The nitrogen regulator PipX acts *in cis* to prevent operon polarity.

Raquel Cantos[1], Jose I. Labella[1], Javier Espinosa and Asunción Contreras*

Departamento de Fisiología, Genética y Microbiología, Universidad de Alicante, Alicante, Spain.

[1]These authors contributed equally to this work.

[2]Lead contact.

*Correspondence: Asunción Contreras contrera@ua.es

Universidad de Alicante, Ctra. San Vicente del Raspeig s/n 03690 Spain. +34965909348

SUMMARY

Cyanobacteria, phototrophic organisms performing oxygenic photosynthesis, must adapt their metabolic processes to important environmental challenges, like those imposed by the succession of days and nights. Not surprisingly, certain regulatory proteins are found exclusively in this phylum. One of these unique factors, PipX, provides a mechanistic link between signals of carbon/nitrogen and of energy, transduced by the signaling protein PII, and the control of gene expression by the global nitrogen regulator NtcA. Here we report a new regulatory function of PipX: enhancement *in cis* of *pipY* expression, a gene encoding a universally conserved protein involved in amino/keto acid and Pyridoxal phosphate homeostasis. In *S. elongatus* and many other cyanobacteria these genes are expressed as a bicistronic *pipXY* operon. Despite being *cis*-acting, polarity suppression by PipX is nevertheless reminiscent of the function of NusG paralogues typified by RfaH, which are non-essential operon-specific bacterial factors acting in *trans* to upregulate horizontally-acquired genes. Furthermore, PipX and members of the NusG superfamily share a TLD/KOW structural domain, suggesting regulatory interactions of PipX with the translation machinery. Our results also suggest that the *cis*-acting function of PipX is a sophisticated regulatory strategy for maintaining appropriate PipX-PipY stoichiometry.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1758-2229.12688

INTRODUCTION

Cyanobacteria are phototrophic organisms that perform oxygenic photosynthesis and require the assimilation of ammonia for autotrophic growth. The nitrogen interaction network of *Synechococcus elongatus* PCC7942 (hereafter *S. elongatus*), the paradigm for nitrogen regulation in cyanobacteria, responds to signals of energy and Carbon/Nitrogen (C/N) balance (Huergo et al., 2013). The hubs of the network are the PII signal transduction protein, one of the most conserved signaling proteins (Chellamuthu et al., 2013; Merrick, 2014; Forchhammer and Luddecke, 2016), and the small cyanobacterial factor PipX, identified by its ability to bind PII (Burillo et al., 2004; Espinosa et al., 2006). PII and PipX mediate protein-protein interactions with regulatory targets that include transcriptional regulators, enzymes and transporters involved in nitrogen and/or carbon assimilation. They both display distinct localization patterns during diurnal cycles, co-localizing at dark periods in an energy dependent manner (Espinosa et al., 2018).

PII perceives metabolic information by competitive binding of ATP or ADP and by synergistic binding of ATP and 2-oxoglutarate (2-OG). The PII trimer has three binding sites for ATP/ADP (in some species AMP) and 2-OG (Fokina et al., 2010; Palanca et al., 2014). PII binds to N-acetyl-L-glutamate kinase (NAGK), stimulating its activity and promoting nitrogen storage as arginine in cyanobacteria and plants (Burillo et al., 2004; Heinrich et al., 2004; Llacer et al., 2007), and to the biotin carboxyl carrier protein (BCCP) of acetyl-CoA carboxylase (ACCase), inhibiting its activity to control acetyl-CoA levels in organisms encoding PII (Feria Bourrellier et al., 2010; Gerhardt et al., 2015; Hauf et al., 2016). PII-dependent inhibition of nitrate transport is known to occur after addition of ammonium to nitrate-containing cultures (Lee et al., 1998), a function that, although poorly characterized at the molecular level, depends on the NrtC subunit of the nitrate transporter (NRT) (Kobayashi et al., 1997). When abundant, 2-OG binds to MgATP-complexed PII, triggering conformational changes that prevent the interaction of PII with either NAGK or PipX (Espinosa et al., 2006; Llacer et al., 2007; Zeth et al., 2014). In the absence of 2-OG, only the ATP/ADP ratio and concentration of ADP directs the competitive interaction of PII with these targets *in vitro* (Luddecke and Forchhammer, 2015). PipX increases the affinity of PII for ADP, and, conversely, the interaction between PII and PipX is highly sensitive to fluctuations in the ATP/ADP ratio (Zeth et al., 2014).

PipX provides a mechanistic link between PII signaling and gene expression depending on the global transcriptional regulator NtcA. PipX uses the same surface, which involve the tudor-like domain (TLD) (Llacer et al., 2010) to bind to either 2-OG-bound NtcA, stimulating DNA binding and transcriptional activity, or to 2-OG-free PII to form PII-PipX complexes (Tanigawa et al., 2002; Vazquez-Bermudez et al., 2002; Llacer et al., 2010; Zhao et al., 2010). PII-PipX complexes interact with the transcriptional regulator PlmA, presumably to decrease transcriptional activity (Labella et al., 2016). When considering the affinities between PipX and partners NtcA (Kd of 85 nM) (Forcada-Nadal et al., 2014) and PII (Kd of 7 µM in the absence of 2-OG and nucleotides) (Llacer et al., 2010) and the relative abundances of the individual proteins in *S. elongatus* (in molar terms of subunits PII is 14 times more abundant than PipX which is 4 times more abundant than NtcA) (Guerreiro et al., 2016; Labella et al., 2016), it appears that, under a wide range of conditions, most of the PipX protein would be forming part of proteins complexes with PII and/or NtcA.

PII and PipX are likely to play unknown roles in adaptation to environmental situations faced by photosynthesizing cyanobacteria. In particular, genetic analyses are consistent with a greater complexity of the protein-protein interaction network involving PII and PipX regulators and thus suggest that they may be integrating multiple signaling pathways. Particularly interesting is the phenomenon described as the toxicity of PipX in the absence of PII, which implicate both proteins in the regulation of yet unknown cyanobacterial processes (Espinosa et al., 2009; Espinosa et al., 2010; Laichoubi et al., 2011; Laichoubi et al., 2012; Chang et al., 2013; Espinosa et al., 2014).

Knowledge of the nitrogen regulatory interaction network of cyanobacteria has largely benefited from the “guilty by association” principle implicit in approaches relying on protein interactions (Burillo et al., 2004; Espinosa et al., 2006; Llacer et al., 2007; Llacer et al., 2010; Laichoubi et al., 2011; Labella et al., 2016) or gene neighborhood analysis (Labella et al., 2017). In *S. elongatus* and many other cyanobacteria, *pipX* is co-transcribed with the downstream gene *pipY*, encoding a canonical member of the widely distributed COG0325/PLPBP protein family, regulatory Pyridoxal phosphate (PLP)-binding proteins involved in amino/keto acid and PLP homeostasis (Labella et al., 2017; Tremiño et al., 2017).

Although the function of COG0325/PLPBP proteins remains largely enigmatic, their presence affects the levels of amino/keto acids (Ito et al., 2013; Prunetti et al., 2016), signals which are relevant in the context of nitrogen regulation and PipX mediated interactions. Previous studies (Labella et al., 2017), aimed towards characterization of PipY and its connection with PipX in the nitrogen interaction network, showed that both proteins make individual contributions to resistance against PLP-targeting antibiotics DCS or BCDA and to expression of a common set of transcripts in *S. elongatus*, thus placing PipX and PipY within the same genetic pathway.

A most intriguing finding of our previous studies was that inactivation of *pipX* resulted in dramatic reduction of *pipY* gene expression (Labella et al., 2017) despite the fact that the mutant construct was specifically designed to prevent polarity, that is, loss of function or expression of the gene located downstream the inactivated one. A null *pipX* mutant was constructed by precisely substituting its coding sequence with that of the *cat* gene, thus maintaining the intergenic region and the ribosome binding site of *pipY* intact. Although the *cat* reporter gene was appropriately expressed in this *pipX* mutant construct, PipY levels decreased significantly (Labella et al., 2017), indicating polarity. Since attempts to obtain evidence for protein-protein interactions between PipX and PipY that would indicate stabilization of PipY by complex formation with PipX were unsuccessful, the role played by PipX in preventing polarity of the *pipXY* operon was unclear.

