonate concentration and incident light intensity showed the most consistent response across the gene clusters. The expression of all five genes in cluster 1 showed a significant ($p < 0.05$) or marginally significant ($p < 0.10$) negative correlation with the bicarbonate concentration, whereas none of the other gene expression patterns were correlated with bicarbonate (**Table 6.1; Figure S6.1, Supplementary Information**). Of these five gene expression patterns, *cmpA* and *flv4* were strongly correlated with the bicarbonate concentration, whereas *bicA*, *sbtA* and *ccmR* showed a weaker correlation.

Expression of *bicA*, *sbtA* and *ccmR* (cluster 1) showed a weak but significant positive correlation with the incident light intensity, whereas expression of *rbcX*, *ccmM* and *isiA* (cluster 4) showed a significant negative correlation with the incident light intensity (**Table 6.1; Figure S6.2, Supplementary Information**). Furthermore, the expression of several genes (*cmpA*, *bicA*, *ccmR*, *ccmR2*, *mcyB* and *flv4*) increased significantly from 4:45 AM to 7:00 AM, *i.e.*, at first daylight (**Figure 6.3; Table S6.1, Supplementary Information**) before the C_i concentration was depleted by the bloom (**Figure 6.2B**).

Table 6.1. Pearson correlation coefficients between gene expression patterns and environmental variables (*n*=8).

Gene	Pearson correlation coefficient					
	CO_2(aq)	Bicarbonate[§]	pH	Temperature	Light[§]	Dissolved oxygen
Cluster 1:						
cmpA	ns	-0.82**	ns	0.69*	ns	0.68*
bicA	ns	-0.62$^{\#}$	ns	ns	0.69*	ns
sbtA	ns	-0.67*	ns	ns	0.70*	ns
ccmR	ns	-0.65$^{\#}$	ns	ns	0.67*	ns
flv4	-0.62$^{\#}$	-0.84**	ns	0.72*	ns	0.70*
Cluster 2:						
ccmR2	ns	ns	ns	ns	ns	ns
Cluster 3:						
chpX	ns	ns	ns	ns	ns	ns
chpY	ns	ns	ns	ns	ns	ns
ccaA	ns	ns	ns	ns	ns	ns
mcyB	ns	ns	ns	ns	ns	ns
Cluster 4:						
rbcX	ns	ns	ns	ns	-0.92**	ns
ccmM	ns	ns	ns	ns	-0.94**	ns
isiA	ns	ns	ns	ns	-0.83**	ns
gvpC	ns	ns	ns	ns	ns	ns

$^{\#}$ $p < 0.10$ (two-tailed); * $p < 0.05$ (two-tailed); ** $p < 0.01$ (two-tailed); ns: not significant ($p > 0.10$)
[§]Scatterplots of gene expression versus bicarbonate concentration and incident light intensity are shown in Figure S6.1 and Figure S6.2.

Discussion

Inorganic carbon dynamics

Our lake study shows that cyanobacterial blooms can deplete the CO_2(aq) concentration and increase pH, and bring about substantial diel variation in pH and concentrations of dissolved inorganic carbon and oxygen. Similar observations have been made in other lakes with dense phytoplankton blooms (Maberly, 1996; Ibelings and Maberly, 1998). For instance, Maberly (1996) monitored the inorganic carbon concentrations of Lake Esthwaite Water (UK) for an entire year, and found diel pH fluctuations up to 1.8 units and diel variation in CO_2(aq) up to ~30 µmol L^{-1}. During dense phytoplankton blooms, CO_2(aq) concentrations could remain undersaturated with respect to the atmospheric CO_2 level for several consecutive weeks.

Depletion of the CO_2(aq) concentration by dense cyanobacterial blooms creates a pCO_2 gradient across the air-water interface, which implies that these lakes act as a sink for atmospheric CO_2 (Balmer and Downing, 2010). More specifically, the CO_2 influx (g_{CO2}) from the atmosphere into a lake can be estimated from the difference between the expected concentration of CO_2(aq) in equilibrium with the atmosphere (according to Henry's law) and the measured concentration of CO_2(aq) (Siegenthaler and Sarmiento, 1993; Cole *et al.*, 2010):

$$g_{CO2} = v\,(K_H\,pCO_2 - CO_2(aq)) \quad (1)$$

where v is the gas transfer velocity across the air-water interface, K_H is the solubility constant of CO_2 gas in water (K_H = 0.0375 mol L^{-1} atm^{-1} at 21.5°C; Weiss, 1974), and pCO_2 is the partial pressure of CO_2 in the atmosphere. Assuming a typical gas transfer velocity of v = 0.02 m h^{-1} (Crusius and Wanninkhof, 2003; Cole *et al.*, 2010), an atmosphere with a pCO_2 of 390 ppm and an average CO_2(aq) of 0.2 µmol L^{-1} (**Figure 6.2B**), the daily CO_2 influx into Lake Kennemermeer would amount to ~7 mmol m^{-2} d^{-1}. This is a substantial CO_2 influx, which will fuel further development of the cyanobacterial bloom. Moreover, the above calculation indicates that a doubling of the atmospheric CO_2 concentration to 780 ppm will cause a near-doubling of the CO_2 influx into the lake to ~14 mmol m^{-2} d^{-1}, if we assume that dense cyanobacterial blooms still deplete the dissolved CO_2 concentration. These results are in line with model predictions that rising atmospheric CO_2 concentrations will stimulate cyanobacterial blooms in eutrophic lakes (Schippers *et al.*, 2004; Verspagen *et al.*, 2014).

