Genetic responses to carbon and nitrogen availability in *Anabaena*

Antonia Herrero* and Enrique Flores

Instituto de Bioquímica Vegetal y Fotosíntesis, CSIC and Universidad de Sevilla,
Américo Vespucio 49, E-41092 Seville, Spain

*For correspondence: herrero@ibvf.csic.es

Outline:

Originality-significance statement

Heterocyst-forming cyanobacteria such as *Anabaena* can acclimate to thrive in environments with different sources of (mainly) inorganic nitrogen and different levels of inorganic carbon. The responses to C and N supplies are orchestrated by global (NtcA and PacR) and pathway-specific transcription factors, and they involve regulation not only of assimilatory pathways but also of other physiological processes such as defense against oxidative damage. These regulatory responses are essential for the cyanobacterial mode of living as oxygenic phototrophs.

Summary

Heterocyst-forming cyanobacteria are filamentous organisms that perform oxygenic photosynthesis and CO_2 fixation in vegetative cells and nitrogen fixation in heterocysts, which are formed under deprivation of combined nitrogen. These organisms can acclimate to use different sources of nitrogen and respond to different levels of CO_2. Following work mainly done with the best studied heterocyst-forming cyanobacterium, *Anabaena*, here we summarize the mechanisms of assimilation of ammonium, nitrate, urea and N_2, the latter involving heterocyst differentiation, and describe aspects of CO_2 assimilation that involves a carbon concentration mechanism. These processes are subjected to regulation establishing a hierarchy in the assimilation of nitrogen sources – with preference for the most reduced nitrogen forms– and a dependence on sufficient carbon. This regulation largely takes place at the level of gene expression and is exerted by a variety of transcription factors, including global and pathway-specific transcriptional regulators. NtcA is a CRP-family protein that adjusts global gene expression in response to the C-to-N balance in the cells, and PacR is a LysR-family transcriptional regulator (LTTR) that extensively acclimates the cells to oxygenic phototrophy. A cyanobacterial-specific transcription factor, HetR, is involved in heterocyst differentiation, and other LTTR factors are specifically involved in nitrate and CO_2 assimilation.

Introduction

Cyanobacteria are a phylogenetically coherent group of bacteria characterized by their ability to perform oxygenic photosynthesis (Stanier and Cohen-Bazire, 1977; Giovannoni *et al.*, 1988). Nonetheless, some cyanobacterial non-photosynthetic mutants have evolved as obligate endosymbionts (Zehr *et al.*, 2017). Although non-photosynthetic bacteria such as Melainabacteria and Sericytochromatia have been included into the phylum "cyanobacteria" (Soo *et al.*, 2017), in our opinion they constitute different groups of organisms (see also Pakrasi and Zehr, 2017). Therefore, photoautotrophy is the characteristic mode of growth of cyanobacteria, although numerous strains in this phylum can use sugars supporting either chemoheterotrophic or photoheterotrophic growth (Rippka *et al.*, 1979). Regarding nitrogen (N) sources, cyanobacteria are generally able to assimilate with high affinity inorganic N compounds and a simple organic compound such as urea, and many strains can fix atmospheric nitrogen (N_2) (Flores and Herrero, 2005). These properties make many cyanobacteria able to grow in oligotrophic environments. Nonetheless, the genomes of cyanobacteria encode numerous transporters for more complex organic compounds including amino acids, which widens the trophic options of these organisms and is consistent with the idea that microbes are prepared to extensively take up substrates that become available (Hobbie and Hobbie, 2013).

The enzyme that catalyzes N_2 fixation, nitrogenase, is very sensitive to oxygen, making protection of nitrogenase against oxygen an important physiological issue in cyanobacteria (Gallon, 1981). Many strains separate nitrogen fixation and oxygenic photosynthesis in time, restricting the former to the dark period in diel cycles, and many filamentous cyanobacteria separate the two processes spatially restricting N_2 fixation to specialized cells (Fay, 1992). This is particularly the case of heterocyst-forming strains, in which N_2 fixation under oxic conditions is restricted to differentiated cells called heterocysts. The heterocyst-forming cyanobacteria are a monophyletic group of organisms (Giovannoni *et al.*, 1988; Shih *et al.*, 2013), implying that the heterocyst evolved only once, likely over 2,000 million years ago (Tomitani *et al.*, 2006). There are two major groups of heterocyst-forming cyanobacteria, those in which cell division always takes place perpendicular to the long axis of the filament (Section IV cyanobacteria or Nostocales) and those in which cell division can take place also in a different angle producing branched filaments (Section V cyanobacteria or Stigonematales) (Rippka *et al.*, 1979; Castenholz, 2001). The existence of these two

groups of heterocyst-forming cyanobacteria is supported by phylogenetic analysis (Shih *et al.*, 2013; Schirrmeister *et al.*, 2015; Mareš, 2017; Ponce-Toledo *et al*., 2017). The heterocyst-forming strains for which most physiological and molecular information is available belong to the genera *Anabaena* and *Nostoc*, which are included in Section IV and in some taxonomic classifications are distinguished by the ability to form hormogonia (small motile filaments frequently made of small cells) in *Nostoc* spp. but not in *Anabaena* spp. (Rippka *et al.*, 1979). Strains ascribed to these genera are however interspersed in the phylogenetic tree of Section IV cyanobacteria (Shih *et al.*, 2013; Schirrmeister *et al.*, 2015). Studying a few model strains has permitted to advance in the understanding of some basic aspects of this group of organisms, but it should be noted that some traits may be strain-specific. The strain for which more molecular information is available is *Anabaena* sp. PCC 7120 (hereafter, *Anabaena*). This strain is sometimes referred to as *Nostoc*, but it does not form evident hormogonia and behaves as a planktonic organism, which in the botanical literature is associated to *Anabaena* spp. rather than to *Nostoc* spp.

Here we will review the genetic responses of Section IV cyanobacteria –mainly *Anabaena*– to the availability of sources of the major bioelements carbon and nitrogen. Whereas the most abundant source of C is atmospheric CO_2, the levels of CO_2 that reach a cyanobacterial filament may vary depending on physicochemical parameters such as temperature-dependent solubility of CO_2 in water or pH-dependent accumulation of bicarbonate and carbonate. (Unfortunately, the regulatory effect of sugars that can be assimilated by many Section IV cyanobacteria [Rippka *et al.*, 1979] has not been generally investigated.) Nitrogen, on the other hand, can be assimilated in several different chemical forms. Whereas N_2 will be generally available as a major component of the atmosphere, its fixation is restricted to conditions in which no sufficient combined N is available. Ammonium, nitrate, nitrite and urea are common N sources that may be found in natural habitats in a range of concentrations up to micromolar, and the cyanobacteria respond to changes in the concentrations of these nutrients through mechanisms that will be described below. (Lab experiments frequently require the use of millimolar concentrations of nutrients, but the uncovered regulatory mechanisms appear to be relevant under environmental conditions.) Finally, as we shall discuss, the assimilation of C and N are linked with each other, and the relative abundance or scarcity of one element is perceived by the cells depending on the availability of the other element.

Ammonium assimilation and NtcA-dependent N control

Cyanobacteria readily utilize ammonium, which is a preferred N source. Thus, when ammonium is available in sufficient amounts, it determines repression of genes encoding proteins for the assimilation of alternative N sources, which are expressed upon ammonium deprivation. Figure 1 presents a scheme of the pathways for assimilation of combined N and inorganic C in *Anabaena*.

Assimilation of ammonium

Ammonium ions are easily assimilated because ammonium is directly incorporated into carbon skeletons to produce organic N compounds. Solutions of ammonium always contain ammonia (pK_a $[NH_4^+/NH_3]$ = 9.3), and ammonia readily permeate biological membranes. Nonetheless, ammonium translocators of the Amt family that act as ammonium scavengers (Wacker *et al.* 2014) are widely distributed in bacteria, permitting the incorporation into the cells of ammonium that may be found at submicromolar concentrations in the environment. The Amt proteins form trimers (Khademi *et al.*, 2004), and the *Anabaena* genome contains three *amt* genes clustered together raising the possibility that their encoded products form heterotrimers (Paz-Yepes *et al.*, 2008). However, although the transcription of the three genes is similarly increased in response to ammonium withdrawal, they are expressed at disparate levels, with *amt1* being expressed at the highest levels (ca. 10:100:1 for *amt4, amt1, amtB*, respectively; Flaherty *et al.*, 2011). This is consistent with the known role of Amt1 as the main ammonium transporter in cyanobacteria (Montesinos *et al.*, 1998). It is possible that the Amt trimer of *Anabaena* is mainly composed of Amt1 with small amounts of Amt4 and AmtB modulating its activity.

Ammonium is incorporated into organic compounds by the glutamine synthetase-glutamate synthase pathway, whose operation was elegantly demonstrated by Wolk *et al.* (1976) following the assimilation of ^{13}N-labeled N_2 in *Anabaena cylindrica*. In this pathway, ammonium is incorporated into glutamate producing glutamine in an ATP-dependent reaction catalyzed by glutamine synthetase (GS), and the amido group of glutamine is then transferred to 2-oxoglutarate (2-OG) by glutamine:oxoglutarate amido transferase (GOGAT) in a reaction that requires reducing power. Most cyanobacterial GOGAT enzymes, including those from *Anabaena* strains, use reduced ferredoxin as an

electron donor (Martín-Figueroa *et al.*, 2000), thus linking the assimilation of ammonium to photosynthesis.

In whole filaments of *Anabaena*, whereas the expression of the *glnA* gene encoding GS is increased in response to ammonium withdrawal, expression of *glsF* encoding GOGAT remains unaltered (Martín-Figueroa *et al.*, 2000). However, *glsF* is repressed in the heterocysts, in which expression of *glnA* is increased as compared to that in the vegetative cells. Whereas, to the best of our knowledge, the molecular basis of *glsF* repression has not been investigated, expression of *glnA* in *Anabaena* has been shown to take place from a complex promoter region (Tumer *et al.*, 1983), in which one promoter activated by the transcription factor NtcA (see below) is specifically responsible for expression in the heterocysts (Valladares *et al.*, 2004). Cyanobacterial GS is further regulated at the post-transcriptional level by specific proteinaceous inactivating factors (García-Domínguez *et al.*, 1999). In *Anabaena*, the 67-amino acid residue inactivating factor IF7A is encoded by the *gifA* gene, which is repressed by NtcA under N-limiting conditions and in the heterocysts, thus ensuring that GS is not inactivated under N deficiency or in the N_2-fixing cells (Galmozzi *et al.*, 2010).

NtcA-dependent N control

The negative regulation of pathways for the assimilation of alternative N sources exerted by ammonium takes place mainly at the level of gene expression. In *Anabaena*, studied pathways that are subjected to repression by ammonium include the nitrate assimilation system (nitrate and nitrite transport and intracellular reduction to ammonium), urea uptake, ammonium translocation and incorporation into glutamine by GS, and N_2 fixation by nitrogenase (including heterocyst differentiation; see below). At the molecular level, pathway regulation is mainly exerted through transcriptional regulation exerted by the NtcA protein.

