Expanding the direct HetR regulon in Anabaena sp. strain PCC 7120

Patrick Videau  
University of Hawaii

Shuisong Ni  
Miami University

Orion S. Rivers  
University of Hawaii

Blake Ushijima  
University of Hawaii

Erik A. Feldmann  
Washington University School of Medicine in St. Louis

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation
https://digitalcommons.wustl.edu/open_access_pubs/2279

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Authors
Expanding the Direct HetR Regulon in Anabaena sp. Strain PCC 7120


Published Ahead of Print 27 December 2013.

Updated information and services can be found at:
http://jb.asm.org/content/196/5/1113

**SUPPLEMENTAL MATERIAL**
Supplemental material

**REFERENCES**
This article cites 53 articles, 30 of which can be accessed free at: http://jb.asm.org/content/196/5/1113#ref-list-1

**CONTENT ALERTS**
Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: http://journals.asm.org/site/misc/reprints.xhtml
To subscribe to another ASM Journal go to: http://journals.asm.org/site/subscriptions/
Expanding the Direct HetR Regulon in *Anabaena* sp. Strain PCC 7120

Patrick Videau,* Shuison Ng,* Orion S. Rivers,* Blake Ushijima,* Erik A. Feldmann,* Loralyn M. Cozy,* Michael A. Kennedy,* Sean M. Callahan*  
Department of Microbiology, University of Hawaii, Honolulu, Hawaii, USA; Department of Chemistry and Biochemistry, Miami University, Oxford, Ohio, USA; Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri, USA

In response to a lack of environmental combined nitrogen, the filamentous cyanobacterium *Anabaena* sp. strain PCC 7120 differentiates nitrogen-fixing heterocyst cells in a periodic pattern. HetR is a transcription factor that coordinates the regulation of this developmental program. An inverted repeat-containing sequence in the hepA promoter required for proheterocyst-specific transcription was identified based on sequence similarity to a previously characterized binding site for HetR in the promoter of hetP. The binding affinity of HetR for the hepA site is roughly an order of magnitude lower than that for the hetP binding site. A BLAST search of the *Anabaena* genome identified 166 hepA-like sites that occur as single or tandem sites (two binding sites separated by 13 bp). The vast majority of these sites are present in predicted intergenic regions. HetR bound five representative single binding sites in *vitro*, and binding was abrogated by transversions in the binding sites that conserved the inverted repeat nature of the sites. Binding to four representative tandem sites was not observed. Transcriptional fusions of the green fluorescent protein gene *gfp* with putative promoter regions associated with the representative binding sites indicated that HetR could function as either an activator or repressor and that activation was cell-type specific. Taken together, we have expanded the direct HetR regulon and propose a model in which three categories of HetR binding sites, based on binding affinity and nucleotide sequence, contribute to three of the four phases of differentiation.

In 1961, Monod and Jacob postulated that differentiation was the sustained change in gene expression leading to a change in morphology (1). This definition has held true for many developmental programs, including endospore formation in *Bacillus* and *Micrococcus* species, aerial mycelium and spore formation in *Streptomyces* species, and the formation of stalk cells in *Caulobacter* species. To mediate the changes in gene expression necessary for the production of these specialized structures, each of these model organisms relies on a global or master regulator of differentiation (Spo0A in *Bacillus subtilis* [2, 3], FruA in *Micrococcus xanthus* [4, 5], AdpA in *Streptomyces griseus* [6, 7], and CtrA in *Caulobacter crescentus* [8, 9]). These regulators directly interact with the promoters of a few to many hundred genes, termed regulons, to either activate or repress transcription, coordinating the process of differentiation. While much work has focused on describing the regulons of the aforementioned regulators, comparatively little is known about the regulon of the master regulator of heterocyst differentiation in *Anabaena* sp. strain PCC 7120 (hereinafter called *Anabaena*).

*Anabaena* is a filamentous cyanobacterium that responds to low levels of combined nitrogen by differentiating specialized heterocyst cells that provide a microoxic environment for the fixation of dinitrogen by the oxygen-labile nitrogenase complex (reviewed in references 10, 11, and 52). Heterocysts are morphologically distinct cells that develop at semiregular intervals and are separated by approximately 10 to 20 photosynthetic vegetative cells, resulting in a 1-dimensional pattern along filaments. This differentiation process results in a change in the transcription of roughly 1,500 genes, which is facilitated by HetR, the master regulator of differentiation (12). A hetR deletion in *Anabaena* results in the inability to develop heterocysts, whereas overexpression yields supernumerary heterocysts even under nitrogen-replete conditions (13, 14).