Owing to the high pH induced by the cyanobacterial bloom, bicarbonate and carbonate were the most abundant inorganic carbon species in Lake Kennemermeer, whereas the CO_2(aq) concentration was more than three orders of magnitude lower. The amplitude of the daily variation in CO_2(aq) concentration was less than 0.3 µmol L^{-1}, whereas the bicarbonate concentration varied more than 400 µmol L^{-1} during the 24-hour period. Hence, cyanobacteria in these dense blooms will be strongly dependent on bicarbonate uptake to cover the high C demands of their photosynthetic activity, and may adapt to the large diel

fluctuations in bicarbonate availability by adjusting the production of their different bicarbonate uptake systems.

C_i uptake genes

Earlier we reported that *Microcystis* isolates from Lake Pehlitzsee (Germany) and Lake Volkerak (the Netherlands) contained both *sbtA* and *bicA+sbtA* strains, suggesting that these two C_i uptake genotypes may coexist (Sandrini *et al.*, 2014). Our current results show that all three *Microcystis* C_i uptake genotypes were present in Lake Kennemermeer, with relative abundances of ≥15% each, demonstrating that the three genotypes can co-occur in the same lake. The high relative abundance of *bicA+sbtA* strains (62% of the *Microcystis* population), containing both high-affinity and low-affinity bicarbonate uptake systems, indicates that strains that can adapt to the strongly fluctuating bicarbonate concentrations have a selective advantage under C_i-limited conditions.

In contrast to most laboratory studies, our gene expression data reflect the average gene expression of the entire *Microcystis* population rather than the expression of a single strain. Nevertheless, our lake study shows distinct patterns of gene expression. In cyanobacteria, the expression of many genes is under the control of the circadian clock genes *kaiA*, *kaiB* and *kaiC* (Ishiura *et al.*, 1998; Iwasaki *et al.*, 2002), and it is difficult to disentangle the role of environmental variation versus an internal circadian clock based on field observations alone. However, *kai*-mutants of the cyanobacterium *Synechococcus* PCC 7002 have shown that several CCM genes are *kai*-independent cycling genes (Ito *et al.*, 2009). Therefore, it is plausible that CCM genes of *Microcystis* are *kai*-independent as well. Our field data show that diel changes in the expression of the C_i uptake genes were associated with diel variation in C_i and/or light availability. This is in agreement with many laboratory experiments, which have demonstrated that the expression of cyanobacterial CCM genes can be strongly affected by changes in C_i and light availability (Hihara *et al.*, 2001; McGinn *et al.*, 2003; Woodger *et al.*, 2003; Wang *et al.*, 2004; Schwarz *et al.*, 2011; Burnap *et al.*, 2013; Sandrini *et al.*, 2015a). We will therefore focus our interpretation of the expression data on diel variation in these environmental factors.

The expression of *cmpA*, encoding for a subunit of the high-affinity bicarbonate transporter BCT1, was negatively correlated with the bicarbonate concentration in the lake (**Table 6.1**). This pattern is in agreement with laboratory studies with *Microcystis* and other cyanobacteria, where *cmpA* expression was also strongly down-regulated at elevated C_i conditions (Sandrini *et al.*, 2015a; 2015c), and up-regulated under C_i-limiting conditions (Woodger *et al.*, 2003; Wang *et al.*, 2004; Schwarz *et al.*, 2011). Bicarbonate concentrations decreased at daytime and increased at night, which resulted in a high *cmpA* expression in the late afternoon and a low *cmpA* expression at the end of the night. The large amplitude of the

diel changes in *cmpA* expression suggests that the BCT1 enzyme is largely degraded at night and resynthesized each day. Expression of *cmpA* was not significantly correlated with light availability. However, bicarbonate uptake by BCT1 is ATP-dependent and ATP synthesis is driven by the light reactions of photosynthesis, which may explain the increase in *cmpA* expression at first daylight when the bicarbonate concentration was still high (**Figures 6.2B and 6.3A**). This explanation is also in agreement with previous laboratory studies under C_i-limited conditions, which showed increased *cmpA* expression at elevated light levels in several other cyanobacteria (Woodger *et al.*, 2003; McGinn *et al.*, 2004).

Expression of the bicarbonate uptake genes *bicA* and *sbtA* was also negatively correlated with the bicarbonate concentration (**Table 6.1**). In addition, their expression was positively correlated with the incident light intensity (**Table 6.1**). The expression of *bicA* and *sbtA* fluctuated in tandem (**Figure 6.3A**), likely because the *Microcystis* population was dominated by *bicA*+*sbtA* strains, in which both genes are located on the same operon (Sandrini *et al.*, 2014). However, the diel variation in expression of *bicA* and *sbtA* had a much smaller amplitude than *cmpA*. The sodium concentration in the lake (12.7±0.4 mmol L^{-1}) allowed near-maximum activity of the sodium-dependent bicarbonate uptake systems BicA and SbtA, since both uptake systems have half-saturation constants of 1-2 mmol L^{-1} sodium (Price *et al.*, 2004; Du *et al.*, 2014). Therefore, the sodium concentration is unlikely to limit the amplitude of *bicA* and *sbtA* expression. Laboratory studies have shown that the response of *bicA* and *sbtA* expression to changes in C_i availability varies among *Microcystis* strains (Sandrini *et al.*, 2015c). In some strains, *bicA* and *sbtA* expression is strongly down-regulated at elevated C_i conditions, whereas in other strains these genes are constitutively expressed. Hence, whether the observed low-amplitude variation in our field data reflects a generic pattern for many *Microcystis* blooms or a specific pattern for the *Microcystis* populations in our lake study remains to be investigated. Price *et al.* (2013) suggested that BicA and SbtA are simply inactivated by darkness and remain safeguarded for renewed activity at dawn, which may explain the low amplitude variation in gene expression, although the exact mechanisms of inactivation and re-activation remain to be revealed. Indeed, recently it was shown that SbtB is involved in the inhibition of SbtA during darkness (Du *et al.*, 2014).

