



The LysR-type transcription factor PacR is a global regulator of photosynthetic carbon assimilation in *Anabaena*

| | |
|-------------------------------|---|
| Journal: | <i>Environmental Microbiology and Environmental Microbiology Reports</i> |
| Manuscript ID: | EMI-2014-1511.R1 |
| Manuscript Type: | EMI - Research article |
| Journal: | Environmental Microbiology |
| Date Submitted by the Author: | 22-Jan-2015 |
| Complete List of Authors: | Picossi, Silvia; Consejo Superior de Investigaciones Científicas, Instituto de Bioquímica Vegetal y Fotosíntesis Flores, Enrique; Consejo Superior de Investigaciones Científicas, Instituto de Bioquímica Vegetal y Fotosíntesis Herrero, Antonia; Consejo Superior de Investigaciones Científicas, Instituto de Bioquímica Vegetal y Fotosíntesis |
| Keywords: | bacteria, environmental signal/stress responses, gene expression/regulation |
| | |

SCHOLARONE™
Manuscripts

23 **Summary**

24 **Cyanobacteria perform water-splitting photosynthesis and are important primary**
25 **producers impacting the carbon and nitrogen cycles at global scale. They fix CO₂**
26 **through ribulose biphosphate carboxylase/oxygenase (RuBisCo) and have**
27 **evolved a distinct CO₂ concentrating mechanism (CCM) that builds high CO₂**
28 **concentrations in the vicinity of RuBisCo favoring its carboxylase activity.**
29 **Filamentous cyanobacteria such as *Anabaena* fix CO₂ in photosynthetic**
30 **vegetative cells, which donate photosynthate to heterocysts that rely on a**
31 **heterotrophic metabolism to fix N₂. CCM elements are induced in response to**
32 **inorganic carbon limitation, a cue that exposes the photosynthetic apparatus to**
33 **photodamage by over-reduction. An *Anabaena* mutant lacking the LysR-type**
34 **transcription factor All3953 grows poorly and dies under high light. The *rbcL***
35 **operon encoding RuBisCo is induced upon carbon limitation in the wild type but**
36 **not in the mutant. CHIP-Seq analysis was used to globally identify All3953 targets**
37 **under carbon limitation. Targets include, besides *rbcL*, genes encoding CCM**
38 **elements, photorespiratory pathway, photosystem- and electron transport-**
39 **related components, and factors, including flavodiiron proteins, with a**
40 **demonstrated or putative function in photoprotection. qRT-PCR analysis of**
41 **selected All3953 targets showed regulation in the wild type but not in the mutant.**
42 **All3953 (PacR) is a global regulator of carbon assimilation in an oxygenic**
43 **photoautotroph.**

44

45

46 Introduction

47 As the organisms that developed oxygenic photosynthesis, cyanobacteria have played
48 a crucial role in Earth's history and the evolution of life in our planet. Indeed, the
49 production of O₂ as a result of cyanobacterial activity was responsible for the oxidation
50 of the Earth's atmosphere about 2.5-2.3 billion years ago (Lyons *et al.*, 2014).
51 Furthermore, all the extant plastids of eukaryotic algae and plants are of cyanobacterial
52 origin. Cyanobacteria were the first organisms to link the activity of the two types of
53 photosystems (PSI and PSII), which allowed the generation of high electrochemical
54 potential, and to combine them with a H₂O-splitting complex. Nowadays, most
55 cyanobacteria are phototrophs relying on oxygenic photosynthesis to generate ATP
56 and reducing equivalents for the fixation of CO₂ and the assimilation of inorganic
57 nitrogen. Indeed, they are responsible for an important fraction of the primary
58 productivity in the Earth's oceans, where they are important CO₂ and N₂ fixers, thus
59 impacting the C and N cycles at a global scale (Knoll 2008; Price *et al.*, 2008).

60 The enzyme responsible for the bulk of CO₂ fixation in the biosphere is ribulose-
61 1,5-bisphosphate carboxylase/oxygenase (RuBisCo), which has a relatively low affinity
62 for CO₂ and, moreover, it can also accept O₂ as a substrate. Compensating for this
63 relatively low performance, RuBisCo is considered the most abundant enzyme on
64 Earth. As a carboxylase, RuBisCo catalyzes the first step of the Calvin-Benson-
65 Bassham (CBB) cycle, i.e., the incorporation of atmospheric CO₂ into ribulose-1,5-
66 bisphosphate to give two molecules of 3-phosphoglycerate. As an oxygenase, it
67 catalyzes the incorporation of O₂ into ribulose-bisphosphate, which produces 2-
68 phosphoglycolate that leads to photorespiration with a subsequent loss of fixed C and
69 energy. To increase the efficiency of CO₂ fixation, cyanobacteria have developed a
70 distinct CO₂ concentrating mechanism (CCM) constituted by inorganic carbon (C_i)
71 transporters that incorporate bicarbonate and CO₂ into the cell, and a proteinaceous
72 compartment, the carboxysome, where RuBisCo, together with carbonic anhydrase, is

73 confined (Price *et al.*, 2008; Cameron *et al.*, 2014). Three high-affinity bicarbonate
74 transporters (the ABC-type Cmp, also called BCT1, and the Na⁺-dependent BicA and
75 SbtA), and two CO₂ transporters (the high-affinity NDH-I₃ and the low-affinity NDH-I₄)
76 have been identified in different cyanobacteria (see Price, 2011). Through CCM, C_i in
77 the form of CO₂ can concentrate in the vicinity of cyanobacterial RuBisCo, to allow high
78 specific activity for production of 3-phosphoglycerate to levels much higher than in
79 plants (Cameron *et al.*, 2014). Indeed, components of cyanobacterial CCM have been
80 transformed in tobacco, with the result of improved specific activity of CO₂ fixation,
81 which represents a step towards improved photosynthesis in plants (Lin *et al.*, 2014).

82 In unicellular cyanobacteria, CCM elements are regulated by C_i availability.
83 Especially genes encoding C_i transporters are induced, whereas the structural genes
84 for carboxysome components and the *rbcl/S* genes encoding RuBisCo are only
85 moderately responsive (see Price *et al.*, 2008; Cameron *et al.*, 2014). In chemotrophic
86 bacteria the process of CO₂ fixation and the response to C_i limitation are usually
87 controlled by LysR-type transcriptional regulators (LTTRs). The genes encoding the
88 enzymes of the CBB cycle are usually found in clusters regulated by CbbR factors,
89 which constitute a sub-family of LTTRs (Gibson and Tabita, 1996). In unicellular
90 cyanobacteria, a number of CbbR homologs have been characterized. CmpR is an
91 activator of the *cmp* genes (Nishimura *et al.*, 2008), whereas CcmR (*aka* NdhR) acts as
92 a transcriptional repressor of multiple genes encoding other C_i transporters (e.g., Figge
93 *et al.*, 2001; Wang *et al.*, 2004). A third type of CbbR-like protein, the activator of the
94 RuBisCo genes, has not yet been identified in cyanobacteria.

95 Filamentous heterocyst-forming cyanobacteria, of which *Anabaena* sp. PCC
96 7120 (hereafter *Anabaena*) is a model organism, additionally have the capacity for cell
97 differentiation to turn some O₂-evolving photosynthetic cells of the filament into
98 heterocysts, which are heterotrophic cells specialized in the fixation of atmospheric
99 N₂. Thus, *Anabaena* is a truly pluricellular bacterium with different cell types specialized
100 in different nutritional tasks that exchange nutrients and regulators and contribute to

101 the performance of the filament as the organism unit (see Flores and Herrero, 2010).
102 Nitrogen assimilation and heterocyst differentiation in *Anabaena* are regulated by the
103 global transcription factor NtcA, which responds to the cellular C-to-N ratio, and the
104 heterocyst-specific transcription factor HetR (see Herrero *et al.*, 2013). The *Anabaena*
105 genomic sequence includes three genes annotated as CbbR-like LTTRs (Kaneko *et al.*,
106 2001). Of these, *all0862* has been identified as *cmpR*, and its product activates the
107 expression of the *cmp* operon (*alr2877-alr2880*) and the *cmpR* gene itself in response
108 to C_i limitation (López-Igual *et al.*, 2012). Notably, this regulation is effected in
109 combination with NtcA, thus revealing a mode of co-regulation by C and N availability
110 (López-Igual *et al.*, 2012). Furthermore, in *Anabaena* the *rbcLXS* operon encoding
111 RuBisCo, which is moderately induced under C_i limitation (López-Igual *et al.*, 2012), is
112 repressed in the heterocysts (Madan and Nierzwicki-Bauer, 1993), a regulation likely
113 exerted by NtcA (Ramasubramanian *et al.*, 1994; Picossi *et al.*, 2014).

114 Here we have identified the CbbR-homolog All3953 as the activator of the
115 RuBisCo-encoding operon in *Anabaena*. We have determined the All3953 regulon by
116 ChIP-Seq, which has revealed that All3953 is a global regulator for C_i assimilation
117 genes and genes involved in protection of the photosynthetic apparatus against
118 oxidative damage that are regulated by C_i availability.

119

120 **Results**

121 *All3953, an RbcR-like factor in Anabaena sp. PCC 7120*

122 To gain insight into the LTTR All3953 in *Anabaena*, the expression of the *all3953* gene
123 under different growth conditions was analyzed by northern blot. Three bands of
124 hybridization corresponding to transcripts of ca. 1.9, 1.6 and 1.3 kb, respectively, which
125 appeared similarly represented before and after C_i limitation could be observed (Supp.
126 Fig. 1A). On the other hand, *all3953* expression did not significantly respond to N
127 depletion (Supp. Fig. 1B). The latter result was consistent with previous global

128 transcriptional studies (Ehira and Ohmori, 2006; Flaherty *et al.*, 2011) and with the
129 lack of binding associated to *all3953* of the N-control transcriptional regulator NtcA
130 upon combined nitrogen depletion (Picossi *et al.*, 2014).

131

132 *Isolation and characterization of an all3953 mutant*

133 To study the role of *All3953*, a mutant strain bearing an inactivated version of *all3953*
134 was constructed (see Experimental procedures). In strain CSS74 most of the *all3953*
135 gene was deleted and the C.S3 gene cassette (encoding Sm/Sp resistance) was
136 introduced to facilitate segregation and maintenance of the mutation in *Anabaena*
137 (Supp. Fig. 2A). As a control, strain CSS77 bearing the C.S3 gene cassette in the
138 *Anabaena* plasmid alpha was also constructed. For *cis* complementation, a wild-type
139 version of *all3953* was transferred to strain CSS74 in an integrative plasmid,
140 generating strain CSS74C (Supp. Fig. 2A).

141 Strain CSS74 exhibited poor growth under standard growth conditions with
142 ammonium as a nitrogen source, and it formed short filaments in liquid medium,
143 whereas strain CSS74C behaved similarly to CSS77 (not shown). To quantify the
144 deleterious effect of the *all3953* mutation, growth rates were calculated in liquid
145 medium under different illumination and C_i-supply conditions. Growth rate of the control
146 strain CSS77 was highest (0.892 days⁻¹) under high light (HL) and high carbon (HC),
147 and was about 30% lower under the other tested conditions (Fig. 1A). In contrast,
148 growth of strain CSS74 was severely affected under HL HC conditions (growth rate ca.
149 70% lower than that of the control) (Fig. 1A), under which it ended up dying after about
150 5 days (Fig. 1B). Under HL and low carbon (LC) or low light (LL) HC, the defect was
151 close to 30% with regard to the control, although after prolonged incubation the mutant
152 was more severely affected under HL LC. The defect was the smallest (ca. 15%) under
153 LL LC conditions (Fig. 1A,B). In solid medium, growth of strain CSS74 was similar, and
154 similar to that of the control, in the presence of ammonium, nitrate or no combined

155 nitrogen under LL LC conditions, whereas under HL LC a severe growth defect was
156 observed with regard to the control with any of the tested nitrogen sources (not shown).
157 In summary, the lack of *All3953* had a deleterious effect on growth, especially under
158 HL (and HC) conditions.

159 To further characterize the growth defect of the *all3953* mutant, the rate of
160 oxygen evolution using CO₂ as a final electron acceptor was measured under different
161 illumination conditions in the CSS74 mutant in relation to the control strain CSS77.
162 When exponentially-growing cells were incubated for 24 h under LL HC, the oxygen
163 evolution rate was slightly lower in strain CSS74 in comparison to CSS77 (88 and
164 105 μmol O₂·[mg Chl]⁻¹·h⁻¹, respectively). Under HL the difference between the two
165 strains was larger (167, for CSS77, and 110, for CSS74, μmol O₂·[mg Chl]⁻¹·h⁻¹).

166

167 *Effect of all3953 mutation on rbcLXS expression*

168 To test the effect of the *all3953* mutation on the expression of the *rbcLXS* operon,
169 northern blot analysis was performed with RNA isolated from cells of the control and
170 mutant strains grown with HC and transferred to LC conditions. After 1 h incubation
171 with LC, a ca. 2-fold increase in the amount of the *rbcLXS* transcripts could be
172 observed in the control strain. No induction could be detected in the mutant (Fig. 2). In
173 the complemented strain CSS74C the expression of *rbcLXS* increased in LC similarly
174 to the control (Fig. 2). These results indicated that the induction of the *rbcLXS* operon
175 upon C_i deficiency was dependent on *All3953* and that the defect in strain CSS74 was
176 exclusively due to the lack of *All3953*.

177

178 *ChIP-Seq analysis of the All3953 targets*

179 To determine the DNA targets of *All3953* at a genomic level, we used chromatin
180 immunoprecipitation followed by high-throughput sequencing analysis. To this end, we
181 constructed a strain (CSS57) expressing from the *all3953* promoter a version of

182 All3953 C-terminally fused to TAP-tag (Rigaut *et al.*, 1999), as well as a control strain
183 (CSS68) expressing the TAP-tag alone under the control of the *all3953* promoter
184 (Supp. Fig. 2B, see Experimental procedures). Immunoprecipitation was carried out
185 using cells of strains CSS57 and CSS68 grown with ammonium as the nitrogen source
186 under HC conditions and incubated for 3 h with ammonium under LC conditions.

