Genome-wide identification and characterization of Fur binding sites in the cyanobacteria *Synechocystis* sp. PCC 6803 and PCC 6714

Matthias Riediger[1], Miguel A. Hernández-Prieto[2], Kuo Song[1], Wolfgang R. Hess[1,*], Matthias E. Futschik[3,4,*]

[1]University of Freiburg, Faculty of Biology, Institute of Biology III, D-79104, Germany;

[2]ARC Centre of Excellence for Translational Photosynthesis & School of Life and Environmental Sciences, The University of Sydney, NSW 2006, Australia;

[3]Sysbiolab, Centre of Marine Sciences (CCMAR), University of Algarve, 8005-139, Faro, Portugal

[4]MRC London Institute of Medical Sciences (LMS), Faculty of Medicine, Imperial College London, London, W12 0NN, UK

Running title: The Fur regulon in *Synechocystis* spp.

*Corresponding authors:

Wolfgang R. Hess, University of Freiburg, Faculty of Biology, Institute of Biology III, D-79104 Freiburg, Schänzlestr. 1, Germany; Tel. +49-761-2032796; FAX: +49-761-2032745; Email: wolfgang.hess@biologie.uni-freiburg.de

and

Matthias E Futschik, MRC London Institute of Medical Sciences (LMS), Faculty of Medicine, Imperial College London, London, W12 0NN, UK; Tel. +44-2083838024; Email: m.futschik@imperial.ac.uk

Abstract

The Ferric uptake regulator (Fur) is crucial to both pathogenic and nonpathogenic bacteria for the maintenance of iron homeostasis as well as the defense against reactive oxygen species. Based on datasets from the genome-wide mapping of transcriptional start sites and transcriptome data, we identified a high confidence regulon controlled by Fur for the model cyanobacterium *Synechocystis* sp. PCC 6803 and its close relative, strain 6714, based on the conserved strong iron starvation response and Fur binding site occurrence. This regulon comprises 33 protein-coding genes and the sRNA IsaR1 that are under the control of 16 or 14 individual promoters in strains 6803 and 6714, respectively. The associated gene functions are mostly restricted to transporters and enzymes involved in the uptake and storage of iron ions, with few exceptions or unknown functional relevance. Within the *isiABC* operon, we identified a previously neglected gene encoding a small cysteine-rich protein, which we suggest calling, IsiE. The regulation of iron uptake, storage and utilization ultimately results from the interplay between the Fur regulon, several other transcription factors, the FtsH3 protease and the sRNA IsaR1.

Keywords: cyanobacteria, expression profiling, iron homeostasis, non-coding RNA, transcriptional regulation

1. Introduction

Ferrous iron (Fe^{2+}) is an important and often essential enzymatic co-factor for many enzymes. Therefore, most organisms have developed dedicated molecular acquisition mechanisms for ferrous as well as ferric (Fe^{3+}) iron, which can be enzymatically reduced to Fe^{2+}. At the same time, the uptake of iron needs to be tightly controlled because of the tendency of free intracellular iron to participate in harmful Fenton chemistry-dependent reactions, which can result in oxidative damage of cellular components including DNA.[1] Due to this duality of iron as an essential prosthetic group and as potential source of cellular damage, the amount of intracellular iron is typically controlled by an intricate regulatory network.[2] The tight control of iron is of special importance for photosynthetic organisms, which require large quantities of iron for the photosynthetic apparatus.

Cyanobacteria are the only prokaryotes performing oxygenic photosynthesis. The photosynthetic apparatus is particularly rich in iron-containing cofactors. Accordingly, cyanobacteria have multiple mechanisms for iron uptake and sophisticated systems for the regulation of iron metabolism. *Synechocystis* sp. PCC 6803 (from here: *Synechocystis* 6803) was the first phototrophic and the third organism overall for which a complete genome sequence was determined,[3] takes up DNA spontaneously and can be easily manipulated by homologous recombination and CRISPR-based tools.[4–6] The genome of *Synechocystis* 6803 has been extensively curated by the research community[7] and dedicated databases have been developed providing information on the regulation of gene expression at RNA level,[8] the composition of protein-protein and RNA-protein complexes,[9,10] and the intracellular localization of proteins.[11] Thus, *Synechocystis* 6803 has become one of the most popular cyanobacterial models. Bacteria sense cytosolic iron (Fe^{2+}) in a concentration-dependent manner through its interaction with the transcription factor Fur (Ferric uptake regulator). In its monomeric form, Fur binds free iron, which promotes a dimerization of Fur and a higher binding affinity for specific sites (so called Fur boxes) in the promotor region of genes blocking their transcription.[12] The genome of *Synechocystis* 6803 encodes three proteins that resemble the *E. coli* Fur protein. Of these proteins, only Fur encoded by *sll0567* seems to be directly involved in iron sensing, while the paralogs PerR (*slr1738*) and Zur (*sll1937*) have been related to oxidative stress[13] and zinc sensing[14], respectively.

Genome-level analyses showed that the expression of a large number of genes is directly or indirectly affected by iron limitation[15–19] but it has remained unknown which of these genes are directly controlled by Fur.

Based on multiple transcriptomic data sets for iron-depleted cultures of *Synechocystis* 6803,[15,17] we defined a set of genes which responded to changes in extracellular iron concentration in a highly robust manner and identified the putative Fur-binding site shared by all promoters of these genes. *Synechocystis* sp. PCC 6714 (from here: *Synechocystis* 6714) is a strain closely related to *Synechocystis* 6803 indicated by a 16S rDNA identity of 99.4% and 2838 shared protein-coding genes,[20] making it ideal for comparative analyses. Moreover, *Synechocystis* 6714 is also an established laboratory strain used by the research community.[21–23] Here, we used it to cross-validate the relative localization of the identified motifs and the regulation of the predicted Fur-controlled genes. The transcriptomic dataset for *Synechocystis* 6714 was generated from samples cultivated under the identical conditions as used for strain 6803.[24] Based on the two datasets of experimentally mapped transcription start sites (TSSs) and the validated Fur box, we identified the high confidence Fur regulon, consisting of at least 33 protein-coding genes and one sRNA gene, whose Fur regulation is conserved across the two strains, as well as ten or seven genes that are strain-specifically regulated in strain 6803 or 6714, respectively.

2. Materials and methods

2.1. Motif finding

Three different components from the MEME Suite (v.5.1.1; available at https://meme-suite.org/meme/) were used for the Fur box definition and identification of potential Fur binding sites.[25] First, we used MEME to discover the Fur binding motifs based on overrepresentation in the selected promotor regions for each strain individually, applying default settings but allowing repetitions of motifs. As input sequences, we selected the promoter regions of 11 high confidence Fur-regulated genes with differential expression in *Synechocystis* 6803 and 6714, using a window of 200 nt up- and downstream of the respective TSS. After obtaining two very similar palindromic motifs for each strain, the analysis was repeated with the combined set of promoter

sequences of the two strains while restricting the search for palindromic sequences. Next, we used TOMTOM,[26] to compare our motifs to experimentally verified Fur boxes from cyanobacteria[27–29] and motifs stored in the prokaryotic database CollecTF.[30] Finally, we used FIMO[31] to identify binding sites on a genome-wide scale based on the Fur motifs, matching the default p-value settings of ≤ 1.0E-04. For a more exhaustive comparison between the two strains, we also kept all predicted Fur boxes with less stringent p-value settings of ≤ 1.0E-03 to include specific cases, when a Fur box fell under the default threshold in just one of the strains.