The aim of this work was to get insights into the complex regulatory connections between the *pipX* and *pipY* genes in *S. elongatus*, using functional and *in silico* approaches. Here we show that prevention of polarity at the *pipXY* operon is independent of *pipY* or its gene products, requires both coding and non-coding *pipX* sequences, and importantly, coding sequences are also required *in cis*. While the later finding reveals an unexpected and novel mechanism for prevention of polarity, the TLD/KOW (Kyrpides et al., 1996) domain of PipX establishes a structural and functional link between this unique cyanobacterial protein and NusG paralogues, in particular with $NusG^{SP}$ proteins, a group of non-essential operon-specific bacterial factors typified by *E. coli* RfaH (Belogurov et al., 2009). We also discuss *in silico* evidence of the tight regulatory connections between *pipX* and *pipY*, two genes belonging to the conserved cyanobacterial genome core.

RESULTS AND DISCUSSION

The *pipX* gene enhances expression of downstream loci

In a previous study we showed that inactivation of the *pipX* gene had a dramatic impact on the protein levels of PipY, dropping to 5.7±1.4 % of the wild-type levels, as determined by quantitative western-blots (Labella et al., 2017). However, the decrease in *pipY* transcripts observed by RT-PCR appeared smaller, suggesting that the effects on *pipY* expression were mainly exerted at the protein level rather than at the transcript level. To independently confirm these results in a quantitative manner, we now performed qRT-PCR in *pipX* and wild-type *S. elongatus* strains. Details of the corresponding strains are shown in Fig. 1. The results indicated that *pipX* inactivation reduced the transcript levels to a 30±0.4 % (Fig. 2A), in close agreement with the qualitative RT-PCR assays.

To get further insights into the basis and molecular determinants of this phenomenon, we next investigated whether the positive regulatory effect of *pipX* over *pipY* gene expression depended on interactions between the two gene products. If that was the case, replacing the *pipY* gene by a reporter gene would prevent the inferred activation function of PipX.

S. elongatus strain CK1XY is a virtually wild-type strain with a CK1 cassette insertion 72 nucleotides upstream of *pipX* coding sequences (Fig. 1) that results in 3-4 fold higher levels of both PipX and PipY proteins (Espinosa et al., 2010; Labella et al., 2017). Increased protein production from the *pipXY* operon is driven by the presence of a strong promoter on CK1 (Vazquez-Bermudez et al., 2000). Given the relatively inefficient immunodetection of PipY from cell extracts of wild-type *S. elongatus*, the use of strain CK1XY improves detection of PipY in routine western assays. Thus, the CK1XY strain was used to obtain precise substitutions of *pipX* (CK1*cat*Y) or *pipY* (CK1X*cat*) genes with the *cat* coding sequence (Fig. S1), following the same procedures used for the equivalent wild-type constructions and strains (Labella et al., 2017). As expected, inactivation of *pipX* in strain CK1XY resulted in failure to detect PipX in standard western blots and, importantly, in a significant decrease in the PipY levels (compare CK1XYand CK1*cat*Y in Fig. 2A, right panel), thus confirming previous results with the wild-type and *pipX* mutant strains (Labella et al., 2017).

To determine whether the stimulatory role of *pipX* could also be exerted over an unrelated gene occupying the *pipY* locus in *S. elongatus*, we compared the levels of chloramphenicol resistance conferred by the *cat* coding sequences when it replaced either the *pipX* or *pipY* coding sequences, using wild-type, CK1XY, and derivative strains in parallel. As shown in Fig. 2A (central panel), the *cat* gene conferred chloramphenicol resistance to both the wild-type and CK1XY strains in each of the two positions, but resistance was significantly higher when the *cat* gene replaced the downstream *pipY* locus (CK1X*cat*) compared with its insertion in the upstream *pipX* locus (CK1*cat*Y).

Therefore, our results with the *cat* reporter demonstrate that the stimulatory effect of *pipX* over *pipY* gene expression is independent of *pipY* or its gene product and that the *pipX* gene functions as a suppressor of operon polarity. The finding that the *pipX* gene is not specific to PipY but can affect a foreign ORF being translated at this location also confirms that the contribution of PipX to PipY levels does not involve physical interaction between the two proteins.

***pipX* acts in *cis* to increase PipY levels**

To gain insights into the role of PipX in polarity suppression, we tested the ability of ectopically expressed PipX to rescue operon polarity in a *pipX* null mutant. To this end we introduced a P*trc*::*pipX* translational fusion and the *lacI*^{q} gene at the *S. elongatus* Neutral Site I (NSI) (Bustos and Golden, 1992) (Fig. 1) to ensure enough intracellular accumulation of PipX in strain CK1*cat*Y. We subsequently analyzed PipX and PipY levels in the resulting strain (CK1*cat*Y/1^{S}Ptrc-PipX, Fig. 2A).

The levels of PipX detected in strain CK1*cat*Y/1^{S}Ptrc-PipX without the inducer isopropyl β-D-1-thiogalactopyranoside (IPTG) were similar to those detected in strain CK1XY (Fig. 2A), that is, 3-4 fold higher than the levels found in the wild-type strain (Espinosa et al., 2010; Labella et al., 2017). On the other hand, addition of IPTG resulted in toxicity to cells (Fig. S2), a result reflecting the toxicity of PipX when the PipX/PII ratio is increased above a certain threshold (Espinosa et al., 2009; Laichoubi et al., 2012; Espinosa et al., 2018).

Remarkably, PipY levels remained undetectable in CK1*cat*Y/1^{S}Ptrc-PipX, indicating that ectopic expression of PipX did not complement the phenotypic defect of the *pipX* null mutant. This result implies that the polarity suppression effect is exerted by *pipX in cis*, a phenomenon that can be explained by two

alternative possibilities. Either, only DNA or RNA sequences at or near the *pipXY* operon were involved, or the PipX protein acts *in cis*, that is, locally. While in the former case we have no candidates for the cellular factors that would specifically recognize *cis*-acting sequences within the *pipXY* operon to suppress polarity, in the later one there is a straightforward explanation for the local action of PipX: sequestration into PII and/or NtcA complexes.

The stability of known complexes involving PipX and the relative abundance of the proteins implicated (Llacer et al., 2010; Forcada-Nadal et al., 2014; Guerreiro et al., 2016; Labella et al., 2016) provide support for the idea that once PII-PipX and NtcA-PipX complexes are formed, PipX would be trapped. Fluctuation of C/N and energy signals would drive partner swapping, changing the relative proportion of PII-PipX and NtcA-PipX (Espinosa et al., 2006; Espinosa et al., 2007), but still preventing PipX from making interactions with additional partners. Additional support for this idea comes from the finding that PII, which appears to be 14 times more abundant than PipX (Guerreiro et al., 2016; Labella et al., 2016) prevents PipX toxicity in *S. elongatus* (Espinosa et al., 2009; Espinosa et al., 2010; Chang et al., 2013), a phenomenon further suggesting that PII binding is counteracting the interaction of PipX with yet unidentified low affinity partners.

Local overproduction of PipX does not suffice to increase PipY levels

The get further insight into the mechanism by which *pipX* increases *pipY* expression *in cis*, we investigated whether, when maintaining the *pipXY* gene organization as in wild-type *S. elongatus*, there was a simple correlation between the expression levels of PipX and PipY proteins. To this end we generated strain 1^{S}Ptrc-PipXY (Fig. 1), containing P*trc*::*pipXY* preceded by the *lacI*^{q} gene at the Neutral Site I (NSI). In a previous work we had obtained the related strains 1^{S}Ptrc-PipX and 1^{S}Ptrc-PipY, used for independent overexpression of PipX and PipY in *S. elongatus* (Labella et al., 2017). Due to strong translation signals, the Ptrc derivatives P*trc*::*pipX* or P*trc*::*pipY* ensure enough intracellular accumulation of the cloned genes even in the absence of the IPTG inducer ((Labella et al., 2017) and Fig. 2B), and we wondered whether in strain 1^{S}Ptrc-PipXY, which shares with strain CK1XY a total of 1215 bp including the complete *pipX* and *pipY* ORFs and 22 bp from the ORF of the dowstream gene *sepF* (Fig. 1), there was still coordinated expression,

that is, overproduction of both PipX and PipY. It should be noted that in all these Ptrc derivative strains the *pipXY* locus remained intact, thus contributing to the total levels of intracellular PipX or PipY proteins.