187 The analysis of the sequences resulted in a total of 142 All3953 binding
188 regions, of which 127 were located in the chromosome, 10 in plasmid alpha, three in
189 plasmid beta and two in plasmid gamma. Each binding region was ascribed to one or
190 two genes according to the location (midpoint) of the region, and the relative location
191 with respect to the assigned gene was also indicated (Table 1 and Supp. Table 1). A
192 total of 175 genes were ascribed to the 142 binding regions. The binding regions were
193 mostly located upstream of the ascribed genes (72%), whereas 21% were intragenic
194 and 7% were located downstream of genes. The results of the ChIP-Seq analysis are
195 available at GEO accession number GSE58861
196 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE58861>).

197 The 175 ascribed genes were classified according to their functional category
198 (Table 2). Remarkably, there were 19 genes encoding proteins related to
199 photosynthesis and respiration, and 21 genes encoding regulatory proteins, including
200 All7179, a SigB homolog. The rest were mostly genes encoding hypothetical or
201 unknown proteins (42%), but also genes encoding proteins involved in translation, in
202 biosynthesis of amino acids and cofactors, prosthetic groups and carriers, in transport,
203 and in other cellular processes. Table 3 highlights All3953-binding regions related to
204 photosynthesis and respiration among which, confirming our results of gene
205 inactivation, the gene encoding the large subunit of the RuBisCo (*rbcL*; binding region
206 #37) is included. The fact that a high number of genes involved in photosynthesis and
207 C fixation, including *rbcL*, were identified as targets of RbcR suggests that this protein
208 is a global transcription factor for photosynthetic C assimilation.

209 A CisFinder analysis of the primary consensus motif was carried out based on
210 142 high-confidence ChIP-Seq peak sequences (Fig. 3A). The consensus motif found
211 has a dyad-symmetry architecture and matches the consensus of LysR-recognition
212 binding sites (RBS) (T-N₁₁-A) (Maddocks and Oyston, 2008), as well as the consensus
213 binding sites proposed for CbbR factors (TNA-N_{7/8}-TNA). The primary motifs identified
214 in the central 100 nt of the binding regions are indicated in Supp. Table 1 (for some
215 binding regions more than one motif have been identified).

216

217 *Expression analysis of some All3953 targets in Anabaena*

218 To corroborate our ChIP-Seq analysis and to support the notion that All3953 is indeed
219 a global regulator for C fixation genes, the expression levels of some of the
220 photosynthesis- and C-fixation-related target genes, its response to C_i limitation and its
221 dependence on All3953 was further analyzed by qRT-PCR (Fig. 4). Strains CSS77 and
222 CSS74 were grown in ammonium and HC under standard light conditions (80
223 $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) at 30°C up to the exponential phase. They were then transferred to medium
224 with ammonium under LC conditions. As previously shown by northern analysis,
225 transcription levels of the *all3953* gene did not significantly change after C_i deprivation
226 in the control strain CSS77. As expected, *all3953* transcript levels were not detectable
227 in CSS74, corroborating that the *all3953* mutation was segregated in this strain. The
228 *rbcL* gene (*alr1524*) was 4.6-fold induced at 3 h after transfer to LC in strain CSS77,
229 whereas no induction was observed in CSS74, thus corroborating the dependence on
230 All3953. Interestingly, under HC the expression of *rbcL* in the mutant was higher than
231 in the control strain, suggesting that besides as an activator under LC, All3953 could
232 be acting as a repressor of *rbcL* under HC conditions.

233 ORF *all4446* (*flv4*) was highly induced (about 100-fold), and its level was
234 maximum 1 h after the shift to LC (Fig. 4). In the *all3953* mutant, the basal transcription
235 of *all4446* in HC was about 8-fold lower than in the control strain, and it was only

236 slightly induced (2-fold) upon the shift to LC. *all3891* (*flv1A*) was induced about 6-fold 3
237 h after the shift to LC in the control strain, whereas only a 2-fold induction was
238 observed in the *all3953* mutant. *all1304* (bicarbonate transporter homolog) and *alr4156*
239 (NdhF homolog) were both highly induced (up to 30- and 20-fold, respectively) upon
240 transfer of strain CSS77 to LC. In contrast, no induction of *alr4156* and only a small
241 induction of *all1304* took place in strain CSS74. Expression of *alr4592* (*psbAIII*)
242 increased about 5-fold upon C_i deprivation in CSS77, but did not appreciably change in
243 CSS74. The *alr0223* (NdhA homolog) gene was about 2-fold induced under C_i
244 deprivation in CSS77 but not in CSS74. Finally, the expression of *alr1004* (alanine-
245 glyoxylate aminotransferase) was repressed by 5-fold under LC conditions in CSS77
246 but not in CSS74. These results confirm that expression of the above studied genes is
247 regulated, either positively or negatively, by All3953.

248

249 Discussion

250 We have identified the LTTR All3953 as the activator of the RuBisCo-encoding genes
251 in the cyanobacterium *Anabaena* sp. PCC 7120. All3953 appears to activate the *rbcL*
252 operon under C_i limitation and to repress it when C_i is abundant. An LTTR factor
253 regulating the expression of the *rbcL* operon has not, to our knowledge, been
254 described in any cyanobacterium. The expression of *all3953* does not respond to C_i
255 limitation (Supp. Fig. 1) and, indeed, no binding region of All3953 was found ascribed
256 to *all3953*. Thus, the *all3953* gene seems to belong to the non-autoregulated LTTRs.
257 All3953 shares 28 and 26% identical residues with NdhR from *Synechocystis* sp. PCC
258 6803 and *Synechococcus* sp. PCC 7002 (CcmR), respectively.

259 By CHIP-Seq analysis of a strain bearing a TAP-tagged version of All3953, we
260 have determined 142 All3953-bound regions at 3 h after transfer from high to low C_i
261 conditions, which have been assigned to 177 genes. Apart from genes encoding
262 unknown or hypothetical proteins, the larger category is of genes encoding regulatory

263 proteins, including the transcriptional regulators Alr0353 (a LTTR) and All4500 (CRP-
264 like), the two-component response regulator All3348, the two-component sensor
265 histidine-kinase All1145 and the group 2-sigma 70-type sigma factor All7179 (Supp.
266 Table 1). Interestingly, when comparing to *Synechocystis* sigma factors, All7179
267 (SigB4) is more similar to SigB, which along with SigD has been shown to be important
268 for PSII recovery in this unicellular cyanobacterium (Pollari *et al.*, 2009). All3953 also
269 binds upstream of genes *patS* and *hetN*, whose products regulate heterocyst
270 differentiation (Supp. Table 1). The fact that All3953 binds to the promoter region of
271 genes encoding other regulatory proteins suggests a wide role of this protein in the
272 physiology of the organism.

273 In the promoter region of the *rbcL* operon we have found three putative binding
274 sites for All3953, Box I, Box II, and Box III (Fig. 3B) that resemble the consensus
275 recognition sequence found by Cisfinder analysis (Fig. 3A). It is conceivable that, like in
276 some other LTTRs these boxes combine repression and activation sites. In this regard,
277 binding of All3953 to Box III, overlapping gene promoter elements, could be related to
278 the repression of *rbcL* observed under high C_i (Fig. 4). On the other hand, as
279 mentioned above, the *rbcL* operon is repressed in heterocysts by the global
280 transcriptional regulator NtcA, for which a binding site is found overlapping the operon
281 TSP (Ramasubramanian *et al.*, 1994; Picossi *et al.*, 2014) (Figure 3B). It is conceivable
282 that NtcA binding in these differentiated cells interfere with All3952-mediated activation.

283 Besides being the activator of the RuBisCo genes, All3953 targets include
284 genes involved in other processes related to carbon assimilation, such as C_i transport
285 (*all1304*, encoding an homolog of the BicA bicarbonate transporter; *alr4156* encoding a
286 homolog of the NdhF3 subunit of the NDH-1₃ CO₂ uptake system; and *alr0869*
287 encoding a homolog of the NdhF4 subunit of the NDH-1₄ CO₂ uptake system);
288 components of the carboxysome shell (*all0868*, putative *ccmK*) and 2-
289 phosphoglycolate metabolism (*alr1004* and *alr2873*, possibly related to
290 photorespiration [Eisenhut *et al.*, 2008]). Notably, All3953 targets include also a

291 number of genes encoding photosystem components, such as *alr5154* (*psaA*,
292 encoding the PSI core protein PsaA), *alr3727* (*psbAII*, encoding a component of form II
293 of the PSII core protein PsbA [D1]), *alr4592* (*psbAIII*, encoding another component of
294 form II of PsbA) and *alr1216* (PSII 12 kD extrinsic protein PsbU), and genes related to
295 PS activity. In the latter group are *alr4149* (biliverdin reductase, putatively involved in
296 phycobilisome -PSII antenna- synthesis), and genes that can participate in
297 photosynthetic electron transfer, such as *alr0223* and *alr0348* (*ndhA* and *ndhD*,
298 subunits of other putative NADH dehydrogenases), *alr1576* (dehydrogenase subunit),
299 *all0737* (thioredoxin reductase), *all1365* (CytM cytochrome), *all4148* (ferredoxin I),
300 *all3891* and *all4446* (flavodiiron proteins Flv1A and Flv4, respectively). (Besides in CO₂
301 uptake, NdhF3 and NdhF4 can also participate in electron transfer.)

302 Reports on gene expression regulation by C_i are scarce for *Anabaena*.
303 However, in the unicellular cyanobacterium *Synechocystis* sp. PCC 6803
304 transcriptomic analysis has already shown down-regulation of some genes encoding
305 polypeptides of PSI and PSII complexes as well as of phycobilisome components,
306 upon transfer to C_i limitation conditions, likely as an adaptation to lower assimilatory
307 power demand, and up-regulation of some PSII core polypeptides, interpreted as
308 adaptation to conditions of shortage of electron acceptors that could lead to
309 photodamage and increased turnover of PS core components (Wang *et al.*, 2004). Our
310 analysis extends the array of photosynthetic genes responding to C_i limitation,
311 remarkably to include the core PSI reaction center *psaA* gene. Moreover, except for
312 *alr1004* that responds negatively, all the *Anabaena* photosynthetic genes mentioned
313 above increase expression upon the shift to C_i limitation.

314 Noteworthy, for the majority of photosynthetic gene targets of All3953 in
315 *Anabaena*, a function in protection against reactive oxygen species, which can be
316 generated by C_i limitation or exposure to HL, has either been described or could be
317 predicted. Thus, the *psbAII* and *psbAIII* genes are induced under HL in the unicellular
318 *Synechococcus* sp. PCC 7942, and cells cultured under HL showed more form II, and

319 less form I (encoded by *psbAI*), of D1 compared to cells under LL (Schaefer and
320 Golden, 1989). Regarding flavodiiron proteins genes, *flv1A*, *flv3A*, and specially *flv2*
321 and *flv4* of *Anabaena* have been shown up-regulated in vegetative cells in low C_i , and
322 *flv1A* and *flv3A* also in high light, whereas *flv1B* and *flv3B* are expressed exclusively in
323 heterocysts (Ermakova *et al.*, 2013). Whereas Flv1A and Flv3A appear involved in
324 photoreduction of oxygen to water by removing excess electrons from PSI through
325 NAD(P)H dehydrogenases, Flv2 and Flv4 could have a role in photoprotection of PSII
326 under low C_i (Ermakova *et al.*, 2013). Regarding photorespiration, it has also been
327 considered to have a role in removal of excess O_2 (Eisenhut *et al.*, 2008). To the best
328 of our knowledge, the regulator responsible for the response to C_i availability of any
329 photosynthetic gene has not been identified in cyanobacteria (oxygenic phototrophs).
330 Our ChIP-Seq and expression analysis indicate that All3953 is a regulator of
331 photosynthetic genes in *Anabaena*.

332 The growth rate of *Anabaena* is highest under HL HC conditions (Fig. 1),
333 implying that this cyanobacterium has mechanisms to get profit of HC while
334 counteracting HL stress. The *all3953* mutant strain CSS74 exhibited a growth defect in
335 all the conditions tested, but especially under HL, where it ends-up dying (Fig. 1). This
336 shares the idea that in *Anabaena* All3953 is required to cope with HL stress. The effect
337 of the lack of All3953 seems more detrimental in relation to impaired photoprotection
338 than to impaired C_i scavenging (preference of the mutant for LL LC over HL HC
339 conditions). Indeed, even in LL, the mutant seems to perform slightly better with LC
340 than with HC (Fig. 1). Although LC could suppose a limitation of electron acceptors, an
341 increased rate of photorespiration under LC together with the fact that some of the
342 All3953 targets that could cope with excess oxygen are residually induced upon the
343 transfer from HC to LC in the CSS74 mutant (Fig. 4) could contribute to the preference
344 of this strain for LC over HC, especially under HL.

345 Our results show that in *Anabaena* the responses to C_i availability include
346 regulation of genes encoding elements of CCM and RuBisCo, but also of

347 photosynthetic genes to adjust generation of assimilatory power while preserving the
348 photosynthetic apparatus from oxidative damage, which is specially relevant in
349 oxygenic phototrophs. Because All3953 is a transcriptional regulator globally
350 coordinating these responses, we have named it PacR (Photosynthetic assimilation of
351 carbon Regulator).

