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

Supplementary Information

1. METHODS

Growth conditions. Cultures of *Anabaena* 7120 wild type, *hetR* mutant 216 (1) (bearing point mutation S179N) and *ntc*A mutant CSE2 (2) were grown photoautotrophically at 30^{o}C in BG11$_0$C (BG11 (3) lacking $NaNO_3$ and supplemented with 10 mM $NaHCO_3$) containing 6 mM NH_4Cl plus 12 mM *N*-tris (hydroxymethyl) methyl-2-aminoethanesulfonic acid (TES)-NaOH buffer (pH 7.5), bubbled with a mixture of CO_2 and air (1 % v/v). 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.

Fluorescence microscopy. Micrographs of filaments subjected to nitrogen step-down on plates were taken four days after plating. The accumulation of GFP was analyzed by laser confocal microscopy as described previously (4). GFP was excited at 488 nm by an argon ion laser, and the fluorescent emission was monitored by collection across windows of 500-540 nm (GFP imaging) and 630-700 nm (cyanobacterial autofluorescence).

Plasmid construction. The promoter region of NsiR1 (repeat 6; P6) was amplified using a 99-nt oligonucleotide (REPEAT for the wild-type version, REPEAT* for the mutated version; **Table S11**) and oligonucleotides Cla-repeat (containing a *Cla*I site) and repeat-Sal (containing a *Sal*I site) (**Table S11**). The resulting *Cla*I-*Sal*I fragments were cloned between the *Cla*I and *Xho*I sites of promoter-probe pCSAM201 (4), yielding pSAM264 (wild-type version) or pSAM269 (mutated version).

Preparation and analysis of RNA. Total RNA was isolated using hot phenol as described (5) with modifications. Frozen pellets corresponding to 50 ml of culture were suspended in 300 μl resuspension buffer (0.3 M sucrose, 10 mM sodium acetate, pH 4.5) and 100 μl 250 mM Na_2-EDTA (pH 8.0), 400 μl lysis buffer (2%

SDS, 10 mM sodium acetate, pH 4.5) and 1 ml of hot (65 °C) acid phenol were added. Samples were briefly vortexed and incubated at 65 °C for 2.5 min (this was repeated twice). The suspension was centrifuged at 12000 x g for 5 min and the upper phase was collected and sequentially extracted with hot phenol, hot phenol:chloroform (1:1) and chloroform. Finally, one volume of isopropanol was added and the RNA was precipitated in liquid nitrogen. The pellet was washed with 70% ethanol, dissolved in H_2O and treated with RNase-free DNase, according to the instructions of the Ambion DNA-free™ kit. Northern blot hybridization was carried out according to standard procedures (6) after the separation of RNAs in 1 % denaturing agarose gels, using ^{32}P labeled PCR fragments as probes. Alternatively, RNAs were separated in urea-polyacrylamide gels (7). In the case of small, putatively noncoding transcripts, ^{32}P-labelled oligonucleotides (**Table S11**) complementary to a region close to the 5´end were used as probes. Oligonucleotides were labeled using γ-^{32}P-ATP and polynucleotide kinase. PCR fragments were ^{32}P- labeled using a Ready-to-Go DNA-labeling kit (General Electric Healthcare). Hybridization to *rnpB* (8) or *trnL-UAA* was used as a loading and transfer control. Images of radioactive filters were obtained and analyzed using a Cyclone storage phosphor imaging system and OptiQuant image analysis software (Packard). Primer extension analysis of 5´ends was carried out as described previously (9).

Deep transcriptome sequencing. The cDNA libraries were prepared by vertis Biotechnologie AG, Germany (http://www.vertis-biotech.com/), and analyzed on an Illumina sequencer by Beckman Coulter Genomics, Danvers, MA, as previously described (10). In brief, total RNA was enriched for primary transcripts by treatment with Terminator™ 5'phosphate-dependent exonuclease (Epicentre), which degrades RNAs with a 5'P (processed RNAs) but not primary transcripts with a 5'PPP RNA (10). Then, 5'PPP RNA was cleaved enzymatically using tobacco acid

pyrophosphatase (TAP), a 5'-Solexa RNA linker oligonucleotide (**Table S11**) was ligated to the 5'P of the 'de-capped' RNA and 1st-strand cDNA synthesis started using random primers. This treatment reduced the amount of reads for the rRNA operons to 35.3 to 51.7% compared to ~90% previously observed for the same organism (11). The 5'-linkers possessed unique tetranucleotide tags for each of the four libraries (**Table S11**). The 2nd strand cDNA synthesis was primed with a biotinylated antisense 5'-Solexa primer, after which cDNA fragments were bound to streptavidin beads. Beads-bound cDNA was blunted and 3' ligated to a Solexa adapter. The cDNA fragments were amplified by 18 - 24 cycles of PCR. For Illumina sequencing (75 bp read length), the cDNAs were pooled in equal amounts and from the pool the cDNA in the size range of 150 - 450 bp was eluted from a preparative agarose gel. A total of 24.312.062 reads was obtained. Based on the tetranucleotide tags, 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 genome or one of the six large plasmids of *Anabaena* 7120.

Computational methods. Sequences of the *Anabaena* 7120 chromosome and plasmids, accession numbers BA000020, BA000019, AP003602, AP003603, AP003604, AP003605, AP003606, were downloaded on the 15th of October 2010 from the NCBI ftp-server. Sequence reads with blast hits to the ribosomal clusters were filtered out, as well as reads <18 nt. Remaining reads were mapped to the genome using the *segemehl* algorithm (12), with default parameters. Output was saved in gff format. Reads were binned within a section of 5 nt, whereby the start of such a window was determined by the first read occurring in 5‘ to 3‘ direction. All reads within the 5 nt window were considered to belong to one TSS. When≥2 reads were found, all reads within the window were considered to belong to a potential TSS

and the position from where the most reads started as the initial TSS. However, we noticed a few cases of initiation of transcription from a broader window. Therefore, this dataset was clustered, allowing the combination of initial TSS which were not more than 5 nt apart. The position from where the most reads started within this window was assigned as TSS as presented here. All reads starting within this final window would then be counted and given as number of reads associated with this TSS. For the final dataset, we set the minimum number of reads to 50.

Normalization and ratio calculation. Numbers of reads associated with a TSS were considered separately for the four different samples. To make the calculation of ratios between pairs of TSS possible for all TSS, single pseudocounts were added to the numbers of reads. If the read sum of the reference and the found matches in the samples reached 50 or more (including pseudocounts) the TSS was included in our investigation(**Table S1**). Normalization was performed according to Robinson and Oshlack 2010 (TMM normalization (13)). The TMM factor calculation was implemented into the TSS annotation pipeline using python 2.6 and used with library size normalization to million base pairs. Trimming for the TMM factor was done with 5% for the absolute intensity (A) and 30% for the log-fold-ratios (M). Ratio values were calculated as $\log_2$ fold changes:

R1) $$\log_2 \frac{sam(y)/sam(N)}{ref(y)/ref(N)},$$

where y is the number of reads for a chosen TSS and N the effective library size of the respective set (ref=reference, sam=sample). The sample WT-0 was defined as reference to which the other three samples were compared. The resulting ratios were classified and filtered into DEF and DIF categories of regulated promoters as outlined below.

Filters for classification into DEF and DIF categories of regulated TSS.

A: Reads WT-0

C: Reads WT-8

G: Reads *hetR*-8

T: Reads *hetR*-0

abs(X) = absolute read value of X

$\log_2 ratioX$ = normalized M-value between reference and X according to equation R1

General rules were applied to every category filter. In addition, every category had specific rules, which were only applied to that category filter. The category filter is a boolean expression. If a TSS fulfils the conditions of the category filter, it is considered to be part of that category. The variable x sets different thresholds for the DIF/DEF categories (x=3 was used in **Tables S3, S4, S5, S8** and **S9**; x=1 in the analysis leading to **Fig. S5*B***).

DIF category

General Rules

1) $-1.0 \le \log_2 ratioG \le 1.0$
2) $abs(A-C) \le 10$
3) $abs(\log_2 ratioC - \log_2 ratioG) \ge 3.0$
4) $abs(C-G) \ge 50$
5) $abs(\log_2 ratioC - \log_2 ratioT) \ge 3.0$
6) $abs(C-T) \ge 50$
7) $abs(A-T) \ge 10$

DIF+ class rules

1) $\log_2 ratioC \ge x$
2) $\log_2 ratioC > \log_2 ratioG$
3) $\log_2 ratioC > \log_2 ratioT$

DIF+ class filter:

[(1 OR 2) OR (3 AND 4)] AND [(5 AND 6) OR 7]
AND 8 AND 9 AND 10

DEF category

General Rules

1) $-2.0 \le \log_2 ratioT \le 2.0$

DEF+ class rules

2) $\log_2 ratioC \ge x$
3) $\log_2 ratioG \ge x$
4) $\log_2 ratioC - \log_2 ratioT \ge 2.0$
5) $\log_2 ratioG - \log_2 ratioT \ge 2.0$

DEF+ class filter:

[1 OR (4 AND 5)] AND 2 AND 3

DEF- class rules

2) $\log_2 ratioC \le x$
3) $\log_2 ratioG \le x$
4) $\log_2 ratioC - \log_2 ratioT \le -2.0$
5) $\log_2 ratioG - \log_2 ratioT \le -2.0$

DEF- class filter:

[1 OR (4 AND 5)] AND 2 AND 3

Determination of NtcA binding sites and -10 elements. Possible -10 elements were searched at positions 6-8 nucleotides upstream of all characterized TSS and scored according to a specific PSSM derived from this dataset (**SI Dataset S1, Table S1**). 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 nucleotides of the $GTAN_8TAC$ motif centered at position 22/23 upstream of the -10 element, and hence were considered useful for construction of the PSSM. Six additional experimentally defined sites included in **Table S2** but not among the 81 sites above were also used in the construction of the matrix.