HetR acts as a transcriptional regulator that functions early in the regulatory cascade governing differentiation. Recent work mapping all of the transcriptional start sites (TSSs) in *Anabaena* has identified 209 TSSs that are differentially regulated in wild-type and hetR mutant strains; expression from these TSSs was >8-fold higher in the wild type than in a hetR mutant strain (15). The regulation by HetR of many of these TSSs is likely indirect. HetR has been shown to bind to large DNA fragments (>150 bp) from the promoters of patS, hepA, hetR (16), and pknE (17), as well as to 29-bp and 40-bp DNA fragments derived from the promoters of hetP (18) and hetZ (19), respectively, in *vitro*, suggesting direct regulation of the transcription of these genes.

The ability of HetR to bind DNA is likely required for its regulatory function. The initial structure of the *Fischerella* HetR dimer, shown to complement an *Anabaena* hetR mutant, has been solved and displays four domains (20). Two flap domains extend outwards from the sides of the structure and are thought to mediate protein-protein interactions. The hood domain, which includes both C termini, likely interacts with a diffusible peptide (21, 22) that is derived from two inhibitors of differentiation (PatS and HetN) and promotes the degradation of HetR (23). The N termini create a DNA-binding domain containing helix-turn-helix motifs. Recently, cocrytallization of *Fischerella* HetR bound to a 21-bp DNA fragment based on that from the hetP promoter has identified the necessary protein-DNA interactions that confer DNA binding specificity to HetR (12). Most strikingly, the interaction of Glu71 with three consecutive cytosines during DNA

Received 22 November 2013 Accepted 23 December 2013
Published ahead of print 27 December 2013
Address correspondence to Sean M. Callahan, scallaha@hawaii.edu.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/JB.01372-13.
Copyright © 2014, American Society for Microbiology. All Rights Reserved. doi:10.1128/JB.01372-13

binding defines the requirement of an inverted repeat-containing sequence with CCC-N$_2$GGG at its core. An allele of HetR with Glu71 mutated was unable to bind DNA in vitro or complement an *Anabaena hetR* mutant strain, illustrating the absolute requirement of this amino acid for proper HetR function.

Clearly defining the HetR regulon would provide insight not only into the exact function of HetR but also into the cascade of events driving cellular differentiation. Here, we report the identification of a 17-bp inverted repeat-containing sequence in the *hepA* promoter that was bound by HetR in vitro and necessary for transcription in vivo. Additional HetR binding sites were uncovered by similarity to the *hepA* site, representatives of which were bound by HetR and used in transcriptional fusions to show that HetR can act as either an activator or repressor. These results suggest complex regulation of the HetR regulon.

**MATERIALS AND METHODS**

**Bacterial strains and growth conditions.** The growth of *Escherichia coli* and *Anabaena* sp. strain PCC 7120 (wild type) and its derivatives, the concentrations of antibiotics, the induction of heterocysts in medium lacking a source of combined nitrogen, and the conditions for photomicroscopy were as previously described (18, 24). Plasmids were introduced into *Anabaena* strains by conjugation from *E. coli* as previously described (25).

**Plasmid and strain construction.** The strains and plasmids used in this study are listed in Table 1. The primers used in this study are listed in Table S1 in the supplemental material. The integrity of all PCR-derived constructs was verified by sequencing. The transcriptional promoter fusions to the green fluorescent protein gene (*gfp*) were designed to include at least the nearest TSS from the chosen genes as defined by Mitschke et al. (15). Plasmid and strain construction is detailed in the supplemental material.

**Purification of recombinant HetR and DNA binding assays.** Recombinant HetR was overexpressed and purified from BL21(DE3) cells as previously described (22). Electrophoretic mobility shift assays (EMSAs) were performed essentially as described previously except that SYBR green, instead of ethidium bromide, was utilized according to the manufacturer’s instructions for imaging gels to estimate $K_D$ (dissociation constant) (22). The preparation of the 29-bp DNA fragments used for binding experiments is detailed in the supplemental material.

