FEMS Microbiology Letters, 366, 2019, fnz077

doi: 10.1093/femsle/fnz077
Advance Access Publication Date: 16 April 2019
Minireview

MINIREVIEW – Physiology & Biochemistry

Organization and regulation of cyanobacterial *nif* gene clusters: implications for nitrogenase expression in plant cells

Teresa Thiel*

Department of Biology, University of Missouri-St. Louis, One University Blvd., St. Louis, MO 63121, USA

*Corresponding author: Department of Biology, University of Missouri-St. Louis, One University Blvd, St. Louis, MO 63121 USA. Tel: 01-314-516-6208; E-mail: thiel@umsl.edu

One sentence summary: This review describes the organization of nitrogenase genes and their regulation by the transcription factor CnfR via its highly conserved binding sites in three different types of cyanobacteria, *Anabaenavariabilis* ATCC 29413, *Leptolyngbya boryana* dg5 and *Cyanothece* sp. ATCC 51142 as well as attempts to express the *nif* genes in heterologous cyanobacteria.

Editor: Hermann Bothe

ABSTRACT

For over 50 years scientists have considered the possibility of engineering a plant with nitrogen fixation capability, freeing farmers from their dependence on nitrogen fertilizers. With the development of the tools of synthetic biology, more progress has been made toward this goal in the last 5 years than in the previous five decades. Most of the effort has focused on nitrogenase genes from *Klebsiella oxytoca*, which has complex gene regulation. There may be advantages in using nitrogenase genes from cyanobacteria, which comprise large polycistronic gene clusters that may be easier to manipulate and eventually express in a plant. The fact that some diatoms have a cyanobacterial nitrogen fixing organelle further supports the idea that a cyanobacterial nitrogenase gene cluster may function in a newly-engineered, cyanobacterial-based plant organelle, a nitroplast. This review describes recent attempts to express the *nif* genes from *Anabaena variabilis* ATCC 29413, *Leptolyngbya boryana* dg5 and *Cyanothece* sp. ATCC 51142 in heterologous cyanobacteria in the context of the organization of the nitrogenase genes and their regulation by the transcription factor CnfR via its highly conserved binding sites.

Keywords: nitrogenase; cyanobacteria; heterocysts; regulation; CnfR

INTRODUCTION

Engineering a plant to fix nitrogen could provide enormous benefits to society in terms of increased food production, environmental protection and reduction in fossil fuel consumption. The critical step of moving nitrogenase genes from a diazotrophic strain to a nondiazotroph was first demonstrated more than 45 years ago in the bacterium *Escherichia coli*, (Dixon and Postgate 1972) but only relatively recently has an *E. coli* strain with *nif* genes produced an active nitrogenase (Wang *et al.* 2013). Because of the regulatory complexity of the *nif* gene cluster source, *Klebsiella oxytoca*, considerable manipulation of the genes was needed to optimize expression (Gouy, Guindon and Gascuel 2010; Temme, Zhao and Voigt 2012; Wang *et al.* 2013; Smanski *et al.* 2014; Yang *et al.* 2018). An even bigger challenge will be to synthesize functional nitrogenase, which is exquisitely sensitive to oxygen, in a eukaryote. Experiments using yeast mitochondria (Buren *et al.* 2017) and plant mitochondria (Allen *et al.* 2017) as hosts have resulted in the expression of individual Nif proteins, but not an active nitrogenase. Cyanobacteria are the evolutionary ancestors of chloroplasts, so there has also been

Received: 7 March 2019; **Accepted:** 11 April 2019

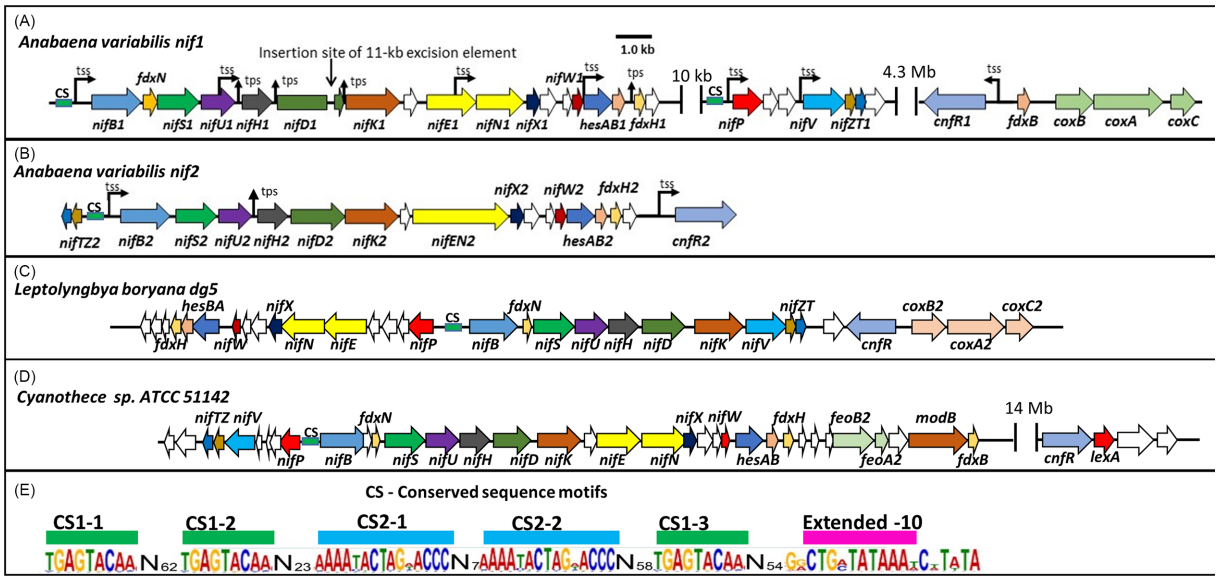

Figure 1. Maps of *nif* gene clusters and *cnfR* genes from IMG-JGI genome database (Chen *et al.* 2019). Arrows in panels A and B indicate published transcription start sites (tss) and transcription processing sites (tps) (Pratte *et al.* 2015). Green boxes labelled CS shown upstream of *nifB* and *nifP* refer to the conserved motifs shown in panel E. Conserved motifs, panel E, calculated using MEME (Bailey *et al.* 2009).

interest in expressing *nif* genes from a nitrogen-fixing cyanobacterial strain in a cyanobacterial strain that lacks the *nif* genes, which has been successful (Liu *et al.* 2018; Tsujimoto *et al.* 2018). The advantage of this system is that cyanobacterial cells are likely to require less modification for high level expression of nitrogenase than mitochondria. In addition cyanobacteria might function as a nitrogen fixing endosymbiont, similar to the intracellular cyanobacterial endosymbionts found in nitrogen fixing diatoms in the family Rhopalodiaceae (Floener and Bothe 1980; Prechtl *et al.* 2004; Kneip *et al.* 2007; Nakayama *et al.* 2011). These cyanobacterial spheroid bodies are phylogenetically most similar to the genus Cyanothece; however, they have a much reduced genome that has the genes for nitrogen fixation, but lacks the genes for photosynthesis (Nakayama *et al.* 2014).