2. SUPPLEMENTARY DATA FILES

http://www.cyanolab.de/Supplementary7120.html

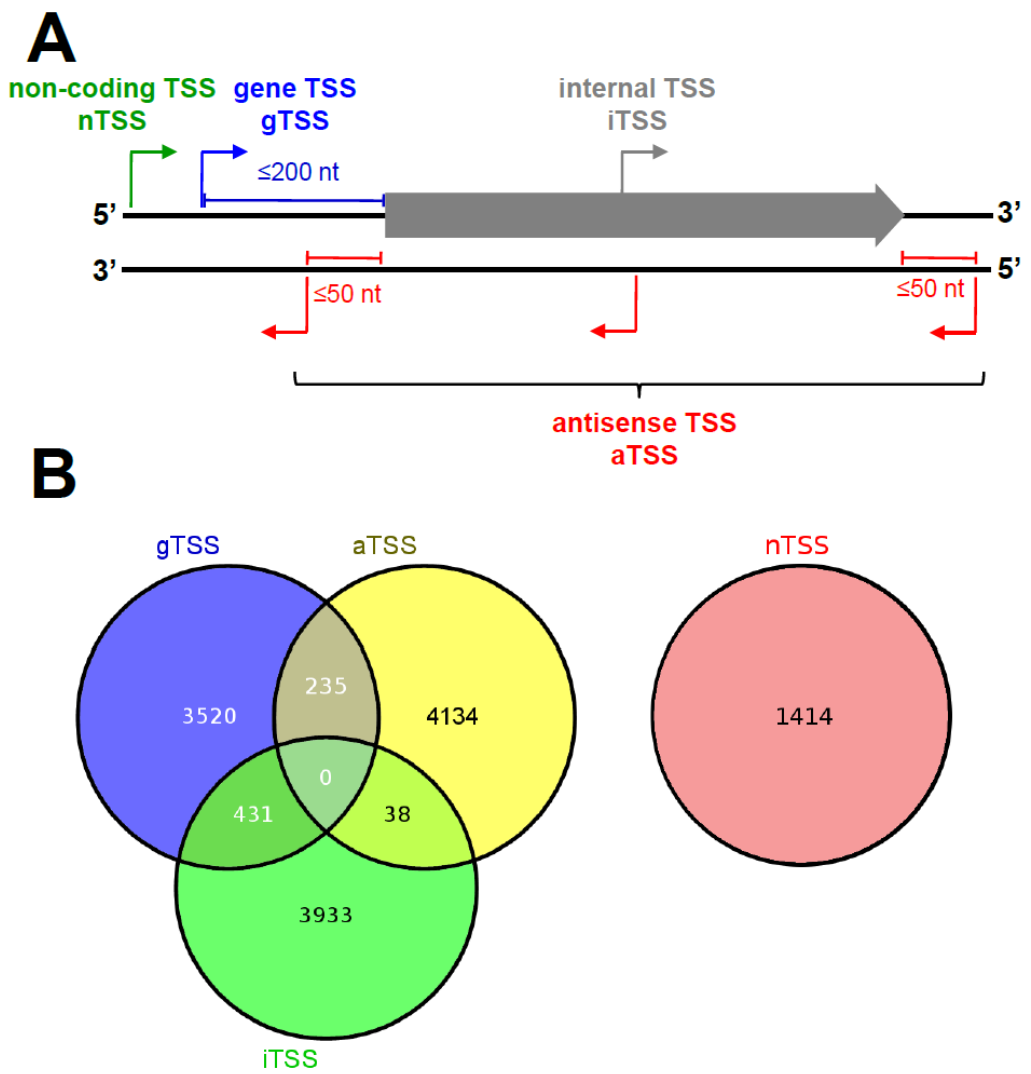
Supplementary data file 1 “Ana7120TSS.gbk“. Genbank file of the *Anabaena* 7120 chromosome with all annotated TSS.

Supplementary data files 2-7 “alpha.gbk“ to "zeta.gbk". Genbank files of *Anabaena* 7120 plasmids with all annotated TSS.

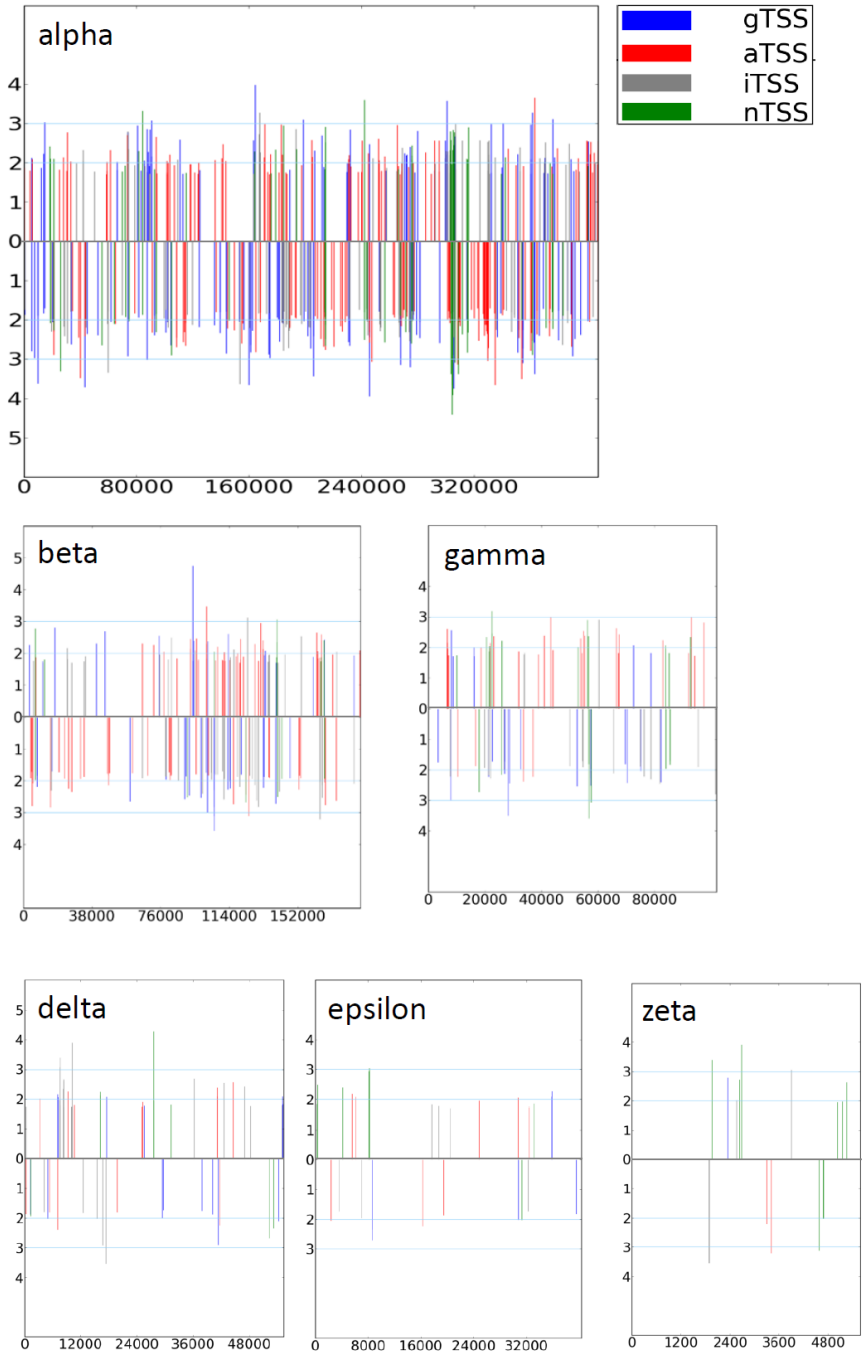
Supplementary data file 8 "Ana7120 tRNAs". The tRNA transcriptome of *Anabaena* 7120.

3. SUPPLEMENTARY FIGURES

Supplementary Figure S1. Details of annotation and classification of the TSS dataset into gTSS giving rise to mRNA, aTSS producing asRNA, iTSS for internal sense transcripts and nTSS for candidate ncRNAs. (*A*) A TSS was classified as gTSS if the TSS was located 0-200 nt upstream of a gene. TSS located antisense or ≤50 bp 5' or 3' of an annotated gene were classified as antisense (aTSS). TSS positioned within a coding sequence were classified as internal TSS (iTSS). TSS not entering any of the previous categories were considered putative ncRNA starts (nTSS). (*B*) Overlaps between the different categories of TSS. Some TSS associate with multiple categories according to **Fig. S1*A***. Thus, 235 of the 4186 gTSS were actually located antisense and 431 were in sense orientation within another annotated gene.

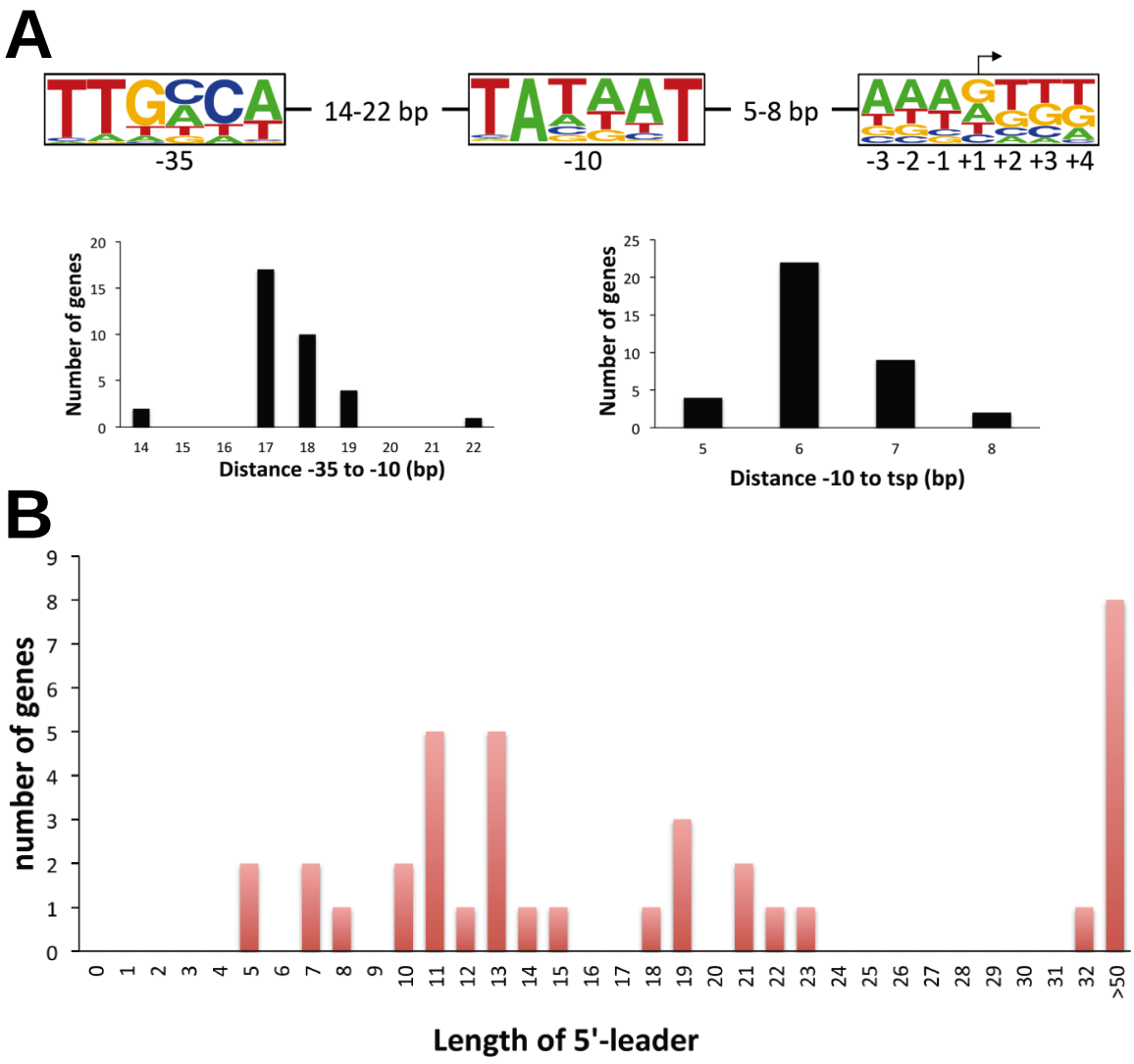


Supplementary Figure S2. Occurrence of TSS along linear plots of the six *Anabaena* 7120 plasmids. Each plasmid is drawn along the x-axis and its length indicated in nucleotides. Mapped TSSs for the forward strand are plotted above the main axis and for the reverse strand below. The number of sequence reads is given on the y-axis (log_{10} scale).The location of each of the TSS according to **Fig. S1*A*** served for its classification, color-coded as indicated in the legend.