**Fluorescence quantification.** Green and red fluorescence from *Anabaena* was quantified with a GloMax-Multi Jr detection system (Promega) using the blue and red fluorometer kits, respectively, according to the manufacturer’s instructions. Initially, standard curves of the optical density at 750 nm (OD$_{750}$) versus both green and red autofluorescence were generated for each tested *Anabaena* strain containing *pAM1956* (26). The optimal OD$_{750}$ for measurement of both green and red fluorescence was found to be 0.1 for all strains, with an acceptable range of 0.05 to 0.15; beyond 0.15, the correlation between autofluorescence and OD$_{750}$ was nonlinear. For *gfp* fluorescence measurements, 20-ml liquid cultures of each tested strain containing a transcriptional fusion to *gfp* were grown in triplicate to an OD$_{750}$ of 0.5 to 0.8. Aliquots from each culture were diluted to an OD$_{750}$ of 0.1 in either BG-11 or BG-11b (BG-11 without a source of fixed nitrogen) to a final volume of 1 ml, and red and green fluorescence was measured. Red autofluorescence measurements were used to validate the OD$_{750}$ of 0.1. Using the standard curves, green background fluorescence was subtracted from the tested strains containing a transcriptional fusion to *gfp* to yield a measurement of green fluorescence in fluorescence standard units (FSU).

**Phylogenetic analysis.** Individual groups of the single and tandem *hepA*-like binding sites, generally 29 bp and 59 bp, respectively, were aligned with ClustalW. A maximum likelihood tree was constructed for each group using the generalized time-reversible algorithm (27) and 1,000 bootstrap replicates were performed using MEGA5 (28).

### Table 1 Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristic(s)$^a$</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Anabaena</em> strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCC 7120</td>
<td>Wild type</td>
<td>Pasteur Culture Collection</td>
</tr>
<tr>
<td>UHM101</td>
<td>ΔpatA</td>
<td>36</td>
</tr>
<tr>
<td>UHM103</td>
<td>ΔhetR</td>
<td>24</td>
</tr>
<tr>
<td>UHM244</td>
<td>ΔhetR/E71N</td>
<td>12</td>
</tr>
<tr>
<td>UHM269</td>
<td>Δsr0068:::1Sp$/^/$Sm'</td>
<td>This study</td>
</tr>
<tr>
<td>UHM270</td>
<td>Δoll1321::1Sp$/^/$Sm'</td>
<td>This study</td>
</tr>
<tr>
<td>UHM273</td>
<td>Δoll1100::1Sp$/^/$Sm'</td>
<td>This study</td>
</tr>
<tr>
<td>UHM274</td>
<td>Δoll1748::1Sp$/^/$Sm'</td>
<td>This study</td>
</tr>
<tr>
<td>UHM275</td>
<td>Δoll1347::1Sp$/^/$Sm'</td>
<td>This study</td>
</tr>
<tr>
<td>UHM276</td>
<td>Δoll5218::1Sp$/^/$Sm'</td>
<td>This study</td>
</tr>
<tr>
<td>UHM277</td>
<td>Δoll8206::1Sp$/^/$Sm'</td>
<td>This study</td>
</tr>
<tr>
<td>UHM278</td>
<td>Δoll746::1Sp$/^/$Sm'</td>
<td>This study</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pAM504</td>
<td>Shuttle vector for replication in <em>E. coli</em> and *Anabaena; Km' Nm'</td>
<td>53</td>
</tr>
<tr>
<td>pAM1956</td>
<td>Shuttle vector pAM504 with promoterless <em>gfp</em></td>
<td>26</td>
</tr>
<tr>
<td>pRL278</td>
<td>Suicide vector; Km' Nm'</td>
<td>49</td>
</tr>
<tr>
<td>pUW9</td>
<td>Source of Sp'/Sm' Ω interposon</td>
<td>54</td>
</tr>
<tr>
<td>pET28b+hetR</td>
<td>Expression vector for purifying polystyridine epitope-tagged HetR; Km'</td>
<td>22</td>
</tr>
<tr>
<td>pPJAV127</td>
<td>pAM504 with <em>P_ hepA</em> transcriptionally fused to <em>gfp</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV128</td>
<td>pAM504 with <em>P_ alr1347</em> transcriptionally fused to <em>gfp</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV132</td>
<td>pAM504 with <em>P_ alr1347</em> transcripationally fused to <em>gfp</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV210</td>
<td>pAM504 with <em>P_ hepA</em> with a mutated HetR binding site</td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV203</td>
<td>pAM1956 with <em>P_ patA</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV204</td>
<td>pAM1956 with <em>P_ patA</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV205</td>
<td>pAM1956 with <em>P_ patA</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV206</td>
<td>pAM1956 with <em>P_ patA</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV207</td>
<td>pAM1956 with <em>P_ patA</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV208</td>
<td>pAM1956 with <em>P_ patA</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV209</td>
<td>pAM1956 with <em>P_ patA</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV210</td>
<td>pAM1956 with <em>P_ patA</em></td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV216</td>
<td>pRL278 used to make UHM277</td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV213</td>
<td>pRL278 used to make UHM278</td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV214</td>
<td>pRL278 used to make UHM275</td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV215</td>
<td>pRL278 used to make UHM273</td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV216</td>
<td>pRL278 used to make UHM274</td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV217</td>
<td>pRL278 used to make UHM270</td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV218</td>
<td>pRL278 used to make UHM276</td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV219</td>
<td>pRL278 used to make UHM269</td>
<td>This study</td>
</tr>
<tr>
<td>pPJAV220</td>
<td>Suicide plasmid used to replace <em>alr1347</em> with an Sp'/Sm' Ω interposon insertion</td>
<td>This study</td>
</tr>
</tbody>
</table>