352

353 **Experimental procedures**

354 *Strains*

355 *Anabaena* sp. strain PCC 7120 was grown photoautotrophically at 30°C with illumination
356 ($80 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) in liquid BG11₀ medium (Rippka *et al.*, 1979) supplemented with 3 mM
357 NH_4Cl , 6 mM TES buffer and 10 mM NaHCO_3 and bubbled with a mixture of CO_2 and
358 air (1% v/v) (HC). Other conditions used were no NaHCO_3 supplement and bubbling
359 with air (LC); $12 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ (LL); $175 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ (HL). For the mutants generated in this
360 work, antibiotics were used at the following concentrations: Sm, $2 \mu\text{g ml}^{-1}$; Sp,
361 $2 \mu\text{g ml}^{-1}$; and Nm, $25 \mu\text{g ml}^{-1}$ for bubbled cultures; and Sm, $5 \mu\text{g ml}^{-1}$; Sp, $5 \mu\text{g ml}^{-1}$;
362 and Nm, $40 \mu\text{g ml}^{-1}$ for cultures in solid medium.

363

364 *Strain construction*

365 To construct a mutant of the *all3953* gene, the 5' and 3' end of the gene, along with the
366 flanking regions, were PCR amplified using chromosomal DNA of PCC 7120 as the
367 template and primers all3953-14 (BglII) and all3953-15 (Sall), and primers all3953-16
368 (Sall) and all3953-17 (PstI), respectively (all primers are specified in Supp. Table 2).

369 The PCR products were digested with Sall and ligated. The resulting mixture was used
370 as a template for overlapping PCR with primers all3953-14 and all3953-17. The new
371 PCR product was digested with BglII and PstI and ligated to the BglII-PstI-digested
372 pRL271 (Black *et al.*, 1993), obtaining plasmid pCSS161. Plasmid pCSS161 was
373 digested with SpeI and the 2-kb Sm^r Sp^r gene cassette C.S3, excised with XbaI from

374 pCSE120 [S.K3/L.HEH2 (BamHI)/C.S3 (BamHI); nomenclature as in (Elhai and Wolk,
375 1988)], was inserted obtaining plasmid pCSS162. Plasmid pCSS162 was transferred to
376 strain PCC 7120 by conjugation (Elhai *et al.*, 1997). Exconjugants resistant to Sm and
377 Sp, which had the $\Delta all3953::C.S3$ construct integrated by double recombination were
378 selected, obtaining strain CSS74. The segregation of the mutation was tested by PCR
379 (Supp. Fig. 2) with primers all3953-14, all3953-15, all3953-17 and all3953-20.

380 To construct a control strain expressing Sm^r and Sp^r plasmid pCSS163, a
381 derivative of plasmid pCSEL24 (Olmedo-Verd *et al.*, 2006) containing the C.S3 gene
382 cassette, was transferred to *Anabaena* by conjugation. Exconjugants that had the
383 pCSS163 integrated in the alpha plasmid of *Anabaena* were selected, obtaining the
384 strain CSS77 (Supp. Fig. 2).

385 To complement the *all3953* mutation of strain CSS74, a DNA fragment
386 encompassing the whole *all3953* gene and sequences upstream from it was amplified
387 by PCR using the primer pair all3953-24/all3953-25, both including EcoRI sites, and
388 strain PCC 7120 DNA as the template. This fragment was cloned in the EcoRI site of
389 the mobilizable Nm^r encoding vector pRL424 (Elhai and Wolk, 1988) producing plasmid
390 pCSS164, which was transferred to strain CSS74 by conjugation followed by selection
391 for Nm^r. The genomic structure of the exconjugants in the *all3953* region (Supp. Fig. 2)
392 was corroborated by PCR.

393 To construct a strain expressing All3953-C-TAP, the *all3953* gene (including the
394 upstream region) was PCR-amplified with primers all3953-11 and all3953-12 and DNA
395 of PCC 7120 as the template. The TAP-tag was PCR-amplified with primers TAPtag-1
396 and TAPtag-2 using DNA of plasmid pBS1479 as the template (Puig *et al.*, 2001). The
397 two PCR products were digested with Sall and ligated, after which the ligation product
398 was used as the template for an overlapping PCR using primers all3953-11 and
399 TAPtag-2. The PCR product was digested with PstI and ligated to the mobilizable
400 vector pCSV3 (Valladares *et al.*, 2011) digested with PstI, rendering plasmid pCSS107.

401 To construct a control strain with the TAP-tag under the control of the *all3953*
402 promoter, a 0.4-kb region upstream of *all3953* was PCR-amplified using primers
403 all3953-11 and all3953-18 and DNA of pCSS107 as the template. The PCR product
404 was digested with Sall and ligated to the PCR-amplified TAP-tag digested with Sall,
405 after which the ligation product was used as the template for an overlapping PCR using
406 primers all3953-11 and TAPtag-2. The PCR product was digested with PstI and ligated
407 to PstI-digested pCSV3 to give plasmid pCSS157. Plasmids pCSS107 and pCSS157
408 were transferred by conjugation to strain PCC 7120 and single Sm^r Sp^r recombinants
409 were selected, obtaining strain CSS57 and CSS68, respectively. Western blots using
410 Peroxidase-Anti-Peroxidase Soluble Complex (PAP, Sigma-Aldrich) were performed to
411 ensure that the two strains expressed the TAP-tag (Supp. Fig. 2B).

412

413 *Chromatin immunoprecipitation*

414 Cells of strains CSS57 growing exponentially (3-5 $\mu\text{g Chl}\cdot\text{ml}^{-1}$) in the light (80
415 $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) in medium supplemented with 2 $\mu\text{g}\cdot\text{ml}^{-1}$ Sm and Sp, in HC conditions were
416 incubated with LC for 3 h. Formaldehyde was then added to the cultures to a final
417 concentration of 1%, and the cultures were incubated for 15 min. Glycine was added at
418 125 mM final concentration and the incubation was continued for 5 min to stop the
419 fixing reaction. The cells were then filtered, washed with cold TBS (20 mM Tris-HCl, pH
420 7.4, 140 mM NaCl) and collected in tubes (25 ml of culture per tube). The pellets were
421 frozen in liquid nitrogen and stored at -20°C until used. Pellets corresponding to about
422 25 ml of culture were resuspended in 500 μl of lysis buffer (50 mM HEPES-KOH, pH
423 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate,
424 supplemented with Mini EDTA-free protease inhibitor cocktail [Roche]) and, after
425 addition of 150 μl of glass beads (acid-washed, 425-600 μm [Sigma]), the cells were
426 broken in a multivortexer at 2,000 rpm for 1 h at 4°C. The cell lysates were collected by
427 centrifugation and the extracts were subjected to sonication to shear the DNA to about
428 300-bp fragments (60 cycles of 10 s, 20 s ice, 15% amplitude, in a Branson Digital

429 Sonifier). After centrifugation to eliminate cell debris, the whole-cell extracts were
430 stored at -20°C or immediately used for immunoprecipitation.

431 Immunoprecipitation of DNA was carried out as described (Picossi *et al.*, 2014),
432 with some modifications. Whole-cell extracts were prepared at $4\text{ mg}\cdot\text{ml}^{-1}$ of total protein
433 with lysis buffer (in $500\ \mu\text{l}$ total volume). A $50\text{-}\mu\text{l}$ sample was taken as the input sample,
434 and the extracts were incubated with $15\ \mu\text{l}$ IgG-conjugated Dynabeads (about $6\ \mu\text{g}$
435 IgG) at 4°C with rotation for 12-14h. The washes of the Dynabeads, as well as the
436 elution of the immunoprecipitated material, the crosslinking reversion and the isolation
437 of the DNA were performed as in (Picossi *et al.*, 2014).

438

439 *Massive sequencing of the immunoprecipitated DNA*

440 Input and ChIP DNA samples were sent for sequencing to the Functional Genomics
441 Core Facility of the Institute for Research in Biomedicine, Barcelona (Spain). Next
442 generation sequencing was carried out using Illumina's sequencing technology. ChIP
443 DNA Sample Prep Kit (Illumina) was used for library preparation. Libraries were loaded
444 at $8\ \text{pM}$ concentration into the flow cell using the Cluster Station running recipe V7 with
445 the Single-Read Cluster Generation Kit v4 (all Illumina). The flow cell was loaded into
446 the Genome Analyzer II and samples were sequenced for 120 nucleotides from a
447 single end using the Sequencing Kit v5 and recipe v8 (all Illumina). Manufacturer's
448 recommendations were strictly followed. Illumina sequencing data were pre-processed
449 with the standard Illumina pipeline version 1.5 and sequences were aligned to the PCC
450 7120 genome (<http://genome.microbedb.jp/cyanobase/Anabaena>) with the Bowtie
451 software 0.12.5 (Langmead *et al.*, 2009). The analysis of the results were carried out
452 using the Triform algorithm (Kornacker *et al.*, 2012) as in (Picossi *et al.*, 2014). The
453 sequences in the ChIP samples of strain CSS68 (TAP control) were used as the
454 background of the sequences found for strain CSS57 (All3953-TAP), and thus to
455 determine the specific binding regions of All3953-TAP in the genome of *Anabaena*.

456 The binding regions were visualized and analyzed using the UCSC Microbial Genome
457 Browser (Schneider *et al.*, 2006). They were ascribed to one or two genes, in case it
458 was not possible to ascribe them to only one, and classified as upstream of the gene, if
459 the midpoint of the binding region was located upstream of the start of the gene,
460 internal, if the midpoint of the binding region was inside the gene, or downstream, if the
461 midpoint of the binding region was located downstream of the end of the gene to which
462 it had been ascribed.

463

464 *Northern and qRT-PCR analyses*

465 Isolation of total RNA from *Anabaena* was done as described previously (Mohamed
466 and Jansson, 1989). Northern analysis was performed as described previously (López-
467 Igual *et al.*, 2012).

468 For qRT-PCR, 750 ng of DNA-free RNA samples were used for all the PCR
469 primer pairs. For the RT reaction, the Quantitech Reverse transcription kit (Qiagen),
470 with the Random Hexamer Primer mix (100 ng per sample) (Bioline) was used. The
471 cDNA produced was diluted 7.5 times to use 2 μ l of cDNA per PCR reaction. PCR was
472 done using the Quantimix Easy SYG Kit (Biotools) (SYBR green I) in a iCycler iQ Multi-
473 Color Real Time PCR Detection System (Bio-Rad). The abundance of a transcript in
474 the RNA sample was calculated as: $\text{abundance} = 2^{[\text{Ct}(\text{sample}) - \text{Ct}(\text{control})]}$, where the
475 RNA sample of the control strain CSS77 in HC condition (0) was used as the control.

476

477 *Oxygen evolution*

478 2-ml samples of exponentially grown cultures in HC LL or HC HL conditions were used
479 to measure O_2 evolution with an O_2 electrode calibrated with culture medium and
480 $\text{Na}_2\text{S}_2\text{O}_4$ as the reducing agent. O_2 production was measured in the light (400
481 $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) after a seven-minute incubation in the dark.

482

483 **Acknowledgments**

484 This work was supported by grants BFU2010-17980 and BFU2013-44686-P from the
 485 Spanish government, co-financed by FEDER. The authors are grateful to K. Kornacker
 486 for carrying out the Triform and CisFinder analyses, and to H. Auer and I. Pons, from
 487 the Functional Genomics Core Facility of the IRB, Barcelona (Spain); to F. Monje-
 488 Casas and Yagut Allahverdiyeva-Rinne for critical reading of the manuscript and to M.
 489 J. Huertas and A. Torrado for help with O₂ production measurements.

490

491 **REFERENCES**

492

- 493 Black, T.A., Cai, Y., and Wolk C.P. (1993) Spatial expression and autoregulation of
 494 *hetR*, a gene involved in the control of heterocyst development in *Anabaena*. *Mol*
 495 *Microbiol* **9**: 77-84.
- 496 Cameron, C.J., Sutter, M., and Kerfeld, C.A. (2014) The carboxysome: Function,
 497 structure and cellular dynamics. *The Cell Biology of Cyanobacteria*, eds. Flores E,
 498 Herrero A (Caister Academic Press, Norfolk), pp. 171-188.
- 499 Ehira, S., and Ohmori, M, (2006) NrrA, a nitrogen-responsive response regulator
 500 facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC
 501 7120. *Mol Microbiol* **59**: 1692-1703.
- 502 Elhai, J., Veprikskiy, A., Muro-Pastor, A.M., Flores, E, and Wolk, C.P. (1997) Reduction
 503 of conjugal transfer efficiency by three restriction activities of *Anabaena* sp. strain
 504 PCC 7120. *J Bacteriol* **179**: 1998-2005.
- 505 Eisenhut, M., Ruth, W., Haimovich, M., Bauwe, H., Kaplan, A., and Hagemann, M.
 506 (2008) The photorespiratory glycolate metabolism is essential for cyanobacteria and
 507 might have been conveyed endosymbiotically to plants. *Proc Natl Acad Sci USA*
 508 **105**: 17199-17204.
- 509 Elhai, J., and Wolk, C.P. (1988) A versatile class of positive-selection vectors based on
 510 the nonviability of palindrome-containing plasmids that allows cloning into long
 511 polylinkers. *Gene* **68**: 119-138.
- 512 Ermakova, M., Battchikova, N., Allahverdiyeva, Y., and Aro, E.M. (2013) Novel
 513 heterocyst-specific flavodiiron proteins in *Anabaena* sp. PCC 7120. *FEBS Lett* **587**:
 514 82-7.
- 515 Figge, R.M., Cassier-Chauvat, C., Chauvat, F., and Cerff, R. (2001) Characterization
 516 and analysis of an NAD(P)H dehydrogenase transcriptional regulator critical for the
 517 survival of cyanobacteria facing inorganic carbon starvation and osmotic stress. *Mol*
 518 *Microbiol* **39**: 455-68.
- 519 Flaherty, B.L., Van Nieuwerburgh, F., Head, S.R., and Golden, J.W. (2011) Directional
 520 RNA deep sequencing sheds new light on the transcriptional response of *Anabaena*
 521 sp. strain PCC 7120 to combined-nitrogen deprivation. *BMC Genomics* **12**: 332.
- 522 Flores, E., and Herrero, A. (2010) Compartmentalized function through cell
 523 differentiation in filamentous cyanobacteria. *Nat Rev Microbiol* **8**: 39-50.
- 524 Gibson, J.L., and Tabita, F.R. (1996) The molecular regulation of the pentose
 525 phosphate pathway in proteobacteria and cyanobacteria. *Arch Microbiol* **166**: 141-
 526 150.
- 527 Herrero, A., Picossi, S., and Flores, E. (2013) Gene Expression during heterocyst
 528 differentiation. *Adv Bot Res* **65**: 281-329.