Understanding the regulation of the large *nif* gene clusters in cyanobacteria, including the similarities and differences in the mechanisms that control *nif* gene expression levels, will be important in determining the best *nif* cluster that may eventually allow a plant to fix its own nitrogen. The best characterized cyanobacterial *nif* clusters include those from *Anabaena variabilis* ATCC 29413 (Thiel and Pratte 2014), *Leptolyngbya boryana* strain dg5 (Tsujimoto, Kamiya and Fujita 2014; Tsujimoto, Kamiya and Fujita 2016; Tsujimoto *et al.* 2018) and *Cyanothece* sp. ATCC 51142 (Stockel *et al.* 2011; Liu *et al.* 2018) (Fig. 1D).

Nitrogen fixation in cyanobacteria

Anabaena variabilis ATCC 29413 is a filamentous cyanobacterium that responds to nitrogen deficiency by differentiating specialized cells called heterocysts that fix atmospheric nitrogen under aerobic growth conditions (Kumar, Mella-Herrera and Golden 2010; Muro-Pastor and Hess 2012; Herrero, Picossi and Flores 2013; Flores *et al.* 2018). In contrast, the filamentous strain *L. boryana* and the unicellular strain *Cyanothece* sp., like most non-heterocyst forming strains, temporally separate nitrogen fixation and oxygen-evolving photosynthesis to protect nitrogenase from oxygen (Misra and Tuli 2000; Schneegurt *et al.* 2000; Stockel *et al.* 2011). *Cyanothece* sp. is unusual because, unlike *L. boryana*, it can fix nitrogen well in the dark in an aerobic environment (Colon-Lopez, Sherman and Sherman 1997).

Heterocysts, which comprise 5%–10% of the cells in a filament, maintain a microoxic environment via the structure of their cell wall, their lack of oxygen-evolving photosystem II, and their high rate of respiration, all features that protect the oxygen-labile nitrogenase from oxygen (Murry, Horne and Benemann 1984; Walsby 1985; Murry and Wolk 1989; Valladares *et al.* 2003; Walsby 2007). *Anabaena variabilis* is unusual because it has three nitrogenases: a heterocyst-specific, Mo-nitrogenase encoded by the *nif1* genes, a heterocyst-specific V-nitrogenase encoded by the *vnf* genes, and a second Mo-nitrogenase encoded by the *nif2* genes that is made only in anaerobic vegetative cells (Thiel 2004; Thiel and Pratte 2014). Expression of the *nif1* cluster of *A. variabilis* depends on the differentiation of heterocysts, which is controlled by a cascade of regulatory factors including NtcA, HetR, NrrA and DevH (Muro-Pastor and Hess 2012; Herrero, Picossi and Flores 2013; Flores *et al.* 2018). NtcA is not only important in heterocyst differentiation, but also more generally mediates global responses, especially to changes in the carbon to nitrogen ratio in the cell that are typically sensed by the levels of 2-oxoglutarate (Forchhammer 2004; Herrero and Flores 2018). When 2-oxoglutarate levels are high and fixed nitrogen levels are low, PipX and 2-oxoglutarate bind to NtcA, which activates genes that respond to nitrogen deprivation leading to heterocyst differentiation (Valladares, Flores and Herrero 2008; Valladares *et al.* 2011; Espinosa *et al.* 2014; Picossi, Flores and Herrero 2014; Herrero and Flores 2018). NtcA is required for expression of the *nif2* nitrogenase in anaerobic vegetative cells of *A. variabilis* (Thiel and Pratte 2001) by activating expression of the transcription factor, CnfR2, which activates *nifB2* (described below). In contrast, another key regulator, HetR, is not required for the expression of the *nif2* nitrogenase in anaerobic vegetative cells of *A. variabilis* (Thiel, unpublished).

The organization of the large *nif1* and *nif2* nitrogenase gene clusters in *A. variabilis* is very similar, as are the *nif* gene clusters in filamentous nitrogen-fixing cyanobacteria without heterocysts, like *L. boryana*, and unicellular strains, such as *Cyanothece*

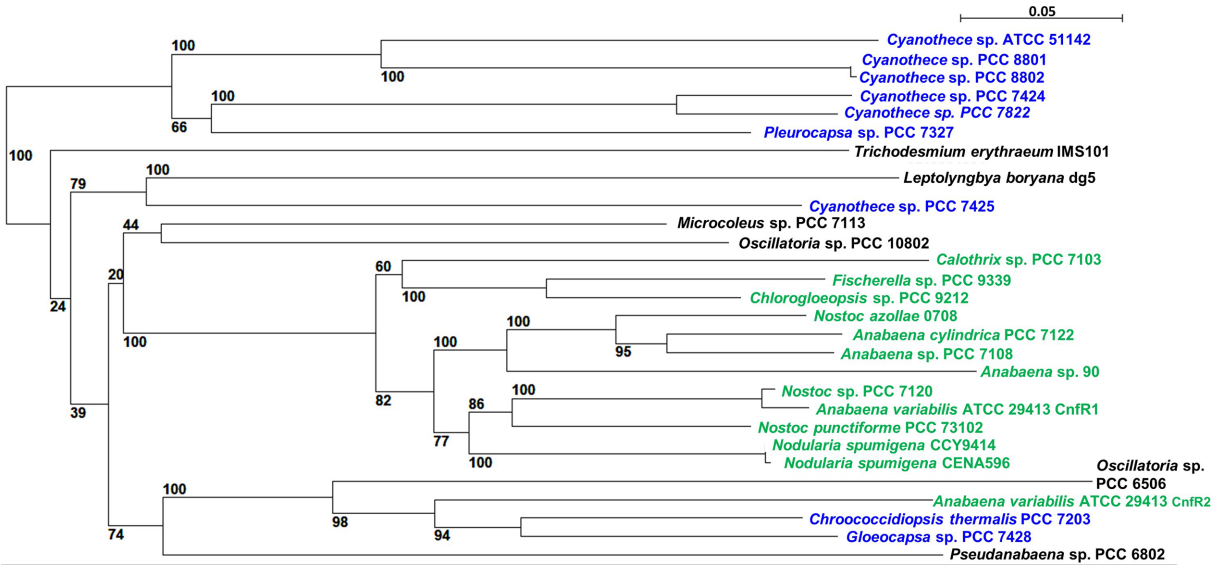

Figure 2. Phylogenetic tree of representative cyanobacterial CnfR deduced proteins. Heterocyst-forming strains are shown in green, filamentous non-heterocyst-forming strains in black and unicellular strains in blue. This BioNJ distance tree was constructed from CnfR sequences obtained from JGI-IMG (Chen *et al.* 2019) using the program SeaView (Gouy *et al.* 2010) with 100-replicate bootstrap values.