As shown in Fig. 2B, 1^{S}Ptrc-PipXY provided strong expression of PipX but did not increase PipY detection above wild-type levels. Furthermore, the addition of IPTG increased the levels of PipX or PipY proteins from 1^{S}Ptrc-PipX or 1^{S}Ptrc-PipY, respectively, and only of PipX from 1^{S}Ptrc-PipXY, while the levels of PipY remained too low to be detected, as in wild-type (Fig. 2C). Therefore, the results indicate that the presence of *pipX* coding sequences upstream *pipY* is not enough for the coordinated expression of the operon genes and thus additional *cis*-acting sequences, probably located within the 72 nucleotides long 5´UTR which is present in CK1XY but not in 1^{S}Ptrc-PipXY, must play a regulatory role. Whether this region is directly targeted by cell factors and/or contributes to the mRNA secondary structure remains to be investigated.

Functional differences between PipX and other operon-specific bacterial factors enhancing expression of downstream genes

To get additional insight into the involvement of PipX protein in polarity suppression at the *pipXY* operon, we searched the literature for operon-specific bacterial factors that function to increase expression of downstream genes. Reported examples belong to a group of non-essential operon-specific proteins, all specialized paralogues of NusG ($NusG^{SP}$). They include the proteins RfaH from *E. coli* (Belogurov et al., 2010, Burmann et al., 2012), LoaP from *Bacillus amyloliquefaciens* (Goodson et al., 2017), TaA from *Micrococcus xantus* and the UpxY family of factors of Bacteroides (Chatzidaki-Livanis et al., 2009). Importantly, they all act *in trans* as processive transcriptional antiterminators, indicating the novelty of the *cis*-acting behavior of PipX.

To prevent interference with the essential NusG factor, $NusG^{SP}$ action is restricted to their specific targets and recruitment to targets is achieved by binding to operon polarity suppressor (*ops*) sequences (Belogurov et al., 2009; Santangelo and Artsimovitch, 2011; NandyMazumdar and Artsimovitch, 2015; Zuber et al., 2018). However, we showed here that the up regulation of the (downstream) *pipY* gene by PipX only take place *in cis*, suggesting that the polarity suppression function of PipX is probably restricted to just

its own operon (see also discussion below). Furthermore, the synthesis of PipX at, or very close to, its site of action would eliminate the need for high affinity recognition sequences.

Activation of target genes by the multifunctional NusG proteins involves recruiting RNA polymerase (RNAP) to target operons as well as interference with transcription termination. However, our results indicating that inactivation of *pipX* has a far more pronounced effect on PipY protein levels than on *pipY* transcript levels suggest that suppression of operon polarity by PipX is mainly exerted at the level of translation. In addition, the apparent lack of intrinsic terminators in the *pipXY* intergenic region (Vijayan et al., 2011) does not necessarily exclude the possibility of transcript termination at this site. In this context, the existence of uncharacterized mechanisms of transcription termination in cyanobacteria has been previously suggested (Washio et al., 1998).

Although PipX, RfaH and other $NusG^{SP}$ proteins are all operon-specific bacterial factors that function to increase expression of downstream genes, the later proteins specifically target horizontally transferred genes. These are recognized in the genomes as xenologs based on G+C content, phylogenetic similarity, codon usage and small number of the highly iterated palindrome 1 motif (Delaye et al., 2011). In contrast, both *pipX* and *pipY* form part of the conserved/cyanobacterial core gene set (Delaye et al., 2011; Shih et al., 2015), which is at odds with the idea of constitutive activation of *pipY* by PipX. Instead, the available data is best reconciled with PipX specifically regulating *pipY* expression in response to particular intracellular signals.

PipX and NusG family proteins share a TLD/KOW domain involved in operon polarity

Interestingly, the N-terminal domain of PipX when in complex with NtcA or PII (Llacer et al., 2010) shows the same fold, a TLD/KOW motif found for the CTD (C-terminal domain) of the NusG family which in both NusG and RfaH is involved in interaction with the ribosome (Burmann et al., 2010). The presence of TLD/KOW domains in PipX and factors regulating operon polarity suggested that they may also be involved in common functions. Since determinants of the interactions between NusG and the NusE/S10 ribosomal protein (Burmann et al., 2010) are known for *E. coli*, we looked for relevant sequence similarities between NusG, RfaH and PipX proteins within TLD/KOW domains.

Important residues for the NusG-NusE interaction, including hydrophobic residues of NusG making close contacts with hydrophobic residues in NusE/S10, are indicated in the weblog shown in Fig. 3. The binding region is provided by the loops between β strands β1 and β2, β3 and β4 and residues from β4. While comparison of weblogs for NusG, RfaH and PipX did not reveal conserved motifs in common, key residues for the interaction between NusG and NusE appeared conserved in PipX and RfaH when the position of the beta sheets of each protein was taken into account, thus giving credit to a hypothetical interaction between PipX and NusE/S10.

To get additional insights into the parallelisms between PipX and NusG paralogs, we next investigated the relatedness between the structures available for TLD/KOW domains of PipX and proteins from the NusG family. To this end, the structures from the TLD/KOW domains of PipX (residues 3-41) monomers in complex with PII (PDB:2XG8) and NtcA (PDB:2XKO) were compared with the NusG structures available in the PDB database (1NZ9,2JVV 2KVQ, 2LQ8, 2MI6, 2XHC, 5TBZ) and RfaH (2LCL). Visual analysis of the results confirmed the similarity between the domain folds of PipX and those of NusG or RfaH, conserving the same beta-sheets spatial disposition, with the main discrepancies being in the loops between beta sheets (Fig. 4A). Sequence independent algorithms for structural alignment were applied on the pairwise comparisons (Fig. 4B). The lowest RMSD values were found between PipX and NusG from *E. coli* and PipX and RfaH (2.07 and 2.58 angstroms average value, respectively), thus indicating a significant structural similarity between their TLD/KOW domains.

Possible involvement of PipX in translation regulation

Transcription-translation coupling in bacteria, used to monitor mRNA quality and to determine if cellular conditions are favorable for translation (Artsimovitch, 2018), is helped by regulatory factors that can interact with both RNAP and ribosomes, affecting gene expression at more than one level. In this context, NusG and NusG^{SP} proteins, normally referred to as anti-terminators, are also regulators of translation and can have complex and diverse effects depending on the particular context (NandyMazumdar and Artsimovitch, 2015). The RfaH CTD binds to the NusE/S10 ribosomal protein and activates translation initiation *in vivo* (Burmann et al., 2012), presumably by recruiting the 30S subunit through direct protein-protein contacts. In the light of the similar structures and features shared by the TLD/KOW domains of

NusG, RfaH and PipX, and the conservation of NusE/S10 proteins ((Gladki et al., 2013), Fig. S3), it is tempting to propose that the TLD/KOW domain of PipX may be involved in similar contacts with the translation machinery.

In NusG and RfaH the same TLD/KOW surface binding to the ribosome interacts in a competitive way with Rho. Although cyanobacteria do not encode Rho factors (D'Heygere et al., 2013), it is tempting to propose that PipX also interacts with the ribosome and that this interaction is regulated by competition with additional cell components. In this context, the nitrogen regulators PII and NtcA, both of them recognizing the same surface of the PipX TLD/KOW, are prime candidates. Such a scenario would ensure that the translation regulation function of PipX was restricted both temporarily and spatially (when and where PipX is being translated) and thus would target only the gene downstream of *pipX*. Thus, translation stimulation of PipY synthesis by PipX would occur during (co-translationally) or immediately after PipX translation, but not after PipX forms complexes with PII or NtcA.