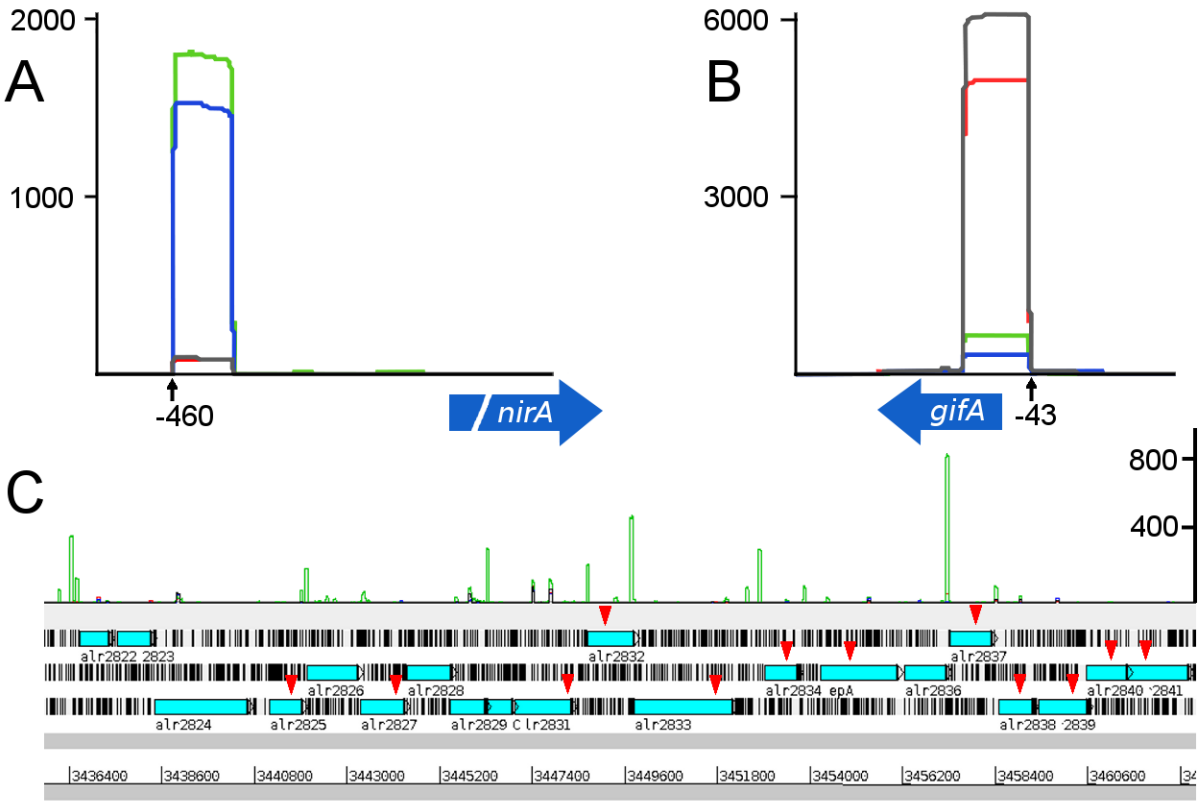


Supplementary Figure S3. Characterization of the tRNA transcriptome.

(*A*) Promoter sequence of tRNA genes based on a total of 37 TSS identified for 35 tRNA genes. Consensus -35, -10 and initiation site are represented with a Weblogo. Letter height is proportional to frequency. From the aligned TSS a consensus -10 element was found located 5-8 bp upstream of TSS. A -35 sequence is found in most tRNA genes at 17-19 bp from the -10 sequence. (*B*) Length of pre-tRNA 5'leader sequences. Length was estimated from the distance between the TSS and the 5'-end of mature tRNA.



Supplementary Figure S4. Paradigm promoters in the different categories defined. Reads associated to several genes with known N- regulation corresponding to the DEF+ (*nirA*, (*A*)), DEF- (*gifA*, (*B*)) or DIF+ (genes in the HEP biosynthesis island, (*C*)) categories (14, 15, 16, 17). The histograms correspond to the WT-0 (red), WT-8 (green), *hetR*-0 (black) and *hetR*-8 (blue) samples. Red triangles in part (*C*) indicate transposon insertions leading to *fox*$^-$ phenotype (16, 17). Scales indicate read numbers.



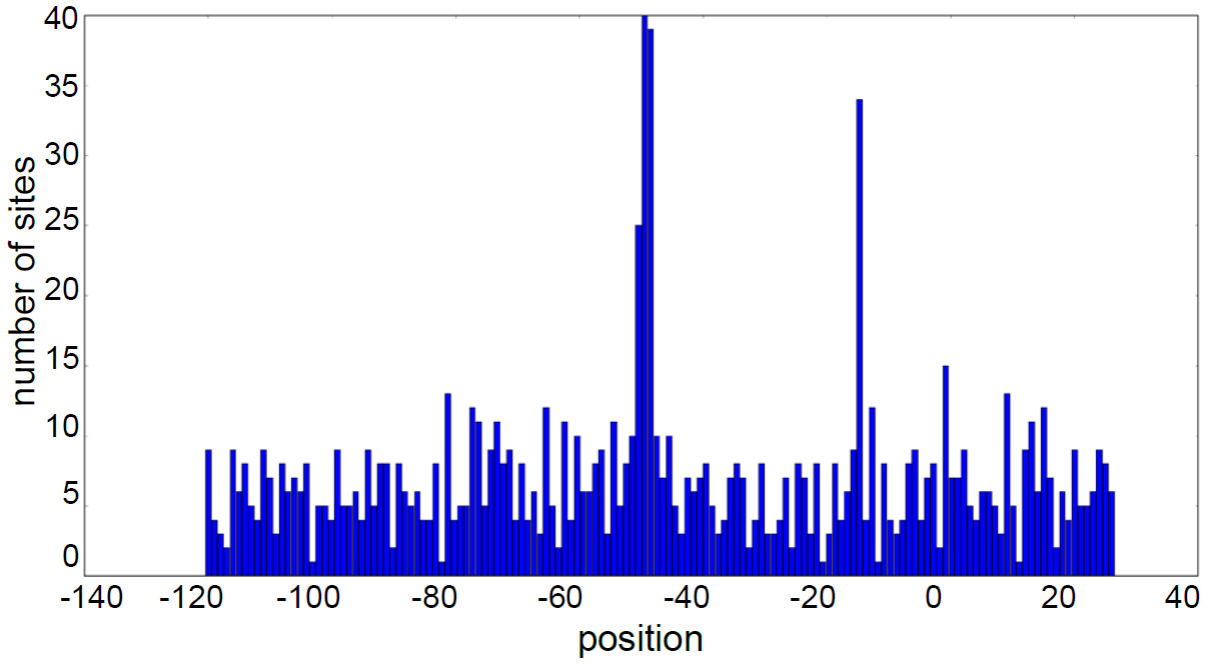
Supplementary Figure S5. (*A*) PSSM obtained from an alignment of sequences for the NtcA binding pattern (positions –49 to –36 with regard to the respective TSS) of 81putatively NtcA-regulated TSS according to **Table S3** and six additional TSS experimentally characterized before.

(*B*) Positions (first nucleotide) of NtcA binding sites identified within the promoters of 965 TSS in the DEF+ category (fold change ≥2, PSSM score≥5.0) in a sliding window approach using the PSSM from part (*A*). The bars indicate the first nucleotide of a putative NtcA binding site and its position with regard to the TSS at position +1.

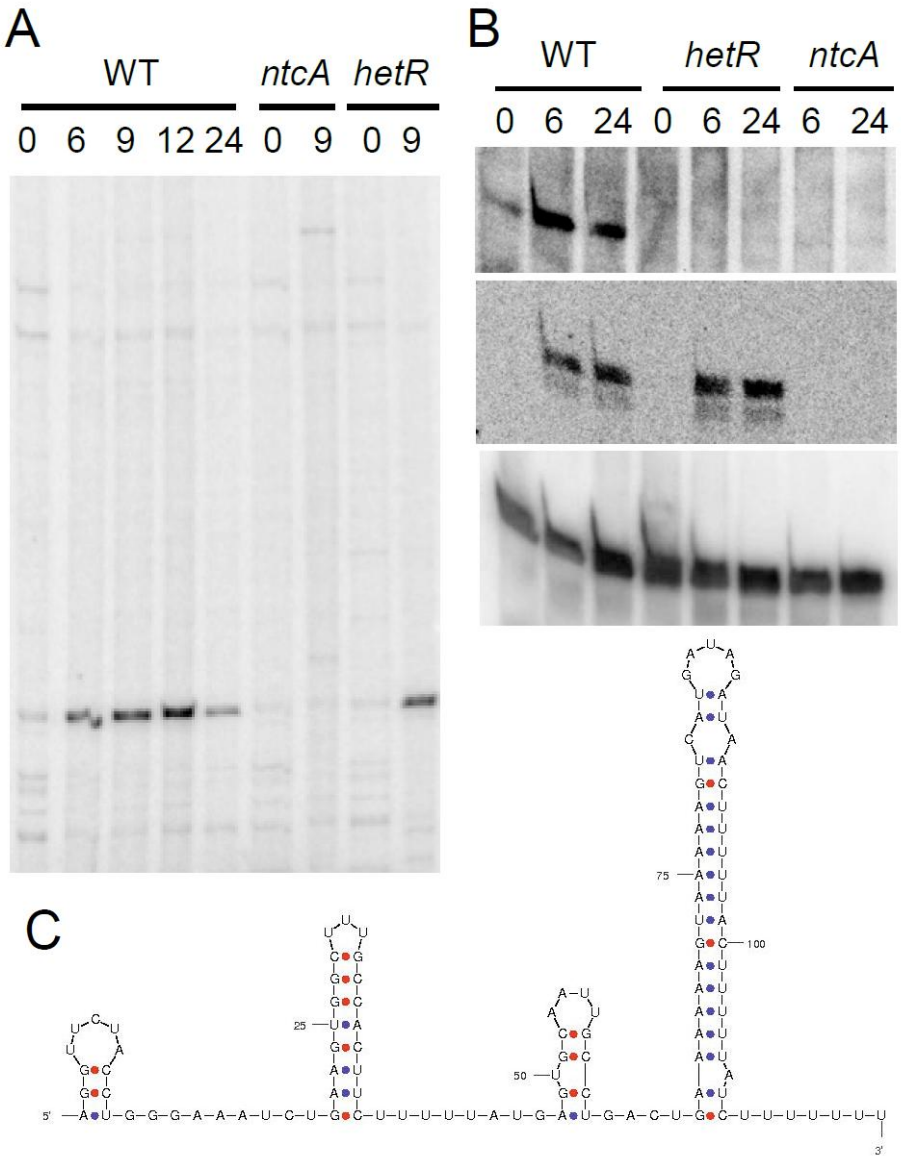
A

	-49	-48	-47	-46	-45	-44	-43	-42	-41	-40	-39	-38	-37	-36
A	-3.16	-1.29	0.84	0.42	-0.66	0.46	0.38	0.11	-0.10	0.16	0.58	-1.16	1.40	-3.75
C	-4.22	-2.64	-1.05	-1.22	1.14	-0.31	-0.64	-0.05	-0.41	-0.76	-0.05	-0.52	-1.90	2.17
G	2.20	-2.64	-0.52	0.24	-1.05	-0.52	-1.05	-0.13	0.24	0.82	-1.90	-0.76	-0.76	-2.64
T	-4.75	1.46	-0.75	-0.10	-0.16	-0.05	0.38	0.01	0.16	-0.66	0.06	1.03	-2.75	-3.75