$^a$ SIp'/Sm', interposon conferring resistance to spectinomycin and streptomycin. Nm, neomycin.

Acetylene reduction assays, thin-layer chromatography, and alcian blue staining. Aerobic and anaerobic acetylene reduction assays were performed as previously described (24, 29). The positive-control Δ*hp86* strain used in the anaerobic assays was a transposon mutant that will be published elsewhere (R. Oshiro, L. M. Cozy, and S. M. Callahan, unpublished data). Lipid extractions were performed 24 h after the induction of...
heterocyst differentiation and run on thin-layer chromatography plates as previously described (30). The plates were developed at 120°C for 1 h. Heterocyst-specific exopolysaccharide was stained with alcian blue as previously described (18).

RESULTS
HetR binds to an inverted repeat in the hepA promoter required for normal transcription. To date, a defined nucleotide sequence required for HetR binding has only been shown for two sites (18, 19). Both sites contain the binding motif CCC-N5-GGG and are located in the promoter regions of the hetP and hetZ genes. To identify additional HetR binding sites in the Anabaena genome, the promoter region DNA sequences of genes whose expression is regulated by HetR were searched for the binding motif CCC-N5-GGG in silico. Within the promoter regions of the patA, patS, hetR, ntcA, hetP, hetZ, trpE, and hepA genes, 14 occurrences of this motif were found. To test for DNA–protein interaction, an electrophoretic mobility shift assay was conducted with HetR and 29-bp DNA fragments containing the 14 CCC-N5-GGG motifs. A shifted species, indicative of HetR binding to the DNA, was present for 29-bp fragments derived from the hetZ, trpE, and hepA promoter regions (Fig. 1A and data not shown). Binding of HetR to the hetZ binding site has previously been shown (19), and the results from the trpE binding site will be published elsewhere. The third fragment bound by HetR contained a 17-bp inverted repeat sequence, herein referred to as the binding site, centered 511 bases upstream from the hepA coding region. The similarly shifted species in the hetP and hepA lanes show that HetR binds to the hepA binding site in vitro (Fig. 1A).

The sequence specificity of HetR for the hetP binding site was assessed by Higa and Callahan (18) by mutation of six nucleotides within the binding site that abrogated specific binding and were later found to be primary sites of interaction between HetR and this DNA (12). To test the sequence specificity of the hepA binding site, the binding of HetR to a 29-bp DNA fragment with transversions at each of the nucleotides comprising the 17-bp hepA binding site core, which maintained the inverted repeat structure, was visualized by EMSA (Fig. 1A; see also Fig. S1 in the supplemental material). Similar to the results for mutated hetP binding site that remained unbound by HetR, no shift was seen with the mutated hepA binding site. While faint bands are present in the lanes containing hetP and hepA mutated binding sites, they are much higher in the gel than the specific shifts seen from HetR bound to the native sites and likely represent nonspecific interactions between the DNA and protein. We conclude that HetR binds specifically to the hepA binding site and suggest that the low stoichiometric ratio of HetR to DNA used in the EMSA (2:1) may represent a physiologically relevant interaction.