NtcA is a transcriptional regulator of universal distribution in the cyanobacterial phylum that orchestrates the circuits of N control. It belongs to the CRP (cAMP receptor protein) family of transcriptional regulators and, similar to other members of this family, consists of an effector-binding N-terminal domain with a β-roll fold, a long dimerization helix, and a C-terminal helix-turn-helix motif for interaction with DNA (Herrero *et al*., 2001; Zhao *et al*., 2010). NtcA binds as a dimer to palindromic DNA sites including the sequence signature $GTAN_8TAC$ (Luque *et al*., 1994; Picossi *et al*., 2014). NtcA binding

to the promoter region of multiple genes related to N assimilation has indeed been detected (e.g., Vázquez-Bermúdez *et al*., 2002; Valladares *et al*., 2008). Besides detailed studies of the regulation by NtcA of a number of genes involved in N assimilation and heterocyst differentiation, a global study of NtcA targets by chromatin immunoprecipitation in *Anabaena* three hours after N step-down revealed more than 2,000 direct targets ascribed to 2,153 genes. The NtcA targets were located in the chromosome and the six plasmids of *Anabaena*, and included genes related to N scavenging and assimilation, but also genes ascribed to other diverse functional categories, among them genes with a predicted regulatory function, suggesting cascades of NtcA-dependent regulation (Picossi *et al*., 2014). Thus, NtcA is indeed a global regulator of cyanobacterial physiology.

NtcA can act as a transcriptional activator or repressor (Luque and Forchammer, 2008). In the most commonly found NtcA-activated promoter, an NtcA-binding site is found upstream from, separated by ca. 22 nucleotides, a -10 promoter determinant, thus conforming to the Class II bacterial-activated promoter. NtcA senses the cellular C-to-N balance by binding 2-OG (Vázquez-Bermúdez *et* al., 2002; Zhao *et al*., 2010). Binding of 2-OG to NtcA has a positive, although variable, effect on NtcA binding to DNA. However, at least *in vitro*, NtcA binding to DNA in the absence of 2-OG can take place. (This contrasts with the requirement of the effector, cAMP, for CRP interaction with DNA at bacterial CRP-regulated promoters.) Beyond this effect on binding to DNA, NtcA and 2-OG are stringently required for the isomerization to the transcriptional Open Complex and, thus, for transcript production at Class II NtcA-activated promoters as studied *in vitro* using the cloned *Anabaena* RNA polymerase (Valladares *et al*., 2008).

Additionally, other types of NtcA-activated promoters have been identified, including Class I promoters, in which an NtcA-binding site is located further upstream from the -10 determinant, and promoters with degenerated binding sites that involve NtcA co-activators. The relative affinity of NtcA for the different promoters, together with the availability of co-activators, are important determinants of the order of promoter activation upon ammonium withdrawal, conceivably playing an important role in the hierarchy of assimilation of alternative nitrogen sources.

Assimilation of combined N sources alternative to ammonium

Other than ammonium, cyanobacteria preferentially utilize inorganic N sources, being nitrate and nitrite excellent nitrogenous nutrients, although some amino acids, including arginine and glutamine, and, especially, urea can also be efficiently assimilated by some cyanobacteria such as *Anabaena*.

Nitrate assimilation

Nitrate can be used as a N source by most heterocyst-forming cyanobacteria. Nitrate is typically taken up by a high-affinity transporter that concentrates nitrate inside the cells to a level that makes it usable by nitrate reductase, which catalyzes the 2-electron reduction of nitrate producing nitrite. Nitrite is then used by nitrite reductase that catalyzes a 6-electron reduction producing ammonium. Cyanobacterial nitrate and nitrite reductases are metalloenzymes that use reduced ferredoxin as electron donor, thus liking nitrate assimilation to photosynthesis (Flores *et al.*, 2005).

As in many other microorganisms, the genes encoding the nitrate assimilation system are clustered together in the genome of *Anabaena* (Cai and Wolk, 1997). The nitrate assimilation structural genes form an operon in which the genes are arranged in a very conserved order (5' to 3'): *nirA* (nitrite reductase), transporter genes, *narB* (nitrate reductase) (Fig. 2). In *Anabaena*, the nitrate uptake system (that can also take up nitrite) is an ABC transporter encoded by genes determining a periplasmic solute-binding protein (*nrtA*) that is distinctly anchored to the cytoplasmic membrane (Maeda *et al*., 2000), a transmembrane domain protein (*nrtB*), a nucleotide-binding domain protein that has a C-terminal extension homologous to NrtA (*nrtC*), and a conventional nucleotide-binding domain protein (*nrtD*) (Frías *et al.*, 1997). To conform the known structure of ABC transporters, NrtB should form a dimer, and one unit of each NrtC and NrtD constitute the nucleotide-binding module. The NrtC protein may have, in addition to its role as an ATPase, a regulatory role mediated by its NrtA-like domain (see below). Interestingly, some heterocyst-forming cyanobacteria such as *Nostoc punctiforme* express a Major Facilitator Superfamily (MFS) nitrate/nitrite transporter instead of the ABC transporter (Aichi *et al.*, 2006).

Nitrate assimilation is regulated at the transcriptional and post-transcriptional levels. In the latter, nitrate uptake is inhibited by ammonium in a process that affects the nitrate (nitrite) transporter. As has been shown for some unicellular cyanobacteria (Kobayashi *et al.*, 2005), the NrtA-like domain of NrtC appears to be required for

253 ammonium-promoted inhibition of nitrate uptake. Additionally, the P_{II} protein (*glnB* gene
254 product –see below) is required for this inhibitory effect to take place (Paz-Yepes *et al.*,
255 2009). An appealing hypothesis is that the non-phosphorylated P_{II} protein (which is the
256 form that accumulates in the presence of ammonium) inhibits the transporter but, to the
257 best of our knowledge, a physical interaction between NrtC and P_{II} has not yet been
258 demonstrated.
259 The expression of the *nirA* operon is subjected to regulation exerted by the global
260 N-control system, and is transcribed from a Class II NtcA-activated promoter (Frías *et*
261 *al.*, 1997, 2000). Additionally, in *Anabaena*, high expression of the *nirA* operon and
262 production of full nitrate and nitrite reductase activities require a number of genes that
263 are clustered with those of the *nirA* operon. Both NtcB (a LysR family transcriptional
264 regulator [LTTR]) and CnaT (a putative glycosyl transferase) boost expression (Frías *et*
265 *al.*, 2000; Frías *et al.*, 2003). Whereas NtcB binds to the *nirA* promoter region upstream
266 of the NtcA-binding site, CnaT acts through an unknown mechanism. The *ntcB* gene is
267 expressed also from a Class II NtcA-activated promoter (Frías *et al.*, 2000). Further, the
268 products of two genes, *nirB* and *narM*, are needed to attain high levels of activity of nitrite
269 reductase and nitrate reductase, respectively, and NirA-NirB and NarB-NarM protein-
270 protein interactions have been found to take place (Frías and Flores, 2010, 2015). High
271 levels of expression of the operon are observed only if nitrate or nitrite (which can be
272 considered physiological inducers) is available. This requirement for nitrate (nitrite) is
273 however not observed in *nirA* or *nirB* mutants, which express the operon at high levels in
274 the absence of the inducer (Frías and Flores, 2010), suggesting a role of the NirA/NirB
275 complex as a direct or indirect repressor of the operon (Fig. 2). Whatever the mechanism,
276 the requirement for nitrate (nitrite) ensures that the *nirA* operon will not be expressed
277 under conditions of severe N deficiency that permit heterocyst differentiation and N_2
278 fixation to proceed (see below).
279
280 <u>Assimilation of organic N</u>
281
282 The genome of *Anabaena* encodes numerous transporters that can mediate the uptake of
283 organic nitrogenous compounds including amino acids and urea. These are mainly ABC
284 transporters that exhibit K_s values in the range of the environmental concentrations of
285 their substrates, from 1 to 50 μM for most amino acids except acidic amino acids for
286 which the K_s values are substantially higher (Montesinos *et al.*, 1995; Pernil *et al.*, 2015),

and about 0.1 μM for urea (Valladares *et al.*, 2002). Thus, assimilation of organic N is likely important in the physiology of *Anabaena*, and the ability of this cyanobacterium to use arginine, glutamine and urea as N sources has been demonstrated with mutants unable to fix N_2 under oxic conditions (Herrero and Flores, 1990; Valladares *et al.*, 2002; Burnat *et al.*, 2014). Although the mechanisms of assimilation of organic N have not been investigated in detail, some relevant pathways can be predicted or have been approached experimentally. Thus, (i) glutamine utilization likely involves GOGAT (described above); (ii) arginine utilization appears to take place through two pathways, including an arginine decarboxylase pathway that synthesizes polyamines (Burnat *et al.*, 2018) and an arginine → proline → glutamate pathway (Burnat and Flores, 2014); and (iii) urea utilization involves urease that hydrolyzes urea to CO_2 and two molecules of ammonium (Valladares *et al.*, 2002).

Bacterial urease is a Ni^{2+}-containing enzyme composed of three subunits, α (UreC, which bears the catalytic site), β (UreB) and γ (UreA), and maturation of urease requires four additional proteins, UreD, UreE, UreF and UreG (Boer *et al.*, 2014). In *Anabaena*, these proteins are encoded in two gene clusters: (i) *alr3666* (*ureD*)-*alr3667* (*ureA*)-*alr3668* (*ureB*)-*alr3670* (*ureC*) and (ii) *alr0733* (*ureE*)-*alr0734* (*ureF*)-*alr0735* (*ureG*) (Kaneko *et al.*, 2001), and a mutant lacking urease activity has been isolated that maps at *ureG* (Valladares *et al.*, 2002). Whereas the urease activity has been found to be expressed constitutively, urea uptake and the *urtABCDE* genes encoding the ABC transporter for urea are subjected to N control by NtcA ensuring that the ability to take up urea actively is expressed only under N deficiency (Valladares *et al.*, 2002).

N_2 fixation and the heterocyst

When no source of combined N is available, *Anabaena*, *Nostoc* and phylogenetically-related cyanobacteria enter into a developmental process in which some vegetative cells of the filament differentiate into heterocysts, which are cells specialized for the fixation of N_2.

Nitrogenase

The enzymatic complex that catalyzes N_2 fixation is the product the three structural genes, *nifD* and *nifK* encoding nitrogenase, which holds the Fe-S cluster known as the P cluster

and the Fe-Mo cofactor in which the nitrogen fixation reaction takes place, and *nifH* encoding nitrogenase reductase, which holds an Fe-S cluster, binds Mg^{2+}-ATP and transfers electrons to nitrogenase (Rubio and Ludden, 2008). In turn, electrons are received from an electron carrier such as ferredoxin or flavodoxin. Notably, alternative nitrogenases that use an Fe-V cofactor or an Fe-Fe cofactor are known (Mus *et al.*, 2018). As in most N_2-fixing bacteria, in *Anabaena* the *nifHDK* genes form an operon and are clustered with other genes encoding nitrogenase maturation proteins and electron donors including ferredoxins (Flores *et al.*, 2015). Whereas most heterocyst-forming cyanobacteria contain only the nitrogenase that is confined to heterocysts (Fe-Mo nitrogenase), *Anabaena variabilis* strain ATCC 29413 is unique among well-characterized cyanobacteria in that it expresses three nitrogenases (Thiel and Pratte, 2014). In addition to the Fe-Mo nitrogenase expressed in heterocysts, this cyanobacterium expresses an Fe-Mo nitrogenease in vegetative cells under anoxic conditions and an Fe-V nitrogenase in heterocysts under conditions of Mo deficiency.

