Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120

Jan Mitschke[a], Agustín Vioque[b], Fabian Haas[a], Wolfgang R. Hess[a,1] and Alicia M. Muro-Pastor[b]

[a]University of Freiburg, Faculty of Biology, Schänzlestr. 1, D-79104 Freiburg, Germany

[b]Instituto de Bioquímica Vegetal y Fotosíntesis, Consejo Superior de Investigaciones Científicas and Universidad de Sevilla, E-41092 Seville, Spain

[1]Corresponding author: Wolfgang R. Hess; Phone: +49-761-203-2796

Fax: +49-761-203-2445; Email: Wolfgang.Hess@biologie.uni-freiburg.de

Keywords: prokaryotic cell differentiation, antisense transcription, heterocyst, nitrogen fixation, differential RNA-seq

Abbreviations: asRNA, antisense RNA; dRNA-seq, differential RNA-seq; HEP, heterocyst envelope polysaccharides; ncRNA, noncoding RNA; PSSM, position-specific scoring matrix; TAP, tobacco acid pyrophosphatase; TEX, Terminator™ 5'phosphate-dependent exonuclease; TSS, transcriptional start site

ABSTRACT

The fixation of atmospheric N_2 by photosynthetic cyanobacteria is a major source of nitrogen in the biosphere. In Nostocales, such as *Anabaena* sp. PCC7120, this process is spatially separated from oxygenic photosynthesis and occurs exclusively in heterocysts. Upon nitrogen step-down, these specialized cells differentiate from vegetative cells in a synchronized process that is controlled by two major regulators: NtcA and HetR. However, the regulon controlled by these two factors is only partially defined, and several aspects of the differentiation process have remained enigmatic. Using dRNA-seq, we experimentally define a genome-wide map of more than 10,000 transcriptional start sites (TSS) of *Anabaena* sp. PCC7120, a model organism for the study of prokaryotic cell differentiation and N_2 fixation. By analyzing the adaptation to nitrogen stress, our global TSS map provides insight into the dynamic changes that modify the transcriptional organization at a critical step of the differentiation process. We identify more than 900 TSS with an absolute fold change in response to nitrogen deficiency of at least eight. From these, at least 209 were under control of HetR, whereas at least 158 other TSS were not, but potentially directly controlled by NtcA. Most of the promoters activated during the switch to N_2 fixation were previously unknown, thereby adding hundreds of protein-coding genes and non-coding transcripts to the list of potentially involved factors. These data experimentally define the NtcA regulon and the DIF+ motif, a palindrome at or close to position -35 that seems essential for heterocyst-specific activation of expression of certain genes.

\body

Introduction

Cyanobacteria are oxygen-producing, photosynthetic organisms that are responsible for approximately half of the global CO_2 fixation. In addition, many cyanobacteria are able to perform N_2 fixation, a process that is extremely sensitive to oxygen. To protect the nitrogenase complex from photosynthetically-evolved oxygen, many filamentous strains, including *Anabaena* sp. PCC7120 (also known as *Nostoc* sp. PCC7120, from here on *Anabaena* 7120), differentiate heterocysts, a specialized cell type devoted to N_2 fixation (1, 2). Heterocyst differentiation is integrated into a series of physiological responses that take place when a source of combined nitrogen is not available. Those responses are globally controlled by NtcA, a transcriptional regulator of the cyclic AMP receptor protein (CAP) family (3-5). NtcA-mediated regulation involves the binding of NtcA to a consensus binding site with the sequence $GTAN_8TAC$ (3). In the absence of ammonium, the preferred nitrogen (N) source, NtcA (5) activates the expression of genes required for alternative assimilation pathways, such as those encoding nitrate and nitrite reductases (3). NtcA also acts as a transcriptional repressor of some genes, such as the *gif* gene, which encodes the glutamine synthetase inactivating factor (5). NtcA is also required for heterocyst differentiation and subsequent N_2 fixation (6, 7). A key factor activated by NtcA is HetR, a master regulator of many genes involved in the differentiation process (8, 9). HetR binds to DNA (10) and folds into an unusual structure (11). Although HetR was hypothesized to control the expression of hundreds of genes, only a single 17-base pair palindrome has been identified as a binding site (12). Moreover, HetR exerts positive feedback on NtcA expression (13), but it is not known how the double-positive feedback between both factors is terminated at a later step in the differentiation cascade.

During the last three decades, fewer than 100 transcriptional start sites (TSS) have been mapped in *Anabaena* 7120 on a gene-by-gene basis, primarily associated with highly expressed genes or genes regulated by nitrogen availability. An analysis of the *Anabaena* 7120 transcriptome during vegetative cell growth and in response to N deprivation has been recently published (14). However, the RNA samples were not enriched for primary transcripts, causing ~90% of the reads at each time point to correspond to processed rRNA and tRNA transcripts and rendering the direct identification of TSS difficult (14).

To complement the scarce existing information, our study utilized a differential RNA-seq (dRNA-seq) approach, which is selective for the 5’ ends of primary transcripts (15, 16) and allows the comprehensive determination of the transcriptional organization of a genome. Pretreatment of bacterial RNA with Terminator^{TM} 5' phosphate-dependent exonuclease (TEX) specifically degraded transcripts with a 5' P (processed RNAs) (16). Based on this approach, we present a genome-wide map of 13,705 candidate TSS that were experimentally mapped for the chromosome and the six plasmids of *Anabaena* 7120. Analyzing the transcriptional changes that occur upon nitrogen step-down in both the wild-type (WT) and a *hetR* mutant lead to the identification of all promoters controlled directly or indirectly by HetR, to a precise definition of the NtcA regulon and to the DIF+ motif, a previously unnoticed sequence element involved in heterocyst-specific expression. The availability of its annotated primary transcriptome will greatly facilitate the use of this genetically tractable organism as a model for prokaryotic cell differentiation and N_2 fixation.

Results

Large-scale mapping of primary 5’ ends using dRNA-seq

RNA samples were obtained from WT and *hetR* mutant strains that were grown in the presence of ammonia (WT-0 and *hetR*-0) or subjected to N step-down for 8 h (WT-8 and *hetR*-8). In total, 24,312,062 sequence reads (up to 75 nt long) were analyzed, and 1.6 billion bases of cDNA were mapped to the *Anabaena* 7120 chromosome and its six large plasmids. In the four samples, between 48.3 and 64.7% of the reads did not correspond to rRNAs. The complete dataset was taken to identify possible TSS, based on a minimum number of 50 sequencing reads associated with an RNA 5' end. An example of the data obtained is shown for the gene encoding HetR in **Fig. 1*A***. All four previously described TSS, including that at -271 whose induction is heterocyst-specific (13, 17, 18), were identified by a total of 6,900 reads in our dataset.

We identified 12,797 putative chromosomal TSS and 908 putative TSS on the six plasmids (**Table 1**). From these, 4,186 TSS were located within a distance of 200 nt upstream of an annotated gene (gTSS, mostly mRNAs); 4,172 TSS in inverse orientation (or ≤50 bp 5' or 3') to annotated genes (aTSS), suggesting antisense transcription; and 1,414 TSS of potential ncRNAs in intergenic spacers (nTSS). In addition, 3,933 TSS in sense orientation were located internally within annotated genes (iTSS). For consistency, this classification (**Fig. S1*A***), solely based on location, was used throughout. Therefore, some of the TSS here categorized as gTSS may actually give rise to ncRNAs and some of the nTSS may rather drive the transcription of genes with long 5'UTRs (see also comments in **Table S1**). Because for 704 TSS an association with more than one category was possible (**Fig. S1*B***), we prioritized gTSS over aTSS and iTSS and all remaining TSS were automatically categorized as nTSS. A global overview of the distribution of TSS is given in **Fig. 1*B*** (chromosome) and **Fig. S2** (plasmids). The exact positions of all putative TSS are indicated in **Supplementary Data Files 1** (chromosome) and **2-7** (plasmids) and in

Table S1. To benchmark, we compared our data with a set of 93 TSS previously reported for 60 different genes or operons from 59 independent studies (**Table S2**). From the previously reported TSS, 81 were confirmed with 69 of these being associated with 50 or more reads. In addition, TSS that had remained unnoticed were observed for several genes, including *glnA, ntcA, patS* and *rbcL*.