The hepA gene encodes a component of an ABC transporter involved in the construction of the heterocyst-specific exopolysaccharide layer (31). A transcriptional fusion of the hepA promoter with luxAB showed that transcription localized to proheterocyst cells (32). To assess the function of the HetR binding site in the hepA promoter (PhepA), a transcriptional fusion of PhepA to gfp was constructed. Similar to the results for the published PhepA-luxAB fusion, fluorescence from the PhepA-gfp fusion localized exclusively to proheterocyst cells (cells committed to differentiation that have not completed morphogenesis) in the wild-type and patA mutant strain backgrounds, and no fluorescence was observed in a hetR deletion mutant (Fig. 2A and G and data not shown). The same mutation of the hepA binding site that abrogated specific HetR binding in vitro was introduced into the PhepA-gfp fusion. With a mutated HetR binding site, no fluorescence was observed from the PhepA-gfp fusion in any strain background under any condition tested (Fig. 2B and G and data not shown), suggesting that the interaction of HetR with the proposed binding site in PhepA is necessary for transcriptional activation in proheterocysts.

HetR binds to high- and low-affinity sites in vitro. The master regulators of cellular differentiation in B. subtilis, M. xanthus, S. griseus, and C. crescentus have been shown to interact with specific DNA binding sites with either high or low affinities to control the timing of developmental programs (3, 6, 33, 34). To investigate the relative affinity of HetR for the hetP and hepA binding sites, the approximate $K_d$ was determined by EMSA. The $K_d$ is defined as the concentration of ligand (the hetP and hepA 29-bp DNA fragments) at which half is bound by protein. Since HetR interacts...
with DNA as a dimer, if half of the DNA is shifted, half of the protein is also bound when a 2:1 ratio of protein to DNA is used. In this assay, 200 nM hetP DNA was fully shifted by HetR (Fig. 3A). Using Sybr green, 100 nM DNA is nearly below the limit of detection, so the $K_d$ of HetR binding to the hetP 29-bp DNA fragment in vitro is less than 200 nM, likely in the 100 nM range if not lower. In contrast, at least 1 $\mu$M hepA DNA is necessary for observable shifting, and complete shifting is not observed at even 4 $\mu$M DNA (Fig. 3B). This suggests that the $K_d$ of HetR and hepA DNA is between 1 and 4 $\mu$M DNA, an order of magnitude higher than the $K_d$ of hetP DNA. Taken together, these data suggest that the hetP 29-bp DNA fragment is a high-affinity binding site and the hepA 29-bp DNA fragment is a low-affinity binding site.

The Anabaena genome contains 166 hepA-like predicted HetR binding sites. As was noted by Du et al. (19), a search of the Anabaena genome for a portion of the hepA binding site yielded 29 exact matches. To identify additional sites, a BLAST search was performed using the 29-bp sequence utilized for HetR binding (35). The search found the same sites as Du et al., as well as 137 more that have a 17-bp core similar to the hepA binding site, totaling 166 hepA-like binding sites (see Dataset S1 in the supplemental material). Of these 166 sites, 85.5% (142 sites) differ from the hepA binding site inverted repeat sequence by two nucleotides or fewer, and 91% (151 sites) occur in intergenic regions. For sites not located within coding regions, 60% (98 sites) are between co-orientated genes, 16% (27 sites) are between convergent genes, and 15% (26 sites) are between divergent genes. These sites are found singly (66 sites) or in tandem (50 sites). In the tandem arrangement, two single binding sites are separated by 13 nucleotides in all but four cases. To determine the levels of similarity within the single and tandem groups of binding sites, phylogenetic trees were constructed. The trees for the single and tandem binding sites, constructed with 29-bp and 59-bp fragments, respectively, each branched into two main groups (see Fig. S2 and S3 in the supplemental material). To investigate the function of these sites, five single and four tandem sites were chosen, generally two from each branch of the respective trees, from promoter regions of co-orientated, probably monocistronic genes as representatives for further study.

HetR binds to single hepA-like binding sites in vitro. The ability of HetR to bind to 29-bp DNA fragments containing representative single hepA-like binding sites derived from the promoter regions of the genes asr0081, all0961, all1321, all1748, and all3746 was assessed by EMSAs (Fig. 1B). A shifted species, indicating binding by HetR, was present for all five representative hepA-like binding sites. While the lacZ negative-control DNA re-
mained largely unbound, the majority of the DNA from the representative fragments appears to be present in shifted bands. Additionally, the five representative hepA-like binding sites shifted higher in the EMSA than the hetP 29-bp DNA fragment. Previous research has shown that HetR binds to DNA as a dimer but can also form tetrameric species when interacting with DNA [12]. It is possible that HetR binds to the five representative hepA-like binding sites differently than to the hetR binding site, and this interaction may include the multimerization of HetR.