Heterocyst differentiation and function

Cell envelope

Heterocyst formation is first noticed by the production of an extra cell wall envelope that is deposited outside of the pre-existing outer membrane. A polysaccharide layer (HEP) is first produced that consists of repeating units of an oligosaccharide that has a tetrasaccharide backbone (→ 3 mannose 1 → 3 glucose 1 → 3 glucose 1 → 3 glucose 1 →) with some of its glucosyl residues substituted by some saccharides (Cardemil and Wolk, 1981). A cluster of *hep* genes encoding proteins involved in the production of this polysaccharide, as well as some non-clustered *hep* genes have been identified (Huang *et al.*, 2005; Wang *et al.*, 2007). Although the mechanism of production is unknown, the presence in *Anabaena* of several *hep* genes encoding proteins homologous to lipopolysaccharide (LPS) biosynthesis proteins suggests that the heterocyst-envelope polysaccharide may be a modified form of LPS or its production may follow a path similar to that of LPS (Huang *et al.*, 2005). Then, heterocyst-specific glycolipids (HGL), which are long-chain polyhydroxy alcohols with a hexose head group, are produced and deposited beneath the polysaccharide layer (Awai *et al.*, 2009). A cluster of *hgl* genes and some non-clustered genes required for the production of the HGL layer have been

355 identified (Fan *et al.*, 2005; Awai and Wolk, 2007; Fiedler *et al.*, 1998). HGL are
356 synthesized by polyketide synthases and at least one glycosyl transferase, and exported
357 by an ATP-driven secretion system formed by DevA (ABC exporter NBD), DevC (ABC
358 transporter TMD), DevB (fusion membrane protein) and HgdD (TolC-like outer-
359 membrane protein) (Staron *et al.*, 2011). It is thought that the HGL layer restricts
360 diffusion of gases, thus representing a barrier to entry of oxygen, and that the HEP layer
361 protects HGL from damage (Awai *et al.*, 2009). Some air has nonetheless to enter the
362 heterocyst to provide the substrate (N_2) for nitrogenase, and hence the simultaneous entry
363 of some O_2 is unavoidable.
364
365 *Intracellular membrane systems – Photosynthesis and respiration*
366
367 Importantly, photosystem II-associated photolysis of water is dismantled in the
368 heterocysts avoiding the intracellular production of O_2 (reviewed by Magnuson and
369 Cardona, 2016). Indeed, heterocyst formation involves a large reorganization of
370 intracellular membranes, which in vegetative cells are found peripherally and in the
371 heterocysts are largely located around the cell poles constituting the so-called *honeycomb*
372 membranes (Lang and Fay, 1971). These membranes are sites of respiration that
373 contribute to reduce the O_2 that enters the heterocyst, and terminal respiratory oxidases
374 exclusively expressed in the heterocysts are known in *Anabaena* (Valladares *et al.*, 2003,
375 2007). O_2 is further reduced in the heterocyst's cytoplasm by flavodiiron proteins
376 specifically expressed in these differentiated cells (Ermakova *et al.*, 2013, 2014), and
377 possibly by L-lactate oxidase (Hackenberg *et al.*, 2011). Respiration in the heterocysts
378 requires a source of reductant that is received from the adjacent vegetative cells. In
379 *Anabaena*, sucrose appears to be the main source of intercellularly transferred reductant,
380 since inactivation of an invertase, InvB, expressed specifically in the heterocysts blocks
381 diazotrophic growth (López-Igual *et al.*, 2010; Vargas *et al.*, 2011; reviewed in Kolman
382 *et al.*, 2015). Invertases irreversibly cleave sucrose into glucose and fructose, and the
383 oxidative pentose phosphate pathway mediates sugar catabolism in the heterocysts
384 (Summers *et al.*, 1995). Reductant is made available to nitrogenase by a heterocyst-
385 specific ferredoxin (FdxH), which is reduced by ferredoxin-NADPH oxidoreductase
386 (FNR) or photosystem I (see discussion in Magnuson and Cardona, 2016). Photosystem
387 I-dependent reduction of FdxH requires donation of electrons by NAD(P)H into an
388 electron transport chain involving NAD(P)H dehydrogenase and cytochrome *b_6-f*

complexes. Energy conservation associated with this electron transport chain and with the operation of respiratory oxidases results in production of ATP that is also needed for nitrogenase function (Magnuson and Cardona, 2016; Valladares *et al.*, 2007; Stal, 2017).

Ammonium assimilation – Intercellular metabolite exchange

In heterocysts, ammonium produced by nitrogenase is incorporated into glutamate by GS producing glutamine (Wolk *et al.*, 1976). The heterocysts, however, lack (or have low levels of) GOGAT, which is present at high levels only in vegetative cells (Thomas *et al.*, 1977; Martín-Figueroa *et al.*, 2000). Hence, glutamate for the operation of GS has to be transferred from vegetative cells to heterocysts, which in turn donate glutamine to vegetative cells (Wolk *et al.*, 1976). The heterocysts also synthesize other amino acids including aspartate and arginine (Wolk *et al.*, 1976), which are used to produce cyanophycin (multi-L-arginyl-poly [L-aspartic acid]), a dynamic polymeric N reserve that accumulates conspicuously at the heterocyst poles (Lang *et al.*, 1972). Cyanophycin is synthesized non-ribosomally by cyanophycin synthetase and degraded sequentially by cyanophycinase, which releases β-aspartyl-arginine, and isoaspartyl dipeptidase, which splits the dipeptide making aspartate and arginine available for metabolism. The dipeptidase is present in vegetative cells at higher levels than in heterocysts, implying that β-aspartyl-arginine is transferred (in addition to glutamine) as a N vehicle from heterocysts to vegetative cells (Burnat *et al.*, 2014). Thus, the reduced C/fixed N exchange that takes place between vegetative cells and heterocysts in the diazotrophic filament appears to include sucrose and glutamate moving from vegetative cells to heterocysts and glutamine and β-aspartyl-arginine moving from heterocysts to vegetative cells (Fig. 3). Nonetheless, other (perhaps many) metabolites may also be translocated from vegetative cells to heterocysts, including alanine, which may serve as a source of reductant in the heterocysts (Jüttner, 1983; Pernil *et al.*, 2010), and some reduced forms of sulfur (Giddings *et al.*, 1981). Metabolite exchange between vegetative cells and heterocysts appears to take place by diffusion through proteinaceous structures known as septal junctions that are located in the intercellular septa and join the adjacent cells in the filament (Nürnberg *et al.*, 2015; Flores *et al.*, 2016; Nieves-Morión *et al.*, 2017).

Regulation of heterocyst distribution in the filament

On a different physiological level, heterocysts transfer regulators to the vegetative cells. Indeed, heterocyst distribution along the filament is not random, and in *Anabaena* growing under N_2-fixing conditions, the heterocysts are semi-regularly spaced, separated by stretches of ca. 10-15 vegetative cells. For establishing this heterocyst distribution pattern, the intercellular movement of heterocyst differentiation inhibitors plays an important role. Among these inhibitors, the PatS morphogen is produced at early stages of differentiation in the specific cells undertaking the process and exported to neighboring cells, thus preventing their differentiation (Yoon and Golden, 1998; Corrales-Guerrero *et al*., 2014). At advanced stages of differentiation, other factors, such as HetN may promote lateral inhibition, thus contributing to maintain the specific heterocyst distribution pattern (Callahan and Buikema, 2001).

Transcriptional regulation of heterocyst differentiation

Heterocyst differentiation requires the action of the products of multiple structural and regulatory genes, which have been identified through the isolation of mutants impaired in the differentiation process and, more recently, by global gene expression studies. Many of the genes needed for heterocyst differentiation are activated at specific stages during the process, only in the differentiating cells (see Herrero *et al*., 2013; Flores *et al*., 2018).

Two transcription factors, NtcA and HetR, are essential for heterocyst formation, so that strains bearing inactivated *ntcA* or *hetR* genes do not show any hint of differentiation upon N step-down. As NtcA, HetR is a DNA-binding factor, and the structure of a HetR dimer conforming a DNA-binding domain that includes the N-terminal helix-turn-helix motifs of both subunits has been published (Kim *et al*., 2011). More recently, HetR tetramers have been detected to occur *in vivo* (Valladares *et al*., 2016). The HetR tetramers are more abundant at intermediate times of heterocyst differentiation, coincident with the period of maximum HetR expression and HetR-dependent activation of transcription, and they are undetected in a mutant that does not form heterocysts. When differentiation is completed, HetR tetramers are detected in heterocysts but not in vegetative cells. Moreover, HetR co-purifies with RNA polymerase from *Anabaena* predominantly in the tetrameric form. These observations support the idea that during heterocyst differentiation the tetrameric HetR can represent an active

455 form of the regulator (Valladares *et al*., 2016). Indeed, besides the dimer, complexes of
456 HetR tetramers with DNA have also been detected *in vitro* (Kim *et al*., 2013).
457 Regarding the sequence of gene regulation events promoting heterocyst
458 differentiation upon N step-down, a rise in the cellular levels of 2-OG effects activation
459 of the low levels of NtcA already present in the filament, triggering activation of the
460 expression of the genes *hetR* and *ntcA* in a mutually-dependent and autoregulatory
461 manner (Muro-Pastor *et al*., 2002). Activation of *hetR* by NtcA appears to take place
462 through NrrA (Ehira and Ohmori, 2006), a response regulator that is expressed from a
463 canonical NtcA-dependent Class II promoter (Muro-Pastor *et al*., 2006). The resulting
464 increased levels of HetR and NtcA are required for the expression of subsequent genes,
465 including those for additional regulators of specific pathways, culminating in expression
466 of nitrogenase activity in the mature heterocyst. As a secondary regulator, NrrA also
467 activates the expression of glycogen catabolism genes and cooperates with NtcA for
468 induction of *sigE* encoding the sigma factor SigE that is involved in the expression of
469 genes operating in the mature heterocyst, including *nifH* and genes of the pentose
470 phosphate pathway (Mella-Herrera *et* al., 2011; Ehira and Ohmori, 2011).
471 Mechanistically, a relevant feature of many genes involved in heterocyst
472 differentiation is that they bear complex promoter regions with several successive
473 promoters (see Herrero *et al*., 2013, and references therein), each with a specific
474 spatiotemporal pattern of expression –with regard to the stage of differentiation and the
475 degree of confinement to the differentiating cells. *In vivo*, promoters of heterocyst-
476 differentiation genes that are activated early upon N step-down and at significant levels
477 in all the cells of the filament conform to Class II NtcA-activated promoters. On the other
478 hand, promoters activated principally in the differentiating cells require both NtcA and
479 HetR. Although purified HetR has been shown to bind alone to the promoter region of
480 several heterocyst-differentiation genes (Huang *et. al*., 2004; Videau *et al*., 2014), it is
481 not known whether HetR alone suffices for transcript production. In the complex
482 promoter region of the *devBCA* operon, involved in the deposition of HGL, besides a
483 distal Class II NtcA-dependent promoter, a proximal promoter is found that, in spite of
484 depending on NtcA and HetR *in vivo*, presents only a very degenerated NtcA-binding
485 sequence. However, *in vitro*, NtcA alone binds to this promoter and leads to production
486 of transcripts in the presence of RNA polymerase and 2-OG. Binding of HetR alone could
487 not be detected, but it enhanced the interaction of NtcA with the promoter DNA and