B



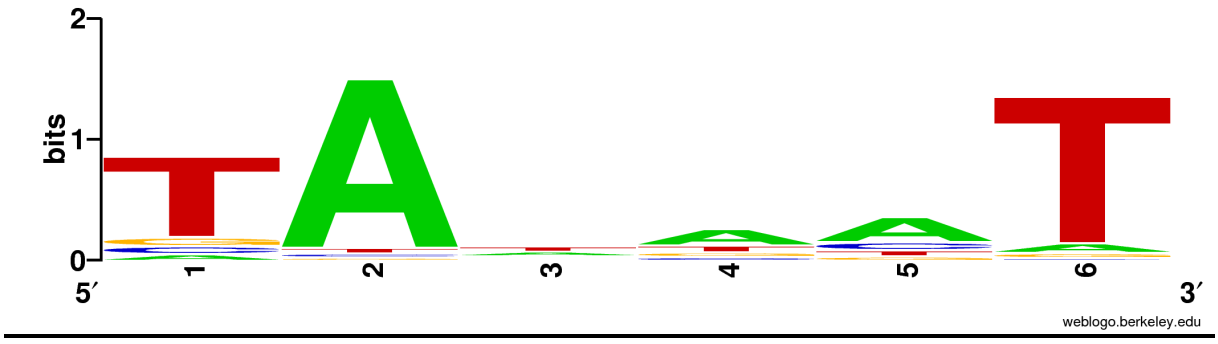
Supplementary Figure S6. Experimental verification of newly identified transcripts classified as antisense or noncoding. RNA was isolated from ammonium-grown cells (lanes labeled 0) or from cells incubated in the absence of combined nitrogen for the number of hours indicated. (*A*) Primer extension analysis 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 (*Top*) and NsiR3 (*Middle*) in *Anabaena* 7120 WT and mutant strains CSE2 (*ntcA*) and 216 (*hetR*). Transcription from position 3141905r (DIF+) produces NsiR2, whereas transcription from position 5452083f (DEF+) produces NsiR3. The samples contained 10 µg of RNA. The *trnL-UAA* transcript (*Bottom*) was used as a loading control. (*C*) Predicted secondary structure of NsiR3, typical for a regulatory RNA.



Supplementary Figure S7. The -10 element of *Anabaena* 7120. Alignment of sequences for the -10 elements (positions –12 to –7 with regard to the respective TSS) of 13705 experimentally determined putative TSS (**Table S1**) yields the numbers of nucleotides for each position (top), the PSSM (middle) and the resulting sequence logo (bottom).

	-12	-11	-10	-9	-8	-7
A	891	12702	4516	7162	7898	827
C	1060	320	2138	1624	2513	231
G	1179	238	1946	1967	1402	413
T	10575	445	5105	2952	1892	12234

	-12	-11	-10	-9	-8	-7
A	-2.18	1.65	0.16	0.82	0.97	-2.29
C	-1.41	-3.13	-0.39	-0.79	-0.16	-3.60
G	-1.25	-3.56	-0.53	-0.51	-1.00	-2.76
T	1.39	-3.18	0.34	-0.45	-1.10	1.60



4. SUPPLEMENTARY TABLES

Suppl. Table S1. Overview on all putative TSS mapped for the chromosome and plasmids alpha, beta, gamma, delta, epsilon and zeta. The TSS are classified as gTSS, aTSS, iTSS or nTSS according to the scheme in **Fig. S1*A***. The annotation gives the respective gene name or systematic ID, or the TSS position in the genome when both were not available. Sequences in the window -13 to -6 were searched for a possible -10 element and the PSSM score for the hexanucleotide motif scoring highest is given.

See "Suppl. Table S1.xlsx" within the *Dataset S1*

Suppl. Table S2. Comparison of dRNAseq-anchored TSS with a set of selected previously determined TSS of *Anabaena* 7120. The list of genes is in alphabetical order. If the TSS was found, the number of reads (pooled from all four samples) is given (b.t., reads are present but below threshold of 50 reads; n.d., not detected). Gene names and references are given for previously determined TSS, followed by the systematic ID, the TSS position (absolute coordinates; if our mapping differed by a few nt, this position is indicated by the number in brackets), the number of reads, orientation (S; f, forward, r, reverse strand), citation (R) and sequence in a window of 51 nt. If more than one TSS was found, the most distally located was numbered TSS1. The respective -10 elements (or extended -10 elements) and TSS as proposed by the authors are underlined and boldface. TSS and -10 elements as defined by us are highlighted in light gray. Previously proposed NtcA binding sites are double underlined, and commented on if experimentally verified. A, alternative TSS for this gene exist which were not previously identified; DIF or DEF indicate the mode of regulation if applicable (see main text).

Gene	ID	Position	dRNA-seq	S	R	Sequence	Comment
apcF	all2327	2807239 (2807238)	(2159)	r	(18)	gggagaaagttaatctcataggggtcaactttgaca**taaaat**cct**taag**T**A**	**DEF-;** overlaps with *glnA* TSS2 (-270; RNA_{IV})
alr0750	alr0750	871876 (871877)	b.t. (21 reads)	f	(19)	atgatttcccctcacaagttcctcaaattattctcc**tataa**acaataga**T**A	
alr1690	alr1690	2018583	714	f	(20)	ctgaaagactattagaggcaaaccaacgcggtttatgg**gaaaat**gtgaat**A**	Readthrough from this TSS results in an asRNA to *furA* (20)
argC	all2498	2999383	156	r	(21)	atcgatttcggataatttatttttggtgttccctaagg**taaaat**tcagt**A**	
argD TSS1	alr1080	1263707	n.d.	f	(21)	ccccggccattaactctttttggatgcgttggtatttctgttcaagagtt**G**	
argD TSS2	alr1080	1263762 (1263763)	(168)	f	(21)	aaaattattttccactattgccacaggtttgaggatata**taatgt**taag**C**T	+A
atp1	all0011	9641	18432	r	(22)	taattgttttttctaaagtattgttaccggatcgtga**tatgat**tcagct**G**	+A
asr0689	asr0689	796925 (796926)	b.t. (16 reads)	f	(23)	ggttattagatgagctaatttgattgacaatcacactc**tagaat**ttata**T**A	+A
atpB TSS1	all5039	6011693	1776	r	(24)	gcagttctctgatgttgtaagtcttaagaatcaatttat**gatata**agtT**G**	+A
atpB TSS2	all5039	6011596 (6011595)	(7483)	r	(24)	aaggaattactaacggaactctgtccccctccggtgt**tagagt**gaaaga**T**G	+A
calA TSS1	alr0946	1099489	6626	f	(25)	atctgtaagttagatttaaaaataaacaactattt**tgctat**tattagtcc**A**	

calA TSS2	alr0946	1100046	n.d.	f	(25)	ccagtaactttatgcaga**tgctaaaaa**gtcaaatactggcatgagacaac**A**	
cnaT	all0601	700173	1097	r	(26)	gtgtctgtcgcattatttttttgaaattgctataaggta**tacgat**tagtca**A**	
conR	all0187	202508	232	r	(27)	accagccaggattactgggaacaatgggctactcatgttgtcatacacac**A**	+A
coxB	alr2514	3021024	185	f	(28)	tatctggaataattcccccaatttttataggca**taagtt**aaaaagttaat**A**	**DIF+**
cox3	alr2729	3327074 (3327077)	(166)	f	(29)	tgttacaattacttttgacaagtactaaaaaaatcct**taagct**aataGatA	**DEF+**
cpcB	alr0528	620574 (620581)	(24732)	f	(30)	tcacaatttgtaacaaaataaggatctatagca**ttgtat**aaac**A**taagctG	**DEF-**
cpcD	asr0531	622781 (622810)	n.d. (580)	f	(30)	cgcgtctatcgac**tagaag**tt**A**ccggaatccgtagcc**ctggct**accccagT	iTSS in *cpcC*
cphA1 TSS1 (-116)	all3879	4681771 (4681796)	n.d. (260)	r	(31)	aaaaaccagagggttagactgaattGatctttaatttat**tttcct**gctgta**G**	TSS reported at pos. -89, -116, -141, -191 and -230 (31)
cphA1 TSS2 (-32)	all3879	4681687	1228	r	this study	aactgtttttcatgaacagtttagcgatcaattccagtaagaaactagaaT	**DEF+**
cphBA1	all3880	4683108 (4683126) (4683096)	n.d. (725) (205)	r	(31)	atttatgttaaatattgtGacaac**taccat**ttaaata**A**ataatctaaccA	TSS for *cphB1* reported at pos. - 339, -357 and -499 (31)
devBCA TSS1 (-704)	alr3710	4478166/69	b.t. (26 reads)	f	(32)	att<u>tgtacagtctgttacc</u>tttacctgaaacagatgaatg**tagaat**ttat**A**	Function of NtcA binding site demonstrated by band shifts and lack of activation in *ntcA* mutant (32)
*devBCA*TSS2 (-454)	alr3710	4478416	559	f	(33)	agtcatctaagttgcagaaaatctttaaaagatttgc**cataat**gaaatct**A**	**DIF+**
devH TSS1 (-406)	alr3952	4769438 (4769441)	865	f	(34)	tctacctttatatagaggaatttttgctttt**tattgt**gctttgctgtt**G**ctG	
devH TSS2 (-157)	alr3952	4769698 (4769700)	8380	f	(34)	ttcgatgaataatatgtaaactttactaccatcattg**tacagt**aggaa**G**tA	
devH TSS3 (-136)	alr3952	4769709 (4769713)	(1072)	f	(34)	atgtaaactttactaccatcattg**tacagt**aggaagtataacaaaa**T**agtA	
devH TSS4 (-112)	alr3952	4769731	n.d.	f	(34)	tcattgtacagtaggaagtataacaaaatagtagagcgcaaatagccctc**T**	
furA	all1691	2020552	3932	r	(20, 35)	taagcaatgtttgttgcaaatcactccaaatatttgt**tatatt**ctcaatt**A**	
gifA	asl2329	2809313	12036	r	(15)	aatcttcaattcc<u>gtagcataagatac</u>agaattcttgc**tatatt**aaatgt**G**	**DEF-;**function of NtcA binding site shown by derepression in *ntcA* mutant (15)
*glnA*TSS1 (-319; RNA_V)	alr2328	2807164 (2807166)	208	f	(36)	tcgcgcattccttcttctcccaatctttgctatc**tatggt**ttgatatt**A**aT	
glnA TSS2	alr2328	2807208 (2807212)	7400	f	(2, 36,	ttaatttggttgcactacgcaccagtaaatttttg**tgctattaa**aa**A**tagA	