The *chpX* and *chpY* genes, encoding hydration subunits of the low-affinity and high-affinity CO_2 uptake system, respectively, were constitutively expressed in our lake study (**Figure 6.3B**). Expression of these genes was not correlated with diel variation in CO_2(aq) concentration or light intensity (**Table 6.1**), despite the NADPH dependence of both CO_2 uptake systems (Mi *et al.*, 1995; Tchernov *et al.*, 2001; Price *et al.*, 2002). A possible reason could be that the diel variation in CO_2(aq) concentrations was too low (from 0.05 to 0.32 µmol L^{-1}) to affect expression of the CO_2 uptake genes. However, both CO_2 uptake genes were

also constitutively expressed in laboratory studies with axenic *Microcystis* strains exposed to a much larger increase in CO_2(aq) concentration (Sandrini *et al.*, 2015a, 2015c). In other cyanobacterial species, constitutive expression was established for the low-affinity CO_2 uptake gene *chpX* whereas expression of the high-affinity CO_2 uptake gene *chpY* was induced at low C_i levels (Woodger *et al.*, 2003; Wang *et al.*, 2004; Eisenhut *et al.*, 2007; Schwarz *et al.*, 2011). Furthermore, Jensen *et al.* (2011) found upregulation of both *chpY* and *chpX* during daytime for a *Synechococcus* mat. Hence, the CO_2 uptake genes of different species respond differently to changing environmental conditions.

Expression of the CCM transcriptional regulator gene *ccmR* was negatively correlated with the bicarbonate concentration, similar to the bicarbonate uptake genes (**Table 6.1**). *Microcystis* lacks the transcriptional regulator gene *cmpR*, which regulates expression of the *cmpABCD* operon in some other cyanobacteria (Omata *et al.*, 2001). Instead, CcmR most likely regulates expression of the *cmpABCD* operon in *Microcystis* (Sandrini *et al.*, 2014, 2015c), which explains the similar gene expression patterns of *cmpA* and *ccmR* in our study. The *ccmR2* gene is located directly upstream of the *bicA/sbtA* operon, and appears to be involved in the transcription of this operon (Sandrini *et al.*, 2014). This gene formed a separate cluster in our analysis (**Figure 6.4**), probably because of a single deviant data point at 10:00 AM. Omitting this single data point, *ccmR2* showed similar diel variation in expression as the bicarbonate uptake genes and *ccmR* (**Figure 6.3A,C**).

Diel changes in expression of other genes

The flavodiiron protein gene *flv4*, which is in one operon with *flv2*, was assigned to the same cluster as the bicarbonate uptake genes and the regulator gene *ccmR* (**Figure 6.4**). Flavodiiron proteins are stress proteins known to be involved in the acclimation to low C_i conditions (Zhang *et al.*, 2009; Allahverdiyeva *et al.*, 2011; Bersanini *et al.*, 2014; Sandrini *et al.*, 2015a), and can also be up-regulated by high light (Zhang *et al.*, 2009). Similarly, in our study *flv4* expression was increased at first daylight and reached high levels when C_i was depleted in the afternoon (**Figure 6.3F**). In contrast, Straub *et al.* (2011) found no significant diel variation in the expression of *flv2* and *flv4* in laboratory experiments with the axenic *Microcystis* strain PCC 7806, probably because their experiments were sparged with 1% CO_2 and hence C_i concentrations were not limiting in their study. Both genes, *flv2* and *flv4*, were shown to be involved in photoprotection of the PSII complex (Zhang *et al.*, 2009). Hence, our study indicates that Flv2/4 proteins are highly active in daytime photoprotection of cyanobacterial blooms.

Interestingly, the iron-stress chlorophyll-binding protein IsiA is also involved in photoprotection (Havaux *et al.*, 2005), but the expression pattern of *isiA* was very different

from *flv4* (**Figure 6.3F**). Expression of *isiA* was high at night but low during daytime (**Figure 6.3F**). In addition to its photoprotective function, IsiA is also involved in the formation of PSI antennae enhancing the light harvesting ability and in the storage of free chlorophyll molecules (Yeremenko *et al.*, 2004; Havaux *et al.*, 2005). Possibly, the turnover of photosystems and presence of free chlorophylls at night might stimulate *isiA* expression for temporary chlorophyll storage.

Expression of the *rbcX*, *ccmM*, *ccaA*, *gvpC* and *mcyB* genes showed only minor diel variation. The RuBisCO gene *rbcX* and carboxysomal gene *ccmM* were negatively correlated with light intensity (**Table 6.1**), which could indicate that cells slightly increase the RuBisCO and carboxysome numbers at night to prepare for cell division, although it was previously shown that *Microcystis* cell division occurs both during daytime and at night (Yamamoto and Tsukada, 2009). Alternatively, the negative correlation might reflect an afternoon photosynthesis dip due to photoinhibition (Ibelings and Maberly, 1998; Wu *et al.*, 2011a) when the incident light intensity exceeded 1,500 µmol photons m^{-2} s^{-1} (**Figure 6.2A**). Expression of the other three genes (*ccaA*, *gvpC*, *mcyB*) did not reveal a significant correlation with any of the environmental variables investigated (**Table 6.1**).