2.2. Comparative TSS pair determination in *Synechocystis* 6803 versus 6714

The assigned TSSs of all transcriptional units (TUs), their transcription profiles and information about conserved non-coding TUs (nTUs) and antisense TUs (aTUs) as well as information about orthologs between the two strains were taken from previous analyses.[17,20,24] However, we re-matched the TSSs to their associated TUs and also allowed multiple TSSs to be assigned to the same TU. In particular, we included all TSS ≤ 300 nt upstream of the assigned main TSS as well as all internal TSS (iTSS) ≤ 20 nt downstream of the start codon as potential TSSs for each TU. If a gene did not have an assigned TSS, all TSSs in a distance ≤ 300 nt upstream from the annotated gene start were considered. In the case of multiple TSSs assigned to a gene, we determined the correlation coefficients of the transcription profiles for each possible TSS combination between the respective orthologs yielding a list of TSSs conserved between the two strains. Subsequently, the correlation coefficients of the transcription profiles of the orthologous TSSs to those of the respective coding regions was determined. The TSS pairs which had the largest correlation coefficients between the strains and those best matching the transcription profiles of the downstream coding regions were selected for the comparison of conserved Fur box occurrence and iron response between the strains.

2.3. Strain construction, cultivation, RNA and protein isolation and detection

Synechocystis 6803 strains overexpressing IsiE with or without a C-terminal triple FLAG epitope tag under control of the *petE* promoter were constructed and cultivated, as previously described for other small genes.[32] The kanamycin resistance cassette from plasmid pET28a was used to delete the *isiE* gene by homologous recombination.

To express IsiE-3xFLAG as a recombinant protein in *E. coli,* the vector backbone was amplified from pET28a using primer pair OEflag-vec-F/R, while the coding sequence of *isiE* was amplified from the genomic DNA of *Synechocystis* 6803 using primer pair OEflag-IsiE-F/R. The two fragments were fused together using AQUA cloning, and the resulting plasmid pOEIsiE-FLAG with ampicillin resistance was transferred into *E. coli* BL21 (DE3) cells for the expression of IsiE-3xFLAG. The primers used to PCR-amplify DNA fragments for the different cloning and mutagenesis steps are listed in **Supplementary Table S1**. Samples for RNA and protein analysis were collected and separated on denaturing formaldehyde-agarose gels or 15% glycine-SDS gels respectively following published protocols.[32] Nylon-membrane transferred RNA samples were hybridized with ^{32}P-labelled single-stranded RNA probes that were generated from amplified DNA fragments or else a directly labelled oligonucleotide, in case of 5S rRNA (**Supplementary Table S1**). FLAG-tagged proteins were detected by Western blot as described.[32]

2.4. Purification of recombinant IsiE-3xFLAG and absorption spectra

E. coli BL21 (DE3) carrying the plasmid pOEIsiE-FLAG was cultured in LB medium at 37^{o}C, 180 rpm to an OD600 of 0.8, and then induced with 0.5 mM IPTG for 4 hours. Cells were collected by centrifugation at 5,000 rpm for 5 min, and washed with lysis buffer (137 mM NaCl, 12 mM phosphate, 2.7 mM KCl, 1X cOmplete™ protease Inhibitor, pH 7.4,). Washed cells were spun again and suspended in 1/10 of the initial culture volume in lysis buffer. Cells were physically disrupted by high pressure using a Constant Systems Cell Disruptor. Cell lysate was centrifuged at 4^{o}C,13,000 rpm for 30 min to remove unbroken cell and membrane fractions. The supernatant was filtered through a 0.22 um pore size filter and used to purify IsiE-3xFLAG by specific binding to ANTI-FLAG® M2 magnetic beads according to the manufacturer's instructions. Absorption spectra were measured at room temperature using Specord® 250 Plus (Analytik Jena) spectrophotometer. The purified IsiE-3xFLAG protein was dissolved in PBS buffer (137 mM NaCl, 12 mM phosphate, 2.7 mM KCl, pH 7.4), while a solution of 3X FLAG peptide using the same buffer was used as blank control.

3. Results

3.1. Selecting a high confidence set of Fur-regulated genes

The transcriptional acclimation of *Synechocystis* 6803 to iron starvation has been analyzed in several studies using different technological platforms.[15,16,18,19,29] The set of genes detected as differentially expressed varied between the studies, likely due to variations in the experimental protocols. While the small overlaps between these studies were challenging for the consolidation of the results, they also provide the opportunity for more robust inference of biological knowledge. Indeed, expression patterns can be more biologically meaningful if they were observed across different experiments and platforms.[33,34] In this way, we aimed at exploiting the small overlap between previous transcriptomics studies to more accurately define a Fur-binding motif. We started with the results of a genome-wide mapping of TSSs and the definition of 4,091 transcriptional units (TUs) representing mono- and multicistronic operons and encompassing coding as well as non-coding genes under ten different growth conditions, including iron starvation.[17] To identify specifically upregulated TSSs, we compared the number of reads obtained under iron limitation for every single TU with those obtained under the nine different growth conditions. Next, genes associated with the identified TSSs were inspected for iron depletion-dependent regulation across the four microarray studies included in a meta-analysis of iron starvation[15] and in CyanoExpress[33]. This crossover led to the identification of 13 TUs that displayed a consistent behaviour upon iron depletion. We added the TU for the sRNA IsaR1, which is an independently validated Fur-controlled gene.[29] The 14 TUs were filtered for their presence, composition and the conservation of the iron depletion response in *Synechocystis* 6714. This led to 11 TUs containing genes which were consistently upregulated under iron limitation, and which clustered together in a meta-analysis of 115 environmental perturbations performed using CyanoEXpress (**Fig. S1A, B**). These genes were also among those showing the highest induction under iron depletion in the RNA-seq data compared to the other nine conditions. (**Fig. S1C**).

3.2. Computational identification of a Fur-binding motif

Using MEME, we detected in the promoter sequences of the 11 selected TUs (**Supplementary Datasets 1 and 2**), a 21 nt long motif in the two strains (**Fig. 1A,**

Table S2 and S3). Both motifs tend to be palindromic with a single central nucleotide, are AT rich, and exhibit only minor divergences in the underlying positional weight matrices. The high motif similarity correlates with the strongly conserved Fur protein sequence, differing in only three amino acids between the two strains (**Fig. S2**). Two substitutions are conservative replacements, while on position 28, a threonine (with uncharged side chain) in strain 6803 corresponds to a positively charged lysine in strain 6714, which might lead to slight differences in the structure of the DNA binding site. This may explain the minor differences between the Fur motifs of the two strains (**Fig. 1A**). The amino acid at this position is frequently replaced, e.g., by a histidine in *Anabaena* sp. PCC 7120 (**Fig. S2**). The MEME analysis using the combined promoter set yielded a *Synechocystis* Fur-binding consensus motif of 23 nt with high statistical confidence (**Fig. 1A**), which was used for all further analyses.

Our original choice of a search window 200 nt up- and downstream of the TSS, respectively, was motivated by the location of a Fur-binding site, 163–186 nt downstream of the TSS of *isiA*, whose relevance for de-repression under iron starvation was demonstrated experimentally.[27,29,35] The location of Fur binding sites that far downstream of the TSS is consistent with Fur functioning as a transcriptional roadblock, which in case of *isiA* leads to a prematurely terminated and clearly detectable transcript.[36,37] However, almost all Fur-regulated genes have a strong tendency to harbor at least one Fur-binding site in a window restricted from 100 nt upstream to 50 nt downstream of the respective TSS (**Fig. 1B**). Therefore, the genome-wide motif search using FIMO was restricted to a 150 nt range to increase specificity. When the *Synechocystis* consensus motif was compared with the transcription factor binding sites in the CollecTF database,[30] 25 motifs were identified with a *p*-value ≤ 0.01, with the top 12 motifs corresponding to Fur box elements in different species of bacteria (**Table S4**).