488 transcript production. Thus, in this promoter HetR acts as an NtcA co-activator (Camargo
489 *et al*., 2012).
490 A global determination by chromatin immunoprecipitation of the HetR targets at
491 an intermediate time during heterocyst differentiation has identified 59 binding sites
492 ascribed to 25 genomic regions, of which only 10 % included recognized heterocyst
493 differentiation genes (Flaherty *et al*., 2014). Because the number of genes involved in
494 heterocyst differentiation that *in vivo* show dependence on HetR is much higher, their
495 requirement may respond to an indirect effect of HetR (e.g., a requirement for increased
496 levels of NtcA or of a secondary, HetR-dependent regulator). In addition, the fact that
497 HetR was detected bound at some vegetative-cell activated promoters implies a function
498 of HetR in vegetative cells, likely as a repressor, as also suggested by the study of *hetR*
499 mutants (see, e.g., López-Igual *et al*., 2010).
500 In *Anabaena*, another regulatory factor, PipX, is maximally expressed in the cells
501 differentiating into heterocysts, at advanced stages of differentiation, dependent on NtcA
502 and HetR, and is required for full activation of late heterocyst-specific genes (Valladares
503 *et al*., 2011). At some NtcA-dependent promoters, PipX has been shown to increase NtcA
504 binding and transcription. Thus, PipX represents another NtcA co-activator (Camargo *et*
505 *al*., 2014).
506 In summary, NtcA alone activates promoters induced early after N step-down in
507 all the cells of the filament. Later, HetR participates in the activation of promoters that
508 exhibit specificity for the differentiating cells, and still later, PipX co-operates with NtcA
509 for transcription activation in near-mature heterocysts. Mechanistically, HetR could
510 promote transcription activation in different ways: cooperating with NtcA, acting alone,
511 or indirectly, by promoting increased levels of NtcA that may be required for activation
512 of late NtcA-activated genes or of secondary regulators. Notably, a sequence of NtcA co-
513 activators, HetR and PipX, appears to play an important role in the establishment of the
514 hierarchical series of transcription activation effecting heterocyst differentiation (Fig 4).
515 Recently, it has been shown that purified HetR is the subject of phosphorylation
516 in vitro (Valladares *et al*., 2016). Phosphorylation interferes with the formation of HetR
517 tetramers and with HetR binding to DNA. Thus, phosphorylation would represent a
518 mechanism for HetR inactivation (Valladares *et al*., 2016). On the other hand, the
519 morphogen PatS is a negative effector of HetR. HetR binds PatS *in vitro* (Feldman *et al*.,
520 2011) and PatS hampers HetR tetramerization (Valladares *et al*., 2016) and inhibits HetR-
521 binding to DNA (Huang *et al*., 2004), whereas *in vivo* HetR tetramers are more

represented in a mutant that does not produce PatS (and HetN) (Valladares *et al*., 2016). Thus, the PatS morphogen and phosphorylation could represent two different mechanisms for negative regulation of the activity of HetR, both acting by interfering with HetR tetramerization. During the first stages of heterocyst differentiation, PatS appears to limit the activity of HetR in the vegetative cells that are neighbors to the differentiating cells, thus limiting heterocyst differentiation. Indeed, coupling of fluctuations of *hetR* expression, influenced by PatS and auto-regulation, along the filament appears to play an important role in the generation of the heterocyst distribution pattern (Corrales-Guerrero *et al*., 2014). In turn, phosphorylation would contribute to inactivation of HetR in the cells that are at advanced stages of differentiation (Valladares *et al*., 2016).

Ci assimilation

Although certain sugars, especially glucose, fructose and sucrose (Rippka *et al.*, 1979), can be assimilated by some heterocyst-forming cyanobacteria, the bulk of carbon is obtained from CO_2/CO_3H^- (generically inorganic carbon, Ci). Inorganic carbon enters the cyanobacterial cell with the aid of active mechanisms and is incorporated into organic material mainly by ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCo), a doorway for the Calvin-Benson-Bassham (CBB) reductive pentose phosphate pathway. A distinct cyanobacterial feature is the expression of a so-called Carbon Concentrating Mechanism (CCM) that enhances the CO_2 concentration in the vicinity of RubisCo, thus increasing the efficiency of C assimilation. The mechanism of Ci assimilation has been less investigated in the Nostocales than in unicellular cyanobacteria such as *Synechococcus* sp. PCC 7942 or *Synechocystis* sp. PCC 6803. Therefore, in this section, we shall first briefly introduce the topic for cyanobacteria in general and then describe what is known for *Anabaena*.

<u>Ribulose bisphosphate carboxylase</u>

RubisCo is considered as one of the most abundant enzymes on Earth, perhaps to compensate for a general low activity that results from low catalytic turnover, low affinity for CO_2 and the concomitant utilization of O_2 as a substrate alternative to CO_2. The carboxylase reaction involves the incorporation of one molecule of CO_2 into one molecule

of ribulose 1,5-bisphosphate (RuBP) to produce two molecules of 3-phosphoglycerate, whereas the oxygenase reaction involves the incorporation of one molecule of O_2 into one molecule of RuBP to produce one molecule of 3-phosphoglycerate and one of 2-phosphoglycolate, which enters photorespiratory pathways resulting in a loss of fixed C (Kaplan *et al*., 2008; Allahverdiyeva *et al*., 2011; Hagemann *et al*., 2013). The type of RubisCo found in cyanobacteria, as in plants and most algae and phototrophic bacteria, is form 1 RubisCo, which is a hexadecameric enzyme composed of eight catalytic large (ca. 55 kDa) and eight small (ca. 14 kDa) subunits (see Tabita *et al*., 2007), although variants of this general theme have been found (see below).

The carbon concentrating mechanism of cyanobacteria

Cyanobacteria have evolved a remarkable physiological system, CCM, whose operation increases the efficiency of CO_2 fixation (Kaplan *et* al., 2008; Price, 2011). Cyanobacterial CCM systems are integrated by CO_2 and HCO_3^- uptake systems, carbonic anhydrases, and a dedicated bacterial micro-compartment, the carboxysome (Kerfeld and Erbilgin, 2015), in which a good part of the cellular RubisCo is confined (Fig. 1). The operation of CCM systems results in a large increase of CO_2 concentration in the proximity of RubisCo within the carboxysome, thus enhancing carboxylase activity. Two classes of cyanobacteria have been distinguished depending on the type of carboxysome they bear (α-cyanobacteria and β-cyanobacteria). Remarkably, the complement of Ci uptake systems and the type of carboxysomes found across cyanobacteria respond to the characteristics of the natural environment in which they have evolved (see below).

CO_2/HCO_3^- uptake systems

Regarding HCO_3^- transport, two secondary transporters (SbtA and BicA) and one ABC transporter (Cmp, a.k.a. BCT1) have been characterized in cyanobacteria. The *cmpABCD* operon encodes the elements of an ABC transporter for HCO_3^-, with CmpA representing a periplasmic substrate-binding protein for HCO_3^- that, notably, is anchored to the cytoplasmic membrane (Maeda *et al*., 2000). Most orthologues of the *cmpABCD* cluster are found in freshwater unicellular and filamentous cyanobacteria (Price *et al*., 2008).

The *sbtA* gene encodes a Na^+-dependent high-affinity HCO_3^- transporter. In the unicellular cyanobacterium *Synechocystis* sp. PCC 6803, SbtA-mediated transport

requires the product of a nearby gene encoding a Na^+/K^+-translocating protein, leading to the proposal that SbtA is energized by $\Delta\mu Na^+$ across the cytoplasmic membrane (Shibata *et al*., 2002). SbtA homologs are present in different unicellular and filamentous β-cyanobacteria (Price *et al*., 2008). The *bicA* gene encodes a Na^+-dependent, low-affinity HCO_3^- transporter of the SulP family (Price *et al*., 2004). Homologs of *bicA* are present in many fresh-water and oceanic cyanobacteria, including the widespread marine *Synechococcus* and *Prochlorococcus* strains. Thus, BicA may represent an important path for carbon acquisition at a global scale (Price *et al*., 2008).

Regarding CO_2 uptake, the NDH-1_3 and NDH-1_4 systems, which are specialized NAD(P)H dehydrogenase (NDH-1) membrane complexes, have been characterized (e.g., Shibata *et al*., 2001). NDH-1_3, comprised of the proteins NdhD3 and NdhF3 (the NDH-1 subunits) and CupA (a.k.a. ChpY), is a low CO_2-inducible CO_2 uptake system with higher maximal rate and affinity than the NDH-1_4 system, which is comprised of NdhD4/NdhF4/CupB (a.k.a. ChpX). At least the NDH-1_3 complex appears restricted to thylakoid membranes (Ohkawa *et al*., 2001). The model for the operation of these CO_2 uptake systems proposes a passive, aquaporine-mediated entrance of CO_2 down a concentration gradient that is maintained by the intracellular hydration of CO_2 to HCO_3^- catalyzed by the Cup proteins, which is linked to $\Delta\mu H^+$ generated by proton translocation across membranes associated to electron transport by the NdhDF subunits (Tchernov *et al*., 2001; Badger and Price, 2003). Both CO_2 uptake systems, NDH-1_3 and NDH-1_4, can be detected in β-cyanobacteria, whereas picoplanktonic marine strains show only NDH-1_4 (*Synechococcus* strains) or none of them (*Prochlorococcus* strains) (Price *et al*., 2008). Notwithstanding the above brief description of the distribution of CO_2/HCO_3^- uptake systems in the cyanobacterial phylum, it should be noted that, as far as we know, the search has not been updated with the current wide collection of available cyanobacterial genomes.

The carboxysome

The uptake of Ci species, either through HCO_3^- transporters or by CO_2 translocation, leads to the accumulation of HCO_3^- in the cytoplasm. Free HCO_3^- enters the carboxysome, which includes RubisCo. All the cyanobacteria so far studied, as well as many chemoautotrophic and a few purple-sulphur phototrophic bacteria possess carboxysomes.

623 Carboxysomes are polyhedral bodies constituted by a central core made of RubisCo and
624 carbonic anhydrase, which converts HCO_3^- into CO_2, encapsulated into a proteinaceous
625 shell that represents a diffusion barrier for CO_2 exit and O_2 entrance, although allowing
626 the movement of HCO_3^-, the RubisCo substrate RuBP, and the product 3-
627 phosphoglycerate. Detailed structural studies of the carboxysome have been performed
628 in recent years revealing a highly-ordered shell structure made of icosahedral facets of
629 hexameric proteins (Kerfeld *et al*., 2005) connected at the vertices by pentameric proteins
630 (Tanaka *et al*., 2008; Sutter *et al*., 2013).
631 All cyanobacterial carboxysomes share the same overall architecture and way of
632 operation. However, two types of carboxysomes have been distinguished, which differ in
633 the type of RubisCo and carbonic anhydrase included, in the components of the
634 carboxysome shell, the lumen organization, and the process of particle assembly. These
635 two carboxysome types have deep phylogenetic divergence in the cyanobacterial phylum.
636 The so-called β-carboxysomes contain form 1-B RubisCo and the CcmM and/or CcaA
637 carbonic anhydrases, with the shell made of the products of *ccm* genes. The *rbcL* and *rbcS*
638 genes (encoding the large and small RubisCo subunits, respectively) are arranged in the
639 *rbc* operon that also includes *rbcX*, encoding a putative RubisCo chaperonin. The *ccm*
640 genes can be spread in the genome, although the core genes including *ccmM* are
641 frequently arranged in one operon, which can be linked to the *rbc* operon and the *ccaA*
642 gene. β-Carboxysomes are widely distributed among the cyanobacterial lineages, being
643 present in fresh-water and marine coastal cyanobacteria (the so-called β-cyanobacteria).
644 In contrast, α-carboxysomes contain form 1-A RubisCo and the CsoSCA carbonic
645 anhydrase, and the shell is made of the products of *cso* genes. The *cso* genes are generally
646 arranged in a single operon that may include also some genes encoding CCM elements
647 and form 1-A RubisCo (*cbbL cbbS*). α-Carboxysomes are found in a group of
648 phylogenetically-related oceanic cyanobacteria (the so-called α-cyanobacteria), which
649 may have acquired the *cso* genes by lateral gene transfer from non-cyanobacterial
650 organisms (for further details see Rae *et* al., 2013; Kerfeld and Melnicki, 2016). In the
651 cyanobacteria studied to date, the elements of the CCM system, as well as RubisCo, are
652 expressed at higher levels under conditions of Ci-limitation.
653
654 CCM of *Anabaena*
655

Anabaena belongs to the group of β-cyanobacteria (Price *et al*., 2008; Kaplan *et al*., 2008), and as such it presents an *rbcLXS* operon. In addition, it bears a gene, *rca*, encoding RubisCo activase that is present only in a few cyanobacteria (Li *et al.*, 1999).