Conclusions

Cyanobacteria are often assumed to be favored at low C_i and high pH conditions, because of the presence of an effective CCM (*e.g.*, Shapiro, 1997). However, recent laboratory studies revealed considerable genetic and phenotypic diversity in the CCM of the ubiquitous harmful cyanobacterium *Microcystis* (Sandrini *et al.*, 2014, 2015c). Some strains perform well at low C_i levels, whereas other strains are much better competitors under high C_i conditions, suggesting that it might be difficult to foretell how natural mixtures of different *Microcystis* strains will respond to changes in C_i availability. Yet, our lake study showed consistent patterns in gene expression. The bicarbonate concentration in the lake showed large diel fluctuations, which may explain the predominance of *bicA+sbtA* strains and the concurrent diel fluctuations in expression of the bicarbonate uptake genes and their regulator genes. Hence, we conclude that the genetic and phenotypic versatility of the CCM of *Microcystis* enables a remarkably flexible response to the large diel fluctuations in C_i conditions often encountered in dense blooms.

Materials and Methods

Study area and sampling

Lake Kennemermeer (52°27'18.5"N, 4°33'48.6"E) is located north-west of Amsterdam, the Netherlands, near the North Sea coast (**Figure 6.1**). The man-made lake has been used as bathing water, but nowadays it is no longer in use as recreational lake because of yearly recurrent problems with harmful cyanobacterial blooms. The lake has a maximum depth of ~1 m, and surface area of ~0.1 km^2. The shallow lake is well mixed by wind throughout the year.

During a dense summer bloom in July 2013, we monitored diel variation in gene expression and environmental conditions. A small boat was used to sample the north side of the lake (**Figure 6.1**) at 13:00, 16:00, 19:00 and 22:45 of July 17, and at 4:45, 7:00, 10:00 and 13:00 of July 18. At each time point, the incident light intensity (photosynthetically active radiation [PAR]) just above the water surface, water temperature, pH and dissolved oxygen at 0.2 m depth were measured in triplicate using a Hydrolab Surveyor with a Datasonde 4a (OTT Hydromet, Loveland, CO, USA). Furthermore, three independent water samples of 5 L each were collected at 0.2 m depth using plastic 10 L tanks. The samples were processed immediately for further analyses.

Cell counts

Lugol's iodine was added to 40 mL aliquots of fresh lake samples (1:100 v/v of a 5% solution) to preserve phytoplankton cells for microscopy. The samples were stored at 4°C until identification. Phytoplankton was counted according to the Utermöhl-method adjusted to the European standard protocol NEN-EN 15204 using a Lyca DM IRB inverted light microscope (Lyca Microsystems BV, Rijswijk, the Netherlands). Phytoplankton was identified to the genus level, and if possible to the species level. Biovolume was estimated from cellular dimensions and geometry (Hillebrand *et al.*, 1999). Individual *Microcystis* cells were counted after disintegrating the colonies with KOH (Kardinaal *et al.*, 2007). All *Microcystis* cells were classified as *Microcystis aeruginosa* (Kützing) following the bacterial nomenclature proposed by Otsuka *et al.* (2001).

Dissolved inorganic carbon and sodium

For analysis of the dissolved inorganic carbon (DIC) and sodium concentration, lake samples were filtered on site using 1.2 µm pore size 47 mm GF/C filters (Whatman GmbH, Dassel, Germany) followed by 0.45 µm pore size 47 mm polyethersulfone membrane filters (Sartorius AG, Goettingen, Germany). The filtrate was transferred to sterile plastic urine analysis tubes (VF-109SURI; Terumo Europe N.V., Leuven, Belgium) which were filled completely (with an inserted needle to release all air), and stored at 4°C until further analysis. A TOC-V_{CPH} TOC

analyzer (Shimadzu, Kyoto, Japan) was used to determine the DIC concentration, with 3-5 technical replicates per sample. Concentrations of dissolved CO_2 (CO_2(aq)), bicarbonate and carbonate were calculated from DIC and the pH and temperature of the lake (Stumm and Morgan, 1996). Sodium ion concentrations were measured using an Optima 8000 ICP-OES Spectrometer (Perkin Elmer, Waltham, MA, USA).

RNA extraction

For RNA extraction, lake samples were filtered on-site with large 90 mm GF/C filters (Whatman GmbH, Dassel, Germany; 1.2 µm pore size) to concentrate phytoplankton biomass. Next, the 90 mm filters were dissected into several pieces using sterile surgical blades, which were transferred to 2 mL tubes that were filled with 1 mL TRIzol (Thermo Fisher Scientific, Waltham, MA, USA). The tubes were shaken vigorously, transported to the lab in a CX 100 Dry Shipper (Taylor Wharton, Theodore, AL, USA), and stored at -80°C until further analysis. RNA was extracted using 0.5 mm bashing beads (Zymo Research, Orange, CA, USA) to facilitate cell disruption, and purified according to Sandrini *et al.* (2015a) with the Direct-Zol™ RNA MiniPrep kit (Zymo Research, Orange, CA, USA) including in-column DNase I digestion. RNA concentrations were quantified using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). All RNA samples had A_{260}/A_{280} and A_{260}/A_{230} values above 1.8.