3.3. Correlation of iron response and Fur box conservation between *Synechocystis* 6803 and 6714

The genome-wide search for potential Fur binding sites (p-value ≤ 1.0 E-04) identified 114 TSSs associated with such elements for *Synechocystis* 6803 (**Table S5)** and 120 TSSs for *Synechocystis* 6714 (**Table S6**), representing ~2% of all verified TSS. The TSSs of both strains were paired according to their distance to the respective

orthologous genes, Fur box occurrence, and similarities in their activities under ten different environmental conditions (**Fig. S3** and **Table S7**).

Under iron starvation, a total of 78 TSSs in strain 6803 and 53 TSSs in strain 6714 were differentially expressed (adj. p-value ≤ 0.05). Among the genes controlled by these TSSs, there were 27 orthologous upregulated genes. Three orthologous protein-coding genes (*slr0888*, *slr0889*, *slr1634*), the *isiA* antisense RNA *isrR*, and the sRNA *pmgR1* showed conserved downregulation. Twenty of the conserved upregulated genes also harbor Fur boxes in both strains (**Fig. 2A**). Of the conserved downregulated genes, only *slr1634* has predicted potential Fur boxes in both strains and might be activated by Fur (**Fig. 2B**). For one gene, *ank*, iron-dependent regulation and a Fur-binding motif was detected only for the homolog in strain 6714 (**Fig. 2C**). The gene pair *isiB/C* exhibited similar regulation in both strains, but a Fur box was only detected in strain 6714 (**Fig. 2C**). Conversely, for the gene pair *exbD3/D4*, a Fur box was only detected in strain 6803,given that the genes are transcribed by readthrough from an upstream gene in strain 6714 (**Fig. 2D**). Finally, we identified a set of eight iron-responsive genes with conserved differential expression but lacking a Fur box in their promoters (**Fig. 2E**). Indeed, their expression is post-transcriptionally controlled by the Fur-regulated sRNA IsaR1 in the case of *sufBCDS*,[29] inversely through co-degradation with the Fur-regulated *isiA* mRNA in the case of IsrR,[36] or by currently unknown alternative mechanisms.

3.4. Defining a high confidence set of genes regulated by Fur and their functions

The presence of a Fur box and the specific differential expression under iron starvation resulting from the respective gene-specific promoter or as part of an operon is summarized for both strains in **Fig. 3**. There are 34 high confidence Fur target genes, including one sRNA gene, which exhibited a significant differential iron starvation response and harbors Fur boxes in both strains. Another ten genes are specific or specifically Fur-regulated only in strain 6803 and seven such genes in strain 6714 (**Fig. 3A**). Functional enrichment analysis identified 23 enriched GO terms, with an adjusted p-value ≤ 0.05, functionally assigned to 32 of these 51 genes (**Fig. 3B**).

To gain more detailed insights, we inspected the chromosomal regions of predicted Fur regulated genes further and found several remarkable examples of conserved or divergent Fur regulation.

3.5. Bidirectional promoters are a hallmark of Fur-mediated regulation

Visual inspection of predicted Fur box occurrences revealed that, besides the respectively assigned TSS, additional iron-regulated TSSs were frequently located in close proximity on the reverse strand, indicating bi-directional transcriptional regulation. This feature is also present in giant iron-responsive gene clusters, ~32 kb and 22 kb in size, which harbor the largest number of Fur-regulated genes (22 and 18), as well as 11 and 7 Fur boxes in the two studied strains, respectively (**Fig. 4**).

The genes located within this iron-responsive gene cluster encode most of the components of the various ferric iron uptake systems. These are the FecB-FecC/D-FecE system that consists of the periplasmic binding proteins FecB; the transmembrane permease proteins, FecC and FecD; and the ATPase FecE that allows the uptake of Fe-siderophores in conjunction with the FhuA and IutA TonB-dependent (TBDT) receptor proteins.[38] IutA was suggested to be renamed to SchT for schizokinen transporter.[39] These uptake systems function by coupling to the inner membrane TonB–ExbB–ExbD protein complex[40] that energizes the transport by building proton motive force. Indeed, one such complex is also encoded by genes of this cluster (*tonB, slr1484; exbB1, sll1404; exbD1, sll1405*),[41] while another set of *exbB/exbD* genes is located in a different location (*exbB1, sll0477; exbD3, sll0478; exbD4, sll0479*).[42]

The central role of these proteins is reflected by the transcriptional organization of their genes: *tonB* and *exbB1D1* are transcribed divergently from a bidirectional promoter that is safeguarded by four clearly discernible Fur boxes in both strains. The distance between the two TSSs are 223 nt in strain 6803 and 203 nt in strain 6714. The fact that in TU36 (in strain 6803) and TU1050 (in strain 6714), short non-coding RNAs possibly originate from this region makes the arrangement even more complex.

Other genes organized within the Fur-regulated operons of the gene cluster are orthologs of the pyochelin siderophore transcriptional activator PchR,[43] the methyltransferase Sll1407, and a few other uncharacterized periplasmic or

transmembrane proteins. Although the gene set differs between strains, with gene *D082_09160* only occurring in strain 6714 (TU1055) and two larger insertions in strain 6803, the positions of Fur boxes and the iron starvation-dependent regulation are strikingly well conserved (**Fig. 4**).

Another instance of bidirectional regulation was discovered for the Fur-regulated genes *sll1878* and *slr1977,* encoding FutC[44] and a nucleoside phosphorylase-I family (cl38914, e value = 1.23e-19) protein. The two major TSSs (TSS2 for both genes) are just spaced 58 nt apart from each other; hence, this is an interesting instance of a bidirectional promoter in the strictest sense (**Fig. S4**). TSS2 for *slr1977* leads to an mRNA with its 5' end already within the coding region, rendering the annotated (and conserved) start codon unlikely to start translation. It is tempting to speculate that transcripts originating at TSS1 allow the translation of *slr1977* from the annotated start codon, whereas TSS2 would lead to an N-terminally truncated, alternative protein. The fact that the *futC-slr1977* arrangement is conserved in both strains supports its possible functionality (**Fig. S4**).

3.6. Extension of the *isiABCD* operon and discovery of *isiE*

An intriguing example of conserved Fur regulation is the *isiABC* operon encoding the iron stress-induced protein A (IsiA) and flavodoxin (IsiB)[45,46] as well as a third annotated gene, *sll0249*, encoding a protein with an alpha/beta hydrolase fold. Previous work indicated that *sll0249* is co-transcribed with *isiAB* under iron-limited conditions and was renamed as *isiC.*[47] Our data suggest that in both strains, three Fur boxes control the transcription of this operon (**Fig. 5A**). One of these is located within the 5'UTR, directly upstream of the conserved *isiA* GTG start codon, as previously reported for strain 6803[27], and here identified also in strain 6714. Hence, a tight control by Fur can be expected and indeed strong transcriptional upregulation was observed under iron starvation starting at the *isiA* TSS.[15] The mRNA coverage extended beyond the *isiABC* genes and included the genes, *ssl0461* and *dfp*, encoding a DUF2555 domain-containing protein and the coenzyme A biosynthesis bifunctional protein CoaBC. It has been suggested independently that these five genes belong to this operon,[48] while *ssl0461* has recently been renamed to *isiD.*[49]