sp. In contrast to the mostly contiguous *nif* genes in *A. variabilis*, the *nif* genes of *L. boryana and Cyanothece* ATCC 51142 comprise two divergently transcribed groups of *nif* genes (Fig. 1C and D). The *nif* regulatory gene *cnfR* is linked to the *nif* cluster in some (e.g. *nif2* in *A. variabilis* and the *nif* cluster of *L. boryana*) but not all clusters (e.g. *nif1* in *A. variabilis* and the *nif* cluster of *Cyanothece* sp.) (Fig. 1A–D). The *cnfR* gene of *Cyanothece* ATCC 51142 appears to be cotranscribed with a transcriptional regulator, *lexA*, located immediately downstream of *cnfR*; however, this not true for other cyanobacteria, including other strains of *Cyanothece* spp. This divergence in *cnfR* among strains of *Cyanothece* spp. is reflected in the phylogeny of the *Cyanothece cnfR* genes (Fig. 2)

nif regulatory proteins

CnfR is a regulatory protein with a C-terminal XRE-type, DNA-binding domain and two N-terminal 4Fe-4S binding sites, similar to bacterial ferredoxins (Liang, Scappino and Haselkorn 1993; Jones, Buikema and Haselkorn 2003; Pratte and Thiel 2016). This gene, found in all nitrogen fixing cyanobacterial genomes, forms a clear monophyletic group only for *nif* systems that function in heterocysts (Fig. 2). The *cnfR* gene of *Anabaena* sp. PCC 7120 was first identified as a frameshift mutation in a gene that was then called *pat-2* which resulted in a strain with poor growth in the absence of fixed nitrogen, a high frequency of heterocysts, and fragmentation of filaments (Liang, Scappino and Haselkorn 1993). The gene, subsequently renamed *patB*, was found to be expressed exclusively in heterocysts late in their development. A *patB* deletion mutant failed to grow in the absence of fixed nitrogen, but had very low levels of nitrogenase activity, about 0.25% of normal activity (Jones, Buikema and Haselkorn 2003). In the filamentous non-heterocystous cyanobacterium *L. boryana*, the *nif* genes (Fig. 1C) were not expressed in a mutant with a deletion of a homolog of *patB*, called *cnfR*, leading to the conclusion that CnfR is the major regulator for *nif* transcription in that strain (Tsujimoto, Kamiya and Fujita 2014). The CnfR protein of *L. boryana* with a deletion in either the conserved CnfR N-terminal iron-sulfur cluster binding region or the C-terminal DNA binding region resulted in loss of transcription of *nifB* (Tsujimoto, Kamiya and Fujita 2016). In contrast, in *Anabaena* sp. PCC 7120, a mutant of CnfR that destroyed the terminal iron-sulfur cluster binding region, by replacement of the N-terminal six cysteines with alanines, grew, albeit slowly, in the absence of fixed nitrogen. A frameshift mutation, which resulted in loss of the C-terminal DNA-binding domain of CnfR in *Anabaena* sp. PCC 7120 also grew slowly in the absence of fixed nitrogen, with about 14% of wild-type nitrogenase activity (Jones, Buikema and Haselkorn 2003).

Anabaena variabilis has two copies of *cnfR* encoding CnfR1, which activates the *nifB1* promoter only in heterocysts and CnfR2, which activates the *nifB2* promoter only in anaerobic vegetative cells (Pratte and Thiel 2016). CnfR1 and CnfR2 do not cluster together phylogenetically. CnfR1 clusters with genes from other heterocyst-specific nitrogenases, while CnfR2 clusters with nitrogenases from filamentous, non-heterocystous cyanobacteria (Fig. 2). Deletion mutations in *cnfR1* or *cnfR2* abolish the expression of *nifB1* or *nifB2*, respectively. The *cnfR1* mutation had little effect on expression of the secondary *nif1* promoters, such as the one inside the *nifE1* gene, or on *hesA1*, which has its own promoter (Fig. 1A). Expression of *nifP* was reduced only 2–3 fold in the *cnfR1* mutant (Pratte and Thiel 2016). CnfR1 is not required for expression of the *vnfDGKEN* genes that encode the structural and scaffolding proteins of the heterocyst-specific V-nitrogenase of *A. variabilis*; however, because the V-nitrogenase requires NifB1, a functional V-nitrogenase cannot be made in a *cnfR1* mutant (Pratte and Thiel 2016). The *vnf* genes are regulated by VnfR1 and VnfR2 (Pratte *et al.* 2013).

In bacterial transcription, the RNA polymerase core enzyme requires a sigma factor for initial binding to the promoter. The *nifB1* and *nifB2* promoters have strikingly conserved sequences upstream of the transcription start site that form a canonical 'extended -10 promoter' (Fig. 1E) (Mitchell *et al.* 2003), but they lack a recognizable -35 region. They represent a type 2 promoter, which typically requires a transcription factor and often uses alternative sigma factor in addition to the normal, housekeeping σ^{70} factor (Imamura and Asayama 2009). Thus, activation of the *nifB1* or *nifB2* promoters may require cell-type-specific

expression of alternative sigma factors. In addition to the normal group 1, housekeeping σ^{70} factor, bacteria have several alternative sigma factors in groups 2 and 3 (Paget 2015). Cyanobacteria typically have 9–12 sigma factors, including the housekeeping SigA (group 1); Sig B, C, D, E (group 2); and Sig F, G, H, I, J (group 3), but they lack the σ^{N} factor often associated with control of nitrogen-stress related genes in other bacteria (Asayama *et al*. 2004; Imamura *et al*. 2006; Imamura and Asayama 2009; Paget 2015). In *Synechocystis* sp. PCC 6803 a mutant of all group 2 sigma factors (ΔsigBCDE) led to poor growth under many stress conditions, including oxidative stress, nitrogen stress, heat, high light or high salt; however, in the absence of environmental stress the deletion strain grew nearly as well as the wild-type strain (Antal *et al*. 2016; Koskinen *et al*. 2016; Hakkila *et al*. 2019). Mutants in any one of the group 2 sigma factors genes had little effect on short-term growth with low levels of fixed nitrogen. A double mutant lacking *sigB* and *sigD* grew poorly under conditions of nitrogen stress and a triple mutant lacking *sigB*, *sigD* and *sigE* recovered very poorly after nitrogen starvation (Antal *et al*. 2016). In *Anabaena* sp. PCC 7120 expression of *sigC*, *sigE* and *sigG* is increased in differentiating heterocysts (Aldea, Mella-Herrera and Golden 2007; Ehira and Miyazaki 2015). SigB and SigC function for expression of some NtcA-dependent, nitrogen-responsive genes (Imamura and Asayama 2009). Both SigC (Ehira and Miyazaki 2015) and SigE (Mella-Herrera *et al*. 2011) are important, but not essential, for expression of the *nif* genes in heterocysts of *Anabaena* sp. PCC 7120. Upregulation of SigC, starting at 3–4 h after nitrogen removal, requires HetR, a key activator of heterocyst differentiation, and transcription of *hetR* requires the global nitrogen regulator NtcA (Huang, Dong and Zhao 2004; Zhang, Chen and Zhang 2009; Herrero, Picossi and Flores 2013). SigE is implicated in the expression of *nif* genes in the unicellular cyanobacterium, *Cyanothece* (Mueller *et al*. 2016). Together, these results suggest that *sigB*, *sigC*, *sigD* and *sigE* probably have overlapping roles in transcribing genes that are important in nitrogen stress and possibly in nitrogen fixation.