Primer development

Primers were developed to study expression of the *Microcystis* genes *bicA*, *sbtA*, *cmpA*, *chpX*, *chpY*, *ccmR*, *ccmR2*, *rbcX*, *ccmM*, *ccaA*, *mcyB*, *gvpC*, *isiA* and *flv4*, using 16S rRNA as 'reference gene'. The complete list of primers is shown in **Table S6.2**. The primers were designed to match the sequences of *Microcystis* strains NIES-843, PCC 7005, PCC 7941, PCC 7806, PCC 9432, PCC 9443, PCC 9701, PCC 9717, PCC 9806, PCC 9807, PCC 9808, PCC 9809, and T1-4. The *bicA*, *sbtA* and *ccmR2* primers were also based on sequences of strains CCAP 1450/10, CCAP 1450/11, HUB 5-2-4, HUB 5-3, NICA-CYA 140, V145 and V163 (Sandrini *et al.*, 2014). Sequences of other freshwater cyanobacteria were used to ensure that the primer design included several mismatches with non-*Microcystis* species. The primers were tested on gDNA (obtained as described below) of 11 different *Microcystis* strains to ensure one target PCR product was formed. These PCR reactions were done with the GoTaq® Hot Start Polymerase kit (Promega Corporation, Madison, WI, USA) according to the supplier's instructions. After an initial denaturation of 2 min at 95°C, 35 cycles were used that consisted of a denaturation step at 95°C for 45 s, an annealing temperature step at 60°C for 30 s and an extension step at 72°C for 3 min. Subsequently, a final extension step at 72°C was used for 5

min. The reactions contained 0.3 µmol L^{-1} primers and 10 ng gDNA in a total reaction volume of 25 µL. Other reaction components were added as instructed by the supplier. Gel electrophoresis showed that only the targeted gDNA sequences were amplified (no by-products were detected) As a negative control, the primers were tested on gDNA of *Anabaena circinalis* CCAP 1403/18, *Aphanizomenon flos-aqua* CCAP 1401/7 and *Planktothrix agardhii* CCAP 1460/1, which did not result in PCR amplification.

RT-qPCR gene expression analysis

To quantify gene expression, cDNA was synthesized by reverse transcription of the RNA samples with Superscript III (Thermo Fisher Scientific, Waltham, MA, USA) according to Sandrini *et al.* (2015a). Subsequently, the qPCR Maxima® SYBR Green Master Mix (2x) (Thermo Fisher Scientific, Waltham, MA, USA) was applied on the cDNA samples according to Sandrini *et al.* (2015a), in an ABI 7500 Real-Time PCR system (Applied Biosystems, Foster City, CA, USA). The two-step cycling protocol was used, with a denaturation temperature of 95°C (15 s) and a combined annealing/extension temperature of 60°C (60 s) during 40 cycles. The reactions contained 0.3 µmol L^{-1} primers and 1 µL of 10 times diluted cDNA from the RT reaction in a total reaction volume of 25 µL. Other reaction components were added as instructed by the supplier. ROX solution was used to correct for any well-to-well variation and melting curve analysis was performed on all measured samples to rule out non-specific PCR products.

Amplification efficiencies of individual runs (*E*) were calculated with LinRegPCR (version 2012.3; Ramakers *et al.*, 2003; Ruijter *et al.*, 2009) and were between 1.8 and 2.0 (**Table S6.2**). Time point 0 (13:00 of the first day) was used as 'reference sample', and 16S rRNA was used as 'reference gene'. Each RT-qPCR plate contained a reference sample and samples with primers targeting 16S rRNA to overcome plate effects. The data were baseline corrected using LinRegPCR, and the same software was used to calculate quantification cycle (*Cq*) values. LinRegPCR did not detect samples without amplification, without a plateau, with a baseline error or noise error, or with deviating amplification efficiencies. Negative control samples did not show significant amplification. Gene expression was quantified as the log2 ratio of the expression at a given time point relative to the mean expression over the 24-hour period using the comparative C_T method (Livak and Schmittgen, 2001). To determine if gene expression varied significantly between time points, one-way analysis of variance (ANOVA) was used with *post-hoc* comparison of the means based on Tukey's HSD test (α = 0.05) using SPSS version 20.0.

Hierarchical clustering was applied to compare the expression patterns of the studied genes. For each gene, time series of the expression values were normalized by the

transformation $(x - \mu)/\sigma$, where x is the original data point, μ is the mean of the time series, and σ is its standard deviation. Thus, all genes obtained normalized expression patterns with mean 0 and standard deviation 1. Hierarchical clustering and heatmap representation of the normalized gene expression data was done using the hclust and heatmap.2 functions of the gplots package in R version 3.0.2. We used the complete linkage clustering method for hierarchical clustering (Everitt *et al.*, 2011).

Quantification of Microcystis genotypes

To quantify the relative abundances of different *Microcystis* genotypes, we applied qPCR on purified gDNA (**Table S6.2**). Lake samples were filtered on-site over 1.2 µm pore size 25 mm GF/C filters (Whatman GmbH, Dassel, Germany), and loaded filters were stored at -20°C. Subsequently, gDNA was extracted using the ZR Fungal/Bacterial DNA MiniPrep™ kit (Zymo Research, Orange, CA, USA) and further purified using the DNA Clean & Concentrator™-25 kit (Zymo Research) according to the supplier's instructions. The gDNA samples were analyzed using a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), which resulted in A_{260}/A_{280} values above 1.8 for all samples.