- 529 Kaneko, T., Nakamura, Y., Wolk, C.P., Kuritz, T., Sasamoto, S., and Watanabe, A.
530 (2001) Complete genomic sequence of the filamentous nitrogen-fixing
531 cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res* **8**: 205-213.
- 532 Knoll, A.H. (2008) Cyanobacteria and Earth history. *The Cyanobacteria: Molecular*
533 *Biology, Genomics and Evolution*, eds. Herrero A, Flores E (Caister Academic
534 Press, Norfolk), pp. 1-19.
- 535 Kornacker, K., Rye, M.B., Handstad, T., and Drablos, F. (2012) The Triform algorithm:
536 improved sensitivity and specificity in ChIP-Seq peak finding. *BMC Bioinformatics*
537 **13**: 176.
- 538 Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009) Ultrafast and memory-
539 efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**:
540 R25.
- 541 Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A.J., and Hanson, M.R. (2014) A
542 faster Rubisco with potential to increase photosynthesis in crops. *Nature* **513**: 547-
543 550.
- 544 López-Igual, R., Picossi, S., López-Garrido, J., Flores, E., and Herrero, A. (2012) N
545 and C control of ABC-type bicarbonate transporter Cmp and its LysR-type
546 transcriptional regulator CmpR in a heterocyst-forming cyanobacterium, *Anabaena*
547 sp. *Environ Microbiol* **14**: 1035-1048.
- 548 Lyons, T.W., Reinhard, C.T., and Planavsky, N.J. (2014) The rise of oxygen in Earth's
549 early ocean and atmosphere. *Nature* **506**: 307-315.
- 550 Madan, A.P., and Nierzwicki-Bauer, S.A. (1993) In situ detection of transcripts for
551 ribulose-1,5-bisphosphate carboxylase in cyanobacterial heterocysts. *J Bacteriol*
552 **175**: 7301-7306.
- 553 Maddocks, S.E., and Oyston, P.C. (2008) Structure and function of the LysR-type
554 transcriptional regulator (LTTR) family proteins. *Microbiology* **154**: 3609-3623.
- 555 Mohamed, A., and Jansson, C. (1989) Influence of light on accumulation of
556 photosynthesis-specific transcripts in the cyanobacterium *Synechocystis* 6803. *Plant*
557 *Mol Biol* **13**: 693-700.
- 558 Nierzwicki-Bauer, S.A., Curtis, S.E., and Haselkorn, R. (1984) Cotranscription of genes
559 encoding the small and large subunits of ribulose-1,5-bisphosphate carboxylase in
560 the cyanobacterium *Anabaena* 7120. *Proc Natl Acad Sci USA* **81**: 5961-5965.
- 561 Nishimura, T., Takahashi, Y., Yamaguchi, O., Suzuki, H., Maeda, S., and Omata, T.
562 (2008) Mechanism of low CO₂-induced activation of the *cmp* bicarbonate transporter
563 operon by a LysR family protein in the cyanobacterium *Synechococcus elongatus*
564 strain PCC 7942. *Mol Microbiol* **68**: 98-109.
- 565 Olmedo-Verd, E., Muro-Pastor, A.M., Flores, E., and Herrero, A. (2006) Localized
566 induction of the *ntcA* regulatory gene in developing heterocysts of *Anabaena* sp.
567 strain PCC 7120. *J Bacteriol* **188**: 6694-6649.
- 568 Picossi, S., Flores, E. and Herrero, A. (2014) ChIP analysis unravels an exceptionally
569 wide distribution of DNA binding sites for the NtcA transcription factor in a
570 heterocyst-forming cyanobacterium. *BMC Genomics* **15**: 22.
- 571 Pollari, M., Ruotsalainen, V., Rantamaki, S., Tyystjarvi, E., and Tyystjarvi, T. (2009).
572 Simultaneous inactivation of sigma factors B and D interferes with light acclimation
573 of the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J Bacteriol* **191**: 3992-
574 4001.
- 575 Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm,
576 M., and Seraphin, B. (2001) The tandem affinity purification (TAP) method: a
577 general procedure of protein complex purification. *Methods* **24**: 218-229.
- 578 Price, G.D., Badger, M.R., Woodger, F.J., and Long, B.M. (2008). Advances in
579 understanding the cyanobacterial CO₂-concentrating-mechanism (CCM): functional
580 components, Ci transporters, diversity, genetic regulation and prospects for
581 engineering into plants. *J Exp Bo.* **59**: 1441-1461.
- 582 Price, G.D. (2011) Inorganic carbon transporters of the cyanobacterial CO₂
583 concentrating mechanism. *Photosynth Res* **109**: 47-57.

- 584 Ramasubramanian, T.S., Wei, T.-F. and Golden, J.W. (1994) Two *Anabaena* sp. strain
585 PCC 7120 DNA-binding factors interact with vegetative cell- and heterocyst-specific
586 genes. *J Bacteriol* **176**: 1214-1223.
- 587 Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) A
588 generic protein purification method for protein complex characterization and
589 proteome exploration. *Na. Biotechnol* **17**: 1030-1032.
- 590 Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y. (1979)
591 Generic assignments, strain stories and properties of pure cultures of
592 cyanobacteria. *J Gen Microbiol* **111**: 1-61.
- 593 Schaefer, M.R., and Golden, S.S. (1989) Differential expression of members of a
594 cyanobacterial *psbA* gene family in response to light. *J Bacteriol* **171**: 3973-3781.
- 595 Scheneider, K.L., Pollard, K.S., Baertsch, R., Pohl, A., and Lowe, T.M. (2006) The
596 UCSC archaeal genome browser. *Nucleic Ac Res* **34**: D407–D410.
- 597 Valladares, A., Rodríguez, V., Camargo, S., Martínez-Nöel, G.M., Herrero, A., and
598 Luque, I. (2011) Specific role of the cyanobacterial PipX factor in the heterocysts of
599 *Anabaena* sp. strain PCC 7120. *J Bacteriol* **193**(5): 1172-1182.
- 600 Wang, H.L., Postier, B.L., and Burnap, R.L. (2004) Alterations in global patterns of
601 gene expression in *Synechocystis* sp. PCC 6803 in response to inorganic carbon
602 limitation and the inactivation of *ndhR*, a LysR family regulator. *J Biol Chem* **279**:
603 5739-5751.
- 604
605
606

607 **Figure legends:**

608 **Fig. 1.** Growth of the *all3953* mutant. A, The growth rate constant ($\mu = \ln 2/t_d$, where t_d is
 609 the doubling time) was calculated from the increase of protein content determined in
 610 0.2 ml samples of cultures. The table shows the mean and standard deviation from 3
 611 independent cultures of each strain and condition. $\Delta all3953$ is strain CSS74; CSS77 is
 612 a control strain that carries the Sm/Sp-resistant determinant in a wild-type background.
 613 B, Samples of cultures were photographed after 5 days of incubation under the
 614 indicated conditions. HL, high light; LL, low light; HC, high carbon; LC, low carbon.

615

616 **Fig. 2.** Expression of *rbcLXS* in the $\Delta all3953$ mutant and complemented strain.
 617 Northern analysis carried out with RNA from strains CSS77 (control) CSS74 ($\Delta all3953$)
 618 and CSS74C (CSS74 complemented) was isolated from cells grown with HC (0) and
 619 incubated for 1h (1) with LC. The membranes were hybridized with an internal
 620 fragment of the *rbcL* gene (upper panels) and, as a loading and transfer control, of the
 621 *rnpB* gene (lower panels). Arrowheads point to the main transcripts detected with the
 622 *rbcL* gene probe (approximate sizes are indicated).

623

624 **Fig. 3.** Consensus All3953 binding sequence and *rbcL* promoter. A, The primary
 625 consensus motif based on 142 high confidence CSS57 ChIP-Seq peak sequences is
 626 shown with indication of the probability of occurrence of each base along the 22-nt
 627 sequence. W is A or T; Y is C or T; B is C, G or T. B, Structure of the *rbcLXS* promoter
 628 region. The transcription initiation point of the operon (+1) and the -10 and -35 boxes
 629 (from Nierzwicki-Bauer *et al.*, 1984) are indicated in red. The NtcA-binding site
 630 (GTAN₈TAC) is indicated in green, and the three putative binding sites for All3953 (Box
 631 I, Box II and Box III) are indicated in blue.

632

633 **Fig. 4.** qRT-PCR analysis of the expression of selected photosynthesis and respiration-
634 related *All3953* gene targets. Transcriptional response of the indicated genes to C_i
635 limitation in the control (CSS77) and $\Delta all3953$ mutant (CSS74) strains was
636 investigated. RNA was isolated from cells grown with 10 mM NaHCO_3 -supplemented
637 medium bubbled with 1% CO_2 in air (0) incubated for 1 h (1) or 3 h (3) in NaHCO_3 -free
638 medium bubbled with air. Bars represent the mean transcript levels (\pm standard
639 deviation) in three independent experiments.

640

1 **The LysR-type transcription factor PacR is a global regulator of photosynthetic**
2 **carbon assimilation in *Anabaena***

3

4 Silvia Picossi, Enrique Flores and Antonia Herrero*

5

6 Instituto de Bioquímica Vegetal y Fotosíntesis, Consejo Superior de Investigaciones
7 Científicas and Universidad de Sevilla, Américo Vespucio 49, E-41092, Seville, Spain.

8

9

10 *Corresponding author. Tel.: +34 954489522. Fax: +34 954460165. E-mail address:

11 herrero@ibvf.csic.es

12

13 Keywords: ChIP; Cyanobacteria; Oxygenic phototrophy; Photoprotection; RuBisCo

14

15 Running title: Photosynthetic carbon assimilation regulator

16

17 Accession link to data:

18 <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=afcwxkacxpydfax&acc=GSE58861>

19

20

21

22

23 **Summary**

24 **Cyanobacteria perform water-splitting photosynthesis and are important primary**
25 **producers impacting the carbon and nitrogen cycles at global scale. They fix CO₂**
26 **through ribulose biphosphate carboxylase/oxygenase (RuBisCo) and have**
27 **evolved a distinct CO₂ concentrating mechanism (CCM) that builds high CO₂**
28 **concentrations in the vicinity of RuBisCo favoring its carboxylase activity.**
29 **Filamentous cyanobacteria such as *Anabaena* fix CO₂ in photosynthetic**
30 **vegetative cells, which donate photosynthate to heterocysts that rely on a**
31 **heterotrophic metabolism to fix N₂. CCM elements are induced in response to**
32 **inorganic carbon limitation, a cue that exposes the photosynthetic apparatus to**
33 **photodamage by over-reduction. An *Anabaena* mutant lacking the LysR-type**
34 **transcription factor All3953 grows poorly and dies under high light. The *rbcL***
35 **operon encoding RuBisCo is induced upon carbon limitation in the wild type but**
36 **not in the mutant. CHIP-Seq analysis was used to globally identify All3953 targets**
37 **under carbon limitation. Targets include, besides *rbcL*, genes encoding CCM**
38 **elements, photorespiratory pathway, photosystem- and electron transport-**
39 **related components, and factors, including flavodiiron proteins, with a**
40 **demonstrated or putative function in photoprotection. qRT-PCR analysis of**
41 **selected All3953 targets showed regulation in the wild type but not in the mutant.**
42 **All3953 (PacR) is a global regulator of carbon assimilation in an oxygenic**
43 **photoautotroph.**

44

45

46 Introduction

47 As the organisms that developed oxygenic photosynthesis, cyanobacteria have played
48 a crucial role in Earth's history and the evolution of life in our planet. Indeed, the
49 production of O₂ as a result of cyanobacterial activity was responsible for the oxidation
50 of the Earth's atmosphere about 2.5-2.3 billion years ago (Lyons *et al.*, 2014).
51 Furthermore, all the extant plastids of eukaryotic algae and plants are of cyanobacterial
52 origin. Cyanobacteria were the first organisms to link the activity of the two types of
53 photosystems (PSI and PSII), which allowed the generation of high electrochemical
54 potential, and to combine them with a H₂O-splitting complex. Nowadays, most
55 cyanobacteria are phototrophs relying on oxygenic photosynthesis to generate ATP
56 and reducing equivalents for the fixation of CO₂ and the assimilation of inorganic
57 nitrogen. Indeed, they are responsible for an important fraction of the primary
58 productivity in the Earth's oceans, where they are important CO₂ and N₂ fixers, thus
59 impacting the C and N cycles at a global scale (Knoll 2008; Price *et al.*, 2008).

60 The enzyme responsible for the bulk of CO₂ fixation in the biosphere is ribulose-
61 1,5-bisphosphate carboxylase/oxygenase (RuBisCo), which has a relatively low affinity
62 for CO₂ and, moreover, it can also accept O₂ as a substrate. Compensating for this
63 relatively low performance, RuBisCo is considered the most abundant enzyme on
64 Earth. As a carboxylase, RuBisCo catalyzes the first step of the Calvin-Benson-
65 Bassham (CBB) cycle, i.e., the incorporation of atmospheric CO₂ into ribulose-1,5-
66 bisphosphate to give two molecules of 3-phosphoglycerate. As an oxygenase, it
67 catalyzes the incorporation of O₂ into ribulose-bisphosphate, which produces 2-
68 phosphoglycolate that ~~can~~ leads to photorespiration with a subsequent loss of fixed C
69 and energy. To increase the efficiency of CO₂ fixation, cyanobacteria have developed a
70 distinct CO₂ concentrating mechanism (CCM) constituted by inorganic carbon (C_i)
71 transporters that incorporate bicarbonate and CO₂ into the cell, and a proteinaceous
72 compartment, the carboxysome, where RuBisCo, together with carbonic anhydrase, is

73 confined (Price *et al.*, 2008; Cameron *et al.*, 2014). Three high-affinity bicarbonate
74 transporters (the ABC-type Cmp, also called BCT1, and the Na⁺-dependent BicA and
75 SbtA), and two CO₂ transporters (the high-affinity NDH-I₃ and the low-affinity NDH-I₄)
76 have been identified in different cyanobacteria (see Price, 2011). Through CCM, Ci in
77 the form of CO₂ can concentrate in the vicinity of cyanobacterial RuBisCo, to allow high
78 specific activity for production of 3-phosphoglycerate to levels much higher than in
79 plants (Cameron *et al.*, 2014). Indeed, components of cyanobacterial CCM have been
80 transformed in tobacco, with the result of improved specific activity of CO₂ fixation,
81 which represents a step towards improved photosynthesis in plants (Lin *et al.*, 2014).