(-270; RNA_{IV})					37)		
glnA TSS3 (-240)	alr2328	2807243	314	f	this study	tttgtgctattaaaaatagattcggcacaaaaacaatctatctgttacttA	
glnA TSS4 (-197; RNA_{III})	alr2328	2807286	n.d.	f	(2, 36, 37)	ctgttacaaggattttatgtcaaagttgacccctatgagattaactttct**C**	Discussed as a processed product (2, 37)
glnA TSS5(-155; RNA_{II})	alr2328	2807326 (2807328)	1209	f	(2, 36, 37)	aaagaaaggt**taatat**tacct**G**tAatccagacgttctgtaacaaagactac	Binding of NtcA to this site activates expression of RNA_I and represses expression of RNA_{II} (36).
glnA TSS6 (-93; RNA_I)	alr2328	2807390 (2807392)	(60)	f	(2, 36, 37)	gtaacaaagactacaaaactgtctaatgtttagaatc**tacgat**atttc**A**gG	Function of NtcA binding site shown by band shifts and lack of activation in *ntcA* mutant (36)
gnd TSS1	alr5275	6294305 (6294306)	(611)	f	(18)	gtgttactagtgagcttttattctttaaattaactct**taatgt**cttctt**T**A	**DEF+;** +A
gnd TSS2	alr5275	6294342 (6294343)	447	f	(18)	cttctttaggggagaagagaaaaactcggtc**taccct**taaaga**T**gaaagaA	Additional gTSS at pos. 6294275, 6294294, 6294349
gor TSS1	all4968	(5931866)	92	r	this study	actctctcctcactgacccagaagtattagaaactcgttaccctgtacttG	iTSS in all4969
gor TSS2	all4968	5931553 (5931564)	(122)	r	(38)	aacccctggaggcggtggatttggttcaaaatgaaaggt**A**a**tag**gggagg**A**	
gor TSS3	all4968	5931496	217	r	(38)	gactgttgacaactgacaattgacaaataacaactgttacattaaatatc**A**	**DEF-**
hanA, hupB	asr3935	4752810 (4752807)	(1309)	f	(39, 40)	ggtttcatctatccacttctggctacaacaccttagaat**gattca**ttGaa**A**	
hesA	all1432	1693413	329	r	(41)	tgtaatcacgctgacagaactatcgtctgattaggagg**tataaa**gt**G**atcA	**DEF+; Fig. 4A**
hetC TSS1 (-571)	alr2817	3427493/ 3427492	b.t. (17 reads)	f	(9, 42)	tctgtaacatgagatacacaatagcatttatatttgctt**tagtat**ctct**CT**	**DEF+**; +A; Function of NtcA binding site shown by band shifts and lack of activation in *ntcA* mutant(9)
hetC TSS2 (-293)	alr2817	3427771	b.t. (22 reads)	f	(33)	aaaaaaataattttcctgatgttttaagaaaattactgttgttataaatt**A**	**DIF+**; +A**; Fig. 3A**
hetF (-403)	alr3546	4273165 (4273166)	343	f	(4)	tgtcttagggaattggtaaattttccctgcccaagacag**tagaat**aatct**A**	
hetP (-557)	alr2818	3431807	72	f	this study	attgagtaataaaattctgataacgggtgacattcatgaaaaaagatagtA	**DIF+;** 5′-ends at positions -727, -545, -208, -177 and -12 bp reported (43)
hetP (-545)	alr2818	3431819	b.t. (25 reads)	f	(43)	aattctgataacgggtgacattcatgaaaaaagatagtataaatactaat**A**	**DIF+;**
hetR TSS1 (-728)	alr2339	2820911	859	f	(42, 44-46)	tttagcaaaagaaaagttaagagtagattaattcatgcaatttaataat**A**	**DEF+**
hetR TSS2 (-696)	alr2339	2820942	1039	f	(42, 44-46)	ttcatgcaattttaataatatgacagataaatatagaa**tattat**gttgtg**A**	

hetR TSS3 (-271)	alr2339	2821366	2030	f	(42, 44-46)	agagcagataagttccggataatagggaaagtccttg**taggtt**acttatt**A**	**DIF+; Fig. 3A**
hetR TSS4 (-184)	alr2339	2821454	2972	f	(42, 44-46)	aatgaggaaattctggcagttattttgtgatttttcgg**taagat**acaagc**A**	
hetZ	alr0099	103033	575	f	(47)	gggtctagcccagcaggtgggattagagaaacatcctg**tatggt**agggaa**A**	**DIF+**
hoxU	alr0762	883760	b.t. (23 reads)	f	(19)	tttgatcaaccttcacaagttcctcaaactcttcacc**tataa**ctcaaggt**G**	
hoxW TSS2 (-44)	all0770	893532	n.d.	r	(48)	atttttctcttacgttaaaagcgggacgatagaaaaaacctcccaaaagc**G**	
hoxW TSS1 (-70)	all0770	893558	n.d.	r	(48)	tctttgtttcctcacgacttcctcaaatttttctct**tacgtt**aaaagcgg**G**	
hupW TSS1	alr1423	1686107 (1686109)	b.t. (48 reads)	f	(48)	cagttaaacatctaccgaatttactggtatttttatgg**tatcat**cact**C**aA	iTSS in alr1422
hupW TSS2	alr1423	1686708	b.t. (12 reads)	f	(48)	gataagaaggcttttaaaatatcctcttacatactcc**tataat**aacaact**G**	
hypF	alr0694	800230	325	f	(23)	attggttaatggtggcggatactgctaattctcggt**tgttaggat**ggagg**A**	iTSS in alr0693
hypC (hupC)	asr0695	802529 (802573)	(920)	f	(23)	aaacag**C**aatttatccctttaagtttacctttccgataatatttattgtA	iTSS in *hypF*
invB	alr0819	942823	236	f	(49)	gtaattattcatattgcatattttcacaaaaagtta**taaatt**attaattt**G**	+A
invB	alr0819	942989 (942986)	128	f	this study	tagttccagaaacaataatggaacgacagataatctaaaaatatatcTtt**A**	**DIF+**
invA	alr1521	1781512	b.t.(17 reads)	f	(49)	actaactcccgacgacttaataatcaaaaattctatga**tagggt**aaagag**G**	
nifB	all1517	1778380 (1778379)	(58)	r	(50)	ttccagagatacaaactgtaatagtcaaagacgactgc**tataaa**tctta**T**A	
nirA (-460)	alr0607	704217 (704218)	3563	f	(14, 26, 51, 52)	tgtagctacttatactattttacctgagatcccgaca**taacct**tagaag**T**A	**DEF+;**Function of NtcA binding site shown by lack of activation in *ntcA* mutant and DNase footprinting (14, 51)
nrrA TSS1 (-93)	all4312	5167860	241	r	this study	aacattttgccaacataggttataaaaaaacgtAgaggtaattgtggctag	
nrrA TSS2	all4312	5167793 (5167792)	10919	r	(53, 54)	agtaacaaagactacaaaaccttgggcatgggcttgt**tacttt**gaaatt**C**A	**DEF+;** Function of NtcA binding site shown by band shifts and lack of activation in *ntcA* mutant (54); **Fig. 4A**
nsiR1	-	4273242	379	r	(4)	catctacaactttatccggattggttagcatctgcccag**aacagt**ttaac**G**	**DIF+;** This is repeat 1 of 12; promoters differ for repeats 2-12 (4); cf. **Fig. 3A**
ntcA TSS1 (-262)	alr4392	5265153	857	f	this study	tagcaaaaatgatgattattaagggcgtttttagaataaaatgtttgttA	
ntcA TSS2 (-242)	alr4392	5265171	157	f	this study	ttaagggcgtttttagaataaaatgtttgttatttcttaagatattttcA	

ntcA TSS3 (-190/-189)	alr4392	5265223 (5265224)	(252)	f	(55)	attacttaatttttctgcaaaaaacctgatggttagagaaagtagaagTA	
ntcA TSS4 (-180)	alr4392	5265235	674	f	(42, 55)	ttttctgcaaaaaacctgatggttagagaaagtagaagtatctatcaagaA	**DIF+**
ntcA TSS5 (-136)	alr4392	5265277	667	f	(42)	tatcaagaactgctagctaaaaatagcaaaagttgggtatcattatgaacA	
ntcA TSS6 (-49)	alr4392	5265364 (5265365)	(62)	f	(42, 55)	acagaaaggttaacggtgctttattgatttttcaggtgtattcttaggtAA	
ntcB	all0602	701249	1339	r	(51)	ctgtaacaaaatctaccaaattggggagcaaaatcagctaacttaattgaA	**DEF+;** Function of NtcA binding site shown by band shifts and lack of activation in *ntcA* mutant (51)
patA TSS1 (-616)	all0521	614365	131	r	(56)	tttatatattatttttatttgcctagataaattggagtaacctagctttA	**DIF+**
patA TSS2 (-306)	all0521	614055	91	r	this study	tgggtttgagccaataactatacgtgtagtacttaaggtactgttagattA	
patS (-314)	asl2301	2771367	n.d.	r	(57)	tttacagagataatccaaggaataggtcaataagaataacaatggtgcaaA	+A
patS (-254)	asl2301	2771307	b.t. (40 reads)	r	this study	gactttaccagtattgactgatttaataaaagtctgctgtaagccttatcA	+A; TSS depends on HetR
patS (-39)	asl2301	2771092	n.d.	r	(57)	tgatgagaagatgcaatcgtagctaatctcagtagatggaaaaagtaaccA	+A
pepC	all4861	5791362 (5791334)	(50)	r	(58)	agtatgggttctgttttatactCtttatctgaatctgctaatttatatccA	
petC TSS1	all0606	704177	249	r	(14)	cgatttcgatacttctaaggttatgtcgggatctcaggtaaaatagtataA	+A
petC TSS2	all0606	704059	474	r	this study	aactaatgagtagctaaaaatgttggtaaaacaaggctagacttttgtttA	
petF TSS1	all4148	4992903	27332	r	(40, 59)	gtggtggatttcaagttatatacttggatttttctcgtagtatcagaaTTG	**DEF-**
petF TSS2	all4148	4992835	4495	r	(40)	tggctaatcccctgagaatagccgcccagttttgcttagcataacttatA	
petH TSS1 (-188)	all4121	4965738	2251	r	(60)	tgactcattattaacattctccacgagacttatcctctaagttagaaggtG	+A; Function of NtcA binding site shown by band shifts and lack of activation in *ntcA* mutant(60)
petH TSS2 (-63)	all4121	4965613	9649	r	(60)	ctttctcagagtgctagttttagtaacgaaaaatccgataaactaaagctG	**DEF+**; +A; derepression in *hetR*-0
pipX TSS1 (-436)	asr0485	580291 (580293)	(335)	f	(61)	ctgtagcaatgcagactgttgttaggaacagttattaggagaatgcgcCtG	**DEF**+; Function of NtcA binding site shown by footprinting and lack of activation in *ntcA* mutant (61); **Fig. 4A**
pipX TSS2 (-388)	asr0485	580339 (580341)	(125)	f	(61)	ctgtgggctattttgctccaacgaaggataattatggcaacatggagaGaA	**DIF+**
pipX TSS3	asr0485	580620 (580621)	543	f	(61)	aggatgagatcaaggtctttacaagctccatttaagggtaaaattgctGA	