The Maxima® SYBR Green Master Mix (2x) kit (Thermo Fisher Scientific, Waltham, MA, USA) was applied to the purified gDNA according to the supplier's instructions in an ABI 7500 Real-Time PCR system (Applied Biosystems, Foster City, CA, USA). The two-step cycling protocol was used, and the reaction settings and melting curve analysis were the same as in the RT-qPCR section (see above). The reactions contained 0.3 µmol L^{-1} primers and 10 ng gDNA from lake samples in a total reaction volume of 25 µL. For data analysis, the LinRegPCR software tool (version 2012.3; Ramakers *et al.*, 2003; Ruijter *et al.*, 2009) and the comparative C_T method (Livak and Schmittgen, 2001) were used.

To determine the relative abundances of the different C_i uptake genotypes, we used the *bicA* gene (primers bicA-F1 and bicA-R1; **Table S6.2**) and the *sbtA* gene (primers sbtA-F1 and sbtA-R1; **Table S6.2**) as 'target genes'. The *bicA*+*sbtA* gene of the *bicA*+*sbtA* strains (primers bicA-F2 and sbtA-R2; **Table S6.2**) served as 'reference gene'. Purified gDNA of the axenic laboratory strains PCC 7005 and PCC 7941 (both *bicA*+*sbtA* strains; Sandrini *et al.*, 2014) served as 'reference samples', to calculate the relative ratios of (1) the *bicA* gene versus the *bicA*+*sbtA* gene, and (2) the *sbtA* gene versus the *bicA*+*sbtA* gene. We note that the *bicA* gene is present in both *bicA* strains and *bicA*+*sbtA* strains, and similarly the *sbtA* gene is present in both *sbtA* strains and *bicA*+*sbtA* strains. Hence, the relative abundances of the different C_i uptake genotypes can be calculated from the above two ratios based on the assumption that the sum of the *bicA* strains, *sbtA* strains and *bicA*+*sbtA* strains equals 100%.

To determine the relative abundance of potentially toxic genotypes, we used *mcyB* (primers mcyB-F and mcyB-R; **Table S6.2**) as 'target gene' and the RuBisCO chaperone gene *rbcX* (primers rbcX-F and rbcX-R) present in all *Microcystis* strains as 'reference gene'. Purified gDNA of the axenic toxic strains PCC 7806 and PCC 7941 was used as 'reference samples'.

Acknowledgements

We thank Veerle M. Luimsta for assistance with the cell counts. This research was supported by the Division of Earth and Life Sciences (ALW) of the Netherlands Organization for Scientific Research (NWO). JMS acknowledges support by the BE-BASIC Foundation.

Supplementary Information

This section contains the supplementary tables and figures of this chapter.

Table S1. One-way analyses of variance testing for significant differences in gene expression between time points.

Gene	df_1, df_2	F	p	Time points							
				13:00 (day 1)	16:00 (day 1)	19:00 (day 1)	22:45 (day 1)	4:45 (day 2)	7:00 (day 2)	10:00 (day 2)	13:00 (day 2)
cmpA	7, 16	53.808	0.000	bc	ef	f	b	a	bc	de	cd
bicA	7, 16	34.031	0.000	cd	de	f	ab	a	bc	ef	cde
sbtA	7, 16	23.886	0.000	bcd	de	e	ab	a	abc	de	cd
chpX	7, 16	1.620	0.200	a	a	a	a	a	a	a	a
chpY	7, 16	0.990	0.473	a	a	a	a	a	a	a	a
ccmR	7, 16	36.693	0.000	cd	de	e	b	a	c	cde	cd
ccmR2	7, 16	22.693	0.000	a	bc	c	b	ab	c	a	c
rbcX	7, 16	7.987	0.000	a	ab	bc	c	bc	c	abc	a
ccmM	7, 16	7.940	0.000	ab	ab	bc	c	c	c	abc	a
ccaA	7, 16	4.870	0.004	b	b	ab	ab	ab	ab	a	a
gvpC	7, 16	10.850	0.000	a	a	a	b	a	a	a	a
mcyB	7, 16	7.317	0.001	c	c	bc	c	ab	c	a	abc
flv4	7, 16	60.401	0.000	bcd	e	e	b	a	bc	d	cd
isiA	7, 16	22.980	0.000	a	bc	cd	d	cd	bc	ab	ab

This table tests differences in gene expression for the data shown in Figure 6.3.
The columns in the table indicate the genes, the degrees of freedom (df_1 and df_2), the value of the F-statistic (F), the corresponding probability (p), and the significant differences in gene expression between time points based on Tukey's HSD *post-hoc* test (α = 0.05). Time points which do not share the same letter are significantly different from each other.

Table S6.2. Overview of *Microcystis* primer pairs used for qPCR analysis.