Finally, although processing of tRNA precursors is fast, 37 TSS were identified unambiguously for 35 tRNA genes. Based on their alignment, a consensus for a constitutive promoter was defined (**Fig. S3**). We observed that the length of the 5' leaders varied from 5 to >200 nt, but most were between 10 and 20 nt (**Fig. S3 and Supplementary data file 8**).

Nitrogen deficiency- versus differentiation-related promoters in the nitrogen stress response of *Anabaena* 7120

To analyze the transcriptional changes induced by nitrogen stress and those specifically leading to the differentiation of heterocysts, we individually compared the numbers of dRNA-seq reads of all TSS identified in RNA isolated from the four different samples. Normalized ratios (fold changes) in the number of reads between the different samples were determined according to established protocols (19).

The comparison of previously described N stress-induced transcriptional responses, e.g., for the *nirA-nrtABCD-narB* cluster (NtcA-activated), the *gifA* gene (NtcA-repressed), and genes in the heterocyst envelope polysaccharide (HEP) island, which are involved in heterocyst maturation, revealed a high degree of consistency (**Fig. S4**). We confirmed the TSS for *nirA* (-460 with respect to the translational start) (20), with a number of reads significantly higher in the samples collected 8 h after N step-down from both the WT and the *hetR* mutant (**Fig. S4*A***). We also confirmed the TSS for *gifA* at position -43 (21), associated with a high

number of reads in the presence of NH_4^+ from both the WT and the *hetR* mutant (**Fig. S4*B***). **Fig S4*C*** shows that transcription of several genes involved in the synthesis of HEP was almost completely restricted to WT under N stress.

Two main groups of TSS were defined that differed in their response to N step-down. The DEF category (deficiency-related changes) includes TSS showing transcriptional changes common to both strains (TSS for *nirA* or *gifA* above are paradigms for this category), whereas the DIF category (differentiation-related changes) includes TSS with transcriptional changes observed exclusively in the WT (e.g., TSS for genes in the HEP island, **Fig S4*C***). The DIF group includes all transcriptional changes that depend on HetR and are thus likely involved in the process of heterocyst differentiation. Additionally, although most changes in the DEF category involved activation (DEF+), we also identified some TSS associated with a decrease in the number of reads upon N step-down (DEF-) (**Tables S3** to **S5).** With a minimum fold change of eight, we identified 129, 28 and 209 TSS in the DEF+, DEF- and DIF+ categories, respectively. Our dataset identifies, for the first time, strongly regulated TSS for many genes with previously described N-dependent regulation or role in heterocyst differentiation, including *amt1* and *amt4*, heterocyst differentiation-related genes *hepA*, *hepB*, *hepN*, and *hepS,* regulatory genes *nirB* and *patB*, or *nblA* involved in phycobilisome degradation (**Table S6)**.

Our dataset allows the identification of multiple TSS in complex promoters (**Fig. 1*A***), providing a powerful approach to the analysis of genes with complex regulation. Two such promoter regions containing several previously unidentified TSS were chosen for further validation by primer extension or northern blot hybridization. Five putative TSS were identified for *nblA* (**Table S6; Fig. 2*A*)**, an N stress-inducible gene required for phycobilisome degradation but not essential for heterocyst differentiation in *Anabaena* 7120 (22). The results presented in **Fig. 2*B***

confirmed all 5' ends identified by dRNA-seq and their activity: TSS1 was not active in the presence of ammonium and NtcA-dependent, TSS2 and TSS5 were also inducible but NtcA-independent, TSS3 was barely detected, and TSS4 was mostly detected in the *ntcA* mutant. Two TSS were identified for *alr3808* **(Table S6; Fig. 2*C*)** encoding a DpsA homologue with known N-dependent regulation (23, 24). Consistent with the dRNA-seq data, two transcripts covering *alr3808* were identified by northern blot (**Fig. 2*D***). The longer transcript, probably originating at position 4601709f, is induced upon N step-down but was not expressed in the *hetR* mutant (therefore categorized as DIF+), whereas the shorter transcript, probably originating at position 4601982f, was also induced in the *hetR* mutant, although at later time points. Thus, not only the positions of TSS but also the regulation observed by dRNA-seq were confirmed by different methodologies.

TSS in the DIF category identify novel HetR-regulated elements and a sequence motif related to heterocyst-specific expression

HetR is the earliest known dedicated regulator involved in the differentiation of functional heterocysts. Therefore, transcriptional responses observed in the WT but not in the *hetR* mutant probably belong to the specific transcriptional program that leads to the differentiation of these cells. These TSS constitute the DIF+ category (see TSS with at least 8-fold change in **Table S5**). Their promoters can be analyzed to identify elements that might play a role in the differentiation process. In fact, the TSS for some genes with known HetR-dependent or heterocyst-specific expression (e.g. *ntcA*, *hetR*, *nsiR1*) appear among the TSS exhibiting the highest fold change in this category. A direct search for the 17-base pair palindrome identified as a HetR binding site (12), or for any other conserved element was unsuccessful. However, we noticed that the promoters for ncRNA NsiR1 (25) were in the DIF+ class. Because

NsiR1 is conserved in Nostocales and transcribed from a tandem array of short repeats, we could compare the promoter regions of 51 repeats from five different strains and found a conserved palindrome 5'TCCGGA at or close to the -35 position. Moreover, the same or a very similar motif is present in several other heterocyst-specific promoters (**Fig. 3*A***). A global search identified this motif at similar position in 58 of the 209 DIF+ promoters (**Table S5;** selected examples in **Fig. 3*A***) when a single mismatch was allowed (**Table S7**). From all remaining 13,496 TSS only 572 also share this motif (and some of those are DIF+ too, but with a fold change <8). Hence the enrichment for this motif within the DIF+ category of promoters is non-random ($P < 2.2e^{-16}$ in a Chi-squared test). We therefore named this sequence the DIF+ motif.

To directly test the functional relevance of the DIF+ motif, the 70 bp promoter from NsiR1 repeat 6 (P6) was placed upstream of a promoter-less GFP. Constructs bearing the intact DIF+ motif expressed green fluorescence exclusively in (pro)heterocysts (**Fig. 3*B***). Hence P6 possesses all the elements for cell-specific expression. In contrast, only very weak, non-heterocyst-specific fluorescence was obtained when the DIF+ motif was replaced by 5'GAATTC (**Fig. 3*B***). Thus the DIF+ motif is required for heterocyst-specific expression of the *nsiR1* promoter.

The NtcA binding site revisited: TSS in the DEF category define the NtcA regulon

We hypothesized that most N step-down-induced responses occurring both in the *hetR* strain and the WT (DEF categories) would not be related to heterocyst differentiation but rather are likely to be NtcA-regulated. The NtcA binding sites in the NtcA-activated promoters overlap in most cases the -35 region and are centered close to position -41.5 (i.e., the first nt is located at -48) with regard to the TSS (3).

Fig. 4*A* shows the promoters of the 20 TSS exhibiting the highest fold change in the DEF+ category. 18 of them contain sequences matching the consensus NtcA binding site at the expected position. Whereas NtcA-dependent activation was previously described for two of them, 580293f (26) and 5167792r (27), the remaining 16 strongly regulated TSS are novel. Thus, the DEF category defines the NtcA regulon at an unprecedented resolution. A position-specific scoring matrix (PSSM) was defined based on an alignment of all promoter regions for TSS in the DEF+ category with at least 8-fold change (**Fig. S5*A***). In addition to the expected conservation of positions 1-3 and 12-14, this analysis indicated that G residues are strongly avoided at positions 5, 7 and 11 of the NtcA binding site, whereas A/C or A/T are somewhat preferred at positions 5-7 (**Fig. 4*B, C***). When all TSS in the DEF+ category were scanned in a sliding window approach for possible NtcA binding sites with a score ≥5, a peak was identified at positions -48 to -49 (first nt of the motif), thereby matching the expected location precisely not only for this dataset with ≥8-fold change (**Fig. 4*D***), but also when a much larger dataset of 965 TSS with ≥2-fold change (**Fig. S5**) was analyzed. This observation indicated a significant proportion of the DEF+ TSS (even with relatively low fold changes) indeed contain NtcA binding sites at positions that are compatible with transcriptional activation and that the PSSM as defined here (**Fig. 4 and S5**) is useful for global searches.