Competition between the ribosome and nitrogen regulators for binding to the same surface of PipX is consistent with the failure of PipX to suppress polarity when expressed from a distant genomic location and supports the notion that only recently translated PipX is "free" to interact locally to regulate *pipY* expression. Finally, although we think it unlikely and we are not aware of intracellular conditions in which sufficient PipX would be liberated from the corresponding complexes with each of the NtcA and PII proteins, we cannot entirely exclude the possibility that under certain environmental conditions PipX may interact with ribosomes at operons other than *pipXY* and thus exert a more global impact on gene expression.

Ribosome pausing is subjected to regulation by accessory factors and environmental conditions (Ansimovitch, 2018) and, in this context, we can not exclude the possibility that, in addition to the proposed function of PipX as a translation regulator, the actual translation of PipX may also play a role in *pipY* expression, thus behaving as the leader peptide exerting attenuation of operons involved in amino acid biosynthesis. The implication of PipY in amino/keto acid homeostasis makes this possibility worth considering, and the challenge would therefore be to identify the specific signals and mechanisms involved.

The unique C-terminal domain of PipX, with no counterpart in other bacterial regulators, may hold a key to understand PipX function(s) in translation regulation. It consists of two alpha-helices, the first of

which contains an R-rich basic patch (Fig. 4C and 4D). Leaving aside the fact that it provides some contacting residues in the NtcA-PipX (Llacer et al., 2010) or PII-PipX-PlmA complexes (Labella et al., 2016), the function of this domain has remained elusive. However, R-rich alpha-helices have been found to provide interaction determinants in non-canonical RNA binding proteins (Jarvelin et al., 2016), and it is tempting to propose that in PipX this motif may be involved in specific interactions with rRNA and/or *pipXY* mRNA, two functions typical of translation regulators.

The tight link between the *pipX* and *pipY* genes in cyanobacteria

The relatively short or non-existent intergenic distances between contiguous *pipX* and *pipY* coding sequences that are found in many cyanobacteria are very remarkable, suggesting co-expression and translational coupling. Interestingly, we have demonstrated PipX-mediated translation stimulation of PipY in *S. elongatus*, where the *pipX-pipY* intergenic distance is relatively large in comparison (24 Nucleotides), and thus translational coupling could not be anticipated for this particular cyanobacterium, a result suggesting an important functional interaction between PipX and PipY in most, if not all, cyanobacteria. However, we also showed that the stimulatory effect of PipX could be exerted over either *pipY* or a reporter gene occupying the *pipY* locus of *S. elongatus* (Fig. 2A) and we wondered whether *pipX* genes may also contribute to the expression of genes other than *pipY* when these are located downstream. To explore this possibility on the bases of genomic information, we compared the length of intergenic distances between *pipX* and either *pipY* or the corresponding non-*pipY* genes placed downstream to and in the same orientation as *pipX*. The corresponding information is provided in supplementary Table 1 and represented in Fig. 5 (solid line).

It is worth noting than when *pipX* and *pipY* genes were linked they were always in the same order and orientation, with *pipX* always upstream (99 out of 99 examples). In contrast, when genes other than *pipY* were downstream *pipX*, in a significant proportion of the cases (9 out of 27), the two genes were found in opposite orientations. Amongst the exceptions in which *pipX* and *pipY* genes are not linked was the model cyanobacterium *Synechocystis* PCC 6803, a heterotrophic organism. In addition, the intergenic distances between *pipX* and contiguous non-*pipY* genes with the same orientation were significantly larger than the ones between *pipX* and *pipY* genes (compare solid and dashed lines in Fig. 5). Taken together, these results

strongly suggest that the most frequent organization of *pipX* in cyanobacteria is as part of bicistronic *pipX*Y operons and that when these genes are not linked, *pipX* tends to be by itself, in monocistronic operons.

To explore the possibility that other regulatory genes may substitute for *pipX* at the operons encoding PipY, we identified the non-*pipX* genes upstream *pipY* in the available cyanobacterial genomes (supplementary Table S2). Since up to 6 very diverse proteins were encoded in the cases in which the upstream gene showed the same orientation as *pipY* (10 out of 24), it follows that the only clear association of *pipY* genes with regulatory genes in cyanobacteria is provided by *pipXY* operons.

Taken together, our results suggest that i) upregulation of PipY levels by PipX is a sophisticated regulatory strategy for maintaining appropriate PipX-PipY stoichiometry and/or controlling PipY protein levels in response to PipX translation, ii) the *cis* acting function of PipX appears to specifically target the *pipY* gene, and iii) the organization of the *pipX* and *pipY* genes as a bicistronic operon are ancestral cyanobacterial features that have been lost in a few cyanobacterial lines. Further work is required to understand the environmental importance of the PipX-PipY balance in cyanobacteria.

EXPERIMENTAL PROCEDURES

Strains and plasmids used in this work are listed in Table 1. Oligonucleotides used to construct plasmids, sequencing and strain verifications are listed in Supplementary Table 3.

Molecular genetic techniques and growth conditions

Cloning procedures were carried out with *Escherichia coli* DH5α, using standard techniques and automated dideoxy DNA sequencing to verify plasmid sequence. *S. elongatus* strains were routinely grown photoautotrophically at 30°C while shaking under constant illumination (40 μmol photons $m^{-2}s^{-1}$) provided by cool white fluorescent lights. Media used were blue–green algae media BG11 ($BG11_0$ plus 17.5 mM $NaNO_3$ and 10mM HEPES/NaOH pH 7.8). For growth on plates the media was solidified by addition of 1.5% (w/v) agar. Plates were routinely incubated at 30°C under constant illumination. *S. elongatus* strains were transformed essentially as described by (Golden and Sherman, 1984). Whenever used, antibiotic concentrations were (concentration in μg ml^{-1}): kanamycin/5, streptomycin/2, and chloramphenicol/5.

To test growth on solid media, exponentially growing cultures were adjusted to 0.5 (DO_{750}nm) before dropping 5 µL of the cell suspensions with the indicated dilutions (1/10 and/or 1/100) onto BG11 or BG11 plus chloramphenicol plates. Plates were photographed after 3 or 5 incubation days.

Construction of plasmids and strains

To overexpress *pipXpipY* from the IPTG-inducible Ptrc promoter, plasmid pUAGC292, carrying P*trc::pipXpipY* transcriptional fusion and *lacI* flanked by the Neutral Site I (NSI) regions of *S. elongatus*, was engineered. The *pipXpipY* sequences were PCR amplified from genomic DNA using primers PipX-OV-2F and PiXXINACR. The amplified product was restricted with *Eco*RI/*Sal*I and cloned into digested pUAGC280, yielding pUAGC292.

To overexpress PipX and PipY in *S. elongatus* from its chromosome region, the wild type strain was transformed with plasmids pUAGC410 and clones were selected on kanamycin plates. The strain CK1XY was transformed with plasmids pUAGC126 or pUAGC127 to inactivate *pipX* or *pipY* producing strains CK1*cat*Y and CK1X*cat* respectively. In this case, clones were selected on plates supplemented with kanamycin and chloramphenicol, and complete segregation was confirmed by PCR analysis with oligonucleotides pairs PipX-BTH-F/PipX5R for *pipX* and inter2060-1F/PipXXinact-R and inter2060-1F/PipX5R for *pipY* inactivations. The CK1*cat*Y and wild-type strains were transformed with pUAGC873 to introduce Ptrc-PipX into the NSI and wild-type was additionally transformed with pUAGC294 or pUAGC292 to overexpress *pipY* or *pipXpipY*, respectively, from Ptrc promoter. Transformants derived from CK1catY strain were selected on plates supplemented with kanamycin, chloramphenicol and streptomycin while wild-type derivatives were selected on streptomycin containing plates. The correct insertion of all constructs into the Neutral Site was confirmed by PCR with primer pairs PTRC99Aseq-F/NSI-1F and 7942NSIA-F/NSI-1R.