Primer name	Sequence 5'→ 3' (length)	Target gene(s)	Locus tag	Accession no. (Genbank)	Expected product size (bp)	Amplification efficiency (*E*)	RT-qPCR on cDNA (1) or qPCR on gDNA (2)	Reference
16S-F	GTCGAACGGGAAT CTTCGGAT (21)	16S rRNA	*IPF_5548*	AM778951.1	157	1.88±0.03	1	Sandrini *et al.*, 2015a
16S-R	GCTAATCAGACGC AAGCTCTTC (22)						1	Sandrini *et al.*, 2015a
cmpA-F	GTTAAACACCCAG GGTAACGGA (22)	*cmpA*	*IPF_2181*	AM778958.1	180	1.87±0.01	1	Sandrini *et al.*, 2015a
cmpA-R	GCTAACCAGTAAC GAATCCAGAAGT (25)						1	Sandrini *et al.*, 2015a
bicA-F1	CAAGCTAACGGTC GCATCAT (20)	*bicA*	*IPF_4911*	AM778949.1	132	1.89±0.01	1,2	Sandrini *et al.*, 2015a
bicA-R1	AGGCACATCACTCA AGTCCA (20)						1,2	Sandrini *et al.*, 2015a
sbtA-F1	CTGGCCTTTTTGA TTGGTGG (20)	*sbtA*	*MAE_62090*	AP009552.1	143	1.89±0.02	1,2	Sandrini *et al.*, 2015c
sbtA-R1	AGGTTGGAATTGC GGATGG (19)						1,2	Sandrini *et al.*, 2015c
bicA-F2	TCAAGACCCATCCT CACCA (19)	*bicA*	*IPF_4911*	AM778949.1	264	1.86±0.02	2	This study
sbtA-R2	CCACCAATCAAAAA GGCCAG (20)	*sbtA*	*MAE_62090*	AP009552.1			2	This study
chpX-F	CCTGTCAAGTCCT CCTCTCAT (21)	*chpX*	*IPF_1842*	AM778957.1	113	1.90±0.02	1	Sandrini *et al.*, 2015a
chpX-R	TTCAGGATACCCAC TACCTCG (21)						1	Sandrini *et al.*, 2015a
chpY-F	ATATCGCCAAAATG CCGACC (20)	*chpY*	*IPF_1545*	AM778958.1	114	1.80±0.01	1	Sandrini *et al.*, 2014
chpY-R	GACATCATCCGCA CCTGTTC (20)						1	Sandrini *et al.*, 2014
ccmR-F2	CCTACCGTCTCAA CCCAAGT (20)	*ccmR*	*IPF_1549*	AM778958.1	109	1.88±0.01	1	Sandrini *et al.*, 2014
ccmR-R	ACAGTAATTCCTGA CCCGCTT (21)						1	Sandrini *et al.*, 2014
ccmR2-F	TCCTTGGGATAAA CCACATACCA (23)	*ccmR2*	*IPF_2166*	AM778949.1	204	1.88±0.02	1	Sandrini *et al.*, 2014
ccmR2-R	TTTTCTCGACCAT GGCATCAC (21)						1	Sandrini *et al.*, 2014
rbcX-F	CGGATCATGACGG TAAGAGAACA (23)	*rbcX*	*IPF_2531*	AM778933.1	157	1.87±0.01	1,2	Sandrini *et al.*, 2015a
rbcX-R	ATTCCGATGTCTC TGGTTGACT (22)						1,2	Sandrini *et al.*, 2015a
ccmM-F	AAGTCCACACCTTC TCTAACCTC (23)	*ccmM*	*IPF_5695*	AM778933.1	118	1.88±0.01	1	Sandrini *et al.*, 2014
ccmM-R	CTGTCGTCGCCAA TGTGAA (19)						1	Sandrini *et al.*, 2014
ccaA1-F	ACTCCTGCGGTTA ATACTGTGG (22)	*ccaA*	*IPF_5538*	AM778919.1	97	1.89±0.02	1	Sandrini *et al.*, 2014
ccaA1-R	GATAAATGCGATCA GCTTGGGAG (23)						1	Sandrini *et al.*, 2014
mcyB-F	ATCCCATGCTCAGA GACGTT (20)	*mcyB*	*IPF_375*	AM778952.1	163	1.87±0.02	1,2	Sandrini *et al.*, 2014
mcyB-R	AGATGTCCGCAGG GATTCAT (20)						1,2	Sandrini *et al.*, 2014
gvpC-F	GTAATTGAGGACAA CCCCATGC (22)	*gvpC*	*MAE_37620*	AP009552.1	153	1.89±0.02	1	This study
gvpC-R	TGCCTGTTCTTGC GCTTG (18)						1	This study

Table S6.2 continued.

isiA-F	CTTTAGGCTTTGG AGTCGGAGA (22)	*isiA*	*IPF_5322*	AM778872.1	117	1.89±0.01	1	This study
isiA-R	GGTGAAATAAGGC TCCTGCTC (21)						1	This study
flv4-F	GATCCCCACGAAG TCAGAGA (20)	*flv4*	*IPF_2586*	AM778929.1	166	1.89±0.01	1	This study
flv4-R	GTTCATCTTCCCC ACCTCCT (20)						1	This study

Locus tags are based on the genomes of *Microcystis* strain PCC 7806 (IPF) and NIES-843 (MAE).
Amplification efficiencies of the different primer pairs are based on n = 27-86 amplification curves.