NtcA binding sites incompatible with transcriptional activation (eventually repressing transcription) can be located closer to, and even downstream of, the TSS. To find such elements, we scanned sequences surrounding all TSS in the DEF category in two windows: position -120 to -44 (first nt of motif; activation-compatible sites), and -44 to +41 (repression-compatible sites) and show sites with score ≥5 in **Tables S8 and S9. Fig. 4*E*** shows ten examples for putative NtcA binding sites at repression-compatible positions around TSS in the DEF- category. Two of these

TSS, 2809313r (21) and 2807328f (28), were previously described as NtcA-repressed. TSS 1785466f (*rbcL*), also described as containing an NtcA binding site in a repressor-compatible position (29), is included for comparison.

Novel non-coding RNAs potentially involved in the response to N stress and heterocyst differentiation

Several of the TSS exhibiting the highest fold change in our dataset correspond to antisense (asRNA) or non-coding (ncRNA) transcripts. We have further confirmed transcription from some of them, including their regulation. The aTSS at position 3953418f, a strongly regulated DEF+ promoter, gives rise to an asRNA for gene *all3278*, whose mutation leads to the inability to fix N_2 in the presence of oxygen (30). Primer extension analysis confirmed the dRNA-seq results (**Fig. 5*A***). The initiation of transcription at this position is strongly induced by N step-down, independently of HetR, but depending on NtcA, consistent with the identification of a putative NtcA binding site upstream (**Table S8; Fig. 4*A***).

We also confirmed two strongly regulated nTSS that produce small ncRNAs (**Fig. 5*B***). Transcription from position 3141905r (DIF+) produces NsiR2, whereas transcription from position 5452083f (DEF+) produces NsiR3. The co-regulation of these TSS with well-studied protein-coding genes in these categories suggests that some of the ncRNAs identified here might be involved in the adaptation to N-stress or the differentiation of functional heterocysts.

DISCUSSION

In this study, we have defined a set of more than 10,000 putative TSS for *Anabaena* 7120. We did not require these 5' ends to be linked to a classical -10 element because the differentiation process could involve alternative sigma factors

recognizing different promoter elements, because the quality of data appeared high (3,401 TSS were identified on the basis of more than 300 reads) and because our dataset was experimentally validated in several ways. This dataset confirms most of the previously defined TSS for this organism (**Table S2**), while identifying new N-regulated TSS for the majority of genes previously reported as involved in heterocyst differentiation or adaptation to N-stress (**Table S6**). Additionally, using primer extension and northern blot analysis, we have confirmed several TSS in complex promoter regions (**Fig. 2**) or corresponding to asRNAs or ncRNAs (**Fig. 5**). Although a certain percentage of false positives cannot be excluded, when considering potential -10 elements, 9,885 TSS remain in the dataset at a threshold of +3.0 (**Table S1**). Approximately one third of all TSS were located upstream of an annotated gene, another third were found within annotated genes, while the remaining TSS were on the reverse complementary strand of 2,412 genes, suggesting antisense transcription to 39% of all genes. This number seems high but is consistent with observations for several other bacteria (31). A total of 1,414 TSS located in the intergenic regions more than 200 nt away from any annotated gene indicated a high number of ncRNAs, although some of these nTSS drive the transcription of mRNAs with very long leaders and therefore are functional gTSS (see comments in **Table S1**).

Our data provide unprecedented insight into the complexity of the primary transcriptome of *Anabaena* 7120 under standard growth conditions and at an early step of the heterocyst differentiation process. The use of the *hetR* strain, unable to start the transcriptional program leading to heterocyst differentiation, allowed us to separate transcriptional changes specifically related to this developmental process (DIF) from other N-stress responses that are still observed in the *hetR* mutant (DEF), thus likely unrelated to heterocyst differentiation but rather involved in other aspects of the adaptation to N stress. We thus defined sets of specifically regulated

promoters belonging to the DEF and DIF categories. The use of TSS included in these two categories defined both the NtcA and the HetR regulons. As exemplified by the cases of *hetR* (**Fig. 1*A***), *nblA* or alr*3808* (**Fig. 2**), the use of TEX-treated samples allowed the identification of multiple TSS in a given promoter region. Complex promoter regions with several TSS are commonly found in genes involved in heterocyst differentiation and patterning, probably due to differential TSS use in the two cell types of the filament (e.g., TSS for *ntcA*, *hetR*, *devB*, *hetC* in **Table S2**).

Over the last decades, genes involved in heterocyst differentiation and N_2 fixation were primarily identified by mutagenesis and screening of strains unable to grow in the absence of combined nitrogen (2, 30). Here, comparison of the wild-type to the *hetR* transcriptome upon N step-down yielded the DIF category, i.e. the HetR regulon, which now can be analyzed in search of HetR-dependent TSS corresponding to new and previously unknown genes potentially involved in the differentiation process. This regulon includes, for instance, genes related to cell wall synthesis and/or remodeling (**Fig. 3**), a key aspect of heterocyst differentiation. The DIF category also provides a valuable dataset to identify sequence motifs potentially involved in heterocyst-specific expression. Indeed, we identified the DIF+ motif common to many heterocyst-specifically expressed promoters. It consists of a short palindrome 5' TCCGGA, centered at or close to position -35, suggesting it might be recognized by a specific sigma factor rather than serve as a HetR binding site.

NtcA-mediated regulation is operated by binding to a consensus sequence, $GTAN_8TAC$, first described for strongly regulated promoters in *Synechococcus* (5). Promoters that are directly activated by NtcA contain an NtcA binding site that, in most cases, is centered close to position -41.5 with respect to the TSS (although some NtcA binding sites further upstream are also described). In such promoters, NtcA activates transcription in a manner that resembles CAP-mediated regulation at

Class II promoters. On the other hand, NtcA-mediated repression is operated by interaction with NtcA binding sites at a position that makes binding incompatible with the normal operation of the promoter. In the case of *gifA* from *Anabaena*, one NtcA binding site is centered at position -28.5 (21). The DEF category defined here identifies transcriptional responses that in many cases were directly regulated by NtcA as deduced from the identification of NtcA binding sites located in positions compatible with transcriptional activation (DEF+) or repression (DEF-) with respect to the corresponding TSS (**Tables S8, S9** and **Fig. 4**). Comparison to a computational prediction (32) of the NtcA regulon (**Table S10**) revealed that many of the TSS in the DEF category correspond to previously unknown NtcA-regulated promoters, thereby expanding the known NtcA regulon. The absence of NtcA binding sites in the promoters for several other TSS in the DEF categories (such as TSS1 for *nblA*; **Fig. 2*B* and 4*A***) suggests their expression is either not directly regulated by NtcA or operated by binding to different positions.

Finally, as observed in other cyanobacteria (15, 33, 34), our dataset indicates the abundant transcription of antisense and ncRNAs (e.g., ncRNA T1 in **Fig. 1*B***, associated with >600,000 reads). As previously described for NsiR1, a short ncRNA, whose expression is induced specifically in proheterocysts upon N step-down and belongs into the NtcA/HetR regulon (25), expression of some of these transcripts (**Fig. 5**) is regulated by N availability, suggesting that antisense and non-coding transcripts might be involved in the regulation of nitrogen assimilation and heterocyst differentiation. The annotated primary transcriptome of *Anabaena* 7120 during the transition from ammonium utilization to N_2 fixation will greatly facilitate the use of this organism as a model for prokaryotic cell differentiation and N_2 fixation in an oxygenic phototroph.

METHODS

Full protocols are available in SI Materials and Methods.