While the general arrangement and transcriptional organization of this operon appears to be common to both strains, a second TSS in strain 6803 at a position 17 nt downstream of the *isiA* stop codon seems to separate *isiA* from *isiB*. This TSS is only active during darkness; it does not appear to be active during iron starvation and the *isiABCD-dfp* operon can be transcribed jointly (**Fig. 5A**). Directly downstream of the *isiA* reading frame, a steep reduction is visible in the transcriptome coverage during iron starvation in both strains, and a second descent at the end of the relatively long (404 nt in strain 6803) intergenic spacer. We also noticed with *D082_02010*, a likely short gene interspersed between the *isiA* and *isiB* homologs in strain 6714, but not in strain 6803. D082_02010 is annotated to encode a 45-amino acid protein, although re-examination suggests a 59-amino acid cysteine-rich protein. Such short genes may be artefacts of gene modelling; therefore, we considered the *isiA-isiB* intergenic spacer in strain 6803 and found a 59-codon open reading frame (ORF) as well. Northern analysis showed that the coding sequence is transcribed as part of the joint transcript with *isiA* and *isiB* and that deletion of the coding sequence by a kanamycin resistance cassette also affected the transcription of the downstream located *isiB* (**Fig. 5B**). We therefore cloned this ORF under the control of an inducible promoter and engineered a short sequence encoding a 3x FLAG tag at its C terminus. Upon conjugation into *Synechocystis* 6803 and induction by adding Cu^{2+}, we observed the accumulation of a short protein in three biological replicates (**Fig. 5C**). We conclude that the *isiA-isiB* intergenic spacer contains a previously unknown protein-coding gene that we suggest be named *isiE*. The IsiE amino acid sequence is extremely cysteine-rich (8/59 amino acids in 6803 and 7/59 in 6714). Homologs can be identified in more than 100 cyanobacteria, both in syntenic and in different locations. Multiple sequence alignments show that CXCC/CXXC is the most widely conserved motif among these homologs (**Fig. 5D**). Similar motifs are known from metal-binding proteins and proteins interacting with DNA or RNA including the Fur protein itself containing six cysteine residues, two of which are arranged into a CXXC motif (**Fig. S2**). The corresponding CXXC motif in Fur from *Anabaena* sp. PCC 7120 was shown to be functionally critical.[50] Spectral analysis of the purified recombinant protein supported a possible iron-binding function of IsiE, while structural modelling indicated a close spatial arrangement of the cysteine residues protruding from the IsiE surface (**Fig. S5**).

Upstream of the *isiAEBCD-dfp* operon, a homolog of the ribosome biosynthesis GTPase YlqF/RbgA (Slr0267) is encoded, exhibiting a bidirectional regulation via Fur in strain 6714. Due to its low expression, there was no TSS defined for strain 6803; thus, a possible bidirectional regulation via Fur cannot be excluded in both strains.

3.7. Iron-dependent regulation by gene insertion

Most of the differentially expressed genes are associated with the conservation of a Fur motif. In the case of *ank* encoding an ankyrin-repeat-containing protein, we found a strain-specific upregulation. Closer inspection suggested a gene insertion that conferred Fur regulation to the *ank* gene exclusively in strain 6714 (**Fig. 6**). The inserted gene, *D082_28000*, is transcribed from a bidirectional promoter with twin Fur boxes that function in both directions. *D082_28000* encodes an iron-uptake porin (cl41527, e-value = 0e+00). There is no closely related ortholog in strain 6803, although it has six paralogous porin genes (Sll0772, Sll1271, Slr0042, Slr1908, Slr1841)[38] and Sll1550 (this work).

4. Discussion

Accurate prediction of the Fur regulon is essential for our understanding of iron homeostasis in cyanobacteria. Multiple transcriptomics experiments of iron depletion in *Synechocystis* 6803 by different laboratories indicated a large number of genes, which might be under the control of Fur. However, many of the observed expression changes were likely not a direct consequence of Fur binding but might have been secondary effects obscuring the primary response. To dissect primary (caused by differential Fur binding) and secondary downstream effects in the existing transcriptomic datasets, a faithful definition of the Fur-binding motif is key. Here, we improved the accuracy for Fur box identification, by considering two closely related *Synechocystis* strains and cross-validating the identified sites with their conserved transcriptional response to iron starvation. The results of our analysis captured the direct effects of the iron starvation response of *Synechocystis* mediated by Fur and led to the construction of a high confidence Fur regulon (**Fig. 7**). Furthermore, the detection of differences in the composition of the Fur regulon between the two closely related strains 6803 and 6714 provides a valuable resource for identifying factors

important for strain divergence. While iron deficiency in cyanobacteria directly affects Fe-S cluster biogenesis or general iron homeostasis and therefore a large number of genes,[18,51] the *Synechocystis* Fur regulon seems to be extended towards the regulation of iron uptake. According to their functions, the genes belonging to the Fur regulon can be assigned to four groups: (i) ferric iron transport, (ii) ferrous iron transport systems, (iii) further iron stress-responsive genes and (iv) regulatory factors.

4.1. The ferric iron transport gene cluster

Our analysis indicates that Fur regulates all genes involved in ferric iron transport in *Synechocystis* (**Fig. 4**). Interestingly, two large insertions in strain 6803 led to additional FecB, FhuA and PchR paralogs, plus two more genes lacking direct orthologs in strain 6714. The genomic organization of these insertions (*pchR-fhuA3-fecB3* and *pcrR-fhuA1-fecB2-sll1203-sll1204*), consisting of a transcriptional regulator, a TBDT receptor and a periplasmic siderophore-binding protein, suggests a conserved role (**Fig. 7**). The multiple TBDT receptors and FecB paralogs might exhibit variations in siderophore specificities and allow strain 6803 to be more flexible in utilizing different types of siderophores for iron uptake.[39,41] The role for PchR is not well-understood in cyanobacteria, as well as the reason for its collective duplication in strain 6803 together with siderophore-uptake and -binding proteins. However, PchR in *P. aeruginosa* acts as transcriptional activator of siderophore biosynthesis and export genes, requiring siderophores as co-activator and being also repressed by Fur.[43] *Synechocystis* 6803 does not produce siderophores and no siderophore biosynthesis genes have been detected.[52] Therefore, it is tempting to speculate that the different PchR paralogs in strain 6803 might respond to different siderophores as co-activators for the regulation of genes involved in recycling and secretion of previously imported siderophores. The Sll1407 methyltransferase might also contribute to the regulation of processes involved in siderophore-mediated iron uptake. The co-transcription of *tonB* with the genes encoding periplasmic protein Slr1485 and the ABC-transporter Slr1488, as well as co-transcription of *fecB4* with the genes encoding the ABC-transporter subunits Slr1494 and Slr1493, as well as of *fecB2* with *sll1204* and *sll1203* (strain 6803 only) suggest the existence of additional mechanisms for siderophore uptake and/or secretion that are not described so far. Overall, the ferric iron transport system in strain 6803 seems to be more complex than in strain 6714, except for the occurrence of

D082_09160 upstream of *fecE*. The functionality of multiple small ORFs (*D082_08980, D082_08970* and *D082_09120*) in the ferric iron transport gene cluster of strain 6714 has yet to be validated.