Immunodetecctions assays

15 ml samples of *S. elongatus* cells were harvested by centrifugation at 7300 g for 5 min. The pellet was resuspended in 100 µl lysis buffer (50 mM Tris/HCl, pH 7.4, 4 mM EDTA, 0.5 mM PMSF, 0.5 mM benzamidine, 1 mM DTT) and 100 nm glass beads were added. The mixture was homogenized with two cycles of 60 s in a Minibeadbeater. After centrifugation (5500 g for 5 min), the supernatant fraction (crude

protein extract) was transferred to a new tube. Protein concentrations were estimated by Lowry (Bio-Rad RC DC reagents). Protein immunodetection was performed on protein-loaded PVDF membrane, transferred from acrylamide gels using a wet system. Membrane was blocked with TBS (20 mM Tris, pH 7.5, 500 mM NaCl) solution containing 5 % BSA for 1 h at room temperature, and incubated overnight in TBS solution containing 2 % BSA and the primary antibody (a 1:300 dilution for anti-PipY and 1:5000 for anti-PipX). The membrane was then incubated with ECL rabbit IgG, HRP-linked F(ab')2 fragment (from donkey) (GE Healthcare) (1:150000 dilution). Immunoreactive bands were detected using the Supersignal West Femto (Thermo) and X-ray films. Antisera against PipX (Pineda Antikörper Service, Berlin, Germany) and PipY (Genosphere Biotechnologies, Paris) were produced in rabbits.

RNA analysis

RNA was purified [phenol-chloroform extraction and TURBO DNase (Ambion) digestion] from 50-ml light incubated nitrate-containing cultures (OD750~0.7) of WT and *pipX* (2 biological samples each). RNA was quantified spectrophotometrically. For RT-PCR experiments purified RNA was first PCR verified for no DNA contamination. Then, 90 ng of total RNA was retrotranscribed in a total volume of 30 µl with RevertAid H Minus M-MuL V Reverse Transcriptase (Fermentas) using primers rnpB-R (*rnpB)* or PipX-5R-X (*pipY*). qRT-PCR was performed with 1.5 µl of the corresponding cDNA, the Applied Biosystems StepOne Plus system (Life Technologies) and Maxima SybrGreen/Rox qPCR Master Mix (ThermoScientific). Abundance was calculated via the ΔΔCT method after correction of amplification efficiencies determined by linear regression using the LinReg program and normalized to *rnpB* as endogenous control.

Computational methods

Distances involving *pipX* or *pipY* genes were calculated from 124 cyanobacteria genomes (Supplementary Table 1) as the number of nucleotides in the intergenic region between *pipX* and the downstream gene or *pipY* and the upstream gene (negative numbers indicate overlap). The *pipX* or *pipY* position in each cyanobacterial genome was determined using blastp with the correspondent protein sequence as query. PipX sequences were retrieved from 124 cyanobacterial complete genomes from NCBI data base, NusG and RfaH protein sequences were retrieved from KEGG database using the Kegg Orthology

identifiers K02601 and K05785 respectively. MAFFT online tool was using for protein alignments (Katoh et al., 2017). Weblogos were generated using the online tool WebLogo (Crooks et al., 2004). Structures of NusG from *E. coli* (2JVV state 19, chain A, residues 123-181; 2KVQ chain G, residues 123-181; 5TBZ chain J, residues 126-181, and chain K, residues 126-181), *Thermus termophilus* (1NZ9 state 9, chain A, residues 127-184), *Thermotoga maritima* (2LQ8 chain A, residues 117-177; 2XHC chain A, residues 285-352) and *Mycobacterium tuberculosis* (2MI6 state 9, chain A, residues 1-62; RfaH from *E. coli* (2LCL state 20, chain A, residues 97-162); and PipX from *S. elongatus* (2XG8 chain D, residues 3-41, chain E, residues 3-41, and chain F, residues 3-41; 2XKO chain C, residues 3-41, and chain D, residues 3-41) were downloaded from PDB database. Alpha carbons were aligned and RMSD calculated using Pymol "*super*" command (Schrodinger, 2015).

Acknowledgements

We thank C.V. Racovac and A. Llop for technical support, K. Forchhammer for PipX antiserum, and R. Dixon and V. Rubio for constructive discussions. The authors thank the Spanish Government (MINECO) for supporting grants BFU2012-33364 and BFU2015-66360-P to A.C. and the University of Alicante for PhD fellowship (FPUUA59) to J.I.L.

References

Artsimovitch, I. (2018) Rebuilding the bridge between transcription and translation. *Mol Microbiol* **108**: 467-472.
Belogurov, G.A., Sevostyanova, A., Svetlov, V., and Artsimovitch, I. (2010) Functional regions of the N-terminal domain of the antiterminator RfaH. *Mol Microbiol* **76**: 286-301.
Belogurov, G.A., Mooney, R.A., Svetlov, V., Landick, R., and Artsimovitch, I. (2009) Functional specialization of transcription elongation factors. *EMBO J* **28**: 112-122.
Burillo, S., Luque, I., Fuentes, I., and Contreras, A. (2004) Interactions between the nitrogen signal transduction protein PII and N-acetyl glutamate kinase in organisms that perform oxygenic photosynthesis. *J Bacteriol* **186**: 3346-3354.
Burmann, B.M., Schweimer, K., Luo, X., Wahl, M.C., Stitt, B.L., Gottesman, M.E., and Rosch, P. (2010) A NusE:NusG complex links transcription and translation. *Science* **328**: 501-504.
Burmann, B.M., Knauer, S.H., Sevostyanova, A., Schweimer, K., Mooney, R.A., Landick, R. et al. (2012) An alpha helix to beta barrel domain switch transforms the transcription factor RfaH into a translation factor. *Cell* **150**: 291-303.
Bustos, S.A., and Golden, S.S. (1992) Light-regulated expression of the *psbD* gene family in *Synechococcus* sp. strain PCC 7942: evidence for the role of duplicated *psbD* genes in cyanobacteria. *Mol Gen Genet* **232**: 221-230.
Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res* **14**: 1188-1190.
Chang, Y., Takatani, N., Aichi, M., Maeda, S., and Omata, T. (2013) Evaluation of the effects of P_{II} deficiency and the toxicity of PipX on growth characteristics of the P_{II}-less mutant of the cyanobacterium *Synechococcus elongatus*. *Plant Cell Physiol* **54**: 1504-1514.
Chatzidaki-Livanis, M., Coyne, M.J., and Comstock, L.E. (2009) A family of transcriptional antitermination factors necessary for synthesis of the capsular polysaccharides of Bacteroides fragilis. *J Bacteriol* **191**: 7288-7295.