Growth conditions. Cultures of *Anabaena* 7120 WT, *hetR* mutant 216 (8) and *ntc*A mutant CSE2 (6) were bubbled with an air/CO_2 mixture (1% v/v) and grown photoautotrophically at 30°C in BG11$_0$C medium lacking $NaNO_3$ but containing 6 mM NH_4Cl, 10 mM $NaHCO_3$, and 12 mM *N*-tris (hydroxymethyl) methyl-2-aminoethanesulfonic acid-NaOH buffer (pH 7.5). Four RNA samples were isolated for dRNA-seq analysis from cells taken at T = 0 h (WT-0 and *hetR*-0) and T = 8 h (WT-8 and *hetR*-8) after removing all combined nitrogen from the media.

Preparation and analysis of RNA. Total RNA was isolated using hot phenol (35) with modifications. Northern blot hybridization and primer extension analysis of 5´ ends was performed as described (13, 34, 36).

Deep transcriptome sequencing. The cDNA libraries were prepared by vertis Biotechnologie, Germany (http://www.vertisbiotech.com/) after enrichment for primary transcripts by treatment with TEX (Epicentre) and analyzed on an Illumina sequencer by Beckman Coulter Genomics, Danvers, MA as previously described (16). Based on tetranucleotide tags (**Table S11**), 5.153.094, 4.690.212, 6.398.708 and 5.497.219 from these sequence reads were assigned to the WT-0, the WT-8, the *hetR*-0 and *hetR*-8 populations, respectively, and matched against the sequences of the chromosome or plasmids of *Anabaena* 7120.

Computational methods. Reads <18 nt and those with blast hits to the ribosomal clusters were filtered out. Remaining reads were mapped to the genome using the *segemehl* algorithm (37), with default parameters. Reads were pooled from the four samples and their 5' ends were binned within a 5-nt section. The position within the

window where the most reads began was considered to be the initial TSS. Because we noticed a few cases of initiation of transcription from a broader window, this dataset was clustered to allow the combination of initial TSS, which were not further than 5 nt apart. The position within this window where the greatest number of reads began was considered a TSS when a minimum of 50 reads was associated with it.

For ratio calculation, the number of reads for the four samples were normalized (19), and single pseudocounts were added to make the calculation of ratios possible for all TSS. The resulting ratios were classified and filtered into DEF and DIF categories of regulated promoters. Possible -10 elements were searched 6-8 nt upstream of all putative TSS and scored according to a PSSM derived from this dataset (**Table S1, Fig. S6**). To construct a PSSM for the NtcA binding site, all promoter regions for the 129 TSS in the DEF+ category with at least 8-fold change were aligned. From these, 81 possessed an element matching at least 4 nt of the $GTAN_8TAC$ motif centered at position 22/23 upstream of the -10 element. These elements were used together with six additional experimentally defined sites **(Table S2)** in the construction of the matrix. For the DIF+ motif, the regions -44 to -25 of all mapped TSS were searched in two distinct datasets: One consisted of all 209 TSS in the DIF+ class and the second of the remaining 13,496 putative TSS. Statistical significance of an enrichment for the DIF+ motif in the DIF+ class was tested in a Pearson's chi-squared test.

ACKNOWLEDGMENTS

This work was supported by the DFG program “Sensory and regulatory RNAs in Prokaryotes” SPP1258 (grant HE 2544 4-2), by the BMBF grant 0313921 (WRH) and by the Ministerio de Ciencia e Innovación grants BFU2007-60651 (AV) and BFU2010-14821 (AMP) co-financed by FEDER.

References

1. Flores E, Herrero A (2010) Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat Rev Microbiol* 8:39-50.
2. Kumar K, Mella-Herrera RA, Golden JW (2010) Cyanobacterial heterocysts. *Cold Spring Harbor Persp Biol* 2:1-19.
3. Herrero A, Muro-Pastor AM, Flores E (2001) Nitrogen control in cyanobacteria. *J Bacteriol* 183:411-425.
4. Herrero A, Muro-Pastor AM, Valladares A, Flores E (2004) Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. *FEMS Microbiol Rev* 28:469-487.
5. Luque I, Forchhammer K (2008) Nitrogen assimilation and C/N balance sensing. *The Cyanobacteria: Molecular Biology, Genomics and Evolution*, eds Herrero A & Flores E (Caister Academic Press, Hethersett), pp 335-382.
6. Frías JE, Flores E, Herrero A (1994) Requirement of the regulatory protein NtcA for the expression of nitrogen assimilation and heterocyst development genes in the cyanobacterium *Anabaena* sp. PCC7120. *Mol Microbiol* 14:823-832.
7. Wei TF, Ramasubramanian TS, Golden JW (1994) *Anabaena* sp. strain PCC 7120 *ntcA* gene required for growth on nitrate and heterocyst development. *J Bacteriol* 176:4473-4482.
8. Buikema WJ, Haselkorn R (1991) Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. *Genes & Development* 5:321-330.
9. Black TA, Cai Y, Wolk CP (1993) Spatial expression and autoregulation of *hetR*, a gene involved in the control of heterocyst development in *Anabaena*. *Mol Microbiol* 9:77-84.

10. Huang X, Dong Y, Zhao J (2004) HetR homodimer is a DNA-binding protein required for heterocyst differentiation, and the DNA-binding activity is inhibited by PatS. *Proc Natl Acad Sci USA* 101:4848-4853.
11. Kim YC*, et al.* (2011) Structure of HetR - a novel transcription factor required for heterocyst differentiation in cyanobacteria. *Proc Natl Acad Sci USA* 108:10109-10114.
12. Higa KC, Callahan SM (2010) Ectopic expression of *hetP* can partially bypass the need for *hetR* in heterocyst differentiation by *Anabaena* sp. strain PCC 7120. *Mol Microbiol* 77:562-574.
13. Muro-Pastor AM, Valladares A, Flores E, Herrero A (2002) Mutual dependence of the expression of the cell differentiation regulatory protein HetR and the global nitrogen regulator NtcA during heterocyst development. *Mol Microbiol* 44:1377-1385.
14. Flaherty BL, Van Nieuwerburgh F, Head SR, Golden JW (2011) Directional RNA deep sequencing sheds new light on the transcriptional response of *Anabaena* sp. strain PCC 7120 to combined-nitrogen deprivation. *BMC Genomics* 12:332.
15. Mitschke J*, et al.* (2011) An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. *Proc Natl Acad Sci USA* 108:2124-2129.
16. Sharma CM*, et al.* (2010) The primary transcriptome of the major human pathogen, *Helicobacter pylori*. *Nature* 464:250-255.
17. Buikema WJ, Haselkorn R (2001) Expression of the *Anabaena hetR* gene from a copper-regulated promoter leads to heterocyst differentiation under repressing conditions. *Proc Natl Acad Sci USA* 98:2729-2734.

18. Rajagopalan R, Callahan SM (2010) Temporal and spatial regulation of the four transcription start sites of *hetR* from *Anabaena* sp. strain PCC 7120. *J Bacteriol* 192:1088-1096.

19. Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11:R25.

20. Frías JE, Flores E, Herrero A (1997) Nitrate assimilation gene cluster from the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 179:477-486.

21. Galmozzi CV, Saelices L, Florencio FJ, Muro-Pastor MI (2010) Posttranscriptional regulation of glutamine synthetase in the filamentous cyanobacterium *Anabaena* sp. PCC 7120: differential expression between vegetative cells and heterocysts. *J Bacteriol* 192:4701-4711.

22. Baier K, Lehmann H, Stephan DP, Lockau W (2004) NblA is essential for phycobilisome degradation in *Anabaena* sp. strain PCC 7120 but not for development of functional heterocysts. *Microbiology* 150:2739-2749.

23. Ehira S, Ohmori M (2006) NrrA, a nitrogen-responsive response regulator facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Microbiol* 59:1692-1703.

24. Ow SY*, et al.* (2008) Quantitative shotgun proteomics of enriched heterocysts from *Nostoc* sp. PCC 7120 using 8-plex isobaric peptide tags. *J Proteome Res* 7:1615-1628.

25. Ionescu D, Voss B, Oren A, Hess WR, Muro-Pastor AM (2010) Heterocyst-specific transcription of NsiR1, a non-coding RNA encoded in a tandem array of direct repeats in cyanobacteria. *J Mol Biol* 398:177-188.