4.2. Ferrous iron transport systems

Fur also regulates the ferrous iron transport systems, which are thought to be supplied with Fe^{2+} diffusing through porins from the outer membrane (**Fig. 7**). While the Feo system is exclusively restricted to Fe^{2+} uptake, there are controversies on the ion affinities (Fe^{2+} or Fe^{3+}) of the Fut system.[53] In *Synechocystis*, the ferric iron uptake (Fut) system possesses two periplasmatic iron-binding proteins, FutA1 and FutA2, with distinct functions. While the former is largely responsible for the actual iron transport, the latter may act as metallochaperone to establish an ion gradient to assist in iron influx. It was also proposed that FutA2 is involved in iron reduction, possibly interacting with the siderophore-dependent ferric iron transport system.[54] Notably, only FutB, the integral membrane protein of the Fut system, seems to be not regulated via Fur. The two Fur-regulated porins, D082_28000 (strain 6714 only) and Sll1550/D082_06170, including the co-transcribed periplasmatic protein Sll1549/ D082_06180, might be supplying the ferrous iron (Feo) transport systems with Fe^{2+}. The Feo system surprisingly is restricted to strain 6803.

4.3. Comparison to other cyanobacteria and further iron starvation-responsive genes

Compared to the Fur regulons in more complex cyanobacteria, such as *Anabaena*, the *Synechocystis* Fur regulon is more compact. In *Anabaena* sp. PCC 7120, FurA-binding sites were identified upstream of 215 genes belonging to diverse functional categories including iron homeostasis, photosynthesis and respiration, heterocyst differentiation, oxidative stress defense and light-dependent signal transduction.[55] In contrast, the Fur regulon in *Synechocystis,* as defined here, comprises 33 protein-coding genes and the sRNA IsaR1 that are under the control of 16 and 14 individual promoters in strains 6803 and 6714, respectively. *Anabaena* spp. have a larger number and bigger diversity of iron transporters and proteins involved in the biosynthesis of siderophores. Therefore, it makes sense that the regulon controlled by Fur is larger in these species. Still, our study revealed several novel Fur targets in *Synechocystis* that are not directly involved in iron uptake, such as the bacterioferritin-

associated ferredoxin Ssl2250 (iron storage mechanisms), the nucleoside phosphorylase Slr1977, the ankyrin-repeat-containing protein Slr1109/D082_27980 (potentially involved in heme trafficking to catalase,[56] Fur-regulated only in strain 6714), the ribosome biogenesis GTPase YlqF/RbgA (only strain 6714), proteins encoded by the *isiAEBCD-dfp* operon, and the prolycopene isomerase CrtH in an operon with *isaR1* (**Fig. 7**). Fur is well documented as a repressor of *isiA*.[27,29,35,57] At the end of this work, a search for "*isiA*+*Synechocystis*" on PubMed Central returned close to 300 publications showing that *isiA* is one of the best-studied genes in this organism. Therefore, our discovery of the *isiE* gene in the intergenic spacer between *isiA* and *isiB* demonstrates that the comparative analysis of regulatory mechanisms and regulons can help to improve the fine mapping of even well-studied genomic regions, such as the Fur regulated *isiAEBCD-dfp* operon. While IsiE could not be assigned to any known protein family we noticed that the conserved CXCC motif (**Fig. 5D**) has been characterized as a metal-binding element.[58] Indeed, spectroscopic analysis yielded preliminary evidence that IsiE could be an iron-binding protein (**Fig. S5**). Recently, the discoordination of operon expression during iron starvation was reported for *isiA* and *isiB*, if ferredoxin 2 function was disturbed, and an unknown posttranscriptional regulation was postulated.[59] IsiE might be involved in this process. Alternatively, IsiE might play a role in iron recycling when photosystem I is structurally re-organized at low iron or redox stress conditions and becomes functionally associated with IsiA proteins.[60,61] Photosystem I is essential for the survival of cyanobacteria, requiring large quantities of iron-containing co-factors, thus a small protein such as IsiE might transiently store released iron from damaged photosystem I and release it when new photosystem subunits are being assembled. A similar role is performed by the small high-light induced proteins (HLIPs) in the assembly of chlorophylls to newly synthesized photosystems II.[62]

4.4. Connection between Fur and other transcriptional regulators and regulatory factors

Several Fur targets, namely those exhibiting strong induction during iron starvation, such as the *tonB* system (**Fig. 4**) or *isiA* (**Fig. 5**), contain more than one predicted binding site usually located upstream of the regulated genes, indicating an additive

effect of the binding of multiple Fur molecules. Additionally, Fur-regulated genes can be controlled also by other transcriptional regulators. Again, this is exemplified by the *isiA* promoter, which in addition to Fur is under control of the TetR-family transcriptional regulator PfsR (Sll1392)[63] and RpaB, integrating its expression with the cellular redox status.[35] Most Fur boxes are located within a distance of less than 50 nt upstream of the TSS. Surprisingly, we did not find a Fur box in the proximity of the *fur* gene, which was inferred to be auto-regulated in *Anabaena* sp. PCC 7119.[64] This indicates that the mechanism controlling the transcription of Fur potentially differs in *Synechocystis*. A candidate is PfsR because *fur* transcription increased in a *pfsR* deletion mutant during iron deprivation and recombinant PfsR was shown to bind to a *fur* promoter fragment.[63] Homologs of PfsR are lacking in *Anabaena* strains PCC 7119 as well as PCC 7120 ($E < 10^{-6}$). Hence, the autoregulatory function of Fur present in those species likely was replaced by the epistatic control through PfsR in *Synechocystis* spp.

The interconnection between the regulons of Fur (this study) and those controlled by other iron-related transcription factors, such as SufR,[65] and of secondary regulons, such as those belonging to the Fur-controlled sRNA IsaR1,[29] Sll1408 (PcrR), Slr1489 and Sll1205 (both PchR), assure the physiological regulation of iron metabolism and utilization (**Fig. 7**). In addition, paralogous transcription factors related to oxidative stress responses and other metal stress responses, such as PerR (*slr1738*) and Zur (*sll1937*),[13,14] need to be taken into account. Finally, the Fur regulon in *Synechocystis* 6803 is interconnected via an epistatic control mechanism involving FtsH proteases to photosynthesis and to other major transcriptional regulators. Fur becomes released from DNA when Fe^{2+} is becomes scarce. The released Fur protein is then degraded by the FtsH1/FtsH3 protease complex.[57] Interestingly, the manipulation of FtsH1/3 abundance leads to a drastic reduction in the transcriptional responses to different types of nutrient starvation, mediated not only by Fur, but also by the Pho, NdhR, and NtcA TFs.[66] Hence, the transcriptional regulation of iron, phosphorus, C and N starvation responses appear to be interconnected at a higher level, with the activity of the photosynthetic machinery through FtsH1/3 mediation illustrating the intriguing complexity of regulatory systems in cyanobacteria.

Funding

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; 322977937/GRK2344), the SCyCode research group FOR 2816 (grant DFG HE2544/15-2), by the Federal Ministry of Education and Research (BMBF) program RNAProNet (grant 031L0164B) to WRH and by a China Scholarship Council grant to K.S. Funding was also provided by the Portuguese Fundação para a Ciência e a Tecnologia (Foundation for Science and Technology; PTDC/BIA-MIC/4418/2012, IF/00881/2013, and UID/Multi/04326/2019 –The Centre of Marine Sciences to M.E.F.).

Acknowledgments

We thank Trudi Ann Semeniuk, Sydney, Australia, for critical reading.

Conflict of interest

None declared.