Regarding Ci transporters, genes encoding homologues of all five systems described above for Ci uptake are detected in the genome of *Anabaena* (Kaneko *et* al., 2001; Price *et al*., 2008; Kaplan *et al*., 2008). The encoded products are schematically represented in Fig. 1. Experimental evidence is however only available for a Cmp bicarbonate transport system. ORFs *alr2877-alr2880*, which are homologous to *cmpABCD* of the unicellular cyanobacteria *Synechococcus* sp. PCC 7942 and *Synechocystis* sp. PCC 6803, are expressed together as an operon, at negligible levels in filaments incubated with CO_2-enriched air and induced upon incubation in air (López-Igual *et al.,* 2012). Inactivation of *alr2877* specifically impairs the active transport of HCO_3^- under Ci limitation (López-Igual *et al.,* 2012). Thus, *alr2877-alr2880* correspond to the *Anabaena cmp* operon. However, the *alr2877* mutant still exhibits considerable activity of HCO_3^- transport, which suggests the presence of additional HCO_3^- transporters in *Anabaena*, in accordance to the results of protein homology searching. A putative SbtA orthologue (82 % identity to SbtA from *Synechocystis* sp. PCC 6803) is encoded by *all2134*, and a putative BicA orthologue (65 % identity to BicA from *Synechococcus* sp. PCC 7002) is encoded by *all1304*. Regarding CO_2 uptake systems, gene clusters *alr4156-alr4157-alr4158* and *alr0869-alr0870-alr0871* putatively encode orthologues of the constituents of NDH-1_3 and NDH-1_4 systems, respectively.

In *Anabaena*, *all0862* encodes a LTTR factor and is induced under conditions of Ci limitation. Inactivation of *all0862* abolishes induction of the *cmp* operon upon Ci step-down, and decreases HCO_3^- transport activity in a way similar to that of the *cmp* mutant. Thus, All0862 represents a CmpR factor of *Anabaena*, being the *cmpR* gene positively autoregulated in response to Ci limitation (López-Igual *et al.,* 2012).

In contrast to *cmpR* of *Synechocystis*, which is located upstream, in the opposite orientation of *cmpA*, in the *Anabaena* genome *cmpR* is located downstream, in the same orientation, of a gene cluster putatively encoding carboxysome *ccm* genes (*all0868* to *all0863*). This gene cluster is expressed as an operon also activated by Ci deprivation, at higher levels in the *cmpR* mutant than in the wild type (López-Igual *et al.,* 2012). In addition, CmpR may negatively influence expression of the *rbcL* operon (López-Igual *et al.,* 2012). Upregulation of the *rbcL* operon by Ci limitation is also negatively affected by CmpR in *Synechococcus* sp. PCC 7942 but unaffected in *Synechocystis* sp. PCC 6803

(Omata *et al*., 2001). In summary, in *Anabaena* CmpR activates the *cmpA* operon and the *cmpR* gene itself, and represses the *ccm* and *rbc* operons (López-Igual *et al*., 2012). Finally, although in *Synechococcus* (Nishimura *et al*., 2008) and *Synechocystis* (Daley *et al*., 2012) RuBP and 2-phosphoglycolate have been proposed as effectors of CmpR, no evidence for metabolic effectors is yet available for *Anabaena*.

Regulation by PacR

In *Synechocystis* sp. PCC 6803 and *Synechococcus* sp. PCC 7002, the genes for multiple components of CCM, mainly Ci transporters, are regulated by repression mediated by the LTTR factor NdhR (a.k.a. CcmR) (see Price *et al*., 2008). NdhR binds the effectors 2-OG and 2-phosphoglycolate. Whereas 2-OG increases binding to DNA, thus acting as a NdhR co-repressor, 2-phosphoglycolate (whose cellular levels are high under Ci limitation) decreases binding, thus acting as an inducer of NdhR-repressed genes (Daley *et al*., 2012; Jiang *et al.,* 2018). In contrast, in *Anabaena*, multiple genes encoding CCM elements and the *rbcL* operon are positively regulated under low Ci conditions. Besides the regulation mediated by CmpR described above, a different LTTR factor, PacR (All3953) is involved in regulation of CCM elements (Picossi *et al*., 2015). The *all3953* gene is constitutively expressed with regard to the N and Ci regimes, and its inactivation abolishes *rbcL* induction upon transference from a high- to a low-carbon regime, but leads to increased *rbcL* expression under high carbon. Hence, it has been proposed that regulation by PacR implies binding to repressor and activator sites in the *rbcL* promoter region (Picossi *et al*., 2015).

A global search for PacR targets by chromatin immunoprecipitation at 3 h after Ci step-down has identified 142 bound regions ascribed to genes of different functional categories, including other regulatory genes, which could expand the physiological range of influence of PacR. Besides *rbcL*, targets of PacR include genes for putative CCM elements: Ci transporter homologs (BicA, NdhF3 and NdhF4) and carboxysome shell proteins. Notably, 2-phosphoglycolate metabolism-related genes and genes encoding some components of the photosystem cores and the electronic transport chains are also regulated by PacR (Picossi *et al*., 2015). Some of these photosynthetic targets, including those encoding the flavodiiron proteins Flv1A and Flv4 (Ermakova *et al*., 2013), appear to be involved in avoiding over-reduction of photosynthetic components, which is enhanced under conditions of low Ci availability and could lead to oxidative damage.

Hence, PacR coordinates protection against photodamage and, consistently, growth of a *pacR* mutant is compromised when subjected to high-light intensity (Picossi *et al*., 2015). This observation, together with regulation of CCM elements by PacR, identifies this factor as a transcriptional regulator especially suited for coordination of Ci acquisition in the context of oxygenic phototrophy.

Coordinated regulation of C and N assimilation in *Anabaena*

In *Anabaena*, a number of regulatory circuits have been found to depend on both the C and N regimes. As described above, NtcA (CRP family) orchestrates N-control in response to the 2-OG levels, which are influenced by the metabolic flow of CO_2 assimilation through the CBB cycle and NH_4^+ assimilation through the GS/GOGAT cycle, and NtcA is a global regulator.

CmpR (LTTR) is a pathway-specific regulator of Ci assimilation that activates the expression of the *cmpA* operon (bicarbonate uptake) in response to low Ci availability. The *cmpR* gene is itself positively autoregulated under Ci limitation and, in addition, it is expressed at lower levels in the presence of ammonium than in the absence of combined nitrogen, as well as in an *ntcA* mutant than in the wild type. Indeed, NtcA-binding upstream *cmpR* has been shown *in vitro*. Thus, CmpR is an N- and Ci-responsive regulator (López-Igual *et al*., 2012). Besides that, NtcA also binds to the *cmpA* operon promoter, likely with a repressive effect. On the other hand, when *Anabaena* grows under diazotrophic conditions, the *rbcL* operon is not expressed in the heterocysts (Elhai and Wolk, 1990), and NtcA binding to a putative site overlapping the translation start of *rbcL* could be responsible for this repression (Ramasubramanian *et al*., 1994; Picossi *et al*., 2014). Thus, the *cmpA* and *rbcL* operons and the *cmpR* gene are coordinately regulated by the CmpR and NtcA transcription factors.

Heterocyst differentiation is regulated by NtcA and HetR in response to a deficit of combined nitrogen, triggered by high 2-OG levels, and the differentiation process includes a drastic reorganization of C metabolism, which becomes basically heterotrophic. The NrrA transcription factor that participates in induction of *hetR* during differentiation, after *nrrA* induction by NtcA, is also involved in the activation of genes for glycogen catabolism, glycolysis and the pentose phosphate pathway, with a positive effect on N_2 fixation. Finally, according to Ci availability and illumination conditions, PacR (LTTR) coordinates regulation of CCM and RubisCo, thus influencing 2-OG levels,

as well as of photosynthetic elements, some of which participate in protection against oxidative damage. Thus, the *Anabaena* metabolism is coordinately regulated by the availability of C and N, which determines the C-to-N balance, at a global level (Fig. 5).

Acknowledgements

Work in the authors' laboratory was supported by grants BFU2016-77097-P (AH) and BFU2017-88202-P (EF) from Agencia Estatal de Investigación, Spain, co-financed by the European Regional Development Fund (EU). The authors declare that they have no conflict of interests.

References

Aichi M, Yoshihara S, Yamashita M, Maeda S, Nagai K, Omata T (2006) Characterization of the nitrate-nitrite transporter of the major facilitator superfamily (the *nrtP* gene product) from the cyanobacterium *Nostoc punctiforme* strain ATCC 29133. *Biosci Biotechnol Biochem* 70(11):2682-9.

Allahverdiyeva Y, Ermakova M, EisenhutM, Zhang P, Richaud P, Hagemann M, Cournac L, Aro E-M (2011) Interplay between flavodiiron proteins and photorespiration in *Synechocystis* sp. PCC 6803. *J Biol Chem* 286(27):24007-24014.

Awai K, Wolk CP (2007) Identification of the glycosyl transferase required for synthesis of the principal glycolipid characteristic of heterocysts of *Anabaena* sp. strain PCC 7120. *FEMS Microbiol Lett* 266(1):98-102.

Awai K, Lechno-Yossef S, Wolk CP (2009) Heterocyst envelope glycolipids, in: Wada H, Murata N (Eds.) Lipids in Photosynthesis: Essential and Regulatory Functions, Springer Science + Business Media B.V., pp.179–202.

Badger MR, Price GD (2003) CO_2 concentrating mechanism in cyanobacteria: molecular components, their diversity and evolution. *J Exp Bot* 54(383):609-622.

Boer JL, Mulrooney SB, Hausinger RP (2014) Nickel-dependent metalloenzymes. *Arch Biochem Biophys* 544:142-52.

Burnat M, Flores E (2014) Inactivation of agmatinase expressed in vegetative cells alters arginine catabolism and prevents diazotrophic growth in the heterocyst-forming cyanobacterium *Anabaena*. *MicrobiologyOpen* 3(5):777-92.

Burnat M, Herrero A, Flores E (2014) Compartmentalized cyanophycin metabolism in the diazotrophic filaments of a heterocyst-forming cyanobacterium. *Proc Natl Acad Sci USA* 111(10):3823-8.

Burnat M, Li B, Kim SH, Michael AJ, Flores E (2018) Homospermidine biosynthesis in the cyanobacterium *Anabaena* requires a deoxyhypusine synthase homologue and is essential for normal diazotrophic growth. *Mol Microbiol* (doi: 10.1111/mmi.14006).

Cai Y, Wolk CP (1997) Nitrogen deprivation of Anabaena sp. strain PCC 7120 elicits rapid activation of a gene cluster that is essential for uptake and utilization of nitrate. *J Bacteriol* 179(1):258-66.

Callahan SM, Buikema WJ (2001) The role of HetN in maintenance of the heterocyst pattern in *Anabaena* sp. PCC 7120. *Mol Microbiol* 40(4):941-50.

Camargo S, Valladares A, Flores E, Herrero A (2012) Transcription activation by NtcA in the absence of consensus NtcA-binding sites in an *Anabaena* heterocyst differentiation gene promoter. *J Bacteriol* 194(11):2939-48.

Camargo S, Valladares A, Forchhammer K, Herrero A (2014) Effects of PipX on NtcA-dependent promoters and characterization of the *cox3* promoter region in the heterocyst-forming cyanobacterium *Anabaena* sp. PCC 7120. *FEBS Lett* 588(9):1787-94.