Promoters and other cis-acting elements for *nif* genes

The *nif1* genes, which are heterocyst-specific, even in cells grown under anaerobic conditions (Elhai and Wolk 1990; Thiel *et al*. 1995), are primarily under the control of the promoter for *nifB1*, the first gene in the 15-kb cluster (Fig. 1A) (Ungerer, Pratte and Thiel 2010; Pratte and Thiel 2014). Similarly, the *nif2* genes are under the control of the *nifB2* promoter. There are additional weak promoters located in *nifU1* and *nifE1* and *hesA1* has its own strong promoter (Pratte and Thiel 2014) (Fig. 1A). The relatively high abundance of some transcripts in both the *nif1* and *nif2* operons, most notably *nifH1* and *nifH2*, is the result of the high stability of these transcripts, not a strong promoter. This stability is due to conserved stem-loop structures located at RNA processing sites upstream of these genes (Pratte and Thiel 2014). The RNA transcriptional processing sites are indicated in Fig. 1A and B. While it seems likely that the promoter for *nifB* serves as the primary promoter for the *nif* operons in other cyanobacteria, this has not yet been clearly demonstrated. However, the region upstream of *nifH* in the filamentous non-heterocystous strain, *L. boryana*, showed no promoter activity, in contrast to the region between the *nifB* and *nifP* genes which serves as the promoter region for the two divergently transcribed operons *nifBSUHDKVZT* and *nifPENXW* (Tsujimoto, Kamiya and Fujita 2016) (Fig. 1C).

Conserved cis-acting elements that were identified in a region that extended 300–600 bp upstream of the transcription start sites of *nifB1*, *nifB2* and *nifP* in *A. variabilis* (Vernon, Pratte and Thiel 2017) and upstream of the promoters of *nifB* and *nifP* in *L. boryana* (Tsujimoto, Kamiya and Fujita 2016) are also highly conserved in most diazotrophic cyanobacteria. Three of these conserved motifs (CS1–1, CS1–2 and CS1–3) (Fig. 1E) have been studied the most extensively. The *nifB* promoter region of *L. boryana*, fused to a *luxAB* reporter gene, with *cnfR* was expressed in *Synechocystis* sp. PCC 6803 under the control of the *trc* promoter. In this heterologous expression system, mutation of the *nifB* CS1–3 conserved motif, (corresponding to Region IX in *L. boryana*), abolished transcription from the *nifB* promoter, while mutation of CS1–1 or CS1–2 (corresponding to Region 0 and I in *L. boryana*) increased transcription from the *nifB* promoter compared to a wild-type sequence (Tsujimoto, Kamiya and Fujita 2016). The results of similar experiments done in vivo in *A. variabilis* had opposite results. In the case of expression of *nifB1*, the loss of either CS1–1 or CS1–2 decreased transcription from *nifB1* about 50% compared to a sequence with all three motifs, and loss of both CS1–1 and CS1–2 led to a 75% reduction in expression of *niB1*. Loss of the entire region upstream of CS1–2, including CS1–2, still showed good heterocyst-specific expression from the *nifB1* promoter (Vernon, Pratte and Thiel 2017).

While the conserved regions upstream of *nifB1* and *nifB2* in *A. variabilis* are very similar, they do not function interchangeably. The conserved region upstream of *nifB2* can replace the similar region upstream of *nifB1* for expression of *nifB1* (Fig. 3A), but the converse is not true; replacement of the region of the *nifB2* promoter that includes the CS1–1 motif with the similar upstream region of *nifB1* results in 80% loss of *nifB2* expression, while replacement of the entire conserved region of the *nifB2* promoter with the similar region from *nifB1* resulted in complete loss of *nifB2* expression. High-level expression of *nifB2* requires not only the conserved motif region of *nifB2*, but also the 5' untranslated region as well as sequences within the 5' end of the gene (Fig. 3B) (Vernon, Pratte and Thiel 2017). Despite the similarity in the CnfR1 and CnfR2 proteins, CnfR2 can only activate the *nifB2* conserved region, while CnfR1 activates either *nifB1* or *nifB2*. This has been confirmed by the fact that expression of *cnfR1* from the *cnfR2* promoter in anaerobic vegetative cells results in activation of both *nifB1* and *nifB2*, whereas expression of *cnfR2* from the *cnfR1* promoter in heterocysts activates neither *nifB1* nor *nifB2* (Pratte and Thiel 2016). Thus, it appears that activation of *nifB2* requires not only CnfR2, but also another factor that is present in anaerobic vegetative cells but absent from heterocysts. A hybrid promoter construct that is almost entirely the upstream region of *nifB2*, with only the -10 region of *nifB1*, can be activated by either CnfR1 in heterocysts or by CnfR2 in anaerobic vegetative cells (Fig. 3C) (Vernon, Pratte and Thiel 2017). Therefore, the conserved upstream region of *nifB2* is critical for its expression, and cannot be replaced by the similar region of *nifB1*, while expression from the *nifB1* promoter stringently requires only the extended -10 region of the *nifB1* promoter, but otherwise functions well using the homologous regions of the *nifB2* gene.