26. Valladares A*, et al.* (2011) Specific role of the cyanobacterial PipX factor in the heterocysts of *Anabaena* sp. strain PCC 7120. *J Bacteriol* 193:1172-1182.

27. Muro-Pastor AM, Olmedo-Verd E, Flores E (2006) All4312, an NtcA-regulated two-component response regulator in *Anabaena* sp. strain PCC 7120. *FEMS Microbiol Lett* 256:171-177.

28. Valladares A, Muro-Pastor AM, Herrero A, Flores E (2004) The NtcA-dependent P1 promoter is utilized for *glnA* expression in N_2-fixing heterocysts of *Anabaena* sp. strain PCC 7120. *J Bacteriol* 186:7337-7343.

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

30. Lechno-Yossef S, Fan Q, Wojciuch E, Wolk CP (2011) Identification of ten *Anabaena* sp. genes that, under aerobic conditions, are required for growth on dinitrogen but not for growth on fixed nitrogen. *J Bacteriol* 193:3482-3489.

31. Georg J, Hess WR (2011) cis-antisense RNA, another level of gene regulation in bacteria. *Microbiol Mol Biol Rev* 75:286-300.

32. Novichkov PS*, et al.* (2010) RegPrecise: a database of curated genomic inferences of transcriptional regulatory interactions in prokaryotes. *Nucleic Acids Res* 38:D111-118.

33. Georg J*, et al.* (2009) Evidence for a major role of antisense RNAs in cyanobacterial gene regulation. *Mol Syst Biol* 5:305.

34. Steglich C*, et al.* (2008) The challenge of regulation in a minimal photoautotroph: non-coding RNAs in *Prochlorococcus*. *PLoS Genet* 4:e1000173.

35. Mohamed A, Jansson C (1989) Influence of light on accumulation of photosynthesis-specific transcripts in the cyanobacterium *Synechocystis* 6803. *Plant Mol Biol* 13:693-700.

36. Muro-Pastor AM, Valladares A, Flores E, Herrero A (1999) The *hetC* gene is a direct target of the NtcA transcriptional regulator in cyanobacterial heterocyst development. *J Bacteriol* 181:6664-6669.

37. Hoffmann S*, et al.* (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput Biol* 5:e1000502.

38. Vioque A (1992) Analysis of the gene encoding the RNA subunit of ribonuclease P from cyanobacteria. *Nucleic Acids Res* 20:6331-6337.

39. Aldea MR, Mella-Herrera RA, Golden JW (2007) Sigma factor genes *sigC*, *sigE*, and *sigG* are upregulated in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 189:8392-8396.

40. Muro-Pastor AM, Flores E, Herrero A (2009) NtcA-regulated heterocyst differentiation genes *hetC* and *devB* from *Anabaena* sp. strain PCC 7120 exhibit a similar tandem promoter arrangement. *J Bacteriol* 191:5765-5774.

FIGURE LEGENDS

Figure 1. Genome-wide identification of transcriptional start sites (TSS) in *Anabaena* 7120. (*A*) Differential RNA-seq of TEX-treated samples identifies single TSS in complex promoter regions as exemplified by the *hetR* gene. The total number of reads mapped to each 5´ end is indicated for each of the four previously described TSS. (*B*) Distribution of 3401 TSS with ≥300 reads each along a linear plot of the *Anabaena* 7120 chromosome. TSS mapped for the forward strand are plotted above the x-axis, and for the reverse strand below. The number of sequence reads is given on the y-axis (logarithmic scaling). The location of each TSS according to **Fig. S1*A*** served for classification as gTSS (blue), nTSS (green), aTSS (red) or iTSS (grey). Selected TSS for each of the four classes are annotated.

Figure 2. Analysis of genes with multiple TSS identified by dRNA-seq. RNA was isolated from ammonium-grown cells (lanes labeled 0) or from ammonium-grown cells incubated in the absence of combined nitrogen for the number of hours indicated. (*A*) Graphical representation of reads mapped to the promoter of *nblA*. The histograms correspond to the WT-0 (red), WT-8 (green), *hetR*-0 (black) and *hetR*-8 (blue) samples. (*B*) Primer extension analysis of the *nblA* mRNA in WT and mutant strains CSE2 (*ntcA*) and 216 (*hetR*). Samples contained 20 µg of RNA. The oligonucleotide used was complementary to positions +5 to -18 with respect to the translational start of *nblA*. The 5' ends identified by dRNA-seq are numbered 1-5 (for positions, see **Table S1**). (*C*) Graphical representation of reads mapped to the promoter of *alr3808*. (*D*) Northern blot analysis of the *alr3808* mRNA in *Anabaena* 7120 and mutant strain 216 (*hetR*). Samples contained 10 µg of RNA. The probe used was an internal fragment of *alr3808*. *rnpB* (38) was used as a loading control.

Figure 3. Occurrence of a palindrome, 5' TCCGGA, in promoters of the DIF+ category. (*A*) Alignment of the heterocyst-specific promoters for NsiR1 (25), the *hetR* TSS3 (18), *sigC* (39) and *hetC* (36, 40) with selected promoters in the DIF+ category (fold change ≥ 8) containing TCCGGA around position -35 (one mismatch allowed). (*B*) Cell-specific transcription from the wild-type promoter of NsiR1 (P6; upper panels) or a mutated version of P6 carrying GAATTC instead of TCCGGA (lower panels). Images corresponding to red autofluorescence (left panel) and GFP fluorescence (right panel) are shown. White triangles point to proheterocysts. Scale bars, 10 μm.

Figure 4. NtcA-activated and repressed promoters. (*A*) Promoter regions of 20 TSS in the DEF+ category with the highest fold change. Possible -10 elements are highlighted in blue, nucleotides matching the consensus for NtcA binding sites in red, and the underlined nucleotide in each sequence is the TSS. (*B*) Nucleotide frequencies derived for positions 1 to 14 of 87 putative NtcA binding sites. (*C*) Corresponding Weblogo. (*D*) Position of NtcA binding sites identified in a sliding window approach using the PSSM along the promoters in the DEF+ category (fold change ≥8). The bars indicate the first nucleotide of a putative NtcA binding site and its position with regard to the TSS (*E*) Putative NtcA binding sites identified at repression-compatible positions around TSS in the DEF- category.

Figure 5. Experimental verification of newly identified transcripts classified as antisense or non-coding. RNA was isolated from ammonium-grown cells (lanes labeled 0) or from cells incubated in the absence of combined nitrogen for the number of hours indicated. (*A*) Primer extension analysis of the *all3278* asRNA in

Anabaena 7120 WT and mutant strains CSE2 (*ntcA*) and 216 (*hetR*). Samples contained 20 µg of RNA. (*B*) Northern blot analysis of ncRNAs NsiR2 (upper panel) and NsiR3 (middle panel) in *Anabaena* 7120 WT and mutant strains CSE2 (*ntcA*) and 216 (*hetR*). The samples contained 10 µg of RNA. The *trnL-UAA* transcript (lower panel) was used as a loading control. (*C*) Predicted secondary structure of NsiR3.