82 In unicellular cyanobacteria, CCM elements are regulated by C_i availability.
83 Especially genes encoding C_i transporters are induced, whereas the structural genes
84 for carboxysome components and the *rbcL/S* genes encoding RuBisCo are only
85 moderately responsive (see Price *et al.*, 2008; Cameron *et al.*, 2014). In chemotrophic
86 bacteria the process of CO₂ fixation and the response to C_i limitation are usually
87 controlled by LysR-type transcriptional regulators (LTTRs). The genes encoding the
88 enzymes of the CBB cycle are usually found in clusters regulated by CbbR factors,
89 which constitute a sub-family of LTTRs (Gibson and Tabita, 1996). In unicellular
90 cyanobacteria, a number of CbbR homologs have been characterized. CmpR is an
91 activator of the *cmp* genes (Nishimura *et al.*, 2008), whereas CcmR (*aka* NdhR) acts as
92 a transcriptional repressor of multiple genes encoding other C_i transporters (e.g., Figge
93 *et al.*, 2001; Wang *et al.*, 2004). A third type of CbbR-like protein, the activator of the
94 RuBisCo genes, has not yet been identified in cyanobacteria.
95 ~~The identification of a third type of CbbR-like protein, the activator of the RuBisCo~~
96 ~~genes, RbcR, has remained elusive in cyanobacteria.~~

97 Filamentous heterocyst-forming cyanobacteria, of which *Anabaena* sp. PCC
98 7120 (hereafter *Anabaena*) is a model organism, additionally have the capacity for cell
99 differentiation to turn some O₂-evolving photosynthetic cells of the filament into
100 heterocysts, which are heterotrophic cells specialized in the fixation of atmospheric

101 N_2 . Thus, *Anabaena* is a truly pluricellular bacterium with different cell types specialized
102 in different nutritional tasks that exchange nutrients and regulators and contribute to
103 the performance of the filament as the organism unit (see Flores and Herrero, 2010).
104 Nitrogen assimilation and heterocyst differentiation in *Anabaena* are regulated by the
105 global transcription factor NtcA, which responds to the cellular C-to-N ratio, and the
106 heterocyst-specific transcription factor HetR (see Herrero *et al.*, 2013). The *Anabaena*
107 genomic sequence includes three genes annotated as CbbR-like LTTRs (Kaneko *et*
108 *al.*, 2001). Of these, *all0862* has been identified as *cmpR*, and its product activates the
109 expression of the *cmp* operon (*alr2877-alr2880*) and the *cmpR* gene itself in response
110 to C_i limitation (López-Igual *et al.*, 2012). Notably, this regulation is effected in
111 combination with NtcA, thus revealing a mode of co-regulation by C and N availability
112 (López-Igual *et al.*, 2012). Furthermore, in *Anabaena* the *rbcLXS* operon encoding
113 RuBisCo, which is moderately induced under C_i limitation (López-Igual *et al.*, 2012), is
114 repressed in the heterocysts (Madan and Nierzwicki-Bauer, 1993), a regulation likely
115 exerted by NtcA (Ramasubramanian *et al.*, 1994; Picossi *et al.*, 2014).

116 Here we have identified the CbbR-homolog All3953 as the activator of the
117 RuBisCo-encoding operon in *Anabaena*. We have determined the All3953 regulon by
118 ChIP-Seq, which has revealed that All3953 is a global regulator for C_i assimilation
119 genes and genes involved in protection of the photosynthetic apparatus against
120 oxidative damage that are regulated by C_i availability.

121

122 **Results**

123 *All3953, an RbcR-like factor in Anabaena sp. PCC 7120*

124 To gain insight into the LTTR All3953 in *Anabaena*, the expression of the *all3953* gene
125 under different growth conditions was analyzed by northern blot. Three bands of
126 hybridization corresponding to transcripts of ca. 1.9, 1.6 and 1.3 kb, respectively, which
127 appeared similarly represented before and after C_i limitation could be observed (Supp.

128 Fig. 1A). On the other hand, *all3953* expression did not significantly respond to N
129 depletion (Supp. Fig. 1B). The latter result was consistent with previous global
130 transcriptional studies (Ehira and Ohmori, 2006; Flaherty *et al.*, 2011) and with the
131 lack of binding associated to *all3953* of the N-control transcriptional regulator NtcA
132 upon combined nitrogen depletion (Picossi *et al.*, 2014).

133

134 *Isolation and characterization of an all3953 mutant*

135 To study the role of All3953, a mutant strain bearing an inactivated version of *all3953*
136 was constructed (see Experimental procedures). In strain CSS74 most of the *all3953*
137 gene was deleted and the C.S3 gene cassette (encoding Sm/Sp resistance) was
138 introduced to facilitate segregation and maintenance of the mutation in *Anabaena*
139 (Supp. Fig. 2A). As a control, strain CSS77 bearing the C.S3 gene cassette in the
140 *Anabaena* plasmid alpha was also constructed. For *cis* complementation, a wild-type
141 version of *all3953* was transferred to strain CSS74 in an integrative plasmid,
142 generating strain CSS74C (Supp. Fig. 2A).

143 | Strain CSS74 exhibited poor growth under standard growth conditions with
144 | ammonium as a nitrogen source, and it formed short filaments in liquid medium,
145 | whereas strain CSS74C behaved similarly to CSS77 (not shown). To quantify the
146 | deleterious effect of the *all3953* mutation, growth rates were calculated in liquid
147 | medium under different illumination and C₂-supply conditions ~~(Fig. 1)~~. Growth rate of
148 | the control strain CSS77 was highest (0.892 days⁻¹) under high light (HL) and high
149 | carbon (HC), and was about 30% lower under the other tested conditions (Fig. 1A). In
150 | contrast, growth of strain CSS74 was severely affected under HL HC conditions
151 | (growth rate ca. 70% lower than that of the control) (Fig. 1A), under which it ended up
152 | dying after about 5 days (Fig. 1B). Under HL and low carbon (LC) or low light (LL) HC,
153 | the defect was close to 30% with regard to the control, although after prolonged
154 | incubation the mutant was more severely affected under HL LC. The defect was the

155 smallest (ca. 15%) under LL LC conditions (Fig. 1A,B). ~~In solid medium~~ Finally, growth
 156 ~~of strain CSS74 was similar, and similar to that of the control, in the presence of~~
 157 ~~ammonium, nitrate or no combined nitrogen the effect of the N-source on the growth of~~
 158 ~~strain CSS74 was tested in solid medium under LL LC and HL LC conditions, whereas~~
 159 ~~under HL LC in LL LC growth of strain CSS74 was similar, and similar to that of the~~
 160 ~~control, in the presence of ammonium, nitrate or no combined nitrogen (Figure 1C).~~
 161 However, under HL LC a severe growth defect was observed with regard to the control
 162 with any of the tested nitrogen sources (~~not shown~~Fig. 1C). ~~It is worth noting that,~~
 163 ~~because the CSS74 strain formed short filaments, colonies were observed at a lower~~
 164 ~~chlorophyll concentration than in the control strain.~~ In summary, the lack of *All3953* had
 165 a deleterious effect on growth, especially under HL (and HC) conditions.

166 To further characterize the growth defect of the *all3953* mutant, the rate of
 167 oxygen evolution using CO₂ as a final electron acceptor was measured under different
 168 illumination conditions in the CSS74 mutant in relation to the control strain CSS77.
 169 When exponentially-growing cells were incubated for 24 h under LL HC, the oxygen
 170 evolution rate was slightly lower in strain CSS74 in comparison to CSS77 (88 and
 171 105 μmol O₂·[mg Chl]⁻¹·h⁻¹, respectively). Under HL the difference between the two
 172 strains was larger (167, for CSS77, and 110, for CSS74, μmol O₂·[mg Chl]⁻¹·h⁻¹).

173

174 *Effect of all3953 mutation on rbcLXS expression*

175 To test the effect of the *all3953* mutation on the expression of the *rbcLXS* operon,
 176 northern blot analysis was performed with RNA isolated from cells of the control and
 177 mutant strains grown with HC and transferred to LC conditions. After 1 h incubation
 178 with LC, a ca. 2-fold increase in the amount of the *rbcLXS* transcripts could be
 179 observed in the control strain. No induction could be detected in the mutant (Fig. 2). In
 180 the complemented strain CSS74C the expression of *rbcLXS* increased in LC similarly
 181 to the control (Fig. 2). These results indicated that the induction of the *rbcLXS* operon

182 upon C₁ deficiency was dependent on All3953 and that the defect in strain CSS74 was
183 exclusively due to the lack of All3953.

184

185 *ChIP-Seq analysis of the All3953 targets*

186 To determine the DNA targets of All3953 at a genomic level, we used chromatin
187 immunoprecipitation followed by high-throughput sequencing analysis. To this end, we
188 constructed a strain (CSS57) expressing from the *all3953* promoter a version of
189 All3953 C-terminally fused to TAP-tag (Rigaut *et al.*, 1999), as well as a control strain
190 (CSS68) expressing the TAP-tag alone under the control of the *all3953* promoter
191 (Supp. Fig. 2B, see Experimental procedures). Immunoprecipitation was carried out
192 using cells of strains CSS57 and CSS68 grown with ammonium as the nitrogen source
193 under HC conditions and incubated for 3 h with ammonium under LC conditions.

194 The analysis of the sequences resulted in a total of 142 All3953 binding
195 regions, of which 127 were located in the chromosome, 10 in plasmid alpha, three in
196 plasmid beta and two in plasmid gamma. Each binding region was ascribed to one or
197 two genes according to the location (midpoint) of the region, and the relative location
198 with respect to the assigned gene was also indicated (Table 1 and Supp. Table 1). A
199 total of 175 genes were ascribed to the 142 binding regions. The binding regions were
200 mostly located upstream of the ascribed genes (72%), whereas 21% were intragenic
201 and 7% were located downstream of genes. The results of the ChIP-Seq analysis are
202 available at GEO accession number GSE58861
203 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE58861>).

204 The 175 ascribed genes were classified according to their functional category
205 (Table 2). Remarkably, there were 19 genes encoding proteins related to
206 photosynthesis and respiration, and 21 genes encoding regulatory proteins, including
207 All7179, a SigB homolog. The rest were mostly genes encoding hypothetical or
208 unknown proteins (42%), but also genes encoding proteins involved in translation, in

209 biosynthesis of amino acids and cofactors, prosthetic groups and carriers, in transport,
210 and in other cellular processes. Table 3 highlights All3953-binding regions related to
211 photosynthesis and respiration among which, confirming our results of gene
212 inactivation, the gene encoding the large subunit of the RuBisCo (*rbcL*; binding region
213 #37) is included. The fact that a high number of genes involved in photosynthesis and
214 C fixation, including *rbcL*, were identified as targets of RbcR suggests that this protein
215 is a global transcription factor for photosynthetic C assimilation.

216 A CisFinder analysis of the primary consensus motif was carried out based on
217 142 high-confidence ChIP-Seq peak sequences (Fig. 3A). The consensus motif found
218 has a dyad-symmetry architecture and matches the consensus of LysR-recognition
219 binding sites (RBS) (T-N₁₁-A) (Maddocks and Oyston, 2008), as well as the consensus
220 binding sites proposed for CbbR factors (TNA-N_{7/8}-TNA). The primary motifs identified
221 in the central 100 nt of the binding regions are indicated in Supp. Table 1 (for some
222 binding regions more than one motif have been identified).

223

224 *Expression analysis of some All3953 targets in Anabaena*

225 To corroborate our ChIP-Seq analysis and to support the notion that All3953 is indeed
226 a global regulator for C fixation genes, the expression levels of some of the
227 photosynthesis- and C-fixation-related target genes, its response to C_i limitation and its
228 dependence on All3953 was further analyzed by qRT-PCR (Fig. 4). Strains CSS77 and
229 CSS74 were grown in ammonium and HC under standard light conditions (80
230 $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) at 30°C up to the exponential phase. They were then transferred to medium
231 with ammonium under LC conditions. As previously shown by northern analysis,
232 transcription levels of the *all3953* gene did not significantly change after C_i deprivation
233 in the control strain CSS77. As expected, *all3953* transcript levels were not detectable
234 in CSS74, corroborating that the *all3953* mutation was segregated in this strain. The
235 *rbcL* gene (*alr1524*) was 4.6-fold induced at 3 h after transfer to LC in strain CSS77,

236 whereas no induction was observed in CSS74, thus corroborating the dependence on
237 All3953. Interestingly, under HC the expression of *rbcL* in the mutant was higher than
238 in the control strain, suggesting that besides as an activator under LC, All3953 could
239 be acting as a repressor of *rbcL* under HC conditions.