In addition to *nifB1* and *nifB2*, the upstream regions of two other genes in *A. variabilis* have the conserved motifs characteristic of CnfR activation sites: *nifP* and a gene of unknown function, *ava0453*, located not far from *cnfR1*, which is present only in heterocyst-forming cyanobacteria where it is expressed late in heterocyst differentiation (Thiel, unpublished data). The *nifP* (*cysE*) gene encodes a serine O-acetyltransferase, essential for the synthesis of cysteine; however most nitrogen-fixing cyanobacteria have *cysE* as well as *nifP*. In the Proteobacterium, *Azotobacter chroococcum*, the first strain in which *nifP* was

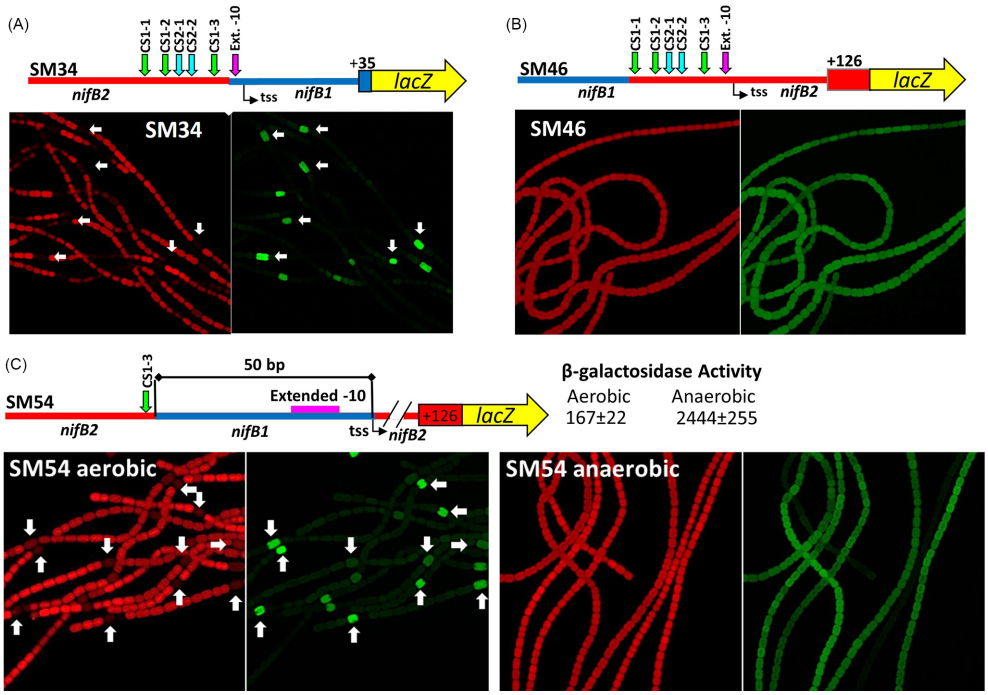

Figure 3. *In situ* expression of *nifB1* and *nifB2* hybrid promoters. (A) Heterocyst-specific expression of a *nifB2/nifB1* hybrid promoter fused to *lacZ*, 24 h after aerobic growth -N. (B) Anaerobic vegetative cell expression of a *nifB1/nifB2* hybrid promoter fused to *lacZ*, 6 h after anaerobic growth -N. (C) Expression of the *nifB2* promoter, modified with the *nifB1* extended -10 region, fused to *lacZ*, 24 h after aerobic growth -N (left panel) or 6 h after anaerobic growth -N (right panel). White arrows indicate heterocysts. Right panels, fluorescein fluorescence from cleavage of 5-dodecanoyl-fluorescein-β-D-galacto-pyranoside by β-galactosidase; left panels, red fluorescence from photosynthetic pigments in cyanobacteria (less in mature heterocysts). Modified from (Vernon *et al.* 2017).

described, mutation of *nifP* resulted in a delay in nitrogen fixation and slower growth of the strain under nitrogen fixing conditions. (Evans *et al.* 1991). In *A. variabilis*, there is a single *nifP* gene, which may serve both the Nif1 and the Nif2 nitrogenases. A *nifP* deletion mutant grows well and fixes nitrogen similarly to the wild-type strain (Thiel, unpublished results). However, in *L. boryana* deletion of *nifP* and the three small orfs immediately downstream reduced nitrogenase to less than 10% of wild-type activity (Tsujimoto, Kamiya and Fujita 2014). In *L. boryana* the conserved upstream region is not sufficient for expression of *nifP*, which also requires an additional 900 bp upstream of *nifP* (Tsujimoto, Kamiya and Fujita 2016).

The role of CnfR in cell-type specific *nif* regulation

In *A. variabilis* cell-type specific expression of *nifB1* (only in heterocyst) or *nifB2* (in anaerobic vegetative cells) depends largely on the cell-type specificity of *cnfR1* or *cnfR2* expression. Heterologous expression of the *cnfR1* gene under the control of the *cnfR2* promoter in anaerobic vegetative cells results in activation of both the *nifB1* and the *nifB2* promoters in these cells. Hence, cell-type specificity of *nifB1* expression is determined solely by the heterocyst-specific expression of *cnfR1*. In contrast, expression of the *cnfR2* gene under the control of the *cnfR1* promoter in heterocysts does not result in expression of either the *nifB1* promoter or the *nifB2* promoter in these cells (Pratte and Thiel 2016). Thus CnfR1 activates only *nifB1* in heterocysts because *cnfR1* is only made in heterocysts and because *nifB2* cannot be activated in heterocysts by CnfR1. In contrast, in anaerobic vegetative cells, *nifB2* can be activated by either CnfR1 or CnfR2, suggesting that there is a factor required for *nifB2* activation that is only present in anaerobic vegetative cells. Therefore, the activation of *nifB2* in anaerobic vegetative cells by CnfR2 results not only from the fact that *cnfR2* is expressed only in those cells, but also that expression of *nifB2* requires another factor that is only made in anaerobic vegetative cells.

Expression of nitrogenase genes in a heterologous host

Anabaena sp. PCC 7120, is very similar to *A. variabilis*, but it lacks the *nif2* cluster. When the *cnfR2* gene of *A. variabilis* was inserted into the chromosome of *Anabaena* sp. PCC 7120, CnfR2 was expressed in anaerobic vegetative cells and activated *nifB2*, indicating that CnfR2 is necessary and sufficient for expression of *nifB2* in anaerobic vegetative cells of *Anabaena* sp. PCC 7120 (Pratte and Thiel 2016) (Fig. 4). When the entire *A. variabilis nif2* cluster (from *nifT2-cnfR2*, Fig. 1B) was transferred to *Anabaena* sp. PCC 7120, it was able to fix nitrogen under anoxic conditions in vegetative cells, but the activity of the Nif2 system was about 5% of that observed for *A. variabilis*. Transcript analysis of this engineered *Anabaena* sp. PCC 7120 strain showed that expression of *cnfR2* was similar to that observed for *A. variabilis*, but that expression of *nifB2* and *nifH2* was 5- to 10-fold lower, leading to the lower nitrogenase activity (Thiel, unpublished). When the *cnfR2* gene, with its own promoter, from *A. variabilis* was integrated in the *Synechocystis* sp. PCC 6803 genome there was little transcription of *cnfR2* and no activation of the *nifB2* promoter under anoxic conditions suggesting that *Synechocystis* 6803 lacks a necessary factor for transcription of *cnfR2* (Thiel, unpublished data).