Chellamuthu, V.R., Alva, V., and Forchhammer, K. (2013) From cyanobacteria to plants: conservation of PII functions during plastid evolution. *Planta* **237**: 451-462.
D'Heygere, F., Rabhi, M., and Boudvillain, M. (2013) Phyletic distribution and conservation of the bacterial transcription termination factor Rho. *Microbiology* **159**: 1423-1436.
Delaye, L., Gonzalez-Domenech, C.M., Garcillan-Barcia, M.P., Pereto, J., de la Cruz, F., and Moya, A. (2011) Blueprint for a minimal photoautotrophic cell: conserved and variable genes in *Synechococcus elongatus* PCC 7942. *BMC Genomics* **12**: 25.
Espinosa, J., Forchhammer, K., and Contreras, A. (2007) Role of the *Synechococcus* PCC 7942 nitrogen regulator protein PipX in NtcA-controlled processes. *Microbiology* **153**: 711-718.
Espinosa, J., Forchhammer, K., Burillo, S., and Contreras, A. (2006) Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA. *Mol Microbiol* **61**: 457-469.
Espinosa, J., Castells, M.A., Laichoubi, K.B., and Contreras, A. (2009) Mutations at *pipX* suppress lethality of PII-deficient mutants of *Synechococcus elongatus* PCC 7942. *J Bacteriol* **191**: 4863-4869.
Espinosa, J., Labella, J.I., Cantos, R., and Contreras, A. (2018) Energy drives the dynamic localization of cyanobacterial nitrogen regulators during diurnal cycles. *Environ Microbiol* **20**: 1240-1252.
Espinosa, J., Castells, M.A., Laichoubi, K.B., Forchhammer, K., and Contreras, A. (2010) Effects of spontaneous mutations in PipX functions and regulatory complexes on the cyanobacterium *Synechococcus elongatus* strain PCC 7942. *Microbiology* **156**: 1517-1526.
Espinosa, J., Rodriguez-Mateos, F., Salinas, P., Lanza, V.F., Dixon, R., de la Cruz, F., and Contreras, A. (2014) PipX, the coactivator of NtcA, is a global regulator in cyanobacteria. *Proc Natl Acad Sci U S A* **111**: E2423-E2430.
Feria Bourrellier, A.B., Valot, B., Guillot, A., Ambard-Bretteville, F., Vidal, J., and Hodges, M. (2010) Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate-regulated by interaction of PII with the biotin carboxyl carrier subunit. *Proc Natl Acad Sci U S A* **107**: 502-507.
Fokina, O., Chellamuthu, V.R., Forchhammer, K., and Zeth, K. (2010) Mechanism of 2-oxoglutarate signaling by the *Synechococcus elongatus* PII signal transduction protein. *Proc Natl Acad Sci U S A* **107**: 19760-19765.
Forcada-Nadal, A., Forchhammer, K., and Rubio, V. (2014) SPR analysis of promoter binding of *Synechocystis* PCC6803 transcription factors NtcA and CRP suggests cross-talk and sheds light on regulation by effector molecules. *FEBS Lett* **588**: 2270-2276.
Forchhammer, K., and Luddecke, J. (2016) Sensory properties of the PII signalling protein family. *FEBS J* **283**: 425-437.
Gerhardt, E.C., Rodrigues, T.E., Muller-Santos, M., Pedrosa, F.O., Souza, E.M., Forchhammer, K., and Huergo, L.F. (2015) The bacterial signal transduction protein GlnB regulates the committed step in fatty acid biosynthesis by acting as a dissociable regulatory subunit of acetyl-CoA carboxylase. *Molecular microbiology* **95**: 1025-1035.
Gladki, A., Kaczanowski, S., Szczesny, P., and Zielenkiewicz, P. (2013) The evolutionary rate of antibacterial drug targets. *BMC Bioinformatics* **14**: 36.
Golden, S.S., and Sherman, L.A. (1984) Optimal conditions for genetic transformation of the cyanobacterium *Anacystis nidulans* R2. *J Bacteriol* **158**: 36-42.
Goodson, J.R., Klupt, S., Zhang, C., Straight, P., and Winkler, W.C. (2017) LoaP is a broadly conserved antiterminator protein that regulates antibiotic gene clusters in Bacillus amyloliquefaciens. *Nat Microbiol* **2**: 17003.
Guerreiro, A.C., Penning, R., Raaijmakers, L.M., Axman, I.M., Heck, A.J., and Altelaar, A.F. (2016) Monitoring light/dark association dynamics of multi-protein complexes in cyanobacteria using size exclusion chromatography-based proteomics. *Journal of proteomics* **142**: 33-44.
Hanahan, D. (1985) Techniques for transformation of *Escherichia coli*. In *DNA cloning*. Glover, D. (ed). Oxford: IRL Press Ltd., pp. 109-135.
Hauf, W., Schmid, K., Gerhardt, E.C., Huergo, L.F., and Forchhammer, K. (2016) Interaction of the Nitrogen Regulatory Protein GlnB (PII) with Biotin Carboxyl Carrier Protein (BCCP) Controls Acetyl-CoA Levels in the Cyanobacterium *Synechocystis* sp. PCC 6803. *Front Microbiol* **7**: 1700.
Heinrich, A., Maheswaran, M., Ruppert, U., and Forchhammer, K. (2004) The *Synechococcus elongatus* PII signal transduction protein controls arginine synthesis by complex formation with N-acetyl-L-glutamate kinase. *Mol Microbiol* **52**: 1303-1314.
Huergo, L.F., Chandra, G., and Merrick, M. (2013) PII signal transduction proteins: nitrogen regulation and beyond. *FEMS microbiology reviews* **37**: 251-283.
Ito, T., Iimori, J., Takayama, S., Moriyama, A., Yamauchi, A., Hemmi, H., and Yoshimura, T. (2013) Conserved pyridoxal protein that regulates Ile and Val metabolism. *J Bacteriol* **195**: 5439-5449.

Jarvelin, A.I., Noerenberg, M., Davis, I., and Castello, A. (2016) The new (dis)order in RNA regulation. *Cell Commun Signal* **14**: 9.
Katoh, K., Rozewicki, J., and Yamada, K.D. (2017) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform*.
Kobayashi, M., Rodriguez, R., Lara, C., and Omata, T. (1997) Involvement of the C-terminal domain of an ATP-binding subunit in the regulation of the ABC-type nitrate/nitrite transporter of the Cyanobacterium *synechococcus* sp. strain PCC 7942. *J Biol Chem* **272**: 27197-27201.
Kyrpides, N.C., Woese, C.R., and Ouzounis, C.A. (1996) KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends Biochem Sci* **21**: 425-426.
Labella, J.I., Cantos, R., Espinosa, J., Forcada-Nadal, A., Rubio, V., and Contreras, A. (2017) PipY, a member of the conserved COG0325 family of PLP-binding proteins, expands the cyanobacterial nitrogen regulatory network. *Frontiers in microbiology* **8**: 1244.
Labella, J.I., Obrebska, A., Espinosa, J., Salinas, P., Forcada-Nadal, A., Tremino, L. et al. (2016) Expanding the Cyanobacterial Nitrogen Regulatory Network: The GntR-Like Regulator PlmA Interacts with the PII-PipX Complex. *Front Microbiol* **7**: 1677.
Laichoubi, K.B., Espinosa, J., Castells, M.A., and Contreras, A. (2012) Mutational analysis of the cyanobacterial nitrogen regulator PipX. *PloS One* **7**: e35845.
Laichoubi, K.B., Beez, S., Espinosa, J., Forchhammer, K., and Contreras, A. (2011) The nitrogen interaction network in Synechococcus WH5701, a cyanobacterium with two PipX and two P(II)-like proteins. *Microbiology* **157**: 1220-1228.
Lee, H.M., Flores, E., Herrero, A., Houmard, J., and Tandeau de Marsac, N. (1998) A role for the signal transduction protein PII in the control of nitrate/nitrite uptake in a cyanobacterium. *FEBS Lett* **427**: 291-295.
Luddecke, J., and Forchhammer, K. (2015) Energy Sensing versus 2-Oxoglutarate Dependent ATPase Switch in the Control of *Synechococcus* PII Interaction with Its Targets NAGK and PipX. *PloS one* **10**: e0137114.
Llacer, J.L., Espinosa, J., Castells, M.A., Contreras, A., Forchhammer, K., and Rubio, V. (2010) Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. *Proc Natl Acad Sci U S A* **107**: 15397-15402.
Llacer, J.L., Contreras, A., Forchhammer, K., Marco-Marin, C., Gil-Ortiz, F., Maldonado, R. et al. (2007) The crystal structure of the complex of PII and acetylglutamate kinase reveals how PII controls the storage of nitrogen as arginine. *Proc Natl Acad Sci U S A* **104**: 17644-17649.
Merrick, M. (2014) Post-translational modification of P II signal transduction proteins. *Front Microbiol* **5**: 763.
NandyMazumdar, M., and Artsimovitch, I. (2015) Ubiquitous transcription factors display structural plasticity and diverse functions: NusG proteins - Shifting shapes and paradigms. *Bioessays* **37**: 324-334.
Palanca, C., Pedro-Roig, L., Llacer, J.L., Camacho, M., Bonete, M.J., and Rubio, V. (2014) The structure of a PII signaling protein from a halophilic archaeon reveals novel traits and high-salt adaptations. *FEBS J* **281**: 3299-3314.
Prunetti, L., El Yacoubi, B., Schiavon, C.R., Kirkpatrick, E., Huang, L., Bailly, M. et al. (2016) Evidence That COG0325 Proteins are involved in PLP Homeostasis. *Microbiology*.
Santangelo, T.J., and Artsimovitch, I. (2011) Termination and antitermination: RNA polymerase runs a stop sign. *Nat Rev Microbiol* **9**: 319-329.
Schrodinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8. In.
Simm, S., Keller, M., Selymesi, M., and Schleiff, E. (2015) The composition of the global and feature specific cyanobacterial core-genomes. *Front Microbiol* **6**: 219.
Tanigawa, R., Shirokane, M., Maeda Si, S., Omata, T., Tanaka, K., and Takahashi, H. (2002) Transcriptional activation of NtcA-dependent promoters of *Synechococcus* sp. PCC 7942 by 2-oxoglutarate in vitro. *Proc Natl Acad Sci U S A* **99**: 4251-4255.
Tremiño, L., Forcada-Nadal, A., Contreras, A., and Rubio, V. (2017) Studies on cyanobacterial protein PipY shed light on structure, potential functions, and vitamin B6 -dependent epilepsy. *FEBS Lett* **591**: 3431-3442.
Vazquez-Bermudez, M.F., Herrero, A., and Flores, E. (2000) Uptake of 2-oxoglutarate in *Synechococcus* strains transformed with the *Escherichia coli kgtP* gene. *J Bacteriol* **182**: 211-215.
Vazquez-Bermudez, M.F., Herrero, A., and Flores, E. (2002) 2-Oxoglutarate increases the binding affinity of the NtcA (nitrogen control) transcription factor for the *Synechococcus glnA* promoter. *FEBS letters* **512**: 71-74.
Vijayan, V., Jain, I.H., and O'Shea, E.K. (2011) A high resolution map of a cyanobacterial transcriptome. *Genome Biol* **12**: R47.
Washio, T., Sasayama, J., and Tomita, M. (1998) Analysis of complete genomes suggests that many prokaryotes do not rely on hairpin formation in transcription termination. *Nucleic Acids Res* **26**: 5456-5463.
Zeth, K., Fokina, O., and Forchhammer, K. (2014) Structural basis and target-specific modulation of ADP sensing by the *Synechococcus elongatus* PII signaling protein. *J Biol Chem* **289**: 8960-8972.