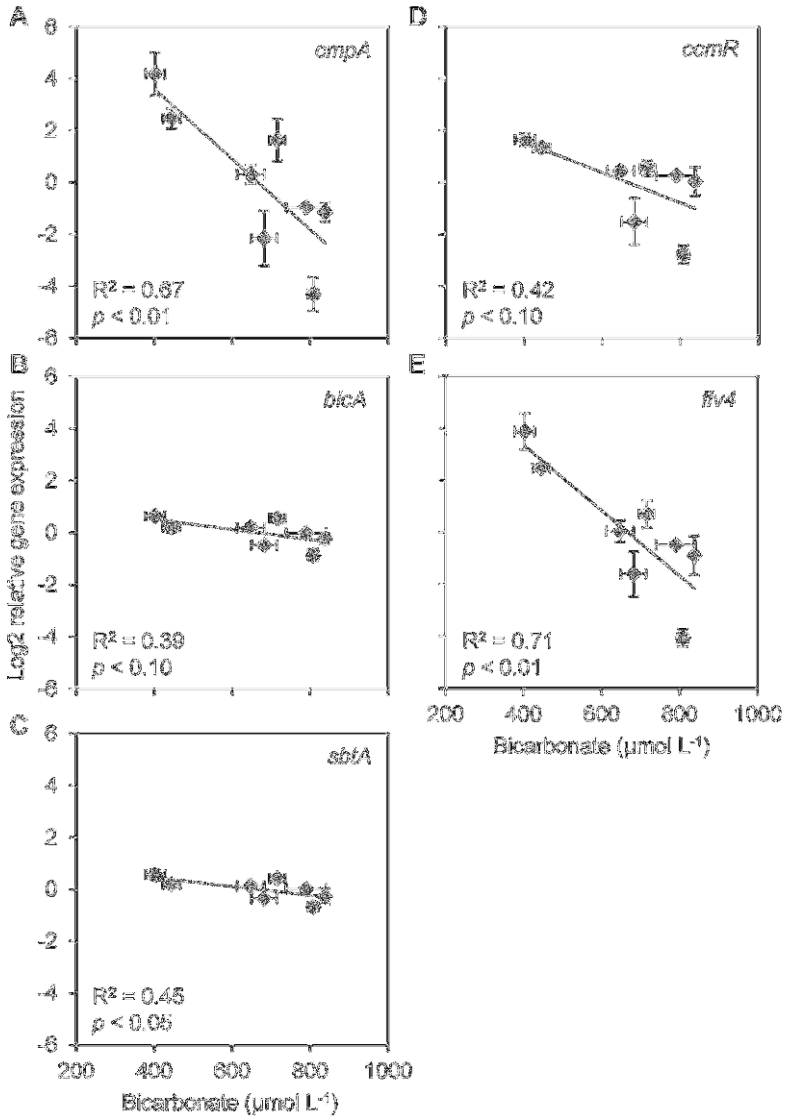

Figure S6.1. Scatter plots of relative gene expression versus bicarbonate concentration, for all genes that showed a significant or marginally significant relationship with bicarbonate. (A-C) Bicarbonate uptake genes *cmpA* **(A)**, *bicA* **(B)** and *sbtA* **(C)**. **(D)** CCM transcriptional regulator gene *ccmR*. **(E)** Photoprotection gene *flv4*. Each data point shows the mean ±s.d. of three independent measurements. The trend lines are based on linear regression (n = 8).

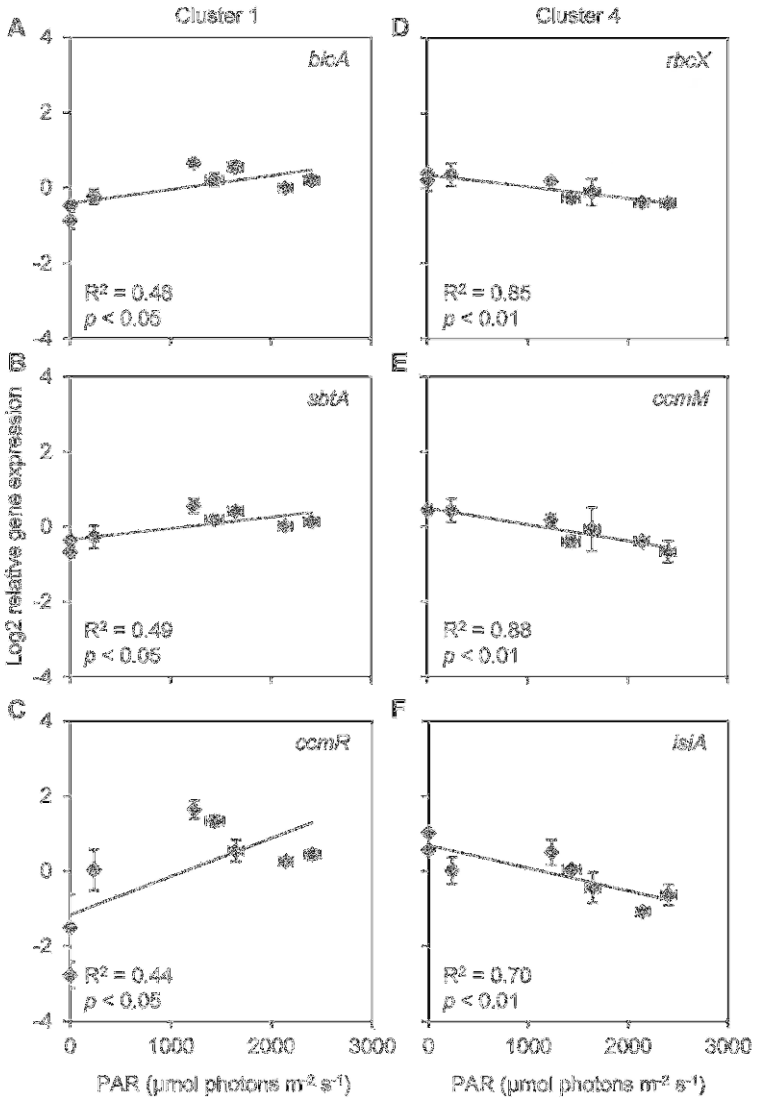

Figure S6.2. Scatter plots of relative gene expression versus incident light intensity, for all genes that showed a significant or marginally significant relationship with light. (A,B) Bicarbonate uptake genes **(A)** *bicA* and **(B)** *sbtA*. **(C)** CCM transcriptional regulator gene *ccmR*. **(D)** RuBisCO chaperone gene *rbcX*. **(E)** Carboxysomal gene *ccmM*. **(F)** Photoprotection gene *isiA*. Each data point shows the mean ±s.d. of three independent measurements. The trend lines are based on linear regression (n = 8).