240 ORF *all4446* (*flv4*) was highly induced (about 100-fold), and its level was
241 maximum 1 h after the shift to LC (Fig. 4). In the *all3953* mutant, the basal transcription
242 of *all4446* in HC was about 8-fold lower than in the control strain, and it was only
243 slightly induced (2-fold) upon the shift to LC. *all3891* (*flv1A*) was induced about 6-fold 3
244 h after the shift to LC in the control strain, whereas only a 2-fold induction was
245 observed in the *all3953* mutant. *all1304* (*bicarbonate transporter homolog*) and *alr4156*
246 (*NdhF homolog*) were both highly induced (up to 30- and 20-fold, respectively) upon
247 transfer of strain CSS77 to LC. In contrast, no induction of *alr4156* and only a small
248 induction of *all1304* took place in strain CSS74. Expression of *alr4592* (*psbAIII*)
249 increased about 5-fold upon C_i deprivation in CSS77, but did not appreciably change in
250 CSS74. The *alr0223* (*NdhA homolog*) gene was about 2-fold induced under C_i
251 deprivation in CSS77 but not in CSS74. Finally, the expression of *alr1004* (*alanine-*
252 *glyoxylate aminotransferase*) was repressed by 5-fold under LC conditions in CSS77
253 but not in CSS74. These results confirm that expression of the above studied genes is
254 regulated, either positively or negatively, by All3953.

255

256 Discussion

257 We have identified the LTR All3953 as the activator of the RuBisCo-encoding genes
258 in the cyanobacterium *Anabaena* sp. PCC 7120. All3953 appears to activate the *rbcL*
259 operon under C_i limitation and to repress it when C_i is abundant. An LTR factor
260 regulating the expression of the *rbcL* operon has not, to our knowledge, been
261 described in any cyanobacterium. The expression of *all3953* does not respond to C_i
262 limitation (Supp. Fig. 1) and, indeed, no binding region of All3953 was found ascribed

263 to *all3953*. Thus, the *all3953* gene seems to belong to the non-autoregulated LTTRs.

264 All3953 shares 28 and 26% identical residues with NdhR from *Synechocystis* sp. PCC
265 6803 and *Synechococcus* sp. PCC 7002 (CcmR), respectively.

266 By ChIP-Seq analysis of a strain bearing a TAP-tagged version of All3953, we
267 have determined 142 All3953-bound regions at 3 h after transfer from high to low C_i
268 conditions, which have been assigned to 177 genes. Apart from genes encoding
269 unknown or hypothetical proteins, the larger category is of genes encoding regulatory
270 proteins, including the transcriptional regulators Alr0353 (a LTTR) and All4500 (CRP-
271 like), the two-component response regulator All3348, the two-component sensor
272 histidine-kinase All1145 and the group 2-sigma 70-type sigma factor All7179 (Supp.
273 Table 1). Interestingly, when comparing to *Synechocystis* sigma factors, All7179
274 (SigB4) is more similar to a homolog of *Synechocystis* SigB, which, along with SigD has
275 been shown to be important for PSII recovery in this unicellular cyanobacterium (Pollari
276 *et al.*, 2009). All3953 also binds upstream of genes *patS* and *hetN*, whose products
277 regulate heterocyst differentiation (Supp. Table 1). The fact that All3953 binds to the
278 promoter region of genes encoding other regulatory proteins suggests a wide role of
279 this protein in the physiology of the organism.

280 In the promoter region of the *rbcL* operon we have found three putative binding
281 sites for All3953, Box I, Box II, and Box III (Fig. 3B) that resemble the consensus
282 recognition sequence found by Cisfinder analysis (Fig. 3A). It is conceivable that, like in
283 some other LTTRs these boxes combine repression and activation sites. In this regard,
284 binding of All3953 to Box III, overlapping gene promoter elements, could be related to
285 the repression of *rbcL* observed under high C_i (Fig. 4). On the other hand, as
286 mentioned above, the *rbcL* operon is repressed in heterocysts by the global
287 transcriptional regulator NtcA, for which a binding site is found overlapping the operon
288 TSP (Ramasubramanian *et al.*, 1994; Picossi *et al.*, 2014) (Figure 3B). It is conceivable
289 that NtcA binding in these differentiated cells interfere with All3952-mediated activation.

290 Besides being the activator of the RuBisCo genes, All3953 targets include
291 genes involved in other processes related to carbon assimilation, such as C_i transport
292 (*all1304*, encoding an homolog of the BicA bicarbonate transporter; *alr4156* encoding a
293 homolog of the NdhF3 subunit of the NDH-1₃ CO₂ uptake system; and *alr0869*
294 encoding a homolog of the NdhF4 subunit of the NDH-1₄ CO₂ uptake system);
295 components of the carboxysome shell (*all0868*, putative *ccmK*) and 2-
296 phosphoglycolate metabolism (*alr1004* and *alr2873*, possibly related to
297 photorespiration [Eisenhut *et al.*, 2008]). Notably, All3953 targets include also a
298 number of genes encoding photosystem components, such as *alr5154* (*psaA*,
299 encoding the PSI core protein PsaA), *alr3727* (*psbAII*, encoding a component of form II
300 of the PSII core protein PsbA [D1]), *alr4592* (*psbAIII*, encoding another component of
301 form II of PsbA) and *alr1216* (PSII 12 kD extrinsic protein PsbU), and genes related to
302 PS activity. In the latter group are *alr4149* (biliverdin reductase, putatively involved in
303 phycobilisome -PSII antenna- synthesis), and genes that can participate in
304 photosynthetic electron transfer, such as *alr0223* and *alr0348* (*ndhA* and *ndhD*,
305 subunits of other putative NADH dehydrogenases), *alr1576* (dehydrogenase subunit),
306 *all0737* (thioredoxin reductase), *all1365* (CytM cytochrome), *all4148* (ferredoxin I),
307 *all3891* and *all4446* (flavodiiron proteins Flv1A and Flv4, respectively). (Besides in CO₂
308 uptake, NdhF3 and NdhF4 can also participate in electron transfer.)

309 Reports on gene expression regulation by C_i are scarce for *Anabaena*.
310 However, in the unicellular cyanobacterium *Synechocystis* sp. PCC 6803
311 transcriptomic analysis has already shown down-regulation of some genes encoding
312 polypeptides of PSI and PSII complexes as well as of phycobilisome components,
313 upon transfer to C_i limitation conditions, likely as an adaptation to lower assimilatory
314 power demand, and up-regulation of some PSII core polypeptides, interpreted as
315 adaptation to conditions of shortage of electron acceptors that could lead to
316 photodamage and increased turnover of PS core components (Wang *et al.*, 2004). Our
317 analysis extends the array of photosynthetic genes responding to C_i limitation,

318 remarkably to include the core PSI reaction center *psaA* gene. Moreover, except for
319 *alr1004* that responds negatively, all the *Anabaena* photosynthetic genes mentioned
320 above increase expression upon the shift to C_i limitation.

321 Noteworthy, for the majority of photosynthetic gene targets of All3953 in
322 *Anabaena*, a function in protection against reactive oxygen species, which can be
323 generated by C_i limitation or exposure to HL, has either been described or could be
324 predicted. Thus, the *psbAll* and *psbAIII* genes are induced under HL in the unicellular
325 *Synechococcus* sp. PCC 7942, and cells cultured under HL showed more form II, and
326 less form I (encoded by *psbAI*), of D1 compared to cells under LL (Schaefer and
327 Golden, 1989). Regarding flavodiiron proteins genes, *flv1A*, *flv3A*, and specially *flv2*
328 and *flv4* of *Anabaena* have been shown up-regulated in vegetative cells in low C_i , and
329 *flv1A* and *flv3A* also in high light, whereas *flv1B* and *flv3B* are expressed exclusively in
330 heterocysts (Ermakova *et al.*, 2013). Whereas Flv1A and Flv3A appear involved in
331 photoreduction of oxygen to water by removing excess electrons from PSI through
332 NAD(P)H dehydrogenases, Flv2 and Flv4 could have a role in photoprotection of PSII
333 under low C_i (Ermakova *et al.*, 2013). Regarding photorespiration, it has also been
334 considered to have a role in removal of excess O_2 (Eisenhut *et al.*, 2008). To the best
335 of our knowledge, the regulator responsible for the response to C_i availability of any
336 photosynthetic gene has not been identified in cyanobacteria (oxygenic phototrophs).
337 Our ChIP-Seq and expression analysis indicate that All3953 is a regulator of
338 photosynthetic genes in *Anabaena*.

339 The growth rate of *Anabaena* is highest under HL HC conditions (Fig. 1),
340 implying that this cyanobacterium has mechanisms to get profit of HC while
341 counteracting HL stress. The *all3953* mutant strain CSS74 exhibited a growth defect in
342 all the conditions tested, but especially under HL, where it ends-up dying (Fig. 1). This
343 shares the idea that in *Anabaena* All3953 is required to cope with HL stress. The effect
344 of the lack of All3953 seems more detrimental in relation to impaired photoprotection
345 than to impaired C_i scavenging (preference of the mutant for LL LC over HL HC

346 conditions). Indeed, even in LL, the mutant seems to perform slightly better with LC
347 than with HC (Fig. 1). Although LC could suppose a limitation of electron acceptors, an
348 increased rate of photorespiration under LC together with the fact that some of the
349 All3953 targets that could cope with excess oxygen are residually induced upon the
350 transfer from HC to LC in the CSS74 mutant (Fig. 4) could contribute to the preference
351 of this strain for LC over HC, especially under HL.

352 Our results show that in *Anabaena* the responses to C_i availability include
353 regulation of genes encoding elements of CCM and RuBisCo, but also of
354 photosynthetic genes to adjust generation of assimilatory power while preserving the
355 photosynthetic apparatus from oxidative damage, which is specially relevant in
356 oxygenic phototrophs. Because All3953 is a transcriptional regulator globally
357 coordinating these responses, we have named it PacR (Photosynthetic assimilation of
358 carbon Regulator).

359

360 **Experimental procedures**

361 *Strains*

362 *Anabaena* sp. strain PCC 7120 was grown photoautotrophically at 30°C with illumination
363 ($80 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) in liquid BG11₀ medium (Rippka *et al.*, 1979) supplemented with 3 mM
364 NH_4Cl , 6 mM TES buffer and 10 mM NaHCO_3 and bubbled with a mixture of CO_2 and
365 air (1% v/v) (HC). Other conditions used were no NaHCO_3 supplement and bubbling
366 with air (LC); $12 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ (LL); $175 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ (HL). ~~For growth on plates illumination~~
367 ~~was $9 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ (LL) or $34 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ (HL).~~ For the mutants generated in this work,
368 antibiotics were used at the following concentrations: Sm, $2 \mu\text{g ml}^{-1}$; Sp, $2 \mu\text{g ml}^{-1}$; and
369 Nm, $25 \mu\text{g ml}^{-1}$ for bubbled cultures; and Sm, $5 \mu\text{g ml}^{-1}$; Sp, $5 \mu\text{g ml}^{-1}$; and Nm,
370 $40 \mu\text{g ml}^{-1}$ for cultures in solid medium.

371

372 *Strain construction*

373 To construct a mutant of the *all3953* gene, the 5' and 3' end of the gene, along with the
374 flanking regions, were PCR amplified using chromosomal DNA of PCC 7120 as the
375 template and primers all3953-14 (BglII) and all3953-15 (Sall), and primers all3953-16
376 (Sall) and all3953-17 (PstI), respectively (all primers are specified in Supp. Table 2).
377 The PCR products were digested with Sall and ligated. The resulting mixture was used
378 as a template for overlapping PCR with primers all3953-14 and all3953-17. The new
379 PCR product was digested with BglII and PstI and ligated to the BglII-PstI-digested
380 pRL271 (Black *et al.*, 1993), obtaining plasmid pCSS161. Plasmid pCSS161 was
381 digested with SpeI and the 2-kb Sm^r Sp^r gene cassette C.S3, excised with XbaI from
382 pCSE120 [S.K3/L.HEH2 (BamHI)/C.S3 (BamHI); nomenclature as in (Elhai and Wolk,
383 1988)], was inserted obtaining plasmid pCSS162. Plasmid pCSS162 was transferred to
384 strain PCC 7120 by conjugation (Elhai *et al.*, 1997). Exconjugants resistant to Sm and
385 Sp, which had the $\Delta all3953::C.S3$ construct integrated by double recombination were
386 selected, obtaining strain CSS74. The segregation of the mutation was tested by PCR
387 (Supp. Fig. 2) with primers all3953-14, all3953-15, all3953-17 and all3953-20.

388 To construct a control strain expressing Sm^r and Sp^r plasmid pCSS163, a
389 derivative of plasmid pCSEL24 (Olmedo-Verd *et al.*, 2006) containing the C.S3 gene
390 cassette, was transferred to *Anabaena* by conjugation. Exconjugants that had the
391 pCSS163 integrated in the alpha plasmid of *Anabaena* were selected, obtaining the
392 strain CSS77 (Supp. Fig. 2).

393 To complement the *all3953* mutation of strain CSS74, a DNA fragment
394 encompassing the whole *all3953* gene and sequences upstream from it was amplified
395 by PCR using the primer pair all3953-24/all3953-25, both including EcoRI sites, and
396 strain PCC 7120 DNA as the template. This fragment was cloned in the EcoRI site of
397 the mobilizable Nm^r encoding vector pRL424 (Elhai and Wolk, 1988) producing plasmid
398 pCSS164, which was transferred to strain CSS74 by conjugation followed by selection
399 for Nm^r. The genomic structure of the exconjugants in the *all3953* region (Supp. Fig. 2)
400 was corroborated by PCR.

401 To construct a strain expressing All3953-C-TAP, the *all3953* gene (including the
402 upstream region) was PCR-amplified with primers all3953-11 and all3953-12 and DNA
403 of PCC 7120 as the template. The TAP-tag was PCR-amplified with primers TAPtag-1
404 and TAPtag-2 using DNA of plasmid pBS1479 as the template (Puig *et al.*, 2001). The
405 two PCR products were digested with Sall and ligated, after which the ligation product
406 was used as the template for an overlapping PCR using primers all3953-11 and
407 TAPtag-2. The PCR product was digested with PstI and ligated to the mobilizable
408 vector pCSV3 (Valladares *et al.*, 2011) digested with PstI, rendering plasmid pCSS107.