(-107)							
pipX TSS4 (-53)	asr0485	580674	167	f	this study	ctcacctgactttatgcaaatattctggcacaagaggtattattgtccgaA	
pipX TSS5 (-23)	asr0485	580704	350	f	this study	caagaggtattattgtccgaatatttgtctttcactgcgaaaaaaattatA	**DIF+; Fig. 3A**
psbAI	alr4866	5796834	21544	f	(62)	catgggtttttagtctagtaaatttgcgtgaattcatg**taaatt**ttatga**G**	
psbB TSS1	all0138	143582 (143579)	(575)	r	(40, 63)	aaacatacgtaggataacttgtaacggcgatctgtgggaa**gataaa**g**C**ctA	
psbB TSS2	all0138	143518 (143512)	(5420)	r	(40, 63)	atcaatggacacgcttgcaaacggttccatcaa**tagtta**cactt**T**ttaaaA	
rbcL TSS1 (-504)	alr1524	1785464 (1785466)	(2865)	f	(40, 64, 65)	ctttcaaagaataacttatgccatttcttga**tatatt**gtgag**A**cAagttac	**DEF-**; NtcA binding site corroborated by band shift and footprinting (64)
rbcL TSS2 (-25)	alr1524	1785945	3987	f	this study	agggcaattgtgagaggaaattgtaccaaaacgtgatttgataagtaaaaA	**DEF-**
rnpB	-	4950813	98738	r	(8)	agaaataagcaaaacagacacaagacaccaacgaagat**tacact**ggagct**A**	
sigA TSS1 (-868)	all5263	6279707	105	r	(66)	atctgtaattttaaatgtcaaagtccttttgaattgat**taaaat**atcgtt**G**	
sigA TSS2 (-800)	all5263	6279640 (6279639)	(147)	r	(66)	ttgtcaagattcagaggttgaaaatctcggcgatacgc**taatgt**gtgaa**T**A	
sigA TSS3 (-605)	all5263	6279444	57	r	(66)	gacctattcaattctccggactaatttatttatgaagg**aaatgt**aaaaat**G**	**DIF+**
sigA TSS4 (-592)	all5263	6279431	n.d.	r	(66)	ctccggactaatttatttatgaaggaaatgtaaaaatgggtgcgttcata**A**	
sigA TSS5 (-328)	all5263	6279167	2221	r	(66)	ttaacacacaaaaagagtgcaatttctgtgaaacaggc**tacaat**cggaag**A**	**DEF-**
sigC	all1692	2022661	4572	r	(67, 68)	aaattagaaaatcgtcaggaaattacttatatacccatg**tagatg**tgact**A**	**DIF+; Fig. 3A**
spsB TSS1	all4376	5242942	1067	r	(69)	tcgtagtatgtatataacaatctttgacatattaaat**taaaat**aaagatt**A**	
spsB TSS2	all4376	5242903	55	r	(69)	aaataaagattactcaacgttcctacataattcccgca**tata**gtggtgtc**A**	
urtA	all1951	2335160(2335159)	b.t. (10 reads)	r	(70)	agtatcaaaaataacaattcaatggttaaatatcaaac**taatat**cacaa**T**C	
woxA, psbO	all3854	4652357 (4652355)	(1857)	r	(40, 71)	agaccagaagtgctttgattttgagttaaaaaaaatcg**tatgat**ctgg**C**tG	

Suppl. Table S3. List of TSS corresponding to the DEF+ class with a normalized ratio (M value) of at least $\log_2 3$

See "Suppl. Table S3.xlsx" within the *Dataset S1*

Suppl. Table S4. List of TSS corresponding to the DEF- class with a normalized ratio (M value) of at least $\log_2 -3$

See "Suppl. Table S4.xlsx" within the *Dataset S1*

Suppl. Table S5. List of TSS corresponding to the DIF+ class with a normalized ratio (M value) of at least $\log_2 3$

See "Suppl. Table S5.xlsx" within the *Dataset S1*

Suppl. Table S6. Selection of newly identified strongly regulated TSS of genes with known N-dependent regulation or previously described role in heterocyst differentiation. The annotation gives the respective gene name and systematic ID, followed by the TSS position and classification. For every TSS, the absolute numbers of reads are given for the four samples, followed by the strand information (f, forward; r, reverse strand), comments and references.

Gene	ID	TSS Position	Reads WT-0	Reads WT-8	Reads hetR-0	Reads hetR-8	S	Comment	Ref.
all2006	all2006	2400767 (DEF+) 2400640 (DEF+)	26 172	3762 954	133 166	3181 599	r	Expression inducible upon nitrogen deprivation (Northern hybridization); **Fig. 4*A***	(72)
all4160	all4160	5007675 (DIF+)	13	172	1	3	r	Polysaccharide synthesis	(73)
all4388	all4388	5261815 (DIF+)	163	1011	15	21	r	Polysaccharide synthesis	(74, 75)
alr2826	alr2826	3441875 (DIF+) 3441999 (DIF+)	2 3	73 187	9 5	2 1	f	HEP biosynthesis island; expression inducible upon nitrogen deprivation (microarray, Northern hybridization)	(16); (72)
alr2832	alr2832	3448677 (DIF+)	1	208	-	-	f	HEP biosynthesis island; expression inducible upon nitrogen deprivation (microarray, Northern hybridization)	(16); (72)
alr2833	alr2833	3449710 (DIF+)	2	467	1	1	f	HEP biosynthesis island; expression inducible upon nitrogen deprivation (microarray, Northern hybridization); **Fig. *3A***	(16); (72)
alr2834	alr2834	3452463 (DIF+) 3452765 (DIF+)	1 1	86 286	- -	- 4	f	HEP biosynthesis island	(16)
alr2837	alr2837	3457231 (DIF+)	52	800	1	9	f	HEP biosynthesis island; **Fig. 3*A***	(16)
alr2838	alr2838	3458381 (DIF+)	16	85	21	21	f	HEP biosynthesis island	(16)
amt1	alr0991	1158166 1158228	134 234	236 667	122 163	148 184	f	Expression inducible upon nitrogen deprivation (Northern hybridization)	(76)
amt4	alr0990	1156305	32	215	9	44	f	Expression inducible upon nitrogen deprivation (Northern hybridization)	(76)
asr1734	asr1734	2086181 (DIF+)	17	972	49	97	f	-	(77)
DPS	alr3808	4601709 (DIF+) 4601982 (DEF+)	11 56	252 1006	1 100	1 887	f	Differentially expressed in response to N-deficiency (shotgun proteomics and Northern hybridization). FurA binds to the promoter in a region located between both TSS. Transcripts confirmed here by Northern (**Figs. 2*A, 3A***).	(53, 78, 79)
glnB	all2319	2793917 (DEF+)	5	675	39	357	r	Expression inducible upon nitrogen deprivation (Northern hybridization)	(80)
hepA	alr2835	3453831/32 (DIF+)	1	94	-	-	f	Expression inducible upon nitrogen step-down; HEP biosynthesis island; one TSS previously described within the coding sequence; **Fig.**	(81); (16); (17)

								3*A*	
hepB	alr3698	4465653 (DIF+)	1	113	-	-	f	HEP biosynthesis	(73)
hepN	alr0117	119143 (DIF+)	59	216	46	41	f	-	(82, 83)
hepS	all2760	3357699 (DIF+)	5	516	6	2	r	Required for heterocyst maturation	(83, 84)
hetL	all3740	4517099	2	51	5	21	r	Overexpression stimulates heterocyst formation	(85)
hetN	alr5358	6401829 6401935	35 6	288 17	6 21	67 18	f	Inducible upon nitrogen deficiency	(86)
hgdD	alr2887	3524849	19	387	6	5	f	-	(75, 87)
nblA	asr4517	5407066 (DEF+) 5407159 (DEF+) 5407261 (DEF+) 5407310 5407397 (DEF+)	- 54 5 13 73	195 584 68 44 1530	8 30 7 7 101	214 594 49 6 735	f	Expression of *nblA* is inducible upon nitrogen deprivation (Northern hybridization). TSS confirmed here by primer extension (**Fig. 2*A***)	(88)
nirB	all0605	703698 (DEF+)	40	173	29	168	r	Expression inducible upon nitrogen deprivation (Northern hybridization)	(14)
patB	all2512	3019637	-	22	2	3	r	Heterocyst-specific expression	(89)
pbpB	alr5101	6077335	44	252	33	37	f	Required for heterocyst maturation	(90)
pknE	alr3732	4506698 (DIF+)	321	1155	113	171	f	Overexpression blocks heterocyst development	(91)
rfbC	alr2830	3446301 (DIF+)	4	298	1	1	f	HEP biosynthesis island	(16)

Suppl. Table S7. List of promoters with a possible DIF+ class element at position -35.

See "Suppl. Table S7.xlsx" within the *Dataset S1*

Suppl. Table S8. Possible NtcA binding sites for TSS in the DEF+ class with a normalized ratio (M value) of at least $\log_2 3$

See "Suppl. Table S8.xlsx" within the *Dataset S1*

Suppl. Table S9. Possible NtcA binding sites for chromosomally located TSS in the DEF- class with a normalized ratio (M value) of at least $\log_2 -3$

See "Suppl. Table S9.xlsx" within the *Dataset S1*

Supplementary Table S10. Comparison of the transcriptome dataset with regulon predictions from bioinformatics analyses. Comparison to the RegPrecise prediction (92) for the NtcA regulon (28 operons, 37 genes). The table displays the respective predicted operon or gene, as listed in the RegPrecise database, the position of the RegPrecise motif relative to the first nt of the coding region, the RegPrecise motif sequence in 5' to 3' orientation, the locus-tag, the TSS position found in this work with position relative to the coding region/ absolute coordinates (c, TSS is on reverse strand; additional TSS described in the literature in non-boldface letters). We moreover give the position of the RegPrecise NtcA motif relative to the TSS found in this work (bold letters in TSS column) and the numbers of sequence reads for the different conditions/strains mapped to the respective TSS (b.t., below threshold of 50 reads; n.d., no TSS identified in our dataset).

We found 14 from 28 sites in our DEF tables but only those with a fold change of >8 were considered sufficient for a redesign of the NtcA binding site (**Fig. 4**). However, many strongly regulated TSS in the DEF category are missing in the RegPrecise database despite exhibiting consensus NtcA binding sites. In fact only one (*nrrA*) of the eighteen NtcA binding sites upstream from TSS exhibiting highest fold change in the DEF+ category (included in Fig. 4A) is predicted in RegPrecise.