The *cnfR* gene of *L. boryana*, under the control of the constitutive *trc* promoter, was expressed in the non-nitrogen fixing cyanobacterium, *Synechocystis* sp. PCC 6803. In this heterologous strain, the *nifB* and *nifP* promoters, fused to a *luxAB* reporter, were expressed only under anaerobic conditions, suggesting that factors in addition to the constitutively expressed CnfR,

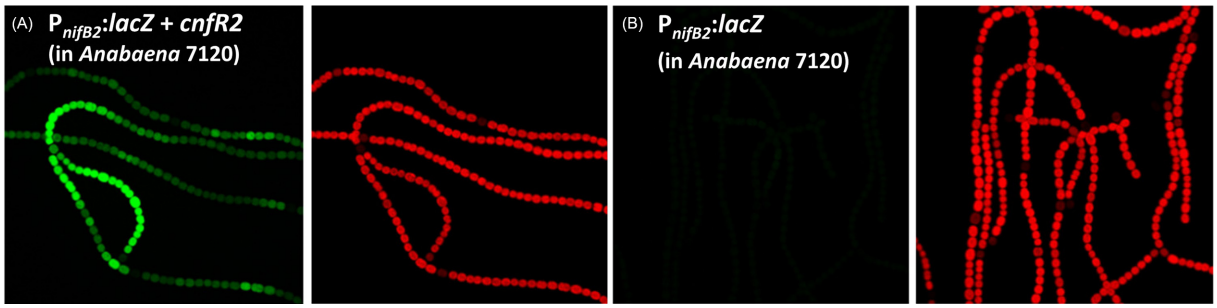

Figure 4. Expression of *nifB2* in *Anabaena* sp. PCC 7120. *In situ* localization of expression of *lacZ*, under the control of the *nifB2* promoter in *Anabaena* sp. PCC 7120 with the *cnfR2* gene of *A. variabilis* (panel A) or without the *cnfR2* gene of *A. variabilis* (panel B). β-galactosidase expression was visualized in cells grown for 24 h under anaerobic conditions -N. Left panels, fluorescein fluorescence from cleavage of 5-dodecanoyl-fluorescein-β-D-galacto-pyranoside by β-galactosidase; right panels, red fluorescence from photosynthetic pigments in cyanobacteria (less in mature heterocysts, which appear dark in the right panel of each set). Modified from (Pratte, Thiel 2016).

apparently present only in anaerobic cells of *Synechocystis* sp. PCC 6803, were also required for *nif* gene expression (Tsujimoto, Kamiya and Fujita 2016). The *nifB* promoter was activated only in the absence of fixed nitrogen, but the *nifP* promoter was weakly activated in cells grown with fixed nitrogen. Recently, the 25 gene *nif* cluster from *fdxB* to *nifT* (Fig. 1C) from *L. boryana* was cloned with *cnfR*, under the control of the *trc* promoter, into a neutral site in the genome of *Synechocystis* 6803 producing strain CN1. Under anaerobic conditions, CN1 produced NifH, NifD and NifK proteins at about 6%–23% of the levels in *L. boryana*; however, nitrogenase levels were only about 0.26% of those of *L. boryana* (Tsujimoto *et al.* 2018).

The first transfer of *nif* genes to *Synechocystis* 6803 with significant nitrogenase expression was recently demonstrated using a region with 24 *nif* genes from *Cyanothece* sp. ATCC 51142, from *nifT* to *hesB* (Fig. 1D). The *nif* genes were integrated into a small, multi-copy plasmid in *Synechocystis* 6803, pCA2.4, creating strain TSyNif-9. Under anaerobic growth conditions, *Synechocystis* TSyNif-9 was able to fix nitrogen at 13%–31% (acetylene reduction versus N_2 reduction values, respectively) of the rate of *Cyanothece* (Liu *et al.* 2018). However, without the increased level of expression provided by expression of the *nif* gene on a plasmid, nitrogenase expression in *Synechocystis* was about 2% of the rate of *Cyanothece* (Liu *et al.* 2018). This suggests that expressing cyanobacterial *nif* genes on a host strain-compatible plasmid is a reasonable strategy for increasing nitrogenase activity in a heterologous host. *Synechocystis* TSyNif-9 lacks the *cnfR* gene, suggesting that, unlike *A. variabilis* or *L. boryana*, in *Cyanothece nif* genes do not require this activator. It would be interesting to see whether the addition of *cnfR* to TSyNif-9 could further increase nitrogenase production. It should be noted that because nitrogenase activity was reported using different units for specific activity in the *Synechocystis* strains expressing the *nif* genes from the *L. boryana* versus those from *Cyanothece*, it is very difficult to compare the actual amounts of nitrogenase activity between these strains (Liu *et al.* 2018; Tsujimoto *et al.* 2018).

CONCLUSIONS

The large contiguous *nif* gene clusters, lacking excision elements, described here for *A. variabilis nif2*, *Cyanothece*, and *L. boryana* show promise for heterologous expression in plants. These *nif* genes comprise large operons, with few promoters. The transcription factor CnfR is the primary activator of *nifB* and *nifP* promoters, but it is likely that other factors, possibly sigma factors, are also involved in regulation. Transcript abundance may be regulated primarily by posttranscriptional processing and differential transcript stability (Pratte, Ungerer and Thiel 2015), but may also be increased by expression of the *nif* genes from a multi-copy cyanobacterial plasmid (Liu *et al.* 2018). It is noteworthy that experiments in different laboratories attempting to express different cyanobacterial *nif* clusters in *Synechocystis* 6803 yielded low, but quite different levels of nitrogenase activity in the engineered strains (Liu *et al.* 2018; Tsujimoto *et al.* 2018). Thus, expressing cyanobacterial *nif* genes in plants appears feasible, but will require better understanding of the regulation of nitrogenase genes in cyanobacteria and optimization of conditions for nitrogenase activity in the host cell. The latter will likely include providing the appropriate redox state of the cell as well as sufficient reductant and ATP, key factors for high nitrogenase activity.

FUNDING

Support for this research was provided by National Science Foundation grant MCB-1818298.

ACKNOWLEDGEMENTS

I thank Brenda Pratte for her help with the research described in this review and in editing this manuscript.

Conflict of interest. None declared.

REFERENCES

Aldea MR, Mella-Herrera RA, Golden JW. Sigma factor genes *sigC*, *sigE*, and *sigG* are upregulated in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 2007;**189**:8392–6.