Cardemil L, Wolk CP (1981) Polysaccharides from the envelopes of heterocysts and spores of the blue-green algae *Anabaena variabilis* and *Cylindrospermun licheniforme*. *J Phycology* 17(3):234-240.

Castenholz, R.W. (2001) Phylum BX. Cyanobacteria (Oxygenic Photosynthetic Bacteria). Bergey's Manual® of Systematic Bacteriology, Vol. 1: The *Archaea* and the Deeply Branching and Phototrophic *Bacteria*. Springer New York, pp. 473-599.

Corrales-Guerrero L, Mariscal V, Nürnberg DJ, Elhai J, Mullineaux CW, Flores E, Herrero A (2014) Subcellular localization and clues for the function of the HetN factor influencing heterocyst distribution in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 196(19):3452-60.

Daley MES, Kappell AD, Carrick MJ, Burnap LR (2012) Regulation of the cyanobacterial CO_2-concentrating mechanism involves internal sensing of $NADP^+$ and α-ketogutarate levels by transcription factor CcmR. *PLoS One* 7(7) e41286.

Ehira S, Ohmori M (2006) NrrA directly regulates expression of *hetR* during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 188(24):8520-5.

Ehira S, Ohmori M (2011) NrrA, a nitrogen-regulated response regulator protein, controls glycogen catabolism in the nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *J Biol Chem* 286(44):38109-14.

Elhai J, Wolk CP (1990) Developmental regulation and spatial pattern of expression of the structural genes for nitrogenase in the cyanobacterium *Anabaena. EMBO J* 9(10)3379-3388.

Ermakova M, Battchikova N, Allahverdiyeva Y, Aro EM (2013) Novel heterocyst-specific flavodiiron proteins in *Anabaena* sp. PCC 7120. *FEBS Lett* 587(1):82-7.

Ermakova M, Battchikova N, Richaud P, Leino H, Kosourov S, Isojärvi J, Peltier G, Flores E, Cournac L, Allahverdiyeva Y, Aro EM (2014) Heterocyst-specific flavodiiron protein Flv3B enables oxic diazotrophic growth of the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *Proc Natl Acad Sci USA* 111(30):11205-10.

827 Fan Q, Huang G, Lechno-Yossef S, Wolk CP, Kaneko T, Tabata S (2005) Clustered genes required for
828 synthesis and deposition of envelope glycolipids in *Anabaena* sp. strain PCC 7120. *Mol Microbiol*
829 58(1):227-43.
830 Fay P (1992) Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev* 56:340-73.
831 Feldmann EA, Ni S, Sahu ID, Mishler CH, Risser DD, Murakami JL, Tom SK, McCarrick RM, Lorigan
832 GA, Tolbert BS, Callahan SM, Kennedy MA (2011) Evidence for direct binding between HetR from
833 *Anabaena* sp. PCC 7120 and PatS-5. *Biochemistry* 50(43):9212-24.
834 Fiedler G, Arnold M, Hannus S, Maldener I (1998) The DevBCA exporter is essential for envelope
835 formation in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Microbiol*
836 27(6):1193-202.
837 Flaherty BL, Van Nieuwerburgh F, Head SR, Golden JW (2011) Directional RNA deep sequencing sheds
838 new light on the transcriptional response of *Anabaena* sp. strain PCC 7120 to combined-nitrogen
839 deprivation. *BMC Genomics* 12:332.
840 Flaherty BL, Johnson DB, Golden JW (2014) Deep sequencing of HetR-bound DNA reveals novel HetR
841 targets in *Anabaena* sp. strain PCC7120. *BMC Microbiol* 14:255.
842 Flores E, Herrero A (2005) Nitrogen assimilation and nitrogen control in cyanobacteria. *Biochem Soc Trans*
843 33(Pt 1):164-7.
844 Flores E, Frías JE, Rubio LM, Herrero A (2005) Photosynthetic nitrate assimilation in cyanobacteria.
845 *Photosynth Res* 83(2):117-33.
846 Flores E, López-Lozano A, Herrero A (2015) Nitrogen fixation in the oxygenic phototrophic prokaryotes
847 (cyanobacteria): the fight against oxygen. In: *Biological Nitrogen Fixation*, Frans J. De Bruijn (ed.),
848 Vol. 2, pp. 879-889.
849 Flores E, Herrero A, Forchhammer K, Maldener I (2016) Septal Junctions in Filamentous Heterocyst-
850 Forming Cyanobacteria. *Trends Microbiol* 24(2):79-82.
851 Flores E, Picossi S, Valladares A, Herrero A (2018) Transcriptional regulation of development in
852 heterocyst-forming cyanobacteria. *Biochim Biophys Acta* doi: 10.1016/j.bbagrm.2018.04.006.
853 Frías JE, Flores E (2010) Negative regulation of expression of the nitrate assimilation *nirA* operon in the
854 heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 192(11):2769-78.
855 Frías JE, Flores E (2015) Induction of the nitrate assimilation *nirA* operon and protein-protein interactions
856 in the maturation of nitrate and nitrite reductases in the cyanobacterium *Anabaena* sp. strain PCC 7120.
857 *J Bacteriol* 197(14):2442-52.
858 Frías JE, Flores E, Herrero A (1997) Nitrate assimilation gene cluster from the heterocyst-forming
859 cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 179(2):477-86.
860 Frías JE, Flores E, Herrero A (2000) Activation of the *Anabaena nir* operon promoter requires both NtcA
861 (CAP family) and NtcB (LysR family) transcription factors. *Mol Microbiol* 38(3):613-25.
862 Frías JE, Herrero A, Flores E (2003) Open reading frame *all0601* from *Anabaena* sp. strain PCC 7120
863 represents a novel gene, *cnaT*, required for expression of the nitrate assimilation *nir* operon. *J Bacteriol*
864 185(17):5037-44.
865 Gallon JR (1981) The oxygen sensitivity of nitrogenase: a problem for biochemists and micro-organisms.
866 *TIBS* 6:19-23.
867 Galmozzi CV, Saelices L, Florencio FJ, Muro-Pastor MI (2010) Posttranscriptional regulation of glutamine
868 synthetase in the filamentous cyanobacterium *Anabaena* sp. PCC 7120: differential expression between
869 vegetative cells and heterocysts. *J Bacteriol* 192(18):4701-11.
870 García-Domínguez M, Reyes JC, Florencio FJ (1999) Glutamine synthetase inactivation by protein-protein
871 interaction. Proc Natl Acad Sci USA 96(13):7161-6.
872 Giddings TH Jr, Wolk CP, Shomer-Ilan A (1981) Metabolism of sulfur compounds by whole filaments and
873 heterocysts of *Anabaena variabilis*. *J Bacteriol* 146(3):1067-74.
874 Giovannoni SJ, Turner S, Olsen GJ, Barns S, Lane DJ, Pace NR (1988) Evolutionary relationships
875 among cyanobacteria and green chloroplasts. *J Bacteriol* 170(8):3584-92.
876 Hackenberg C, Kern R, Hüge J, Stal LJ, Tsuji Y, Kopka J, Shiraiwa Y, Bauwe H, Hagemann M (2011)
877 Cyanobacterial lactate oxidases serve as essential partners in N2 fixation and evolved into
878 photorespiratory glycolate oxidases in plants. *Plant Cell* 23(8):2978-90.
879 Hagemann M, Fernie AR, Espie GS, Kern R, Eisenhut M., Reumann S, Bauwe H, Weber APM (2013)
880 Evolution of the biochemistry of the photorespiratory C2 cycle. *Plant Biol* 15:639–647.
881 Herrero A, Flores E (1990) Transport of basic amino acids by the dinitrogen-fixing cyanobacterium
882 *Anabaena* PCC 7120. *J Biol Chem* 265(7):3931-5
883 Herrero A, Muro-Pastor AM, Flores E (2001) Nitrogen control in cyanobacteria. *J Bacteriol* 183(2):411-
884 25.
885 Herrero A, Picossi S, Flores E (2013) Gene expression during heterocyst differentiation, Adv Bot Res
886 65:281-329.

887 Hobbie JE, Hobbie EA (2013) Microbes in nature are limited by carbon and energy: the starving-survival
888 lifestyle in soil and consequences for estimating microbial rates. *Front Microbiol* 4: 324.
889 Huang X, Dong Y, Zhao J (2004) HetR homodimer is a DNA-binding protein required for heterocyst
890 differentiation, and the DNA-binding activity is inhibited by PatS. *Proc Natl Acad Sci USA*
891 101(14):4848-53.
892 Huang G, Fan Q, Lechno-Yossef S, Wojciuch E, Wolk CP, Kaneko T, Tabata S (2005) Clustered genes
893 required for the synthesis of heterocyst envelope polysaccharide in *Anabaena* sp. strain PCC 7120. *J*
894 *Bacteriol* 187(3):1114-23.
895 Jiang; Y-L, Wang X-P, Sun H, Han S-J, Li W-F, Cui N, Lin G-M, Zhang J-Y, Cheng W, Cao D-D, Zhang
896 Z-Y, Zhang C-C, Chen Y, Zhou C-Z (2018) Coordinating carbon and nitrogen metabolic signaling
897 through the cyanobacterial global repressor NdhR. *Proc Natl Acad Sci USA* 1158(2)403-408.
898 Jüttner F (1983) ^{14}C-labeled metabolites in heterocysts and vegetative cells of *Anabaena cylindrica*
899 filaments and their presumptive function as transport vehicles of organic carbon and nitrogen. *J*
900 *Bacteriol* 155(2):628-33.
901 Kaneko T, Nakamura Y, Wolk CP, Kuritz T, Sasamoto S, Watanabe A, Iriguchi M, Ishikawa A,
902 Kawashima K, Kimura T, Kishida Y, Kohara M, Matsumoto M, Matsuno A, Muraki A, Nakazaki N,
903 Shimpo S, Sugimoto M, Takazawa M, Yamada M, Yasuda M, Tabata S (2001) Complete genomic
904 sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res*
905 8(5):205-13; 227-53.
906 Kaplan A, Hagemann M, Bauwe H, Kahlon S, Ogawa T (2008) In: Herrero A, Flores E (eds.) The
907 Cyanobacteria: Molecular Biology, Genomics and Evolution, Caister Academic Press, Norfolk, UK,
908 pp. 305-334.
909 Kerfeld CA, Erbilgin O (2015) Bacterial microcompartments and the modular construction of microbial
910 metabolism. *Trends Microbiol* 23(1)22-34.
911 Kerfeld CA, Melnicki MR (2016) Assembly, function and evolution of cyanobacterial carboxysomes. *Curr*
912 *Op Plant Biol* 31:66-75.
913 Kerfeld CA, Sawaya MR, Tanaka S, Nguyen CV, 1 Phillips M, Beeby M, Yeates TO (2005) Protein
914 structures forming the shell of primitive bacterial organelles. *Science Reports* 309: 936-938.
915 Khademi S, O'Connell J 3rd, Remis J, Robles-Colmenares Y, Miercke LJ, Stroud RM (2004) Mechanism
916 of ammonia transport by Amt/MEP/Rh: structure of AmtB at 1.35 A. *Science* 305(5690):1587-94.
917 Kim Y, Joachimiak G, Ye Z, Binkowski TA, Zhang R, Gornicki P, Callahan SM, Hess WR, Haselkorn R,
918 Joachimiak A (2011) Structure of transcription factor HetR required for heterocyst differentiation in
919 cyanobacteria. *Proc Natl Acad Sci USA* 108(25):10109-14.
920 Kim Y, Ye Z, Joachimiak G, Videau P, Young J, Hurd K, Callahan SM, Gornicki P, Zhao J, Haselkorn R,
921 Joachimiak A (2013) Structures of complexes comprised of *Fischerella* transcription factor HetR with
922 *Anabaena* DNA targets. *Proc Natl Acad Sci USA* 110(19):E1716-23.
923 Kobayashi M, Takatani N, Tanigawa M, Omata T (2005) Posttranslational regulation of nitrate assimilation
924 in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J Bacteriol* 187(2):498-506.
925 Kolman MA, Nishi CN, Perez-Cenci M, Salerno GL (2015) Sucrose in cyanobacteria: from a salt-response
926 molecule to play a key role in nitrogen fixation. *Life (Basel)* 5(1):102-26.
927 Lang NJ, Simon RD, Wolk CP (1972) Correspondence of cyanophycin granules with structured granules
928 in *Anabaena cylindrica*. *Arch Mikrobiol* l83:313-320.
929 Li LA, Zianni MR, Tabita FR (1999) Inactivation of the monocistronic rca gene in *Anabaena variabilis*
930 suggests a physiological ribulose bisphosphate carboxylase/oxygenase activase-like function in
931 heterocystous cyanobacteria. *Plant Mol Biol* 40(3):467-78.
932 López-Igual R, Flores E, Herrero A (2010) Inactivation of a heterocyst-specific invertase indicates a
933 principal role of sucrose catabolism in heterocysts of *Anabaena* sp. *J Bacteriol* 192(20):5526-33.
934 López-Igual R, Picossi S, López-Garrido J, Flores E, Herrero A (2012) N and C control of ABC-type
935 bicarbonate transporter Cmp and its LysR-type transcriptional regulator CmpR in a heterocyst-forming
936 cyanobacterium, *Anabaena* sp. *Environ Microbiol* 14(4):1035-48.
937 Luque I, Forchhammer K (2008) Nitrogen assimilation and C/N balance sensing. In: Herrero A, Flores E
938 (eds.) The Cyanobacteria: Molecular Biology, Genomics and Evolution, Caister Academic Press,
939 Norfolk, UK, pp. 335-382.
940 Luque I, Flores E, Herrero A (1994) Molecular mechanism for the operation of nitrogen control in
941 cyanobacteria. *EMBO J* 13(12):2862-9. Erratum in: *EMBO J* 13(23):5794.
942 Maeda S, Price GD, Badger, MR, Enomoto C, Omata T (2000) Bicarbonate binding activity of the CmpA
943 protein of the cyanobacterium *Synechococcus* sp. Strain PCC 7942 involved in active transport of
944 bicarbonate. *J Biol Chem* 275(27):20551-20555.
945 Magnuson A, Cardona T (2016) Thylakoid membrane function in heterocysts. *Biochim Biophys Acta*
946 1857(3):309-19.