To test the specificity of HetR interaction with the single binding sites, transversion mutations were introduced at each of the nucleotides comprising the 17-bp binding sites within the 29-bp DNA fragment containing representative single hepA-like binding sites, and binding was assessed (see Fig. S1 in the supplemental material). No shifted bands were observed for any of the mutated single binding sites (Fig. 1C). This indicates that the interaction of HetR with the single binding site DNA fragments requires a specific nucleotide sequence for binding rather than a general 17-bp inverted repeat motif. Together, these data show that HetR specifically binds to the five representative hepA-like binding sites, possibly as a tetramer, in vitro.

The spacing of tandem sites would not sterically hinder binding by HetR. Like the analysis of the representative single binding sites, the binding of HetR to 59-bp DNA fragments containing representative tandem binding sites derived from the promoter regions of alr1000, all1347, and all5218 was assessed by EMSA. No shifted species were observed for any of the tandem binding sites, indicating that HetR did not interact with these DNA fragments under the conditions tested (data not shown). Based on the structure of HetR bound to DNA, it appears that the flap domains that extend outwards from the DNA binding domain mediate the formation of a tetrameric complex and can interact with DNA nonspecifically [12]. A model of HetR bound to a 29-bp fragment based on the hetP binding site showed that the majority of the nucleotides interacted with HetR, and the possibility of HetR interacting with as many as six additional nucleotides was proposed. Because the 59-bp tandem sites investigated here occur as two 29-bp binding sites with an additional base between them, it was possible that the flap domains of two HetR dimers adjacent to one another on a DNA strand could interact with each other. To test this, a synthetic 58-bp DNA fragment containing two 29-bp hetP binding sites aligned head-to-head was assessed for HetR binding. In an EMSA, the 58-bp fragment was shifted in the presence of HetR. The mobility of the shifted species was significantly reduced compared to the mobility of the single sites with HetR shown earlier (Fig. 1), consistent with binding of HetR to this DNA as two adjacent dimers (data not shown). Therefore, steric interference is unlikely to account for the absence of binding between HetR and the representative tandem sites.

HetR either activates or represses transcription from representative promoters. To test whether HetR binding in vitro is indicative of transcriptional control in vivo, transcriptional fusions of the nine promoter regions to gfp were individually introduced into the wild type, a hetR deletion mutant, a hetR(E71N) mutant (having an E-to-N change at position 71 encoded by hetR), and a patA mutant strain. The wild type has a normal level of HetR, the hetR deletion strain is devoid of HetR, and the patA mutant strain has been shown to produce highly elevated HetR levels [36]. The hetR(E71N) mutant contains an allele of HetR in the native chromosomal locus whose protein product is unable to bind DNA in vitro because the E71N mutation abolishes interaction with the core of the HetR binding site consensus sequence [12]. By assessing the fluorescence output from the nine transcriptional fusions at the population and individual cell levels, the effect of HetR on transcription from each promoter fusion was determined.

To account for possible plasmid copy number differences across strain backgrounds, an rrn16Sa promoter fusion, chosen for its robust transcription in Anabaena across various conditions [37], was also assessed and displayed roughly equivalent fluorescence in all strains (see Table S2 in the supplemental material). The fluorescence baseline was determined by measuring the fluorescence from each strain background containing the promoterless copy of gfp on pAM1956 [26]. Each promoter fusion was measured in the four strain backgrounds in the presence (N⁺) and absence (N⁻) of combined nitrogen, and the results for representative promoter fusions are presented in Fig. 2, while the complete data set is included in the supplemental material. Quantification of fluorescence from the ast0081, ast0206, all0961, alr1000, alr1748, alr3746, and all5218 promoter fusions generally showed that transcription was higher in both the hetR deletion and hetR(E71N) mutant strains and lower in the patA mutant strain than in the wild type; this suggests that HetR repressed transcription from these promoters (Fig. 2E, F, and G; see also Table S2 in the supplemental material). Fluorescence from all0961 was observed only under N⁻ conditions, while alr3746 was only tran-
scribed under N\textsuperscript{+} conditions. Fluorescence from the all1321 promoter fusion was almost entirely absent from any strain background under N\textsuperscript{+} or N\textsuperscript{-} conditions. Unlike the other 8 promoter fusions, the alr1347 promoter displayed higher fluorescence levels in the wild type than in either the hetR deletion or hetR(E71N) mutant strain, indicating that HetR contributed to the activation of this promoter (Fig. 2C and D; see also Table S2).