409 To construct a control strain with the TAP-tag under the control of the *all3953*
410 promoter, a 0.4-kb region upstream of *all3953* was PCR-amplified using primers
411 all3953-11 and all3953-18 and DNA of pCSS107 as the template. The PCR product
412 was digested with Sall and ligated to the PCR-amplified TAP-tag digested with Sall,
413 after which the ligation product was used as the template for an overlapping PCR using
414 primers all3953-11 and TAPtag-2. The PCR product was digested with PstI and ligated
415 to PstI-digested pCSV3 to give plasmid pCSS157. Plasmids pCSS107 and pCSS157
416 were transferred by conjugation to strain PCC 7120 and single Sm^r Sp^r recombinants
417 were selected, obtaining strain CSS57 and CSS68, respectively. Western blots using
418 Peroxidase-Anti-Peroxidase Soluble Complex (PAP, Sigma-Aldrich) were performed to
419 ensure that the two strains expressed the TAP-tag (Supp. Fig. 2B).

420

421 *Chromatin immunoprecipitation*

422 Cells of strains CSS57 growing exponentially (3-5 $\mu\text{g Chl}\cdot\text{ml}^{-1}$) in the light (80
423 $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$) in medium supplemented with 2 $\mu\text{g}\cdot\text{ml}^{-1}$ Sm and Sp, in HC conditions were
424 incubated with LC for 3 h. Formaldehyde was then added to the cultures to a final
425 concentration of 1%, and the cultures were incubated for 15 min. Glycine was added at
426 125 mM final concentration and the incubation was continued for 5 min to stop the
427 fixing reaction. The cells were then filtered, washed with cold TBS (20 mM Tris-HCl, pH
428 7.4, 140 mM NaCl) and collected in tubes (25 ml of culture per tube). The pellets were

429 frozen in liquid nitrogen and stored at -20°C until used. Pellets corresponding to about
430 25 ml of culture were resuspended in 500 µl of lysis buffer (50 mM HEPES-KOH, pH
431 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate,
432 supplemented with Mini EDTA-free protease inhibitor cocktail [Roche]) and, after
433 addition of 150 µl of glass beads (acid-washed, 425-600 µm [Sigma]), the cells were
434 broken in a multivortexer at 2,000 rpm for 1 h at 4°C. The cell lysates were collected by
435 centrifugation and the extracts were subjected to sonication to shear the DNA to about
436 300-bp fragments (60 cycles of 10 s, 20 s ice, 15% amplitude, in a Branson Digital
437 Sonifier). After centrifugation to eliminate cell debris, the whole-cell extracts were
438 stored at -20°C or immediately used for immunoprecipitation.

439 Immunoprecipitation of DNA was carried out as described (Picossi *et al.*, 2014),
440 with some modifications. Whole-cell extracts were prepared at 4 mg·ml⁻¹ of total protein
441 with lysis buffer (in 500 µl total volume). A 50-µl sample was taken as the input sample,
442 and the extracts were incubated with 15 µl IgG-conjugated Dynabeads (about 6 µg
443 IgG) at 4°C with rotation for 12-14h. The washes of the Dynabeads, as well as the
444 elution of the immunoprecipitated material, the crosslinking reversion and the isolation
445 of the DNA were performed as in (Picossi *et al.*, 2014).

446

447 *Massive sequencing of the immunoprecipitated DNA*

448 Input and ChIP DNA samples were sent for sequencing to the Functional Genomics
449 Core Facility of the Institute for Research in Biomedicine, Barcelona (Spain). Next
450 generation sequencing was carried out using Illumina's sequencing technology. ChIP
451 DNA Sample Prep Kit (Illumina) was used for library preparation. Libraries were loaded
452 at 8 pM concentration into the flow cell using the Cluster Station running recipe V7 with
453 the Single-Read Cluster Generation Kit v4 (all Illumina). The flow cell was loaded into
454 the Genome Analyzer II and samples were sequenced for 120 nucleotides from a
455 single end using the Sequencing Kit v5 and recipe v8 (all Illumina). Manufacturer's

456 recommendations were strictly followed. Illumina sequencing data were pre-processed
457 with the standard Illumina pipeline version 1.5 and sequences were aligned to the PCC
458 7120 genome (<http://genome.microbedb.jp/cyanobase/Anabaena>) with the Bowtie
459 software 0.12.5 (Langmead *et al.*, 2009). The analysis of the results were carried out
460 using the Triform algorithm (Kornacker *et al.*, 2012) as in (Picossi *et al.*, 2014). The
461 sequences in the ChIP samples of strain CSS68 (TAP control) were used as the
462 background of the sequences found for strain CSS57 (All3953-TAP), and thus to
463 determine the specific binding regions of All3953-TAP in the genome of *Anabaena*.
464 The binding regions were visualized and analyzed using the UCSC Microbial Genome
465 Browser (Schneider *et al.*, 2006). They were ascribed to one or two genes, in case it
466 was not possible to ascribe them to only one, and classified as upstream of the gene, if
467 the midpoint of the binding region was located upstream of the start of the gene,
468 internal, if the midpoint of the binding region was inside the gene, or downstream, if the
469 midpoint of the binding region was located downstream of the end of the gene to which
470 it had been ascribed.

471

472 *Northern and qRT-PCR analyses*

473 Isolation of total RNA from *Anabaena* was done as described previously (Mohamed
474 and Jansson, 1989). Northern analysis was performed as described previously (López-
475 Igual *et al.*, 2012).

476 For qRT-PCR, 750 ng of DNA-free RNA samples were used for all the PCR
477 primer pairs. For the RT reaction, the Quantitech Reverse transcription kit (Qiagen),
478 with the Random Hexamer Primer mix (100 ng per sample) (Bioline) was used. The
479 cDNA produced was diluted 7.5 times to use 2 µl of cDNA per PCR reaction. PCR was
480 done using the Quantimix Easy SYG Kit (Biotools) (SYBR green I) in a iCycler iQ Multi-
481 Color Real Time PCR Detection System (Bio-Rad). The abundance of a transcript in
482 the RNA sample was calculated as: abundance = $2^{[Ct(\text{sample}) - Ct(\text{control})]}$, where the
483 RNA sample of the control strain CSS77 in HC condition (0) was used as the control.

484

485 *Oxygen evolution*

486 2-ml samples of exponentially grown cultures in HC LL or HC HL conditions were used
487 to measure O₂ evolution with an O₂ electrode calibrated with culture medium and
488 Na₂S₂O₄ as the reducing agent. O₂ production was measured in the light (400
489 µE·m⁻²·s⁻¹) after a seven-minute incubation in the dark.

490

491 **Acknowledgments**

492 This work was supported by grants BFU2010-17980 and BFU2013-44686-P from the
493 Spanish government, co-financed by FEDER. The authors are grateful to K. Kornacker
494 for carrying out the Triform and CisFinder analyses, and to H. Auer and I. Pons, from
495 the Functional Genomics Core Facility of the IRB, Barcelona (Spain); to F. Monje-
496 Casas and Yagut Allahverdiyeva-Rinne for critical reading of the manuscript and to M.
497 J. Huertas and A. Torrado for help with O₂ production measurements.

498

499 **REFERENCES**

- 500
501 Black, T.A., Cai, Y., and Wolk C.P. (1993) Spatial expression and autoregulation of
502 *hetR*, a gene involved in the control of heterocyst development in *Anabaena*. *Mol*
503 *Microbiol* **9**: 77-84.
- 504 Cameron, C.J., Sutter, M., and Kerfeld, C.A. (2014) The carboxysome: Function,
505 structure and cellular dynamics. *The Cell Biology of Cyanobacteria*, eds. Flores E,
506 Herrero A (Caister Academic Press, Norfolk), pp. 171-188.
- 507 Ehira, S., and Ohmori, M, (2006) NrrA, a nitrogen-responsive response regulator
508 facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC
509 7120. *Mol Microbiol* **59**: 1692-1703.
- 510 Elhai, J., Vepritskiy, A., Muro-Pastor, A.M., Flores, E, and Wolk, C.P. (1997) Reduction
511 of conjugal transfer efficiency by three restriction activities of *Anabaena* sp. strain
512 PCC 7120. *J Bacteriol* **179**: 1998-2005.
- 513 [Eisenhut, M., Ruth, W., Haimovich, M., Bauwe, H., Kaplan, A., and Hagemann, M.](#)
514 [\(2008\) The photorespiratory glycolate metabolism is essential for cyanobacteria and](#)
515 [might have been conveyed endosymbiotically to plants. *Proc Natl Acad Sci USA*](#)
516 [**105**: 17199-17204.](#)
- 517 Elhai, J., and Wolk, C.P. (1988) A versatile class of positive-selection vectors based on
518 the nonviability of palindrome-containing plasmids that allows cloning into long
519 polylinkers. *Gene* **68**: 119-138.
- 520 Ermakova, M., Battchikova, N., Allahverdiyeva, Y., and Aro, E.M. (2013) Novel
521 heterocyst-specific flavodiiron proteins in *Anabaena* sp. PCC 7120. *FEBS Lett* **587**:
522 82-7.
- 523 Figge, R.M., Cassier-Chauvat, C., Chauvat, F., and Cerff, R. (2001) Characterization
524 and analysis of an NAD(P)H dehydrogenase transcriptional regulator critical for the
525 survival of cyanobacteria facing inorganic carbon starvation and osmotic stress. *Mol*
526 *Microbiol* **39**: 455-68.
- 527 Flaherty, B.L., Van Nieuwerburgh, F., Head, S.R., and Golden, J.W. (2011) Directional
528 RNA deep sequencing sheds new light on the transcriptional response of *Anabaena*
529 sp. strain PCC 7120 to combined-nitrogen deprivation. *BMC Genomics* **12**: 332.
- 530 Flores, E., and Herrero, A. (2010) Compartmentalized function through cell
531 differentiation in filamentous cyanobacteria. *Nat Rev Microbiol* **8**: 39-50.
- 532 Gibson, J.L., and Tabita, F.R. (1996) The molecular regulation of the pentose
533 phosphate pathway in proteobacteria and cyanobacteria. *Arch Microbiol* **166**: 141-
534 150.
- 535 Herrero, A., Picossi, S., and Flores, E. (2013) Gene Expression during heterocyst
536 differentiation. *Adv Bot Res* **65**: 281-329.
- 537 Kaneko, T., Nakamura, Y., Wolk, C.P., Kuritz, T., Sasamoto, S., and Watanabe, A.
538 (2001) Complete genomic sequence of the filamentous nitrogen-fixing
539 cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res* **8**: 205-213.
- 540 Knoll, A.H. (2008) Cyanobacteria and Earth history. *The Cyanobacteria: Molecular*
541 *Biology, Genomics and Evolution*, eds. Herrero A, Flores E (Caister Academic
542 Press, Norfolk), pp. 1-19.
- 543 Kornacker, K., Rye, M.B., Handstad, T., and Drablos, F. (2012) The Triform algorithm:
544 improved sensitivity and specificity in ChIP-Seq peak finding. *BMC Bioinformatics*
545 **13**: 176.
- 546 Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009) Ultrafast and memory-
547 efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**:
548 R25.
- 549 Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A.J., and Hanson, M.R. (2014) A
550 faster Rubisco with potential to increase photosynthesis in crops. *Nature* **513**: 547-
551 550.
- 552 López-Igual, R., Picossi, S., López-Garrido, J., Flores, E., and Herrero, A. (2012) N
553 and C control of ABC-type bicarbonate transporter Cmp and its LysR-type

- 554 transcriptional regulator CmpR in a heterocyst-forming cyanobacterium, *Anabaena*
 555 sp. *Environ Microbiol* **14**: 1035-1048.
- 556 Lyons, T.W., Reinhard, C.T., and Planavsky, N.J. (2014) The rise of oxygen in Earth's
 557 early ocean and atmosphere. *Nature* **506**: 307-315.
- 558 Madan, A.P., and Nierzwicki-Bauer, S.A. (1993) In situ detection of transcripts for
 559 ribulose-1,5-bisphosphate carboxylase in cyanobacterial heterocysts. *J Bacteriol*
 560 **175**: 7301-7306.
- 561 Maddocks, S.E., and Oyston, P.C. (2008) Structure and function of the LysR-type
 562 transcriptional regulator (LTTR) family proteins. *Microbiology* **154**: 3609-3623.
- 563 Mohamed, A., and Jansson, C. (1989) Influence of light on accumulation of
 564 photosynthesis-specific transcripts in the cyanobacterium *Synechocystis* 6803. *Plant*
 565 *Mol Biol* **13**: 693-700.
- 566 Nierzwicki-Bauer, S.A., Curtis, S.E., and Haselkorn, R. (1984) Cotranscription of genes
 567 encoding the small and large subunits of ribulose-1,5-bisphosphate carboxylase in
 568 the cyanobacterium *Anabaena* 7120. *Proc Natl Acad Sci USA* **81**: 5961-5965.
- 569 Nishimura, T., Takahashi, Y., Yamaguchi, O., Suzuki, H., Maeda, S., and Omata, T.
 570 (2008) Mechanism of low CO₂-induced activation of the *cmp* bicarbonate transporter
 571 operon by a LysR family protein in the cyanobacterium *Synechococcus elongatus*
 572 strain PCC 7942. *Mol Microbiol* **68**: 98-109.
- 573 Olmedo-Verd, E., Muro-Pastor, A.M., Flores, E., and Herrero, A. (2006) Localized
 574 induction of the *ntcA* regulatory gene in developing heterocysts of *Anabaena* sp.
 575 strain PCC 7120. *J Bacteriol* **188**: 6694-6649.
- 576 Picossi, S., Flores, E. and Herrero, A. (2014) ChIP analysis unravels an exceptionally
 577 wide distribution of DNA binding sites for the NtcA transcription factor in a
 578 heterocyst-forming cyanobacterium. *BMC Genomics* **15**: 22.
- 579 Pollari, M., Ruotsalainen, V., Rantamaki, S., Tyystjarvi, E., and Tyystjarvi, T. (2009).
 580 Simultaneous inactivation of sigma factors B and D interferes with light acclimation
 581 of the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J Bacteriol* **191**: 3992-
 582 4001.
- 583 Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm,
 584 M., and Seraphin, B. (2001) The tandem affinity purification (TAP) method: a
 585 general procedure of protein complex purification. *Methods* **24**: 218-229.
- 586 Price, G.D., Badger, M.R., Woodger, F.J., and Long, B.M. (2008). Advances in
 587 understanding the cyanobacterial CO₂-concentrating-mechanism (CCM): functional
 588 components, Ci transporters, diversity, genetic regulation and prospects for
 589 engineering into plants. *J Exp Bo.* **59**: 1441-1461.
- 590 Price, G.D. (2011) Inorganic carbon transporters of the cyanobacterial CO₂
 591 concentrating mechanism. *Photosynth Res* **109**: 47-57.
- 592 Ramasubramanian, T.S., Wei, T.-F. and Golden, J.W. (1994) Two *Anabaena* sp. strain
 593 PCC 7120 DNA-binding factors interact with vegetative cell- and heterocyst-specific
 594 genes. *J Bacteriol* **176**: 1214-1223.
- 595 Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999) A
 596 generic protein purification method for protein complex characterization and
 597 proteome exploration. *Na. Biotechnol* **17**: 1030-1032.
- 598 Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y. (1979)
 599 Generic assignments, strain stories and properties of pure cultures of
 600 cyanobacteria. *J Gen Microbiol* **111**: 1-61.
- 601 Schaefer, M.R., and Golden, S.S. (1989) Differential expression of members of a
 602 cyanobacterial *psbA* gene family in response to light. *J Bacteriol* **171**: 3973-3781.
- 603 Schneider, K.L., Pollard, K.S., Baertsch, R., Pohl, A., and Lowe, T.M. (2006) The
 604 UCSC archaeal genome browser. *Nucleic Ac Res* **34**: D407-D410.
- 605 Valladares, A., Rodríguez, V., Camargo, S., Martínez-Nöel, G.M., Herrero, A., and
 606 Luque, I. (2011) Specific role of the cyanobacterial PipX factor in the heterocysts of
 607 *Anabaena* sp. strain PCC 7120. *J Bacteriol* **193**(5): 1172-1182.