RegPrecise prediction	**Motif distance to gene**	**Motifs**	**Locus**	**TSS (relative / absolute)**	**Motif to TSS**	**Reads WT-0h WT-8h 216-8h 216-0h**	**Comments**
urtA-urtB-urtC-urtD-urtE	-116	AGTATCAAAAATAACA	all1951	**-67/2335160c,** -79	-49	b.t.	NtcA activated promoter previously described (70) and confirmed in our dataset by 10 reads (below the applied cutoff of 50), cf. **Table S2**.
rbcL-rbcX-rbcS	-32	AGTAAAAAGAGTGACA	alr1524	**-25/1785945**	-7	1442 732 760 1053	Repressed TSS at 1785945 identified here as DEF- (see **Table S2**). NtcA binding site located at a repression-compatible position.
gor	-48	AATAACAACTGTTACA	all4968	**-23/5931496c,** -91	-25	72 47 28	Repressed TSS at 5931496c previously described (38) and confirmed in our dataset as DEF- (see **Table S2**). NtcA binding site located at a

						70	repression-compatible position.
amt1	-50	TGTAACTAAGTATACA	alr0992			n.d.	Renamed *amtB* (76). No TSS identified in our dataset.
amt1	-250	TGTATTAACTAATACA	alr0990	**-202/1156305**	-48	32 215 44 9	Renamed *amt4* (76). A DEF+ TSS at 1156305 identified in our dataset (see **Table S6**). NtcA binding site located at consensus position for activated promoters.
SYNPCC7002_A0076	-107	CGTATAGCAAAATACA	asr0064	**-112/69876**	+6	3473 588 870 1265	Repressed TSS at 69876 identified in our dataset (**Table S1**). NtcA binding site located at a repression-compatible position.
SYNPCC7002_A2395	-64	GGTAGTCGGAGTTACA	alr4308	**-13/5164427**, -140	-51	58 509 1095 44	Activated TSS at 5164427 identified in our dataset (DEF+, **Table S1**). NtcA binding site located at consensus position for activated promoters
SYNPCC7002_A1954	-250	TGTAGTGATAGCTACA	alr3982			n.d.	No TSS identified in our dataset. NtcA binding sites likely corresponds to the promoter for the divergent all3981.
gifA	-78	CGTAGCATAAGATACA	asl2329	**-40/2809313c**	-38	4999 631 283 6123	NtcA-repressed TSS at 2809313 previously described (15) and confirmed in our dataset as DEF- (see **Table S2 and S4**). NtcA binding site located at a repression-compatible position as shown in **Fig. 4*E***.
glnA	-142	TGTAACAAAGACTACA	alr2328	**-155/2807328,** -91/2807392	+14	725 44 22 418	TSS located at 2807392 and 2807328 previously described (36) and confirmed in our dataset (**Table S2**). NtcA binding site located at a repression-compatible position with respect to TSS at position 2807328 (DEF-,**Table S4)**.
glnB	-214	AGTTACACAGACTACA	all2319	**-89/2793917c,** -29,-61,-71	-125	5 675 357 39	Activated TSS at 2793917 identified in our dataset as DEF+ (see **Table S1, S3** and **S6)**. NtcA binding site not located at expected consensus position as shown in **Fig. 4*A***.
ndh	-146	TGTAGCTTAAATTACT	all1127	**-41/1324785c**	-105	2 20 28 1	Activated TSS at 1324785 identified in our dataset (**Table S1**). NtcA binding site not located at expected consensus position
nrrA	-75	AGTAACAAAGACTACA	all4312	**-25/5167792c**, -93	-50	58 5560 5115 186	NtcA-activated TSS at position 5167792, previously described (54) and confirmed in our dataset (**Fig. 4*A***). NtcA binding site located at consensus position for activated promoters as shown in **Fig. 4*A***.
ntcA	-151 -111	GGTATCATTATGAACA AGTATAGGAAAGTACA	alr4392	**-48/5265365,** -30	-103, -63	1 33 6	Both NtcA binding sites located at non-consensus positions with respect to TSS at 5265365, confirmed in our dataset, and required for induction (fully

						22	analyzed in (93)).
ntcB	-80, -21	TGTAACAAAATCTACC, TGTAATTAAGGCTACA	all0602	**-31/701249**	-49, +11	9 427 880 23	NtcA-activated TSS at position 701249, previously described (51) and confirmed in our dataset. NtcA binding site 1 located at consensus position for activated promoters, the second at a repression-compatible position
SYNPCC7002_ A0264	-52	CGTAGTGTATGTTACA	all1327	**+1/1577077**	-52	52 377 304 52	Activated TSS at 1577077 identified in our dataset (**Table S1**). NtcA binding site located at consensus position for activated promoters.
coxB-coxA-coxC	-291	AGAAGTAGAAGCTACT	alr2514	**-174/3021024**	-117	4 166 10 5	Activated TSS at position 3021024 previously described (28) and confirmed in our dataset (DIF+, **Tables S1** and **S5)**.
fur	-218	TGTAGCTTGAGATTCA	all1691	**-26/2020552c**	-192	421 673 1617 1221	TSS at position 2020552 previously described (35) and confirmed in our dataset (**Tables S1** and **S2**). We are not aware of evidences supporting a function for this NtcA binding site.
alr1690	-91	TGTGAATATAGAAACA	alr1690	**-84/2018583**, -157	-7	362 117 119 116	TSS at position 2018583 previously described (20) and confirmed in our dataset (**Tables S1** and **S2**). We are not aware of evidences supporting a function for this NtcA binding site.
isiA	-37	AGTATTTCTGGCAACA	all4001	**-131/4820249c**	+95	24396 1775 697 25692	Repressed TSS at position 4820249 identified in our dataset as DEF- (see **Table S1** and **S4**). We are not aware of evidences supporting a function for this NtcA binding site.
psbZ	-161	TGTTACAAAAATTAGG	all1258	**-22/1493584c,** -5	-139	3428 1096 1650 4810	Repressed TSS at position 1493584 identified in our dataset as DEF- (**Table S1**). We are not aware of evidences supporting a function for this NtcA binding site.
gltS	-265	TGGAAAAATTAAAACA	alr4344			n.d.	No TSS identified in our dataset.
nifJ	-281	TGAATTTGATGTAACC	alr1911			n.d.	No TSS identified in our dataset.
ndhF-ndhD	-260	TGTCTTATAAAATACA	alr4156	**-122/4999494**	-138	42 20 47 27	TSS at position 4999494 identified in our dataset. We are not aware of evidences supporting a function for this NtcA binding site.
hanA	-279	GGTATTTTTTCCTACT	asr3935	**-147/4752807,** -7	-132	581 103 519 106	TSS at position 4752807 previously described (39) and confirmed in our dataset (**Tables S1** and **S2)**. We are not aware of evidences supporting a function for this NtcA binding site.
trxA	17	AGTTACAGATTCTACT	alr0052	**-12/56630**, -36, -43	+29	12 17 10	We are not aware of evidences supporting a function for this NtcA binding site, located within the reading frame.

						20	
apcF	-233	TGTAGTCTTTGTTACA	all2327	**-115/2807238c,** -29	-118	1042 179 316 622	This NtcA binding site is actually regulating the divergent *glnA* promoter (see above).
icd	-6	TGAAATATGTACAACA	alr1827	**-161/2192172**, -229	+155	25 16 28 38	We are not aware of evidences supporting a function for this NtcA binding site, that actually overlaps the reading frame.

Supplementary Table S11. List of oligonucleotides. Sequences are given in 5'→ 3' orientation.

See "Suppl. Table S11.xlsx" within the *Dataset S1*

5. SUPPLEMENTARY MATERIAL REFERENCES

1. Buikema WJ, Haselkorn R (1991) Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. *Genes & Development* 5:321-330.
2. 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.
3. Rippka R, Deruelles J, B. WJ, Herdmann M, Stanier RY (1979) Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J Gen Microbiol* 111:1-61.
4. 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.
5. 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.
6. Ausubel FM (2011) *Current Protocols in Molecular Biology* (John Wiley & Sons, Inc., Hoboken, NJ).
7. Steglich C*, et al.* (2008) The challenge of regulation in a minimal photoautotroph: non-coding RNAs in *Prochlorococcus*. *PLoS Genet* 4:e1000173.
8. Vioque A (1992) Analysis of the gene encoding the RNA subunit of ribonuclease P from cyanobacteria. *Nucleic Acids Res* 20:6331-6337.
9. 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.
10. Sharma CM*, et al.* (2010) The primary transcriptome of the major human pathogen, *Helicobacter pylori*. *Nature* 464:250-255.
11. 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.
12. Hoffmann S*, et al.* (2009) Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput Biol* 5:e1000502.
13. Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11:R25.
14. 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.
15. 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.
16. Huang G*, et al.* (2005) Clustered genes required for the synthesis of heterocyst envelope polysaccharide in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 187:1114-1123.
17. Zhu J, Kong R, Wolk CP (1998) Regulation of *hepA* of *Anabaena* sp. strain PCC 7120 by elements 5' from the gene and by *hepK*. *J Bacteriol* 180:4233-4242.
18. Curtis SE, Martin JA (1994) The transcription apparatus and the regulation of transcription initiation. *The molecular biology of cyanobacteria*, eds Bryant DA (Kluwer Academic Publishers, Dordrecht, The Netherlands), pp 613-639.
19. Sjöholm J, Oliveira P, Lindblad P (2007) Transcription and regulation of the bidirectional hydrogenase in the cyanobacterium *Nostoc* sp. strain PCC 7120. *Appl Environ Microbiol* 73:5435-5446.
20. Hernández JA*, et al.* (2006) Identification of a *furA* cis antisense RNA in the cyanobacterium *Anabaena* sp. PCC 7120. *J Mol Biol* 355:325-334.
21. Floriano B, Herrero A, Flores E (1994) Analysis of expression of the *argC* and *argD* genes in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 176:6397-6401.
22. McCarn DF, Whitaker RA, Alam J, Vrba JM, Curtis SE (1988) Genes encoding the alpha, gamma, delta, and four F0 subunits of ATP synthase constitute an operon in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 170:3448-3458.
23. Agervald A, Stensjö K, Holmqvist M, Lindblad P (2008) Transcription of the extended *hyp*-operon in *Nostoc* sp. strain PCC 7120. *BMC Microbiol* 8:69.
24. Curtis SE (1987) Genes encoding the beta and epsilon subunits of the proton-translocating ATPase from *Anabaena* sp. strain PCC 7120. *J Bacteriol* 169:80-86.
25. Agervald A, Zhang X, Stensjo K, Devine E, Lindblad P (2010) CalA, a cyanobacterial AbrB protein, interacts with the upstream region of *hypC* and acts as a repressor of its transcription in the cyanobacterium *Nostoc* sp. strain PCC 7120. *Appl Environ Microbiol* 76:880-890.
26. Frías JE, Herrero A, Flores E (2003) Open reading frame *all0601* from *Anabaena* sp. strain PCC 7120 represents a novel gene, *cnaT*, required for expression of the nitrate assimilation *nir* operon. *J Bacteriol* 185:5037-5044.
27. Mella-Herrera RA, Neunuebel MR, Golden JW (2011) *Anabaena* sp. strain PCC 7120 *conR* contains a LytR-CpsA-Psr domain, is developmentally regulated, and is essential for diazotrophic growth and heterocyst morphogenesis. *Microbiology* 157:617-626.
28. Valladares A, Herrero A, Pils D, Schmetterer G, Flores E (2003) Cytochrome *c* oxidase genes required for nitrogenase activity and diazotrophic growth in *Anabaena* sp. PCC 7120. *Mol Microbiol* 47:1239-1249.
29. 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.
30. Belknap WR, Haselkorn R (1987) Cloning and light regulation of expression of the phycocyanin operon of the cyanobacterium *Anabaena*. *EMBO J* 6:871-884.
31. Picossi S, Valladares A, Flores E, Herrero A (2004) Nitrogen-regulated genes for the metabolism of cyanophycin, a bacterial nitrogen reserve polymer: expression and mutational analysis of two cyanophycin synthetase and

cyanophycinase gene clusters in heterocyst-forming cyanobacterium *Anabaena* sp. PCC 7120. *J Biol Chem* 279:11582-11592.