While the quantitative measurements described above provided a population average of transcriptional activity, at the individual cell level, GFP localization was assessed by fluorescence microscopy. Of the nine promoter fusions, only the alr1347 promoter fusion (P\textsubscript{alr1347}–gfp) displayed any heterocyst-specific transcription (Fig. 2C). In all strain backgrounds under N\textsuperscript{-} conditions, P\textsubscript{alr1347}–gfp showed low levels of fluorescence (data not shown). When nitrogen was removed, vegetative cells continued to fluoresce at a low level, while developing heterocysts fluoresced with increased intensity (Fig. 2D and data not shown). Fluorescence from all other promoter fusions was present in both cell types under N\textsuperscript{-} and N\textsuperscript{-} conditions in all strain backgrounds tested (Fig. 2E and F and data not shown). These results suggest that HetR can function as both an activator and repressor of transcription from different promoters.

**alr1000 is required for diazotrophic growth.** To determine the involvement of the nine representative genes in diazotrophic growth, an antibiotic resistance cassette was inserted into each gene, and the growth of the mutant strains on medium lacking a source of combined nitrogen was assessed. Mutation of one gene, alr0961, was unsuccessful, suggesting that this gene is essential. Of the eight remaining strains, only the alr1000 mutant was incapable of growth on medium lacking a combined nitrogen source. This strain developed a wild-type pattern of heterocysts, and none in Trichodesmium cylindrica ATCC 29413, all of which form heterocysts, and none in Trichodesmium erythraeum IMS101, a filamentous, nonheterocystous strain. A BLAST search of the genome by Feb 26, 2014 by Washington University in St. Louis http://jb.asm.org/ Downloaded from

**DISCUSSION**

In this work, we identified an inverted repeat-containing HetR binding site in the hepA promoter and showed its necessity for proheterocyst-specific transcription. The hepA gene product is required for deposition of the exopoly saccharide layer that contributes to the physical barrier precluding oxygen entry into mature heterocysts. Strains lacking functional HepA cannot maintain the microoxic environment within heterocysts needed for the nitrogenase complex to fix nitrogen (29). Transcription of hepA is up-regulated 5 to 10 h after the removal of combined nitrogen and localizes to proheterocyst and heterocyst cells (38, 39). Deletion analysis of the hepA promoter region, utilizing chromosomal hepA::luxAB fusions, identified a region required for the upregulation following nitrogen step-down located between 585 and 414 bp upstream from the translational start site (40). This region begins 7 bp upstream from a TSS that is dependent on a functional copy of hetR for upregulation (15) and contains the HetR binding site identified in this study. Smaller deletions in this region showed that removal of the HetR binding site resulted in a modest reduction in transcription (40); however, our data show a complete loss of gfp fluorescence when the HetR binding site was mutated, indicating abrogation of transcription. While these studies both show an effect on hepA transcription from the loss of the HetR binding site, differences in the reporters and their placement on plasmids or in the chromosome could account for the discrepancy.