Table 1. Overview on the number and types of putative TSS mapped for the chromosome (chr) and plasmids alpha, beta, gamma, delta, epsilon and zeta.

	chr	alpha	beta	gamma	delta	epsilon	zeta	Total
length (nt)	6,413,771	408,101	186,614	101,965	55,414	40,340	5,584	-
# of genes	5430	386	186	90	85	31	5	
gTSS	3955	145	41	24	14	6	1	4186
aTSS	3854	188	73	34	13	8	2	4172
iTSS	3722	113	50	18	21	6	3	3933
nTSS	1266	88	15	22	8	7	8	1414
Total	12797	534	179	98	56	27	14	13705

FIGURE 1

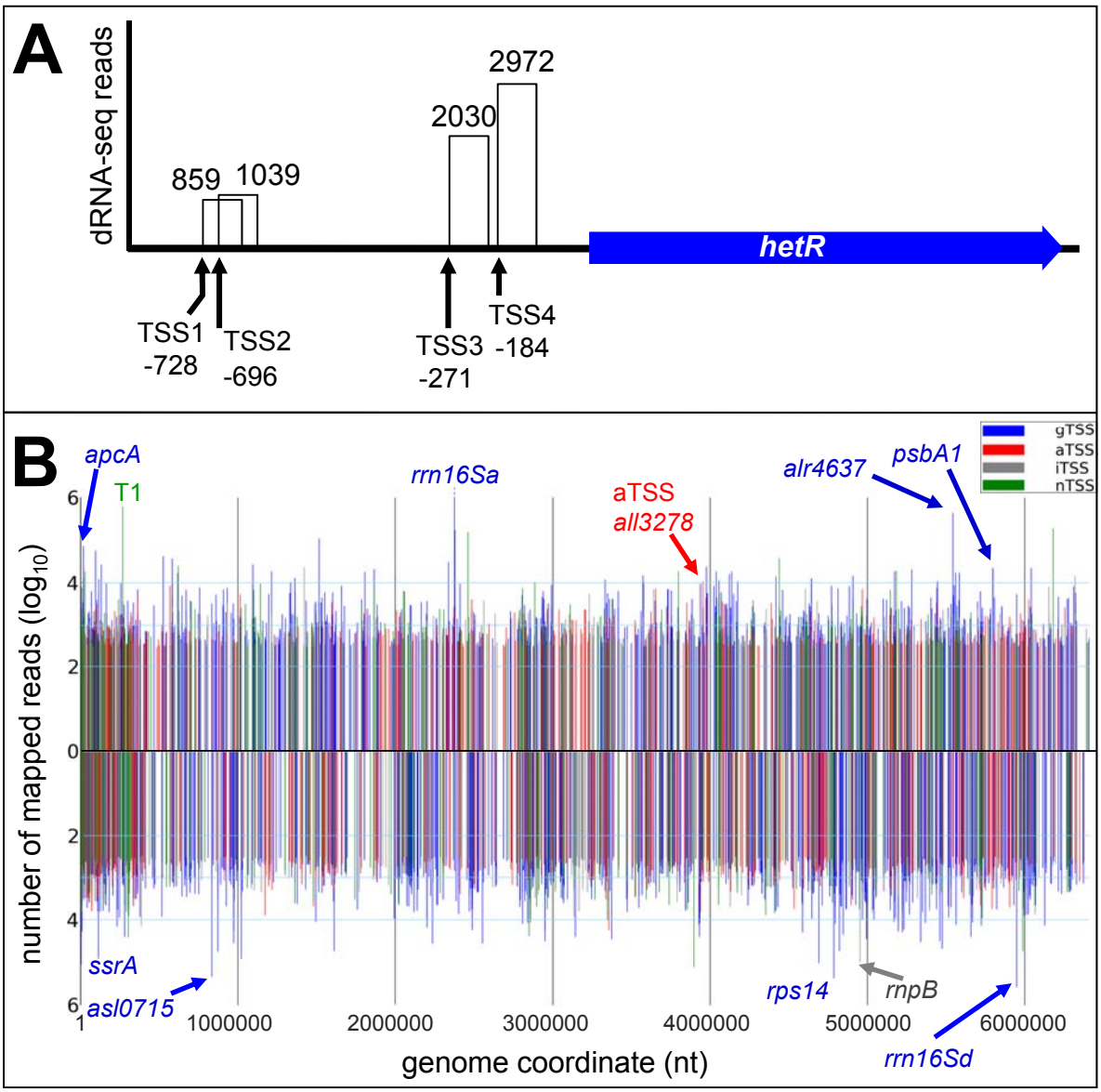

A
dRNA-seq reads
859
1039
2030
2972
hetR
TSS1
-728
TSS2
-696
TSS3
-271
TSS4
-184
B
gTSS
aTSS
iTSS
nTSS
apcA
T1
rrn16Sa
aTSS
all3278
alr4637
psbA1
ssrA
asl0715
rps14
rnpB
rrn16Sd
number of mapped reads (log10)
genome coordinate (nt)
1000000
2000000
3000000
4000000
5000000
6000000

FIGURE 2

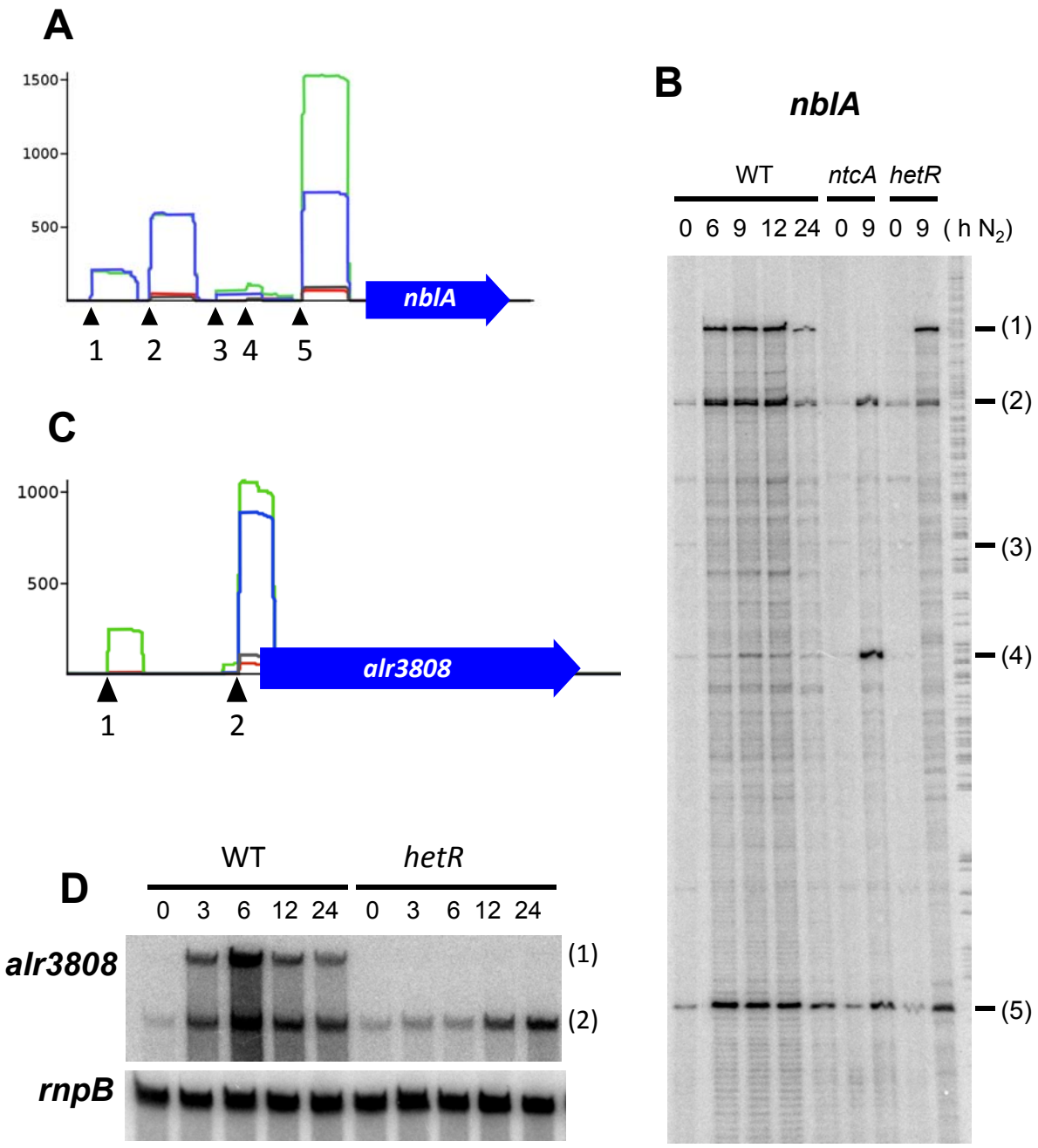

A
1500
1000
500
nblA
1
2
3
4
5
B
nblA
WT
ntcA
hetR
0 6 9 12 24 0 9 0 9 (h N2)
(1)
(2)
(3)
(4)
(5)
C
1000
500
alr3808
1
2
D
WT
hetR
0 3 6 12 24 0 3 6 12 24
alr3808
(1)
(2)
rnpB