Zhao, M.X., Jiang, Y.L., Xu, B.Y., Chen, Y., Zhang, C.C., and Zhou, C.Z. (2010) Crystal structure of the cyanobacterial signal transduction protein PII in complex with PipX. *J Mol Biol* **402**: 552-559.
Zuber, P.K., Artsimovitch, I., NandyMazumdar, M., Liu, Z., Nedialkov, Y., Schweimer, K. et al. (2018) The universally-conserved transcription factor RfaH is recruited to a hairpin structure of the non-template DNA strand. *Elife* **7**.

Figure Legends

Figure 1. Modifications affecting *pipX* and/or *pipY* genes engineered in *S. elongatus*. Schematic representations of the *pipXY* and NSI regions in wild-type and derivative strains. Relevant non-coding sequences are shown for regions between *pipX-pipY*, CK1-*pipXY* or, in the case of Ptrc constructs, upstream the corresponding *pipX* or *pipY* coding sequences, as indicated. The green shadow indicates the 1215 bp region shared by WT, CK1XY and 1^{S}Ptrc-PipXY strains.

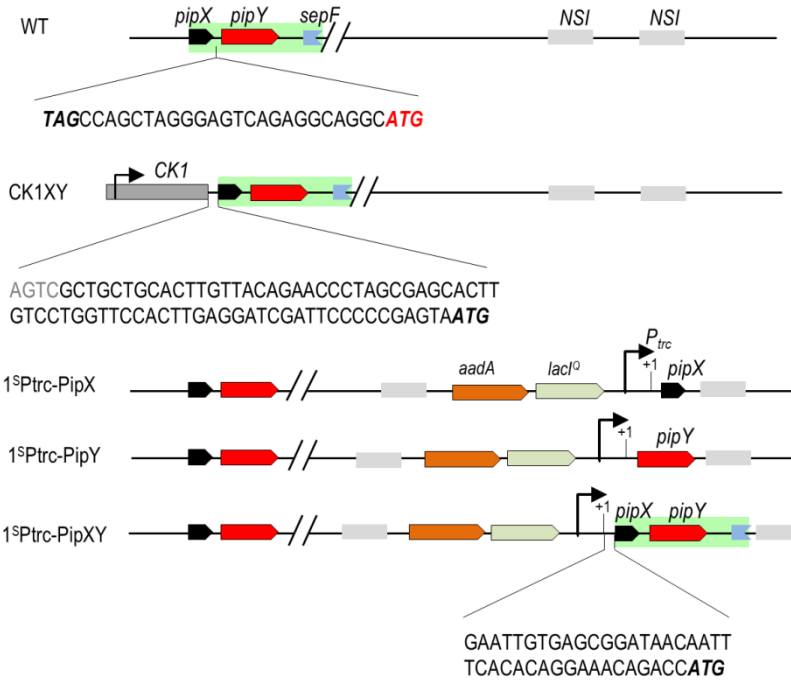
Figure 2. Effects of *pipX* and genomic context on up regulation of *pipY* or *cat* reporter genes. A) and B) Schematic representation of the relevant chromosomal region and expression levels of the indicated products of the corresponding strains. For chloramphenicol assays, cultures were adjusted to 0,5 (DO_{750nm}) prior to drop 5µl of the cell suspension onto BG11 plates (supplemented with 10 mg/L of chloramphenicol when indicated) and pictures were taken after 3 (- Cm) or 5 (+ Cm) days of incubation. Numbers on bottom refer to culture dilutions. Examples of PipX and PipY immunodetections in wild-type and mutant strains (1 or 2 clones, as indicated) are shown alongside strains. In A) Results of qRT-PCR for *pipY* transcripts normalized and relative to the WT values are shown (WT and *pipX* rows). C) Effect of IPTG on the expression of PipX and PipY from Ptrc constructs. Representative Western-blot images were obtained before (-) and 1 hour (+) after IPTG addition. PipX and PipY bands are always indicated by white arrowheads.

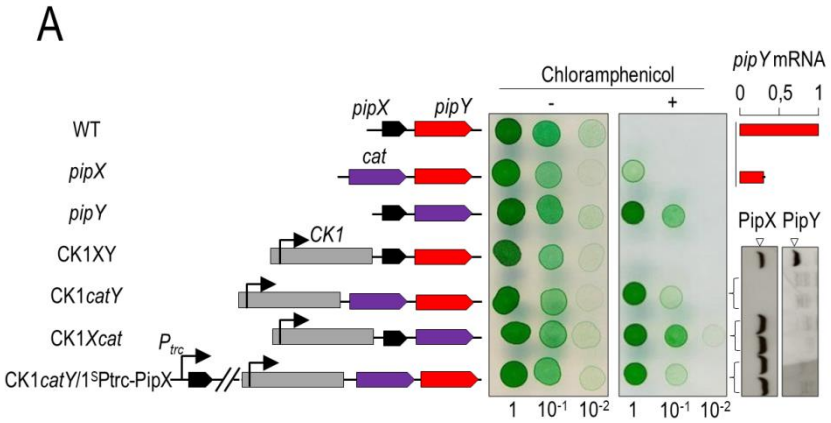
Figure 3. Weblogo comparison of NusG, RfaH and PipX. Sequences from NusG CTD (n = 5052), RfaH CTD (n = 695) and PipX NTD (n = 117) were aligned using MAFFT tool. Weblogos have been cleaned from gaps and aligned according to the MAFFT alignment of their amino acid reference sequences (*E. coli* NusG and RfaH and *S. elongatus* PipX proteins). Residues implicated in the NusG:NusE interaction (Burmann et al., 2010) are indicated with red arrowheads over the NusG weblogo. Conserved (F141) and conservatively replaced (F144; F165; R167) NusG key residues are indicated in the same manner over the

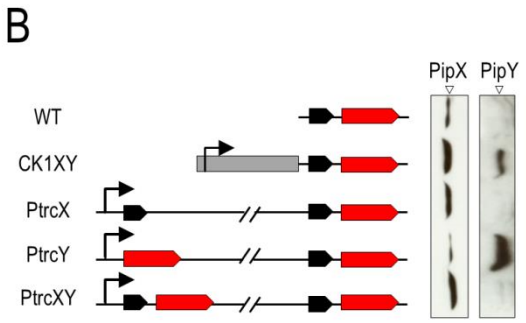
RfaH and PipX weblogos. The position of beta sheets is indicated with colored bars below the corresponding sequences.