References

1. Halliwell, B., and Gutteridge, J. M. C. 2015, *Free Radicals in Biology and Medicine* Fifth Edition. Oxford University Press, Oxford, New York.
2. Morrissey, J., and Bowler, C. 2012, Iron utilization in marine cyanobacteria and eukaryotic algae. *Front. Microbiol.*, **3**, 43.
3. Kaneko, T., Sato, S., Kotani, H., et al. 1996, Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.*, **3**, 109–36.
4. Behler, J., Vijay, D., Hess, W. R., and Akhtar, M. K. 2018, CRISPR-based technologies for metabolic engineering in cyanobacteria. *Trends Biotechnol.*, **36**, 996–1010.
5. Grigorieva, G., and Shestakov, S. 1982, Transformation in the cyanobacterium *Synechocystis* sp. 6803. *FEMS Microbiol. Lett.*, **13**, 367–70.
6. Jansson, C., Debus, R. J., Osiewacz, H. D., Gurevitz, M., and McIntosh, L. 1987, Construction of an obligate photoheterotrophic mutant of the cyanobacterium *Synechocystis* 6803 1: Inactivation of the *psbA* gene family. *Plant Physiol.*, **85**, 1021–5.
7. Fujisawa, T., Narikawa, R., Maeda, S.-I., et al. 2017, CyanoBase: a large-scale update on its 20th anniversary. *Nucleic Acids Res.*, **45**, D551–4.
8. Hernandez-Prieto, M. A., and Futschik, M. E. 2012, CyanoEXpress: A web database for exploration and visualisation of the integrated transcriptome of cyanobacterium *Synechocystis* sp. PCC6803. *Bioinformation*, **8**, 634–8.
9. Riediger, M., Spät, P., Bilger, R., Voigt, K., Macek, B., and Hess, W. R. 2021, Analysis of a photosynthetic cyanobacterium rich in internal membrane systems via gradient profiling by sequencing (Grad-seq). *Plant Cell*, **33**, 248–69.
10. Xu, C., Wang, B., Yang, L., et al. 2021, Global landscape of native protein complexes in *Synechocystis* sp. PCC 6803. *Genomics Proteomics Bioinformatics, S1672-0229(21)00037-1*.
11. Baers, L. L., Breckels, L. M., Mills, L. A., et al. 2019, Proteome mapping of a cyanobacterium reveals distinct compartment organization and cell-dispersed metabolism. *Plant Physiol.*, **181**, 1721–38.
12. Lee, J.-W., and Helmann, J. D. 2007, Functional specialization within the Fur family of metalloregulators. *BioMetals*, **20**, 485.
13. Li, H., Singh, A. K., McIntyre, L. M., and Sherman, L. A. 2004, Differential gene expression in response to hydrogen peroxide and the putative PerR regulon of *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.*, **186**, 3331–45.
14. Barnett, J. P., Millard, A., Ksibe, A., Scanlan, D. J., Schmid, R., and Blindauer, C. A. 2012, Mining genomes of marine cyanobacteria for elements of zinc homeostasis. *Front. Microbiol.*, **3**, 142.
15. Hernández-Prieto, M. A., Schön, V., Georg, J., et al. 2012, Iron deprivation in *Synechocystis*: inference of pathways, non-coding RNAs, and regulatory elements from comprehensive expression profiling. *G3 Bethesda Md*, **2**, 1475–95.
16. Houot, L., Floutier, M., Marteyn, B., et al. 2007, Cadmium triggers an integrated reprogramming of the metabolism of *Synechocystis* PCC6803, under the control of the Slr1738 regulator. *BMC Genomics*, **8**, 350.

17. Kopf, M., Klähn, S., Scholz, I., Matthiessen, J. K. F., Hess, W. R., and Voß, B. 2014, Comparative analysis of the primary transcriptome of *Synechocystis* sp. PCC 6803. *DNA Res.*, **21**, 527–39.
18. Shcolnick, S., Summerfield, T. C., Reytman, L., Sherman, L. A., and Keren, N. 2009, The mechanism of iron homeostasis in the unicellular cyanobacterium *Synechocystis* sp. PCC 6803 and its relationship to oxidative stress. *Plant Physiol.*, **150**, 2045–56.
19. Singh, A. K., McIntyre, L. M., and Sherman, L. A. 2003, Microarray analysis of the genome-wide response to iron deficiency and iron reconstitution in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Physiol.*, **132**, 1825–39.
20. Kopf, M., Klähn, S., Pade, N., et al. 2014, Comparative genome analysis of the closely related *Synechocystis* strains PCC 6714 and PCC 6803. *DNA Res.*, **21**, 255–66.
21. Joset, F. 1988, Transformation in *Synechocystis* PCC 6714 and 6803: Preparation of chromosomal DNA In: Lester Packer, A. N. G., (ed.), *Methods in Enzymology*. Academic Press, pp. 712–4.
22. Kamravamanesh, D., Slouka, C., Limbeck, A., Lackner, M., and Herwig, C. 2019, Increased carbohydrate production from carbon dioxide in randomly mutated cells of cyanobacterial strain *Synechocystis* sp. PCC 6714: Bioprocess understanding and evaluation of productivities. *Bioresour. Technol.*, **273**, 277–87.
23. Ajlani, G., Kirilovsky, D., Picaud, M., and Astier, C. 1989, Molecular analysis of *psbA* mutations responsible for various herbicide resistance phenotypes in *Synechocystis* 6714. *Plant Mol. Biol.*, **13**, 469–79.
24. Kopf, M., Klähn, S., Scholz, I., Hess, W. R., and Voß, B. 2015, Variations in the non-coding transcriptome as a driver of inter-strain divergence and physiological adaptation in bacteria. *Sci. Rep.*, **5**, 9560.
25. Bailey, T. L., and Elkan, C. 1994, Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **2**, 28–36.
26. Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L., and Noble, W. S. 2007, Quantifying similarity between motifs. *Genome Biol.*, **8**, R24.
27. Kunert, A., Vinnemeier, J., Erdmann, N., and Hagemann, M. 2003, Repression by Fur is not the main mechanism controlling the iron-inducible isiAB operon in the cyanobacterium *Synechocystis* sp. PCC 6803. *FEMS Microbiol. Lett.*, **227**, 255–62.
28. González, A., Bes, M. T., Valladares, A., Peleato, M. L., and Fillat, M. F. 2012, FurA is the master regulator of iron homeostasis and modulates the expression of tetrapyrrole biosynthesis genes in *Anabaena* sp. PCC 7120. *Environ. Microbiol.*, **14**, 3175–87.
29. Georg, J., Kostova, G., Vuorijoki, L., et al. 2017, Acclimation of oxygenic photosynthesis to iron starvation is controlled by the sRNA IsaR1. *Curr. Biol.*, **27**, 1425-1436.e7.
30. Kiliç, S., White, E. R., Sagitova, D. M., Cornish, J. P., and Erill, I. 2014, CollecTF: a database of experimentally validated transcription factor-binding sites in Bacteria. *Nucleic Acids Res.*, **42**, D156-160.
31. Grant, C. E., Bailey, T. L., and Noble, W. S. 2011, FIMO: scanning for occurrences of a given motif. *Bioinformatics*, **27**, 1017–8.