608 Wang, H.L., Postier, B.L., and Burnap, R.L. (2004) Alterations in global patterns of
609 gene expression in *Synechocystis* sp. PCC 6803 in response to inorganic carbon
610 limitation and the inactivation of *ndhR*, a LysR family regulator. *J Biol Chem* **279**:
611 5739-5751.
612
613
614

For Peer Review Only

615 **Figure legends:**

616 **Fig. 1.** Growth of the *all3953* mutant. A, The growth rate constant ($\mu = \ln 2/t_d$, where t_d is
 617 the doubling time) was calculated from the increase of protein content determined in
 618 0.2 ml samples of cultures. The table shows the mean and standard deviation from 3
 619 independent cultures of each strain and condition. $\Delta all3953$ is strain CSS74; CSS77 is
 620 a control strain that carries the Sm/Sp-resistant determinant in a wild-type background.
 621 B, Samples of cultures were photographed after 5 days of incubation under the
 622 indicated conditions. ~~C, Growth in solid medium (no NaHCO₃) with 3 mM NH₄Cl (A),~~
 623 ~~17.6 mM NaNO₃ (N), or no combined nitrogen (-N). 5 μ l drops, containing 50, 5, 0.5 or~~
 624 ~~0.05 ng Chl, were spotted, incubated for 14 days and photographed.~~ HL, high light; LL,
 625 low light; HC, high carbon; LC, low carbon.

626

627 **Fig. 2.** Expression of *rbclXS* in the $\Delta all3953$ mutant and complemented strain.
 628 Northern analysis carried out with RNA from strains CSS77 (control) CSS74 ($\Delta all3953$)
 629 and CSS74C (CSS74 complemented) was isolated from cells grown with HC (0) and
 630 incubated for 1h (1) with LC. The membranes were hybridized with an internal
 631 fragment of the *rbcl* gene (upper panels) and, as a loading and transfer control, of the
 632 *mpB* gene (lower panels). Arrowheads point to the main transcripts detected with the
 633 *rbcl* gene probe (approximate sizes are indicated).

634

635 **Fig. 3.** Consensus All3953 binding sequence and *rbcl* promoter. A, The primary
 636 consensus motif based on 142 high confidence CSS57 ChIP-Seq peak sequences is
 637 shown with indication of the probability of occurrence of each base along the 22-nt
 638 sequence. W is A or T; Y is C or T; B is C, G or T. B, Structure of the *rbclXS* promoter
 639 region. The transcription initiation point of the operon (+1) and the -10 and -35 boxes
 640 (from Nierzwicki-Bauer *et al.*, 1984) are indicated in red. The NtcA-binding site

Formatted: Font: Helvetica

Formatted: Font: Not Bold

641 (GTAN₈TAC) is indicated in green, and the three putative binding sites for All3953 (Box
642 I, Box II and Box III) are indicated in blue.

643

644 **Fig. 4.** qRT-PCR analysis of the expression of selected photosynthesis and respiration-
645 related All3953 gene targets. Transcriptional response of the indicated genes to C_i
646 limitation in the control (CSS77) and $\Delta all3953$ mutant (CSS74) strains was
647 investigated. RNA was isolated from cells grown with 10 mM NaHCO₃-supplemented
648 medium bubbled with 1%CO₂ in air (0) incubated for 1 h (1) or 3 h (3) in NaHCO₃-free
649 medium bubbled with air. Bars represent the mean transcript levels (\pm standard
650 deviation) in three independent experiments.

651

Table 1. Results of the ChIP-Seq analysis of AII3953 binding to DNA at 3 h after C_i limitation

| | Binding regions found | Genes ascribed | Position of the binding region with respect to the gene | | |
|------------|-----------------------|----------------|---|----------|------------|
| | | | upstream | internal | downstream |
| Chromosome | 127 | 157 | 118 | 29 | 10 |
| Alpha | 10 | 13 | 8 | 4 | 2 |
| Beta | 3 | 3 | 0 | 3 | 0 |
| Gamma | 2 | 2 | 0 | 1 | 1 |
| Total | 142 | 175 | 126 | 37 | 13 |

*Some of the binding regions were ascribed to more than one gene (see Supp. Table 1).

For Peer Review Only

Table 2. Functional categories of the genes ascribed to the All3953 binding regions.

| Functional category | Number |
|--|--------|
| Amino acid biosynthesis | 5 |
| Biosynthesis of cofactors, prosthetic groups, and carriers | 4 |
| Cell envelope | 2 |
| Cellular processes | 5 |
| Central intermediary metabolism | 3 |
| DNA replication, recombination and repair | 3 |
| Energy metabolism | 4 |
| Other categories | 19 |
| Photosynthesis and respiration | 19 |
| Purines, pyrimidines, nucleosides and nucleotides | 2 |
| Regulatory proteins | 21 |
| Translation | 9 |
| Transport and binding proteins | 4 |
| Unknown and hypothetical proteins | 75 |

Table 3. All3953-binding regions assigned to photosynthesis and respiration-related genes.

| BR | NLQ | LOC | GENE | POSITION | FUNCTION [†] | ST | START | SEQ |
|-----|--------|---------|-----------------|-----------------|--|----|---------|-------------------------|
| 7 | 172,84 | 239694 | <i>alr0223*</i> | upstream | NADH dehydrogenase subunit 1; NdhA | + | 239712 | AGGTATTAGTTTTAACTAATGTT |
| 12 | 38,58 | 398948 | <i>alr0348</i> | upstream close | NADH dehydrogenase subunit 4; NdhD | + | 398912 | AACAATTCCTTAATATAATGTA |
| 19 | 62,81 | 859596 | <i>all0737</i> | upstream | thioredoxin reductase | - | 859596 | ATTCATAAAAGCGTTTTATATC |
| 25 | 238,45 | 997693 | <i>all0868</i> | upstream | CO ₂ concentrating mechanism protein CcmK | + | 997698 | ATGTATAAGTTTTATTAATATG |
| | | | <i>alr0869</i> | upstream | NADH dehydrogenase subunit 5 | | | |
| 27 | 66,90 | 1175262 | <i>alr1004*</i> | upstream | alanine--glyoxylate aminotransferase | + | 1175259 | GTATATAGGCGATCATTATGGC |
| 31 | 36,15 | 1433154 | <i>alr1216</i> | upstream | photosystem II 12 kD extrinsic protein PsbU | + | 1433147 | AAATATTGTGAGCATTAATAAG |
| 34 | 385,11 | 1547013 | <i>all1304*</i> | upstream | bicarbonate transporter | - | 1546894 | GTGCATTTGCAATAGTTATTAT |
| 35 | 31,96 | 1620674 | <i>all1365</i> | downstream | cytochrome CytM | + | 1620645 | CGTAATAAATTTTAATCATCAT |
| 37 | 63,26 | 1785455 | <i>alr1524*</i> | upstream | RbcL | + | 1785446 | ACTTATGCCATTTCTTGATATA |
| 38 | 204,06 | 1843221 | <i>alr1576</i> | upstream far | dehydrogenase subunit | - | 1843218 | AGTAATAACTGCTACTTATTAC |
| 61 | 29,78 | 3499605 | <i>alr2873</i> | internal 3' end | possible glycerate kinase | - | 1843218 | CAAATAAAGTCTAATTTTC |
| 86 | 170,25 | 4499861 | <i>alr3727</i> | upstream | photosystem II protein D1 (<i>psbAII</i>) | + | 4499883 | GTATATATATTTTAGTAATATT |
| 91 | 155,06 | 4693128 | <i>all3891*</i> | upstream | flavoprotein (<i>flv1A</i>) | + | 4693099 | ATTTATAAGTTTTACTTAAGCT |
| 98 | 166,48 | 4993006 | <i>all4148</i> | upstream | ferredoxin I | - | 4993375 | AACCATAAATTTTCTAATAAC |
| 99 | 61,47 | 4999503 | <i>alr4156*</i> | upstream | NADH dehydrogenase subunit 5; NdhF | + | 4999503 | AAAGATAAATTTGCCTTATTTA |
| 108 | 163,25 | 5332512 | <i>all4446*</i> | upstream | flavoprotein (<i>flv4</i>) | - | 5332500 | AATAATAAATTTACTAATAAA |
| 112 | 233,71 | 5489698 | <i>alr4592*</i> | upstream | photosystem II protein D1 (<i>psbAIII</i>) | + | 5489979 | CTATATAGTTTTACTCATATT |
| 121 | 51,04 | 6151146 | <i>alr5154</i> | upstream | photosystem I core protein A1 | + | 6151133 | GGACATAAGTTTTACGAATTGT |

*Genes whose expression has been studied by qRT-PCR

[†]Functions are as specified in cyanobase (<http://genome.microbedb.jp/cyanobase/Anabaena>).

BR: binding region, NLQ: -logQvalue, LOC: chromosome location of the midpoint of the binding region, ST: DNA strand, START: chromosome location of the 5' end of the putative binding sequence of All3953 (SEQ).

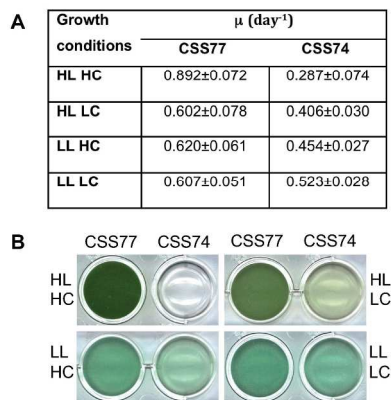
Picossi *et al.*, Fig. 1

Figure 1
262x350mm (300 x 300 DPI)

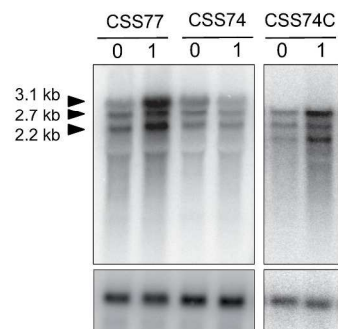
Picossi *et al.*, Fig. 2

Figure 2
262x350mm (300 x 300 DPI)

Picossi *et al.*, Fig. 3

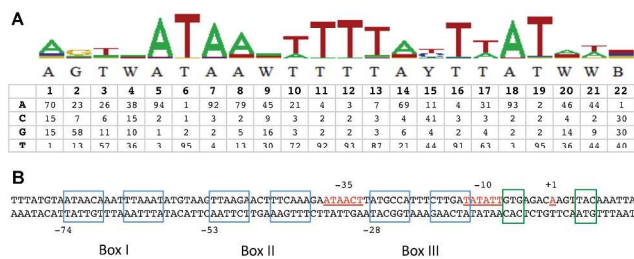


Figure 3
 297x420mm (300 x 300 DPI)

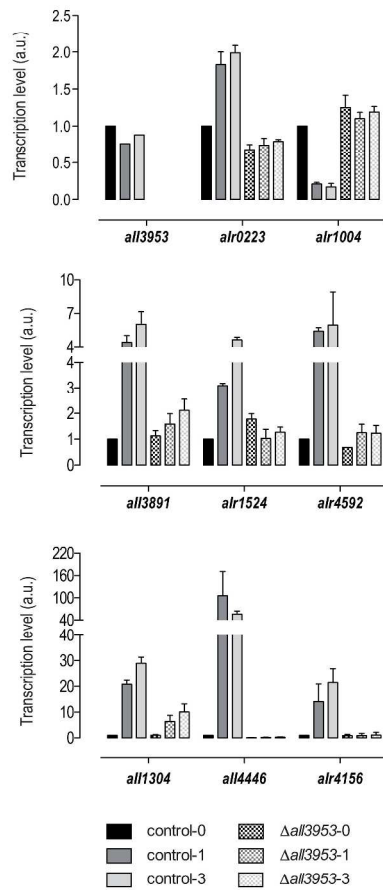


Figure 4
297x420mm (300 x 300 DPI)