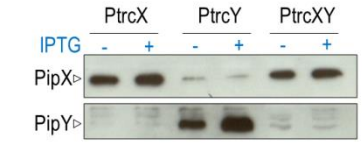
Figure 4. Structural (dis)similarities between PipX and NusG proteins. A) Spatial superposition of $PipX^{3-41}$ (grey, PDB 2XG8:D), NusG CTD from *E. coli* (cyan, PDB 2KVQ:G) and RfaH CTD from *E. coli* (yellow, PDB 2LCL:A). B) Barplot of mean RMSD values between $PipX^{3-41}$ domain, NusG and RfaH TLD/KOW domains. For each comparison, individual values are plotted as points in darker color (Species codes: Ec: *E. coli*, Tt: *Thermus thermophilus*, Tm: *Thermotoga maritima*, Mt: *Mycobacterium tuberculosis*; PDB IDs from left to right: 2JVV, 2KVQ, 5TBZ, 1NZ9, 2LQ8, 2XHC, 2MI6 and 2LCL). C) Structures of whole protein length of NusG from *E. coli* (PDB 5TBZ:J) and PipX from *S. elongatus* (PDB 2XG8:D). N-ter domain of NusG (NGN) is indicated in blue, TLD/KOW domains are indicated in red, C-ter domain of PipX is indicated in grey. D) Weblogo of the CTD of PipX, alpha helices are underlined in grey.

Figure 5. **Distribution of the intergenic distance between *pipX* and the gene downstream**. Solid and dashed lines denote organisms where the gene downstream is *pipY* (n = 99) or non-*pipY* (n = 18) respectively. Only organisms with both genes in the same DNA strand are plotted.









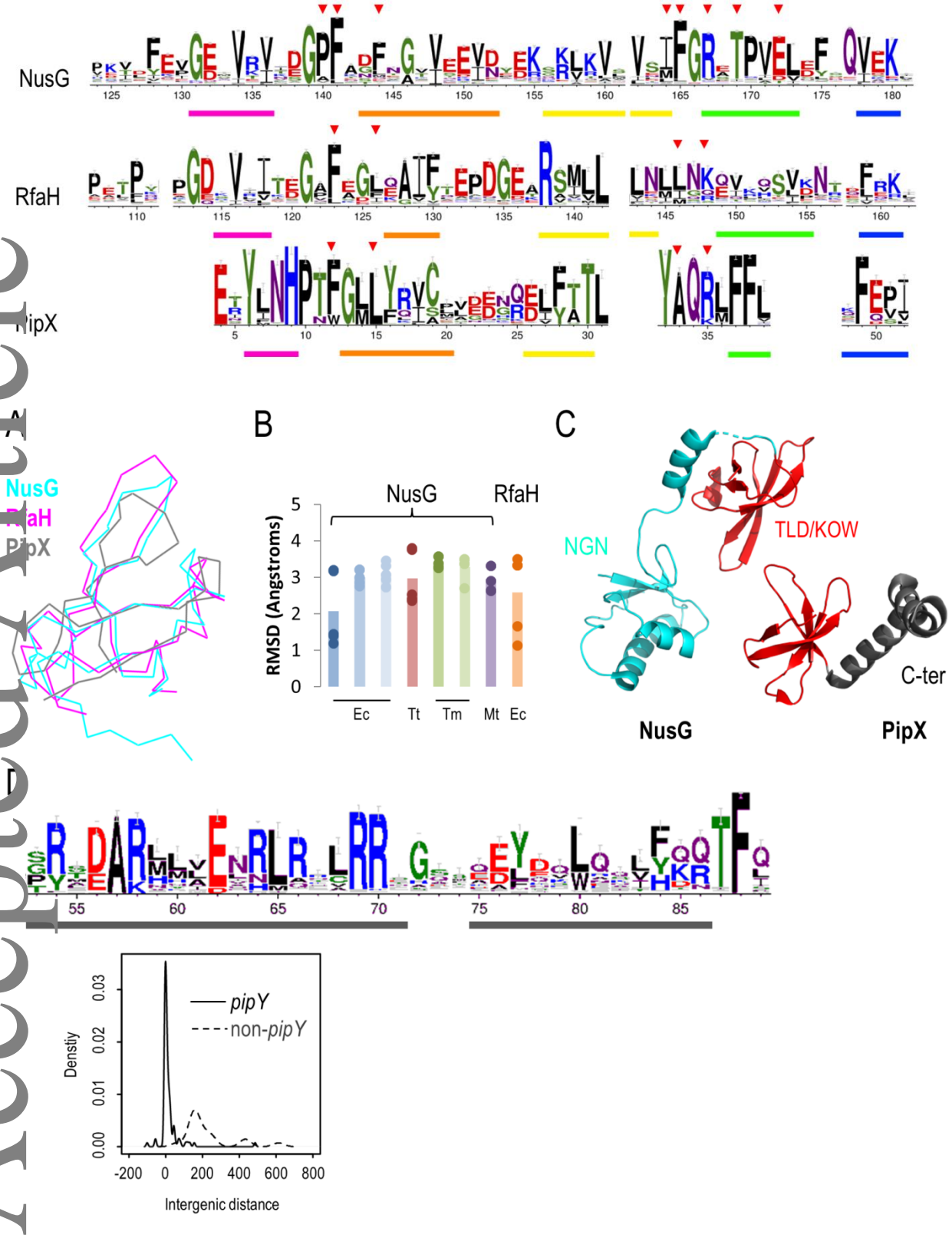


Table 1. Strains and plasmids

Strain	Genotype or relevant characteristics	Source or reference
Escherichia coli DH5α	F$^-$φ80 d*lacZ*ΔM15Δ(*lacZYA-argF*)U169 *endA1 recA1 hsdR17*(r_K^- m_K^+) *deoR thi-1 supE44 gyrA96 relA1*λ^-	(Hanahan, 1985)
WT	*Synechococcus elongatus* WT	Pasteur collection

pipX	PipX$^-$, Φ(P_{pipX}::*cat*), Cm^R	(Labella et al., 2017)
pipY	PipY$^-$, Φ(P_{pipX}::*cat*), Cm^R	(Labella et al., 2017)
CK1XY	PipX PipY(Con) [Φ(C.K1-*pipXY*)], Km^R	(Espinosa et al., 2006)
CK1*cat*Y	PipX$^-$ PipY(Con) [Φ(C.K1-*catpipY*)], Cm^R Km^R	This work
CK1X*cat*	PipY$^-$ PipX(Con) [Φ(C.K1-*pipXcat*)], Cm^R Km^R	This work
CK1*catY*-1^{S}Ptrc-PipX	PipY(Con) [Φ(C.K1-*pipX*)], Φ(P_{pipX}::*cat*), PipX^C, Φ(NSI-P_{trc}::*pipX*), Sm^R Cm^R Km^R	This work
1^{S}Ptrc-PipX	PipX^C, Φ(NSI-P_{trc}::*pipX*), Sm^R	(Labella et al., 2017)
1^{S}Ptrc-PipY	PipY^C, Φ(NSI-P_{trc}::*pipY*), Sm^R	(Labella et al., 2017)
1^{S}Ptrc-PipXY	PipX^C, Φ(NSI-P_{trc}::*pipXpipY*), Sm^R	This work
Plasmid	**Relevant characteristics**	**Reference**
pUAGC126	PipX$^-$, Ap^R, Cm^R	(Labella et al., 2017)
pUAGC127	PipY$^-$, Ap^R, Cm^R	(Labella et al., 2017)
pUAGC410	Φ(C.K1(+)-*pipX*), Ap^R, Km^R	(Espinosa et al., 2006)
pUAGC873	NSI, P_{trc}::*pipX lacI*, Ap^R, Sm^R	(Labella et al., 2017)
pUAGC294	NSI, P_{trc}::*pipY lacI*, Ap^R, Sm^R	(Labella et al., 2017)
pUAGC292	NSI, P_{trc}::*pipXpipY lacI*, Ap^R, Sm^R	This work