32. Baumgartner, D., Kopf, M., Klähn, S., Steglich, C., and Hess, W. R. 2016, Small proteins in cyanobacteria provide a paradigm for the functional analysis of the bacterial micro-proteome. *BMC Microbiol.*, **16**, 285.
33. Hernández-Prieto, M. A., Semeniuk, T. A., Giner-Lamia, J., and Futschik, M. E. 2016, The transcriptional landscape of the photosynthetic model cyanobacterium *Synechocystis* sp. PCC6803. *Sci. Rep.*, **6**, 22168.
34. Russ, J., and Futschik, M. E. 2010, Comparison and consolidation of microarray data sets of human tissue expression. *BMC Genomics*, **11**, 305.
35. Riediger, M., Kadowaki, T., Nagayama, R., Georg, J., Hihara, Y., and Hess, W. R. 2019, Biocomputational analyses and experimental validation identify the regulon controlled by the redox-responsive transcription factor RpaB. *iScience*, **15**, 316–31.
36. Dühring, U., Axmann, I. M., Hess, W. R., and Wilde, A. 2006, An internal antisense RNA regulates expression of the photosynthesis gene *isiA*. *Proc. Natl. Acad. Sci. USA*, **103**, 7054–8.
37. Vinnemeier, J., Kunert, A., and Hagemann, M. 1998, Transcriptional analysis of the *isiAB* operon in salt-stressed cells of the cyanobacterium *Synechocystis* sp. PCC 6803. *FEMS Microbiol. Lett.*, **169**, 323–30.
38. Qiu, G.-W., Lou, W.-J., Sun, C.-Y., et al. 2018, Outer membrane iron uptake pathways in the odel cyanobacterium *Synechocystis* sp. strain PCC 6803. *Appl. Environ. Microbiol.*, **84**.
39. Obando S., T. A., Babykin, M. M., and Zinchenko, V. V. 2018, A cluster of five genes essential for the utilization of diydroxamate xenosiderophores in *Synechocystis* sp. PCC 6803. *Curr. Microbiol.*, **75**, 1165–73.
40. Fischer, E., Günter, K., and Braun, V. 1989, Involvement of ExbB and TonB in transport across the outer membrane of *Escherichia coli*: phenotypic complementation of exb mutants by overexpressed *tonB* and physical stabilization of TonB by ExbB. *J. Bacteriol.*, **171**, 5127–34.
41. Babykin, M. M., Obando, T. S. A., and Zinchenko, V. V. 2018, TonB-dependent utilization of dihydroxamate xenosiderophores in *Synechocystis* sp. PCC 6803. *Curr. Microbiol.*, **75**, 117–23.
42. Jiang, H.-B., Lou, W.-J., Ke, W.-T., Song, W.-Y., Price, N. M., and Qiu, B.-S. 2015, New insights into iron acquisition by cyanobacteria: an essential role for ExbB-ExbD complex in inorganic iron uptake. *ISME J.*, **9**, 297–309.
43. Michel, L., González, N., Jagdeep, S., Nguyen-Ngoc, T., and Reimmann, C. 2005, PchR-box recognition by the AraC-type regulator PchR of *Pseudomonas aeruginosa* requires the siderophore pyochelin as an effector. *Mol. Microbiol.*, **58**, 495–509.
44. Katoh, H., Grossman, A. R., Hagino, N., and Ogawa, T. 2000, A gene of *Synechocystis* sp. strain PCC 6803 encoding a novel iron transporter. *J. Bacteriol.*, **182**, 6523–4.
45. Burnap, R. L., Troyan, T., and Sherman, L. A. 1993, The highly abundant chlorophyll-protein complex of iron-deficient *Synechococcus* sp. PCC7942 (CP43') is encoded by the *isiA* gene. *Plant Physiol.*, **103**, 893–902.
46. Laudenbach, D. E., and Straus, N. A. 1988, Characterization of a cyanobacterial iron stress-induced gene similar to *psbC*. *J. Bacteriol.*, **170**, 5018–26.
47. Kojima, K., Suzuki-Maenaka, T., Kikuchi, T., and Nakamoto, H. 2006, Roles of the cyanobacterial *isiABC* operon in protection from oxidative and heat stresses. *Physiol. Plant.*, **128**, 507–19.

48. Cheng, Y., Zhang, T., Wang, L., and Chen, W. 2020, Transcriptome analysis reveals IsiA-regulatory mechanisms underlying iron depletion and oxidative-stress acclimation in *Synechocystis* sp. strain PCC 6803. *Appl. Environ. Microbiol.*, **86**, e00517-20.
49. Cheng, Y., Zhang, T., Cao, Y., Wang, L., and Chen, W. 2021, New insights into the function of the proteins IsiC and IsiD from *Synechocystis* sp. PCC 6803 under iron limitation. *Appl. Microbiol. Biotechnol.*, **105**, 4693–707.
50. Botello-Morte, L., Bes, M. T., Heras, B., Fernández-Otal, Á., Peleato, M. L., and Fillat, M. F. 2014, Unraveling the redox properties of the global regulator FurA from *Anabaena* sp. PCC 7120: disulfide reductase activity based on its CXXC motifs. *Antioxid. Redox Signal.*, **20**, 1396–406.
51. Gao, F. 2020, Iron–sulfur cluster biogenesis and iron homeostasis in cyanobacteria. *Front. Microbiol.*, **11**, 165.
52. Kranzler, C., Lis, H., Shaked, Y., and Keren, N. 2011, The role of reduction in iron uptake processes in a unicellular, planktonic cyanobacterium. *Environ. Microbiol.*, **13**, 2990–9.
53. Lau, C. K. Y., Krewulak, K. D., and Vogel, H. J. 2016, Bacterial ferrous iron transport: the Feo system. *FEMS Microbiol. Rev.*, **40**, 273–98.
54. Kranzler, C., Lis, H., Finkel, O. M., Schmetterer, G., Shaked, Y., and Keren, N. 2014, Coordinated transporter activity shapes high-affinity iron acquisition in cyanobacteria. *ISME J.*, **8**, 409–17.
55. González, A., Angarica, V. E., Sancho, J., and Fillat, M. F. 2014, The FurA regulon in *Anabaena* sp. PCC 7120: in silico prediction and experimental validation of novel target genes. *Nucleic Acids Res.*, **42**, 4833–46.
56. Flint, A., Sun, Y.-Q., and Stintzi, A. 2012, Cj1386 is an ankyrin-containing protein involved in heme trafficking to catalase in *Campylobacter jejuni*. *J. Bacteriol.*, **194**, 334–45.
57. Krynická, V., Tichý, M., Krafl, J., et al. 2014, Two essential FtsH proteases control the level of the Fur repressor during iron deficiency in the cyanobacterium *Synechocystis* sp. PCC 6803. *Mol. Microbiol.*, **94**, 609–24.
58. Sydor, A. M., Jost, M., Ryan, K. S., et al. 2013, Metal binding properties of Escherichia coli YjiA, a member of the metal homeostasis-associated COG0523 family of GTPases. *Biochemistry*, **52**, 1788–801.
59. Schorsch, M., Kramer, M., Goss, T., et al. 2018, A unique ferredoxin acts as a player in the low-iron response of photosynthetic organisms. *Proc. Natl. Acad. Sci. USA*, **115**, E12111–20.
60. Havaux, M., Guedeney, G., Hagemann, M., Yeremenko, N., Matthijs, H. C. P., and Jeanjean, R. 2005, The chlorophyll-binding protein IsiA is inducible by high light and protects the cyanobacterium *Synechocystis* PCC6803 from photooxidative stress. *FEBS Lett.*, **579**, 2289–93.
61. Ivanov, A. G., Krol, M., Sveshnikov, D., et al. 2006, Iron deficiency in cyanobacteria causes monomerization of photosystem I trimers and reduces the capacity for state transitions and the effective absorption cross section of photosystem I *in vivo*. *Plant Physiol.*, **141**, 1436–45.
62. Hernandez-Prieto, M. A., Tibiletti, T., Abasova, L., Kirilovsky, D., Vass, I., and Funk, C. 2011, The small CAB-like proteins of the cyanobacterium Synechocystis sp. PCC 6803: Their involvement in chlorophyll biogenesis for Photosystem II. *Biochim. Biophys. Acta BBA - Bioenerg.*, **1807**, 1143–51.