Allen RS, Tilbrook K, Warden AC *et al.* Expression of 16 nitrogenase proteins within the plant mitochondrial matrix. *Front Plant Sci* 2017;**8**:287.

Antal T, Kurkela J, Parikainen M *et al.* Roles of group 2 sigma factors in acclimation of the cyanobacterium *Synechocystis* sp. PCC 6803 to nitrogen deficiency. *Plant Cell Physiol* 2016;**57**:1309–18.

Asayama M, Imamura S, Yoshihara S *et al.* SigC, the group 2 sigma factor of RNA polymerase, contributes to the late-stage gene expression and nitrogen promoter recognition in

the cyanobacterium *Synechocystis* sp. strain PCC 6803. *Biosci Biotechnol Biochem* 2004;**68**:477–87.

Bailey TL, Boden M, Buske FA *et al*. MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res* 2009;**37**:W202–8.

Buren S, Young EM, Sweeny EA *et al*. Formation of nitrogenase NifDK tetramers in the mitochondria of *Saccharomyces cerevisiae*. *ACS Synth Biol* 2017;**6**:1043–55.

Chen IA, Chu K, Palaniappan K *et al*. IMG/M v.5.0: An integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res* 2019;**47**:D666–77.

Colon-Lopez MS, Sherman DM, Sherman LA. Transcriptional and translational regulation of nitrogenase in light-dark- and continuous-light-grown cultures of the unicellular cyanobacterium *Cyanothece* sp. strain ATCC 51142. *J Bacteriol* 1997;**179**:4319–27.

Dixon RA, Postgate JR. Genetic transfer of nitrogen fixation from *Klebsiella pneumoniae* to *Escherichia coli*. *Nature* 1972;**237**:102–3.

Ehira S, Miyazaki S. Regulation of genes involved in heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120 by a group 2 sigma factor SigC. *Life (Basel)* 2015;**5**:587–603.

Elhai J, Wolk CP. Developmental regulation and spatial pattern of expression of the structural genes for nitrogenase in the cyanobacterium *Anabaena*. *EMBO J* 1990;**9**:3379–88.

Espinosa J, Rodríguez-Mateos F, Salinas P *et al*. PipX, the coactivator of NtcA, is a global regulator in cyanobacteria. *Proc Natl Acad Sci USA*. 2014;**111**:E2423–30.

Evans DJ, Jones R, Woodley PR *et al*. Nucleotide sequence and genetic analysis of the *Azotobacter chroococcum nifUSVWZM* gene cluster, including a new gene (*nifP*) which encodes a serine acetyltransferase. *J Bacteriol* 1991;**173**:5457–69.

Floener L, Bothe H. Nitrogen fixation in *Rhopalodia gibba*, a diatom containing blue-greenish inclusions symbiotically. *Endocytobiol endosymbiosis cell biol* 1980;**1**:514–52.

Flores E, Picossi S, Valladares A *et al*. Transcriptional regulation of development in heterocyst-forming cyanobacteria. *Biochim Biophys Acta Gene Regul Mech*. 2018. doi.org/10.1016/j.bbagrm.2018.04.006

Forchhammer K. Global carbon/nitrogen control by PII signal transduction in cyanobacteria: From signals to targets. *FEMS Microbiol Rev* 2004;**28**:319–33.

Gouy M, Guindon S, Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 2010;**27**:221–4.

Hakkila K, Valev D, Antal T *et al*. Group 2 sigma factors are central regulators of oxidative stress acclimation in cyanobacteria. *Plant Cell Physiol*. 2019;**60**:436–47.

Herrero A, Flores E. Genetic responses to carbon and nitrogen availability in *Anabaena*. *Environ Microbiol*. 2018. doi.org/10.1111/1462-2920.14370

Herrero A, Picossi S, Flores E. Gene expression during heterocyst differentiation. *Advances in Botanical Research*, Franck C, Corinne C-C (eds). 2013;**65**, pp.281–329. Academic Press.

Huang X, Dong Y, Zhao J. HetR homodimer is a DNA-binding protein required for heterocyst differentiation, and the DNA-binding activity is inhibited by PatS. *Proc Natl Acad Sci USA*. 2004;**101**:4848–53

Imamura S, Asayama M. Sigma factors for cyanobacterial transcription. *Gene Regul Syst Bio* 2009;**3**:65–87.

Imamura S, Tanaka K, Shirai M *et al*. Growth phase-dependent activation of nitrogen-related enes by a control network of group 1 and roup 2 σ Factors in a cyanobacterium. *J Biol Chem* 2006;**281**:2668–75.

Jones KM, Buikema WJ, Haselkorn R. Heterocyst-specific expression of *patB*, a gene required for nitrogen fixation in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 2003;**185**:2306–14.

Koskinen S, Hakkila K, Gunnelius L *et al*. In vivo recruitment analysis and a mutant strain without any group 2 sigma factor reveal roles of different sigma factors in cyanobacteria. *Mol Microbiol* 2016;**99**:43–54.

Kneip C, Lockhart P, Voss C *et al*. Nitrogen fixation in eukaryotes-new models for symbiosis. *BMC Evol Biol* 2007;**7**:55.

Kumar K, Mella-Herrera RA, Golden JW. Cyanobacterial heterocysts. *Cold Spring Harb Perspect Biol* 2010;**2**:a000315.

Liang J, Scappino L, Haselkorn R. The *patB* gene product, required for growth of the cyanobacterium *Anabaena* sp. strain PCC 7120 under nitrogen-limiting conditions, contains ferredoxin and helix-turn-helix domains. *J Bacteriol* 1993;**175**:1697–704.

Liu D, Liberton M, Yu J *et al*. Engineering nitrogen fixation activity in an oxygenic phototroph. *MBio* 2018. doi:10.1128/mBio.01029-18

Mella-Herrera RA, Neunuebel MR, Kumar K *et al*. The *sigE* gene is required for normal expression of heterocyst-specific genes in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 2011; **193**:1823–32.

Misra HS, Tuli R. Differential expression of photosynthesis and nitrogen fixation genes in the cyanobacterium *Plectonema boryanum*. *Plant Physiol* 2000;**122**:731–6.

Mitchell JE, Zheng D, Busby SJ *et al*. Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucleic Acids Res* 2003;**31**:4689–95.

Mueller TJ, Welsh EA, Pakrasi HB *et al*. Identifying regulatory changes to facilitate nitrogen fixation in the nondiazotroph *Synechocystis* sp. PCC 6803. *ACS Synth Biol* 2016;**5**:250–8.

Muro-Pastor AM, Hess WR. Heterocyst differentiation: from single mutants to global approaches. *Trends Microbiol* 2012;**20**:548–57.

Murry MA, Wolk CP. Evidence that the barrier to the penetration of oxygen into heterocysts depends upon two layers of the cell envelope. *Arch Microbiol* 1989;**151**:469–74.