32. Fiedler G, Muro-Pastor AM, Flores E, Maldener I (2001) NtcA-dependent expression of the *devBCA* operon, encoding a heterocyst-specific ATP-binding cassette transporter in *Anabaena* spp. *J Bacteriol* 183:3795-3799.
33. 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.
34. Hebbar PB, Curtis SE (2000) Characterization of *devH*, a gene encoding a putative DNA binding protein required for heterocyst function in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 182:3572-3581.
35. Hernández JA*, et al.* (2006) Interaction of FurA from *Anabaena* sp. PCC 7120 with DNA: a reducing environment and the presence of Mn(2+) are positive effectors in the binding to *isiB* and *furA* promoters. *Biometals* 19:259-268.
36. 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.
37. Tumer NE, Robinson SJ, Haselkorn R (1983) Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen. *Nature* 306:337-342.
38. Jiang F, Hellman U, Sroga GE, Bergman B, Mannervik B (1995) Cloning, sequencing, and regulation of the glutathione reductase gene from the cyanobacterium *Anabaena* PCC 7120. *J Biol Chem* 270:22882-22889.
39. Nagaraja R, Haselkorn R (1994) Protein HU from the cyanobacterium *Anabaena*. *Biochimie* 76:1082-1089.
40. Schneider GJ, Lang JD, Haselkorn R (1991) Promoter recognition by the RNA polymerase from vegetative cells of the cyanobacterium *Anabaena* 7120. *Gene* 105:51-60.
41. Borthakur D, Basche M, Buikema WJ, Borthakur PB, Haselkorn R (1990) Expression, nucleotide sequence and mutational analysis of two open reading frames in the *nif* gene region of *Anabaena* sp. strain PCC7120. *Mol Gen Genet* 221:227-234.
42. 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.
43. 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.
44. 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.
45. Ehira S, Ohmori M (2006) NrrA directly regulates expression of *hetR* during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 188:8520-8525.
46. 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.
47. Zhang W*, et al.* (2007) A gene cluster that regulates both heterocyst differentiation and pattern formation in *Anabaena* sp. strain PCC 7120. *Mol Microbiol* 66:1429-1443.

48. Devine E, Holmqvist M, Stensjo K, Lindblad P (2009) Diversity and transcription of proteases involved in the maturation of hydrogenases in *Nostoc punctiforme* ATCC 29133 and *Nostoc* sp. strain PCC 7120. *BMC Microbiol* 9:53.
49. Vargas W, Nishi C, Giarrocco L, Salerno G (2011) Differential roles of alkaline/neutral invertases in *Nostoc* sp. PCC 7120: Inv-B isoform is essential for diazotrophic growth. *Planta* 233:153-162.
50. Mulligan ME, Haselkorn R (1989) Nitrogen fixation (*nif*) genes of the cyanobacterium *Anabaena* species strain PCC 7120. The *nifB-fdxN-nifS-nifU* operon. *J Biol Chem* 264:19200-19207.
51. Frías JE, Flores E, Herrero A (2000) Activation of the *Anabaena nir* operon promoter requires both NtcA (CAP family) and NtcB (LysR family) transcription factors. *Mol Microbiol* 38:613-625.
52. Frías JE, Flores E (2010) Negative regulation of expression of the nitrate assimilation *nirA* operon in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 192:2769-2778.
53. 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.
54. 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.
55. Ramasubramanian TS, Wei TF, Oldham AK, Golden JW (1996) Transcription of the *Anabaena* sp. strain PCC 7120 *ntcA* gene: multiple transcripts and NtcA binding. *J Bacteriol* 178:922-926.
56. Young-Robbins SS, Risser DD, Moran JR, Haselkorn R, Callahan SM (2010) Transcriptional regulation of the heterocyst patterning gene *patA* from *Anabaena* sp. strain PCC 7120. *J Bacteriol* 192:4732-4740.
57. Yoon HS, Golden JW (2001) PatS and products of nitrogen fixation control heterocyst pattern. *J Bacteriol* 183:2605-2613.
58. Luinenburg I, Coleman JR (1992) Identification, characterization and sequence analysis of the gene encoding phosphoenolpyruvate carboxylase in *Anabaena* sp. PCC 7120. *J Gen Microbiol* 138:685-691.
59. Alam J, Whitaker RA, Krogmann DW, Curtis SE (1986) Isolation and sequence of the gene for ferredoxin I from the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 168:1265-1271.
60. Valladares A, Muro-Pastor AM, Fillat MF, Herrero A, Flores E (1999) Constitutive and nitrogen-regulated promoters of the *petH* gene encoding ferredoxin:NADP+ reductase in the heterocyst-forming cyanobacterium *Anabaena* sp. *FEBS Lett* 449:159-164.
61. 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.
62. Vrba J, Curtis S (1989) Characterization of a four-member *psbA* gene family from the cyanobacterium *Anabaena* PCC 7120. *Plant Mol Biol* 14:81-92.
63. Lang JD, Haselkorn R (1989) Isolation, sequence and transcription of the gene encoding the photosystem II chlorophyll-binding protein, CP-47, in the cyanobacterium *Anabaena* 7120. *Plant Mol Biol* 13:441-457.
64. 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.

65. Nierzwicki-Bauer SA, Curtis SE, Haselkorn R (1984) Cotranscription of genes encoding the small and large subunits of ribulose-1,5-bisphosphate carboxylase in the cyanobacterium *Anabaena* 7120. *Proc Natl Acad Sci USA* 81:5961-5965.
66. Brahamsha B, Haselkorn R (1991) Isolation and characterization of the gene encoding the principal sigma factor of the vegetative cell RNA polymerase from the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 173:2442-2450.
67. Brahamsha B, Haselkorn R (1992) Identification of multiple RNA polymerase sigma factor homologs in the cyanobacterium *Anabaena* sp. strain PCC 7120: cloning, expression, and inactivation of the *sigB* and *sigC* genes. *J Bacteriol* 174:7273-7282.
68. 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.
69. Cumino AC, Marcozzi C, Barreiro R, Salerno GL (2007) Carbon cycling in *Anabaena* sp. PCC 7120. Sucrose synthesis in the heterocysts and possible role in nitrogen fixation. *Plant Physiol* 143:1385-1397.
70. Valladares A, Montesinos ML, Herrero A, Flores E (2002) An ABC-type, high-affinity urea permease identified in cyanobacteria. *Mol Microbiol* 43:703-715.
71. Borthakur D, Haselkorn R (1989) Nucleotide sequence of the gene encoding the 33 kDa water oxidizing polypeptide in *Anabaena* sp. strain PCC 7120 and its expression in *Escherichia coli*. *Plant Mol Biol* 13:427-439.
72. Ehira S, Ohmori M, Sato N (2003) Genome-wide expression analysis of the responses to nitrogen deprivation in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *DNA Res* 10:97-113.
73. Wang Y*, et al.* (2007) Predicted glycosyl transferase genes located outside the HEP island are required for formation of heterocyst envelope polysaccharide in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 189:5372-5378.
74. 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.
75. Maldener I, Hannus S, Kammerer M (2003) Description of five mutants of the cyanobacterium *Anabaena* sp strain PCC 7120 affected in heterocyst differentiation and identification of the transposon-tagged genes. *FEMS Microbiol Lett* 224:205-213.
76. Paz-Yepes J, Merino-Puerto V, Herrero A, Flores E (2008) The *amt* gene cluster of the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 190:6534-6539.
77. Wu X, Lee DW, Mella RA, Golden JW (2007) The *Anabaena* sp. strain PCC 7120 *asr1734* gene encodes a negative regulator of heterocyst development. *Mol Microbiol* 64:782-794.
78. Hernández JA, Pellicer S, Huang L, Peleato ML, Fillat MF (2007) FurA modulates gene expression of *alr3808*, a DpsA homologue in *Nostoc* (*Anabaena*) sp. PCC7120. *FEBS Lett* 581:1351-1356.
79. 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.
80. Paz-Yepes J, Flores E, Herrero A (2009) Expression and mutational analysis of the *glnB* genomic region in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 191:2353-2361.

81. Holland D, Wolk CP (1990) Identification and characterization of *hetA*, a gene that acts early in the process of morphological differentiation of heterocysts. *J Bacteriol* 172:3131-3137.
82. Ning D, Xu X (2004) alr0117, a two-component histidine kinase gene, is involved in heterocyst development in *Anabaena* sp. PCC 7120. *Microbiology* 150:447-453.
83. Lechno-Yossef S, Fan Q, Ehira S, Sato N, Wolk CP (2006) Mutations in four regulatory genes have interrelated effects on heterocyst maturation in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 188:7387-7395.
84. Fan Q*, et al.* (2006) Signal transduction genes required for heterocyst maturation in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 188:6688-6693.
85. Liu D, Golden JW (2002) *hetL* overexpression stimulates heterocyst formation in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 184:6873-6881.
86. Callahan SM, Buikema WJ (2001) The role of HetN in maintenance of the heterocyst pattern in *Anabaena* sp. PCC 7120. *Mol Microbiol* 40:941-950.
87. Moslavac S*, et al.* (2007) A TolC-like protein is required for heterocyst development in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 189:7887-7895.
88. 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.
89. Jones KM, Buikema WJ, Haselkorn R (2003) Heterocyst-specific expression of *patB*, a gene required for nitrogen fixation in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 185:2306-2314.
90. Lázaro S, Fernández-Piñas F, Fernández-Valiente E, Blanco-Rivero A, Leganés F (2001) *pbpB*, a gene coding for a putative penicillin-binding protein, is required for aerobic nitrogen fixation in the cyanobacterium *Anabaena* sp. strain PCC7120. *J Bacteriol* 183:628-636.
91. Saha SK, Golden JW (2011) Overexpression of *pknE* blocks heterocyst development in *Anabaena* sp. strain PCC 7120. *J Bacteriol* 193:2619-2629.
92. Novichkov PS*, et al.* (2010) RegPrecise: a database of curated genomic inferences of transcriptional regulatory interactions in prokaryotes. *Nucleic Acids Res* 38:D111-118.
93. Olmedo-Verd E, Valladares A, Flores E, Herrero A, Muro-Pastor AM (2008) Role of two NtcA-binding sites in the complex *ntcA* gene promoter of the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* 190:7584-7590.