720 63. Cheng, D., and He, Q. 2014, PfsR Is a key regulator of iron homeostasis in
721 *Synechocystis* PCC 6803. *PLOS ONE*, **9**, e101743.
722 64. Bes, M. T., Hernández, J. A., Peleato, M. L., and Fillat, M. F. 2001, Cloning,
723 overexpression and interaction of recombinant Fur from the cyanobacterium
724 *Anabaena* PCC 7119 with isiB and its own promoter. *FEMS Microbiol. Lett.*, **194**,
725 187–92.
726 65. Vuorijoki, L., Tiwari, A., Kallio, P., and Aro, E.-M. 2017, Inactivation of iron-sulfur
727 cluster biogenesis regulator SufR in *Synechocystis* sp. PCC 6803 induces unique
728 iron-dependent protein-level responses. *Biochim. Biophys. Acta Gen. Subj.*, **1861**,
729 1085–98.
730 66. Krynická, V., Georg, J., Jackson, P. J., et al. 2019, Depletion of the essential
731 FtsH1/3 proteolytic complex suppresses the nutrient stress response in the
732 cyanobacterium *Synechocystis* PCC 6803. *Plant Cell*, **31**, 2912–28.
733

Figure Legends

Figure 1. **Definition and relative locations of putative Fur boxes in *Synechocystis* 6803 and 6714. (A)** Sequence logos of putative Fur boxes in strains 6803 and 6714 determined by MEME motif prediction and resulting consensus for both strains (bottom). The promoter sequences given in **Supplementary Datasets 1 and 2** were used as input for the motif prediction. The tabular output of the MEME motif prediction used as basis for the individual sequence logos can be found in **Table S2** and **S3**. The *Synechocystis* consensus Fur box was determined by performing a combined analysis of the promoters of the two strains. **(B)** Graphical output of the MEME motif prediction for 11 promoter regions conserved between *Synechocystis* 6803 (purple, left) and 6714 (orange, right). The Fur boxes of each strain are visualized by rectangular boxes of their corresponding colors. Predicted elements just below the threshold are indicated by less intense coloring. A dashed line boxes the region from -100 to +50 nt, in which most Fur boxes are located. The TU numbers are given as previously defined [17,24].

Figure 2. Scatter plot showing the conservation of expression under iron starvation between the two strains. The plots display the transcription levels at low iron, divided by the median of nine other conditions for strain 6803 (x axis) versus strain 6714 (y axis), and indicate Fur box occurrence. **(A)** Genes with Fur boxes in both strains (p-value ≤ 1.0E-04). **(B)** Genes with potential Fur boxes in both strains (p-value ≤ 1.0E-03). **(C)** Genes with Fur boxes in strain 6803 (p-value ≤ 1.0E-04). **(D)** Genes with Fur boxes in strain 6714 (p-value ≤ 1.0E-04). **(E)** Genes with iron-dependent regulation but lacking Fur boxes. Data points of genes which were differentially expressed during iron starvation (p-value ≤ 0.05) are colored: red = differentially expressed in both strains; purple = differentially expressed only in 6803; orange = differentially expressed only in 6714. Text labels of the *Synechocystis* 6803 orthologs are given for all differentially expressed genes with adjusted p-values ≤ 0.05 in at least one of the strains (except for panel E, adjusted p-value ≤ 0.05 in both strains only).

765

766 **Figure 3. High-confidence Fur target genes and their functions. (A)** Genomic
767 organization of high-confidence Fur target genes, including information of ortholog
768 conservation, relevant TUs, Fur box occurrence and differential expression under iron
769 starvation for both strains. Conserved Fur regulated orthologs are highlighted in red,
770 strain 6803 specific are in purple, and strain 6714 specific in orange. Genes with grey
771 background are likely, but not verifiably Fur targets, since these are linked to instances
772 of gene duplication events in strain 6803 and do not have an assigned TSS (compare
773 with **Fig. 4**). Instances of read-though transcription are highlighted by the vertical
774 arrows. Key examples are further outlined in the genome plots in **Fig. 4–6 and S4**. **(B**)
775 Functional enrichment analysis of the Fur target genes using Gene Ontology (GO)
776 terms. Genes without functionally enriched GO terms are in grey.

777

778 **Figure 4. Fur regulation of the ferric iron transport gene cluster.** The expression
779 under iron starvation is shown for *Synechocystis* **(A)** 6803 and **(B)** 6714. Fur box
780 locations are indicated with triangles. High-confidence Fur-regulated TUs are
781 highlighted: red, if regulated in both strains; purple, if regulated only in strain 6803;
782 and orange, if regulated only in strain 6714. Two insertions of several genes in strain
783 6803 compared to strain 6714 are indicated and an instance of two recombined genes
784 boxed by a dashed line. This figure illustrates that bidirectional transcriptional
785 regulation is a common feature of Fur. The conservation of Fur regulation even after
786 complex recombination events is illustrated, as seen in the separation of *fecC/D* from
787 *fecE/B* in strain 6714.

788

789 **Figure 5. Extension of the Fur-regulated *isiA* operon. (A)** The expression under
790 iron starvation is shown for *Synechocystis* 6803 and 6714. Conserved genomic
791 regions between the strains are boxed light grey. Fur box locations are indicated by
792 red triangles. High-confidence Fur-regulated TUs are highlighted together with gene
793 names and read coverage under iron starvation. The TSS separating *isiA* and *isiB* in
794 strain 6803 and the TSS following *isiC* in both strains play no roles under iron
795 starvation conditions. Therefore, genes of TU1555 and TU1554 were added to the

conserved Fur regulon. The newly defined *isiE* genes (this study) are connected by the bidirectional arrow. **(B)** Northern hybridization for the verification of *isiA, isiE* and *isiB* transcript accumulation in the Δ*isiE* and wild type (WT) strains during a time course of iron starvation. A size marker is given on the right, a 5S rRNA control hybridization is shown underneath. **(C)** Western blot for the detection of FLAG-tagged IsiE in three independent biological replicates C1–C3 but not in the wild type (WT). **(D)** IsiE homologs can be predicted in several cyanobacteria (strain acronyms and Genbank accession numbers are indicated). The multiple sequence alignment shows the conservation of a region containing two cysteine-rich motifs.

Figure 6. Regulation by Fur via gene insertion in strain 6714. The expression under iron starvation is shown for *Synechocystis* **(A)** 6803 and **(B)** 6714. Fur box locations are indicated by orange triangles. High-confidence Fur-regulated TUs are highlighted in orange. The insertion of the two genes between *gltD* and *ank* triggers Fur dependent de-repression of gene expression under iron starvation conditions specifically in strain 6714.

Figure 7. Extent of Fur-dependent regulation in *Synechocystis* strains 6803 and 6714 and functional association of the predicted target genes. Proteins colored in red: regulation is predicted for the orthologs in both strains, only the gene name from *Synechocystis* 6803 is given. Proteins in purple or with orange background: Gene has no ortholog in the other strain, or only the gene in the indicated strain is controlled by Fur (purple, *Synechocystis* 6803; orange: *Synechocystis* 6714). The interaction between the Fur regulon and the sRNA IsaR1-associated subregulon[29] is indicated. The role of secondary regulons controlled by PcrR and PchR homologs in iron homeostasis is currently unknown but might connect to later or additional stress responses. OM, IM and TM; outer, inner and thylakoid membrane.