Murry MA, Horne AJ, Benemann JR. Physiological studies of oxygen protection mechanisms in the heterocysts of *Anabaena cylindrica*. *Appl Env Microbiol* 1984;**47**:449–54.

Nakayama T, Ikegami Y, Ishida K *et al*. Spheroid bodies in rhopalodiacean diatoms were derived from a single endosymbiotic cyanobacterium. *J Plant Res* 2011; **124**:93–97.

Nakayama T, Kamikawa R, Tanifuji G *et al*. Complete genome of a nonphotosynthetic cyanobacterium in a diatom reveals recent adaptations to an intracellular lifestyle. *Proc Natl Acad Sci USA* 2014;**111**:11407–12.

Paget MS. Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules* 2015;**5**:1245–65.

Picossi S, Flores E, Herrero A. ChIP analysis unravels an exceptionally wide distribution of DNA binding sites for the NtcA transcription factor in a heterocyst-forming cyanobacterium. *BMC Genomics* 2014;**15**:22.

Pratte BS, Sheridan R, James JA *et al*. Regulation of V-nitrogenase genes in *Anabaena variabilis* by RNA processing and by dual repressors. *Mol Microbiol* 2013;**88**:413–24.

Pratte BS, Thiel T. Regulation of nitrogenase gene expression by transcript stability in the cyanobacterium *Anabaena variabilis*. *J Bacteriol* 2014;**196**:3609–21.

Pratte BS, Thiel T. Homologous regulators, CnfR1 and CnfR2, activate expression of two distinct nitrogenase gene clusters in the filamentous cyanobacterium *Anabaena variabilis* ATCC 29413. *Mol Microbiol* 2016;**100**:1096–109.

Pratte BS, Ungerer J, Thiel T. Role of RNA secondary structure and processing in stability of the *nifH1* transcript in the cyanobacterium *Anabaena variabilis*. *J Bacteriol* 2015;**197**:1408–22.

Prechtl J, Kneip C, Lockhart P *et al*. Intracellular spheroid bodies of *Rhopalodia gibba* have nitrogen-fixing apparatus of cyanobacterial origin. *Mol Biol Evol* 2004;**21**:1477–81.

Schneegurt MA, Tucker DL, Ondr JK *et al*. Metabolic rhythms of a diazotrophic cyanobacterium, *Cyanothece* sp. strain ATCC 51142, heterotrophically grown in continuous dark. *J Phycol* 2000;**36**:107–17.

Smanski MJ, Bhatia S, Zhao D *et al*. Functional optimization of gene clusters by combinatorial design and assembly. *Nat Biotechnol* 2014;**32**:1241–9.

Stockel J, Jacobs JM, Elvitigala TR *et al*. Diurnal rhythms result in significant changes in the cellular protein complement in the cyanobacterium *Cyanothece* 51142. *PLoS One* 2011;**6**:e16680.

Temme K, Zhao D, Voigt CA. Refactoring the nitrogen fixation gene cluster from *Klebsiella oxytoca*. *Proc Natl Acad Sci USA* 2012;**109**:7085–90.

Thiel T. Nitrogen fixation in heterocyst-forming cyanobacteria. *Genetics and Regulation of Nitrogen Fixing Bacteria*, Klipp W, Masepohl B, Gallon JR, Newton WE (eds). Dordrecht, The Netherlands: Kluwer Academic Publishers, 2004, pp.73–110.

Thiel T, Pratte B. Effect on heterocyst differentiation of nitrogen fixation in vegetative cells of the cyanobacterium *Anabaena variabilis* ATCC 29413. *J Bacteriol* 2001;**183**:280–6.

Thiel T, Pratte B. Regulation of three nitrogenase gene clusters in the cyanobacterium *Anabaena variabilis* ATCC 29413. *Life* 2014;**4**:944–67.

Thiel T, Lyons EM, Erker JC *et al*. A second nitrogenase in vegetative cells of a heterocyst-forming cyanobacterium. *Proc Natl Acad Sci USA* 1995;**92**:9358–62.

Tsujimoto R, Kamiya N, Fujita Y. Transcriptional regulators ChlR and CnfR are essential for diazotrophic growth in nonheterocystous cyanobacteria. *Proc Natl Acad Sci USA*. 2014; **111**:6762–7.

Tsujimoto R, Kamiya N, Fujita Y. Identification of a cis-acting element in nitrogen fixation genes recognized by CnfR in the nonheterocystous nitrogen-fixing cyanobacterium *Leptolyngbya boryana*. *Mol Microbiol* 2016;**101**:411–24.

Tsujimoto R, Kotani H, Yokomizo K *et al*. Functional expression of an oxygen-labile nitrogenase in an oxygenic photosynthetic organism. *Sci Rep* 2018;**8**:7380.

Ungerer JL, Pratte BS, Thiel T. RNA processing of nitrogenase transcripts in the cyanobacterium *Anabaena variabilis*. *J Bacteriol* 2010;**192**:3311–20.

Valladares A, Herrero A, Pils D *et al*. Cytochrome c oxidase genes required for nitrogenase activity and diazotrophic growth in *Anabaena* sp. PCC 7120. *Mol Microbiol* 2003; **47**:1239–49.

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

Valladares A, Rodríguez V, Camargo S *et al*. Specific role of the cyanobacterial PipX Factor in the heterocysts of *Anabaena* sp. strain PCC 7120. *J Bacteriol* 2011;**193**:1172–82.

Vernon SA, Pratte BS, Thiel T. Role of the *nifB1* and *nifB2* promoters in cell-type-specific expression of two Mo nitrogenases in the cyanobacterium *Anabaena variabilis* ATCC 29413. *J Bacteriol* 2017. doi.org/10.1128/JB.00674-16.

Walsby AE. The permeability of heterocysts to the gases nitrogen and oxygen. *Proc R Soc Lond B* 1985;**226**:345–66.

Walsby AE. Cyanobacterial heterocysts: Terminal pores proposed as sites of gas exchange. *Trends Microbiol* 2007;**15**:340–9.

Wang X, Yang JG, Chen L *et al*. Using synthetic biology to distinguish and overcome regulatory and functional barriers related to nitrogen fixation. *PLoS One* 2013;**8**:e68677.

Yang J, Xie X, Xiang N *et al*. Polyprotein strategy for stoichiometric assembly of nitrogen fixation components for synthetic biology. *Proc Natl Acad Sci USA* 2018; **115**:E8509–17.

Zhang JY, Chen WL, Zhang CC. *hetR* and *patS*, two genes necessary for heterocyst pattern formation, are widespread in filamentous nonheterocyst-forming cyanobacteria. *Microbiology* 2009;**155**:1418–26.