FIGURE 3

A

4271758r	TTGATGCAATTTA	TCCGGA	AGACTGTAATTCAAAATAGAACAATTAATTG	NsiR1 (TSS12)	Table S2
2821366f	AGAGCAGATAAGT	TCCGGA	TAATAGGGAAAGTCCTTGTAGGTTACTTATTA	*hetR* (TSS3)	Table S2
2022661r	AATTAGAAAATCG	TCAGGA	AATTACTTATATACCCATGTAGATGTGACTA	*sigC*	Table S2
3427771f	AAAAAAATAATTT	TCCTGA	TGTTTTAAGAAAATTACTGTTGTTATAAATT	*hetC* (TSS2)	Table S2
3453831f	AGAATTAGGTTTA	TCCTGA	AAGAGTAAAAAAAATCCGAATATCCTAAATTA	*hepA* HEP island	Table S6
3449710f	AGCATTATTAGAG	TCCGGA	GAAATCACTTGGGATTATGAAAATATTCGTA	*alr2833* HEP island	Table S6
3457231f	GTTCCGATGTTCA	TACGGA	AAGCACCACAATTTAGCCGTAAGTATGTTTA	*alr2837* HEP island	Table S6
580704f	AGAGGTATTATTG	TCCGAA	TATTTGTCTTTCACTGCGAAAAAAATTATA	*asr0485 (pipX)*	Table S2
4601709f	TTCTGATAATTTT	TCCTGA	GAACACCATTATTTACAAGTAGAGTGTGATG	*alr3808*	Figure 2
3569154	GTTGTGCGCCCTT	CCCGGA	TTGTAGCGATCCGAGTAGAACCTGTTTTTA	*alr2933*	peptidoglycan turnover
5742628r	TCCCACAGCACCT	TCCGGA	AAATTAAAAAAACGCCGGAAAATATTGCA	*all4822*	cell envelope biogenesis
5953754f	TTACACAATTTTA	TCTGGA	TATATACCGCTTCAGGTGGAATAAATTTCTTA	*alr4984*	peptidoglycan-binding
5954131f	AAGTAGACGTTAT	TCCGGA	ATAATTGATTTTTTTCTGTCAAAGGAATTA	*alr4984*	peptidoglycan-binding
518833r	ATCCCCAATTCAA	TTCGGA	ATATTCTATAATCTGGAGAATAATAAATTATC	*all0438*	ser/thr kinase
3072950r	TTAAGAAAAAGTT	TGCGGA	TTTGAGTCCCAATGAGTGATCTAAGTTGTTA	*all2571*	unknown protein
2881051r	AGGTATTTCTGTA	TCCGGA	TAATTAAACTGCCAAGTGTAATTCTTAACA	*asl2397*	unknown protein

B

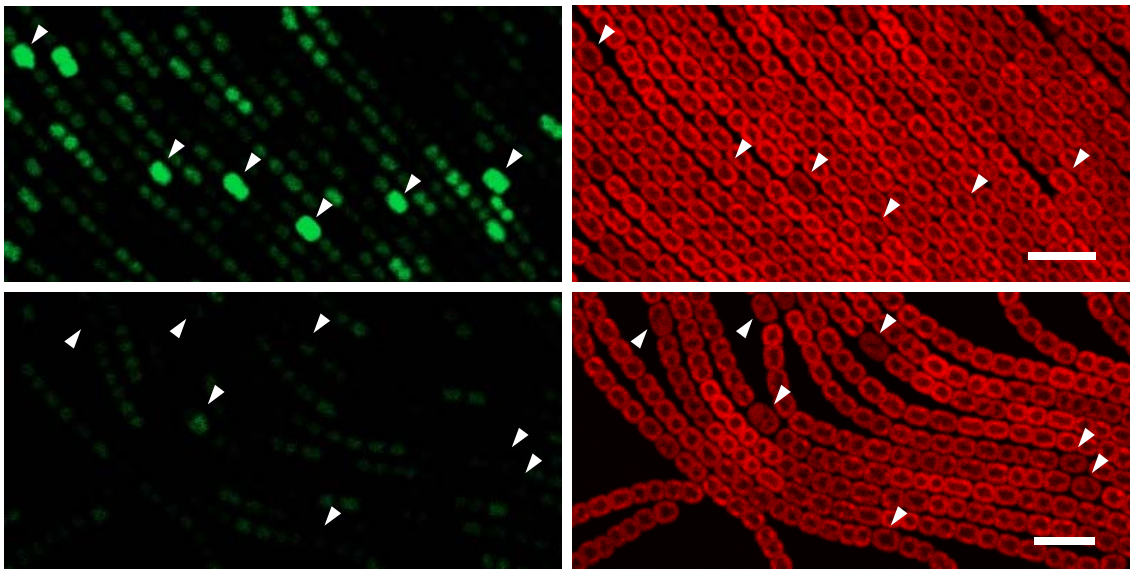


FIGURE 4

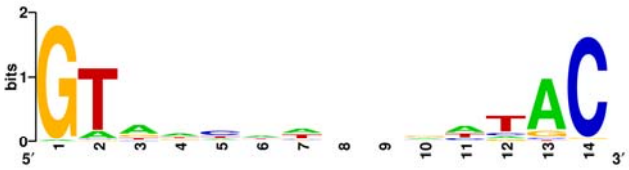
A

2058877f	AGTTTAT	GTAACCTATAAGAC	ATTTTATTTGATACCTCATACTC	TAAAAT	CAAGTA	nTSS
2460393r	TAAAAAA	GTAACTCTTGATAC	ATACGCTTATGAAAACCGCATA	TACCAT	TGAAAAAA	gTSS *asl2052*
580293f	TAAAACT	GTAGCAATGCAGAC	TGTTGTTAGGAACAGTTATTAG	GAGAAT	GCGCCTG	gTSS *asr0485 (pipX)* Table S2
1273249r	TCTTTTG	GGTACAAGATATAC	AAAATAATATTGAGGAATTAGGC	TATCTT	CATATCT	gTSS *all1087*
5547631f	GTTTTTT	GTTGCGTGCTAGAC	ATAACCAGACGGGTGTTTTGATC	CAAACT	CCTGTAA	aTSS *all4644*
2059119f	TTATTTT	GTATTTAACGGGAC	AGTTCTTACTTATCTAGTTAAGT	TTAAAT	AACAATCA	gTSS *alr1713*
2837125f	GTAGATA	GATATCCACAATAC	GGAAGTGTCAGTCTGATACTGG	CAGGCT	AAATTA	gTSS *alr2355*
5731963f	GTTTGTT	GGCGCAACGGCTAC	AGTTTGCTGGCGAGAGACAGGG	GATGAT	GGATTAG	aTSS *all4813*
4907756f	GCAAACT	GAATTGTTTGATAC	GGCAGGATGTGCAGTTTTCTCT	TACCCT	GAGCAAG	gTSS *alr4077*
1693413r	AAAAAAT	GTAATCACGCTGAC	AGAACTATCGTCTGATTAGGAGG	TATAAA	GTGATCA	gTSS *all1432 (hesA)* Table S2
3953418f	TGAGTTA	GTCGCTAAAGCTAC	ATTTTGGCTAACAGTATCCGACT	TATTAT	GAGATTTA	aTSS *all3278* Figure 5I
2400767r	GTTGCTC	GTATATTTCAACAC	GAATTTGATCATTTAGATGGTG	TACTGT	TTATAGA	gTSS *all2006* Table S6
519953f	ACATAAC	GTGTTTTCAGTTAC	AGTTATGCCAGATGCAATTAAGC	CACAAT	GTTGATTA	gTSS *alr0440*
105428r	CATTATG	GTATGAAATAGTAC	AGTTTAAAATTAGTGTTTGCGT	CATCAT	TACGAGA	gTSS *all7614*
1657401r	GAGAGTC	GTAGCATAACACAC	TAAAACTTCTGGAAACAGTAGGT	TAGGCT	TGCCTTA	gTSS *all1395*
3346518f	ATAAACT	GATAGTTATAATAC	TGTTCTCAGAAACGAAAAACTA	TATATT	GAGCATA	nTSS
5248514r	TGTTTTT	GCGATCGGCGATAC	AATTTACACGGGGCAAAAGCTG	GAATAT	GAAGGAA	iTSS *all4379*
5167792r	GGCTAGA	GTAACAAAGACTAC	AAAACCTTGGGCATGGGCTTGT	TACTTT	GAAATTCA	gTSS *all4312 (nrrA)* Table S2
5407066f	CTCAGCAATTTGTTCAACCTGAGCATTTTTCCCATTTGCAACTTGA			TACAAA	TATTTTTA	gTSS *asr4517 (nblA)* Table S6
2793917r	CTTCCTCAACTGCTCATACAGAGCAGATACGGTTAAAAAAAGTTGC			AATTCT	CATAAGTG	gTSS *all2319 (glnB)* Table S6