947 Mareš J, (2017) Multilocus and SSU rRNA gene phylogenetic analyses of available cyanobacterial
948 genomes, and their relation to the current taxonomic system. *Hydrobiologia*. doi: 10.1007/s10750-017-
949 3373-2.
950 Martín-Figueroa E, Navarro F, Florencio FJ (2000) The GS-GOGAT pathway is not operative in the
951 heterocysts. Cloning and expression of *glsF* gene from the cyanobacterium *Anabaena* sp. PCC 7120.
952 *FEBS Lett* 476(3):282-6.
953 Mella-Herrera RA, Neunuebel MR, Kumar K, Saha SK, Golden JW (2011) The *sigE* gene is required for
954 normal expression of heterocyst-specific genes in *Anabaena* sp. strain PCC 7120. *J Bacteriol*
955 193(8):1823-32.
956 Montesinos ML, Herrero A, Flores E (1995) Amino acid transport systems required for diazotrophic growth
957 in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 177(11):3150-7.
958 Montesinos ML, Muro-Pastor AM, Herrero A, Flores E (1998) Ammonium/methylammonium permeases
959 of a cyanobacterium. Identification and analysis of three nitrogen-regulated *amt* genes in *Synechocystis*
960 sp. PCC 6803. *J Biol Chem* 273(47):31463-70.
961 Muro-Pastor AM, Valladares A, Flores E, Herrero A (2002) Mutual dependence of the expression of the
962 cell differentiation regulatory protein HetR and the global nitrogen regulator NtcA during heterocyst
963 development. *Mol Microbiol* 44(5):1377-85.
964 Muro-Pastor AM, Olmedo-Verd E, Flores E (2006) All4312, an NtcA-regulated two-component response
965 regulator in *Anabaena* sp. strain PCC 7120. FEMS *Microbiol Lett* 256(1):171-7.
966 Mus F, Alleman AB, Pence N, Seefeldt LC, Peters JW (2018) Exploring the alternatives of biological
967 nitrogen fixation. *Metallomics* 10(4):523-538.
968 Nieves-Morión M, Mullineaux CW, Flores E (2017) Molecular diffusion through cyanobacterial septal
969 junctions. *mBio* 8(1):e01756-16.
970 Nishimura T, Takahashi Y, Yamaguchi O, Suzuki H, Maeda S, Omata T (2008) Mechanism of low CO_2-
971 induced activation of the *cmp* bicarbonate transporter operon by a LysR family protein in the
972 cyanobacterium *Synechococcus elongatus* strain PCC 7942. *Mol Microbiol* 68(1):98-109.
973 Nürnberg DJ, Mariscal V, Bornikoel J, Nieves-Morión M, Krauß N, Herrero A, Maldener I, Flores E,
974 Mullineaux CW (2015) Intercellular diffusion of a fluorescent sucrose analog via the septal junctions
975 in a filamentous cyanobacterium. *mBio* 6(2):e02109.
976 Ohkawa H, Sonoda M, Shibata M, Ogawa T (2001) Localization of NAD(P)H dehydrogenase in the
977 cyanobacterium *Synechocystis* sp. Strain PCC 6803. *J Bacteriol* 183(16)4938-4939.
978 Omata T, Gohta S, Takahashi Y, Harano Y, Maeda S (2001) Involvement of a CbbR homolog in low CO2-
979 induced activation of the bicarbonate transporter operon in cyanobacteria. *J Bacteriol* 183 (6):1891-
980 1898.
981 Pakrasi H, Zehr JP (2017) Cyanobacteria are oxygenic phototrophic prokaryotes. *Science*
982 (http://science.sciencemag.org/content/355/6332/1436/tab-e-letters).
983 Paz-Yepes J, Merino-Puerto V, Herrero A, Flores E (2008) The *amt* gene cluster of the heterocyst-forming
984 cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 190(19):6534-9.
985 Paz-Yepes J, Flores E, Herrero A (2009) Expression and mutational analysis of the glnB genomic region
986 in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 191(7):2353-61.
987 Pernil R, Herrero A, Flores E (2010) Catabolic function of compartmentalized alanine dehydrogenase in
988 the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 192(19):5165-72.
989 Pernil R, Picossi S, Herrero A, Flores E, Mariscal V (2015) Amino acid transporters and release of
990 hydrophobic amino acids in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120.
991 *Life (Basel)* 5(2):1282-300.
992 Picossi S, Flores E, Herrero A (2014) ChIP analysis unravels an exceptionally wide distribution of DNA
993 binding sites for the NtcA transcription factor in a heterocyst-forming cyanobacterium. BMC Genomics
994 15:22.
995 Picossi S, Flores E, Herrero A (2015) The LysR-type transcription factor PacR is a global regulator of
996 photosynthetic carbon assimilation in *Anabaena. Environ Microbiol* 17(9)3341-3351.
997 Ponce-Toledo RI, Deschamps P, Lopez-Garcia P, Zivanovic Y, Benzerara K, and Moreira D (2017) An
998 early-branching freshwater cyanobacterium at the origin of plastids. *Curr Biol* 27:386-391.
999 Price (2011) Inorganic carbon transporters of the cyanobacterial CO_2 concentrating mechanism. *Photosynth*
1000 *Res* 109:47-57.
1001 Price GD, Woodger FJ, Badger MR, Howitt SM, Tucker L (2004) Identification of a SulP-type bicarbonate
1002 transporter in marine cyanobacteria. *Proc Natl Acad Sci USA* 101(52)18228-18233.
1003 Price GD, Badger MR, Woodger FJ, Long BM (2008) Advances in understanding the cyanobacterial CO_2-
1004 concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic
1005 regulation and prospects for engeenering into plants. *J Exp Bot* 59(7):1441-1461.
1006 Rae BD, Long BM, Badger MR, Price GD (2013) Functions, Compositions, and Evolution of the Two

Types of Carboxysomes: Polyhedral Microcompartments That Facilitate CO_2 Fixation in Cyanobacteria and Some Proteobacteria. *Microbiol Mol Biol Rev* 77(3):357-379.

Ramasubramanian TS, Wei T-F, Golden J-W (1994) Two *Anabaena* sp. strain PCC 7120 DNA-binding factors interact with vegetative cell- and heterocyst-specific genes. *J Bacteriol* 176(5)1214-1223.

Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J Gen Microbiol* 111:1-61.

Rubio LM, Ludden PW (2008) Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Annu Rev Microbiol* 62:93-111.

Schirrmeister BE, Gugger M, Donoghue PC (2015) Cyanobacteria and the Great Oxidation Event: evidence from genes and fossils. *Palaeontology* 58(5):769-785.

Shibata M, Ohkawa H, Kaneko T, Fukuzawa H, Tabata S, Kaplan A (2001) Distinct constitutive and low-CO_2-induced CO_2 uptake systems in cyanobacteria: genes involved and their phylogenetic relationship with homologous genes in other organisms. *Proc Natl Acad Sci USA* (98(20):11789-11794.

Shibata M, Katoh H, Sonoda M, Ohkawa H, Shimoyama M, Fukuzawa H, Kaplan A, Ogawa T (2002) Genes essential to sodium-dependent bicarbonate transport in cyanobacteria. *J Biol Chem* 277(21):18658-18664.

Shih PM, Wu D, Latifi A, Axen SD, Fewer DP, Talla E, Calteau A, Cai F, Tandeau de Marsac N, Rippka R, Herdman M, Sivonen K, Coursin T, Laurent T, Goodwin L, Nolan M, Davenport KW, Han CS, Rubin EM, Eisen JA, Woyke T, Gugger M, Kerfeld CA (2013) Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc Natl Acad Sci USA* 110(3):1053-8.

Soo RM, Hemp J, Parks DH, Fischer WW, Hugenholtz P (2017) On the origins of oxygenic photosynthesis and aerobic respiration in cyanobacteria. *Science* 355(6332):1436-1440.

Stal LJ (2017) The effect of oxygen concentration and temperature on nitrogenase activity in the heterocystous cyanobacterium *Fischerella* sp. *Sci Rep* 7(1):5402.

Stanier RY, Cohen-Bazire G (1977) Phototrophic prokaryotes: the cyanobacteria. *Annu Rev Microbiol* 31:225-74.

Staron P, Forchhammer K, Maldener I (2011) Novel ATP-driven pathway of glycolipid export involving TolC protein. *J Biol Chem* 286(44):38202-10.

Summers ML, Wallis JG, Campbell EL, Meeks JC (1995) Genetic evidence of a major role for glucose-6-phosphate dehydrogenase in nitrogen fixation and dark growth of the cyanobacterium *Nostoc* sp. strain ATCC 29133. *J Bacteriol* 177(21):6184-94.

Sutter M, Wilson SC, Deutsch S, Kerfeld CA (2013) Two new high-resolution crystal structures of carboxysome pentamer proteins reveal high structural conservation of CcmL orthologs among distantly related cyanobacterial species. *Photosynth Res* 118:9-16.