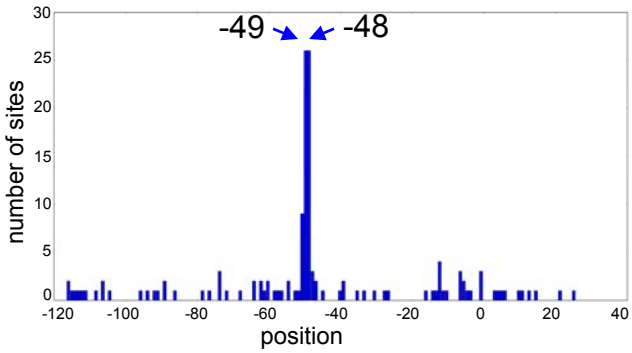
B

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	2	10	47	35	16	36	34	28	24	29	39	11	70	1
C	0	2	8	7	40	14	11	17	13	10	17	12	4	83
G	85	2	17	21	8	12	8	16	21	32	4	10	10	2
T	0	73	15	24	23	25	34	26	29	16	27	54	3	1

C



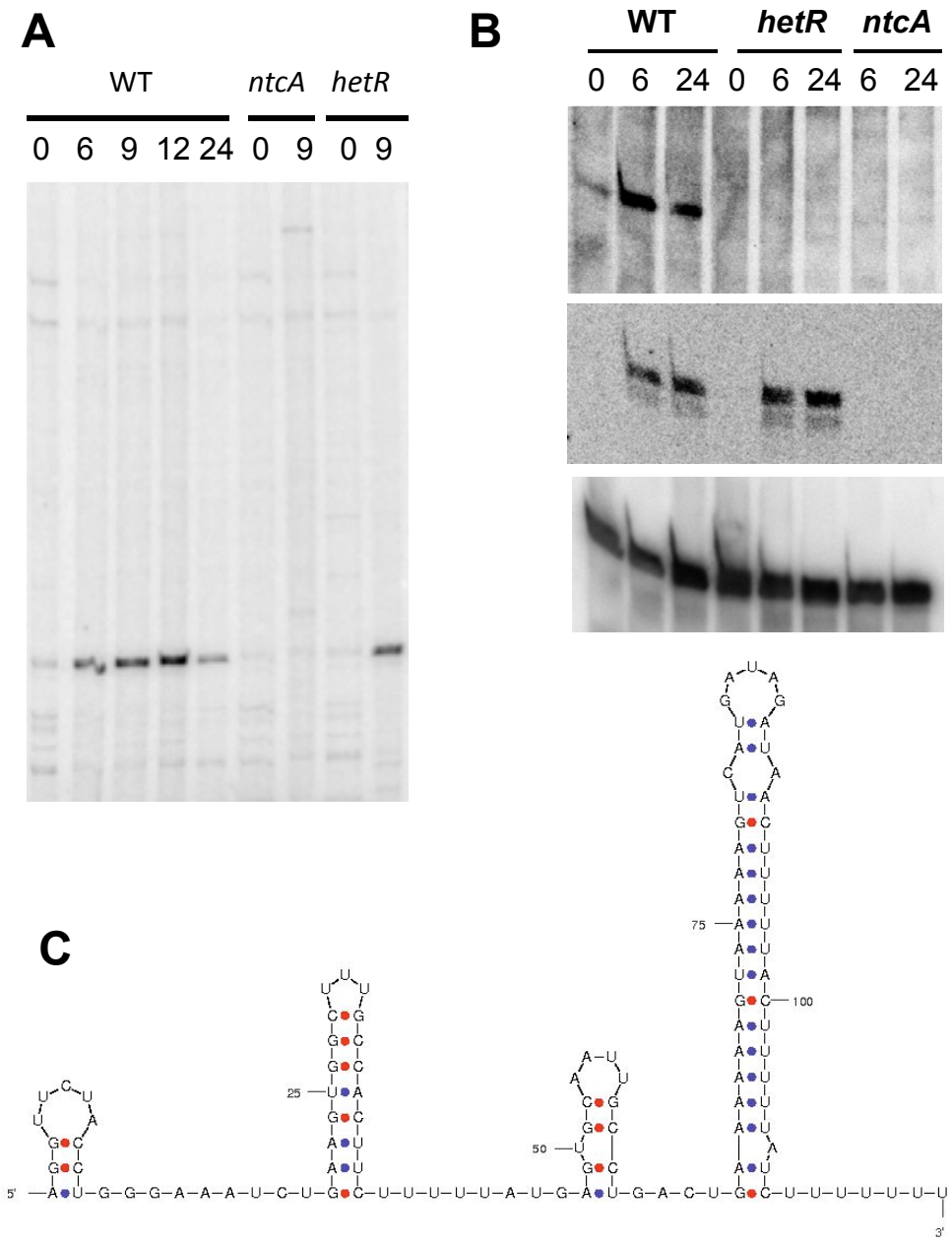
D



E

2828946f	GGATCAATACGTATTTCTATCTAGAAG	GAAAAT	ATAGAGACGAAATCAGCCCTAAACGGCACTAAAGATGAA	gTSS	*alr2346*
3359637f	TTTTGTATTTTCATATACAATTTTTGT	TATTTT	TAGATTAAGATGTTATTTACCTAAATTTAATCAACGTGT	gTSS	*asr2763*
5441031f	TTACTGGAAAAATTAGCGGTTTGCACA	TAGAAT	TATTAGAGACGGTAAAAATTTGTAAACTAGATTGACAAT	gTSS	*alr4548 (psbD)*
5046301r	TTGAGATTTTAGATATGGAATTAAATT	TACAAT	CTAAAATCGGAAAATAAAGTTTCCCTGACAAAGCCCAGG	gTSS	*all4203 (rpl5)*
2809313r	CCGTAGCATAAGATACAGAATTCTTGC	TATATT	AAATGTGTGAAGGGTCAAATCCAATTAATATCACTAGGA	gTSS	*asl2329 (gifA)*
3433648r	GTCTTTTAAAACACGCCCAATCACTGC	CATGAT	GGATTAGTTCATAGTGTTGTTGTTTGTTGTGATAGGCGG	nTSS	
4785763f	CCAGGCAATCATTGTCATCAGCCATGA	GAAAAT	AGCCTTGAGCGTAGCCTTGAATAATAGAGGCGATATCAT	gTSS	*alr3968*
1577153f	CTGTAACATACACTACGAAACTTATGC	TATGTT	AGGAAGAACAGACATAAAGCAGAAAAATTAAGAGGTTAA	gTSS	*asr1328*
2807328f	CTTTTGTGCAGATGTCGAAAGAAAGGT	TAATAT	TACCTGTAATCCAGACGTTCTGTAACAAAGACTACAAAA	gTSS	*alr2328 (glnA)*
1785466f	CAAAGAATAACTTATGCCATTTCTTGA	TATATT	GTGAGACAAGTTACAAATTACGTGGTGTGCAATTTTTTC	gTSS	*rbcL*

FIGURE 5


A
WT
ntcA
hetR
0 6 9 12 24 0 9 0 9
B
WT
hetR
ntcA
0 6 24 0 6 24 6 24
C
5'
3'
25
50
75
100