Tabita RF, Hanson TE, Li H, Satagopan S, Singh J, Chan S (2007) Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs. *Microbiol Mol Biol Rev* 71(4):576-599.

Tanaka S, Kerfeld CA, Sawaya MR, Cai F, Heinhorst S, Cannon GC, Yeates TO (2008) Atomic-level models of the bacterial carboxysome shell. *Science Reports* 319: 1083-1086.

Tchernov D, Helman Y, Keren N, Luz B, Ohad I, Reinhold L, Ogawa T, Kaplan A (2001) Passive entry of CO_2 and its energy-dependent intracellular conversion to HCO_3^- in cyanobacteria are driven by a photosystem I-generated $\Delta\mu H^+$. *J Biol Chem* 276(26):23450-23455.

Thiel T, Pratte BS (2014) Regulation of three nitrogenase gene clusters in the cyanobacterium *Anabaena variabilis* ATCC 29413. *Life (Basel)* 4(4):944-67.

Thomas J, Meeks JC, Wolk CP, Shaffer PW, Austin SM (1977) Formation of glutamine from [^{13}N]ammonia, [^{13}N]dinitrogen, and [^{14}C]glutamate by heterocysts isolated from *Anabaena cylindrica*. *J Bacteriol* 129(3):1545-55.

Tomitani A, Knoll AH, Cavanaugh CM, Ohno T (2006) The evolutionary diversification of cyanobacteria: molecular-phylogenetic and paleontological perspectives. *Proc Natl Acad Sci USA* 103:5442-7.

Tumer NE, Robinson SJ, Haselkorn R (1983) Different promoters for the *Anabaena* glutamine synthetase gene during groth using molecular or fixed nitrogen. *Nature* 306:337-342.

Valladares A, Montesinos ML, Herrero A, Flores E (2002) An ABC-type, high-affinity urea permease identified in cyanobacteria. *Mol Microbiol* 43(3):703-15.

Valladares A, Herrero A, Pils D, Schmetterer G, Flores E (2003) Cytochrome c oxidase genes required for nitrogenase activity and diazotrophic growth in *Anabaena* sp. PCC 7120. *Mol Microbiol* 47(5):1239-49.

Valladares A, Muro-Pastor AM, Herrero A, Flores E (2004) The NtcA-dependent P1 promoter is utilized for *glnA* expression in N_2-fixing heterocysts of *Anabaena* sp. strain PCC 7120. *J Bacteriol* 186:7337-43.

Valladares A, Maldener I, Muro-Pastor AM, Flores E, Herrero A (2007) Heterocyst development and diazotrophic metabolism in terminal respiratory oxidase mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 189(12):4425-30.

Valladares A, Flores E, Herrero A (2008) Transcription activation by NtcA and 2-oxoglutarate of three genes involved in heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 190(18):6126-33.

Valladares A, Rodríguez V, Camargo S, Martínez-Noël GM, Herrero A, Luque I (2011) Specific role of the cyanobacterial PipX factor in the heterocysts of *Anabaena* sp. strain PCC 7120. *J Bacteriol* 193(5):1172-82.

Valladares A, Flores E, Herrero A (2016) The heterocyst differentiation transcriptional regulator HetR of the filamentous cyanobacterium *Anabaena* forms tetramers and can be regulated by phosphorylation. *Mol Microbiol* 99(4):808-19.

Vargas WA, Nishi CN, Giarrocco LE, Salerno GL (2011) Differential roles of alkaline/neutral invertases in *Nostoc* sp. PCC 7120: Inv-B isoform is essential for diazotrophic growth. *Planta* 233(1):153-62.

Vázquez-Bermúdez MF, Flores E, Herrero A (2002) Analysis of binding sites for the nitrogen-control transcription factor NtcA in the promoters of *Synechococcus* nitrogen-regulated genes. *Biochim Biophys Acta* 1578(1-3):95-8.

Videau P, Ni S, Rivers OS, Ushijima B, Feldmann EA, Cozy LM, Kennedy MA, Callahan SM (2014) Expanding the direct HetR regulon in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 196(5):1113-21.

Wacker T, Garcia-Celma JJ, Lewe P, Andrade SL (2014) Direct observation of electrogenic NH_4^+ transport in ammonium transport (Amt) proteins. *Proc Natl Acad Sci USA* 111(27):9995-10000.

Wang Y, Lechno-Yossef S, Gong Y, Fan Q, Wolk CP, Xu X (2007) Predicted glycosyl transferase genes located outside the HEP island are required for formation of heterocyst envelope polysaccharide in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 189(14):5372-8.

Wolk CP, Thomas J, Shaffer PW, Austin SM, Galonsky A (1976) Pathway of nitrogen metabolism after fixation of ^{13}N-labeled nitrogen gas by the cyanobacterium, *Anabaena cylindrica*. *J Biol Chem* 251(16):5027-34.

Yoon HS, Golden JW (1998) Heterocyst pattern formation controlled by a diffusible peptide. *Science* 282(5390):935-8.

Zehr JP, Shilova IN, Farnelid HM, Muñoz-Marín MC, Turk-Kubo KA (2017) Unusual marine unicellular symbiosis with the nitrogen-fixing cyanobacterium UCYN-A. *Nat Microbiol* 2:16214.

Zhao MX, Jiang YL, He YX, Chen YF, Teng YB, Chen Y, Zhang CC, Zhou CZ (2010) Structural basis for the allosteric control of the global transcription factor NtcA by the nitrogen starvation signal 2-oxoglutarate. *Proc Natl Acad Sci USA* 107(28):12487-92.

Figure legends:

Fig. 1. Schematic of the pathways of assimilation of the main sources of combined nitrogen and of inorganic carbon in cells of an undifferentiated filament of *Anabaena*. ABC transporters for nitrate and nitrite (Nrt) and urea (Urt) and an ammonium translocator (Amt) mediate uptake of nitrogenous nutrients. The nitrate reduction system consists of reduced ferredoxin (Fd_{red})-dependent nitrate reductase (NarB) and nitrite reductase (NirA), and the Ni^{2+}-containing urease consists of several subunits (UreABC). Ammonium is incorporated into glutamate by ATP-dependent glutamine synthetase (GS) producing glutamine, whose amido group is transferred to 2-oxoglutarate (2-OG) by reduced ferredoxin-dependent glutamate synthase (GOGAT). 2-OG is a product of carbon assimilation. *Anabaena* is predicted to express the inorganic carbon transport systems common in cyanobacteria: NDH-1_3 (localized in the thylakoid) and NDH-1_4 (which may also be localized in the thylakoid) participate in CO_2 uptake with high and low affinity, respectively, producing bicarbonate; secondary, sodium-dependent SbtA and BicA permeases transport bicarbonate into the cells with high and low affinity, respectively; and the ABC transporter Cmp concentrates bicarbonate in the cells dependent on ATP. Bicarbonate diffuses into the carboxysome where it is dehydrated by carbonic anhydrase (CA), and where CO_2 is combined with ribulose 1,5-bisphosphate (produced by the Calvin-Benson-Bassham cycle –not shown) by RubisCo to produce 3-phosphoglycerate (3-PG) that follows metabolism down to 2-OG. The prevalence of some specific transporters and pathways depends on environmental conditions as explained in the text. Stoichiometries are not depicted.

Fig. 2. Schematic of the nitrate assimilation gene cluster and its regulation in *Anabaena*. The nitrate assimilation structural genes (*nirA* [nitrite reductase], *nrtABCD* [ABC transporter], *narB* [nitrate reductase]) form an operon, whose expression is activated by NtcA bound to 2-oxoglutarate (2-OG) and the LTTR factor NtcB, both of which bind to the *nirA* promoter. Expression of *ntcB* is itself activated by NtcA+2-OG. Full expression of the *nirA* operon also requires the glycosyl transferase-like protein CnaT acting through an unknown mechanism. NirB and NarM are required for *Anabaena* to show full nitrite reductase (NirA) or nitrate reductase (NarB) activity, respectively. In the absence of a substrate (nitrate or nitrite), the NirA/NirB complex appears to negatively affect expression of the *nirA* operon.

Fig. 3. The dissimilar filaments of *Anabaena*. (A) Micrographs showing fragments of filaments of *Anabaena* grown with nitrate or N_2 as nitrogen source. Note the presence of heterocysts only in the latter (white arrowhead, intercalary heterocyst; black arrowhead, terminal heterocyst) (micrographs courtesy of Sergio Camargo, CSIC, Seville). (B) Schematic of filaments of *Anabaena* grown with combined nitrogen or N_2. The homogeneous metabolism in the cells of filaments grown with combined nitrogen ([N], upper filament) and the metabolic specialization in cells of diazotrophic filaments (lower filament; the central cell is a N_2-fixing heterocyst) is highlighted. In the left-hand cell of the upper filament, the role of transcription factors NtcA (with effector 2-oxoglutarate [2-OG]) and PacR in the regulation of nitrogen and carbon assimilation, respectively, is highlighted. In the diazotrophic filament, intercellular exchange of metabolites (sucrose and glutamate in exchange for glutamine and b-aspartyl-arginine) and transfer of regulators (PatS- and HetN-derived morphogens) from heterocysts or prospective heterocysts to adjacent vegetative cells are indicated. HetR plays a role in heterocyst differentiation, and NtcA (with 2-OG) plays a role in heterocyst differentiation and in the mature heterocyst. Note that septal junctions (blue lines) that mediate intercellular molecular exchange are present in all types of cells.

Fig. 4. Upon nitrogen deficiency, the low amount of NtcA present in the cells of filaments growing with combined nitrogen (depicted in black) is activated in response to an increase in the 2-oxoglutarate (2-OG) levels (activated NtcA depicted in red). This should suffice to activate early genes involved in heterocyst differentiation, including the *nrrA* gene. Then, NrrA and NtcA promote activation of the *hetR* gene, initially in clusters of cells. Self and mutual activation of HetR and NtcA results in high levels of both regulators, with an increased single cell specificity promoted by lateral inhibition of the PatS morphogen (not depicted). Then, HetR-dependent genes are activated in specific cells, including the *pipX* gene that is activated at late stages. Activity of HetR will thereafter decay (with a possible role of HetR phosphorylation) and transcription dependent on both NtcA and PipX will predominate in the mature heterocyst. The width of the horizontal bars at the bottom represent cellular levels of NtcA (red), HetR (blue) and PipX (purple) during differentiation.

1169 Fig. 5. Schematic representation of specific and cross regulatory interactions in the
1170 assimilation of nitrogen and carbon and in photosynthesis/high-light stress in *Anabaena*.
1171 NtcA is a global transcriptional regulator that coordinates N assimilation in response to
1172 the 2-oxoglutarate levels, which are largely determined through the activity of the Calvin-
1173 Benson-Bassham (CBB) and GS/GOGAT cycles. NtcB is a pathway-specific regulator
1174 of nitrate assimilation that enhances activation by NtcA of the *nir* operon, and the *ntcB*
1175 gene is activated by NtcA. CmpR is a pathway-specific regulator of Ci assimilation that
1176 activates the expression of the *cmp* operon (bicarbonate uptake) and represses RubisCo
1177 and carboxysome genes in response to low Ci availability; the *cmpR* gene is activated by
1178 NtcA. PacR coordinates regulation of elements of the carbon concentrating mechanism,
1179 RubisCo and photosynthetic elements including some for protection against oxidative
1180 damage, according to Ci availability and illumination conditions. In addition, NtcA
1181 participates in the regulation of the *rbcL* operon during heterocyst differentiation. Thus,
1182 the metabolism of *Anabaena* is globally regulated by the C-to-N balance. CA, carbonic
1183 anhydrase; 2-OG, 2-oxoglutarate; 3-PG, 3-phosphoglycerate.
1184
1185