The site in hepA was used to identify 166 potential HetR binding sites that were similar in sequence. This number of potential HetR binding sites is comparable to the number of binding sites in well-studied regulons, including the cyclic AMP receptor protein (CRP) (41), RpoN (42), and PhoP/Q (43) regulons in Escherichia coli, Salmonella enterica serovar Typhimurium, and Yersinia pestis, respectively, as well as the developmental regulons of Sp oP from Bacillus subtilis (2) and CtrA from Caulobacter crescentus (8). Of the 166 sites, only 13 hepA-like sites were tested, and evidence consistent with real binding of HetR with the single sites was found. While only a subset were tested, three lines of evidence suggest that the hepA-like binding sites are biologically relevant and are not present throughout the Anabaena genome by chance. First, taking into account the 41.2% GC content of the Anabaena genome, the 17-bp core hepA sequence is predicted to occur less than once. Instead, this 17-bp sequence occurs 37 times, 68 times if either member of the A/T pair in the middle of the inverted repeat sequence is allowed. Second, the majority of the hepA-like binding sites are located in intergenic regions. Of the 166 hepA-like binding sites, 91% are present in intergenic regions, compared to 17.4% of the genome that is predicted to not code for protein (44). The HetR binding sites located in open reading frames may regulate transcription, as work on the E. coli RtuR regulon showed that the majority of its binding sites were located within coding regions (45). Third, nucleotide sequences similar to the hepA binding site are present in heterocystous cyanobacteria but absent in a nonheterocystous strain. A BLAST search of the 17-bp hepA palindrome revealed 20 similar sites in Nostoc punctiforme ATCC 29133 and Anabaena azollae 0708, 25 in Anabaena cylindrica PCC 7122, 149 in Anabaena variabilis ATCC 29413, all of which form heterocysts, and none in Trichodesmium erythraeum IMS101, a filamentous, nonheterocystous nitrogen-fixing strain. Trichodesmium can grow diazotrophically, but nitrogen fixation relies on a fluctuating ratio of temporally regulated diazocytes rather than on a fixed ratio of spatially regulated heterocysts (46, 47). The lack of heterocysts could account for the lack of hepA-like HetR binding sites in Trichodesmium. These data are consistent with the biological relevance of the proposed HetR binding sites.

The interaction of HetR with the binding site present in the hepP promoter is the best-characterized example of HetR binding DNA to date. In EMSAs, HetR was shown to interact with the hepP binding site at a 1:1 HetR dimer/DNA ratio, and a supershifted species was observed at a 2:1 HetR dimer/DNA ratio, indicative of a tetrameric HetR complex (22). This HetR-DNA interaction was abrogated by mutation of six bases within the 17-bp binding site palindrome (18). Additionally, crystal structures of HetR bound to DNA based on the hepP binding site elucidated the specific contributions of the amino acid residues that facilitate the interaction of HetR with DNA and described HetR binding to DNA as a tetramer (12). In the work presented here, the K\textsubscript{d} estimation of HetR for the hepP and hepA binding sites showed that HetR interacted with the hepP binding site with a higher affinity than with the hepA site. Based on these differences in affinity and published
stoichiometric ratios used in EMSAs of HetR with additional known sites, we propose that HetR interacts with three categories of binding sites to regulate different phases of heterocyst differentiation: category I is comprised of sites that have been difficult to delimit and define and exhibit low binding affinity in vitro; category II is comprised of well defined sites that share a similar inverted repeat sequence and exhibit an intermediate level of binding affinity in vitro; and category III is comprised of well defined sites that share a similar inverted repeat sequence different from the category II sites and exhibit a high level of binding affinity in vitro.

Category I HetR binding sites are poorly defined. Large DNA fragments, labeled with 32P or biotin, from the promoters of hetR (16, 17, 48), patS (16), and pknE (17) have been shown to interact specifically with HetR in EMSAs. However, stoichiometric ratios ranging from 50:1 to 1,500,000:1 of HetR to labeled DNA were required for shifting. Binding of the PhetR DNA fragment at a 50:1 ratio did not result in complete shifting (17), and shifting of the entire band has only been observed at far higher stoichiometric ratios (48). Binding assays utilizing the method in this study were unable to shift fragments derived from the hetR promoter (S. Ni and M. A. Kennedy, unpublished results). This method is not amenable to large stoichiometric ratios, so a direct comparison is not possible.

Transcription of the hetR, patS, and pknE genes is upregulated by 0.5 (49), 3 (50), and 3 h (17), respectively, following the removal of combined nitrogen and localizes to single cells roughly 6 to 8 h following upregulation. All three genes require a functional copy of hetR for upregulation (16, 17, 49), but sequence determinants that provide specificity for HetR binding to these promoters are unknown. Potential HetR binding sites derived from the hetR and patS promoters were tested for interaction in this work by performing EMSA, but no binding was observed (data not shown). It is possible that additional proteins absent from the EMSAs may interact with HetR to mediate binding or that post-translational modification of HetR may facilitate binding to category I sites. As the upregulation of hetR, patS, and pknE coincides with the patterning phase of differentiation (Fig. 4), additional factors required for high-affinity binding in vitro may also be upregulated during this phase.

The binding sites found in the work described here fall into category II, which consists of sites with a defined palindromic sequence and a binding affinity in vitro that falls between that of category I and III sites. Many category II binding sites are adjacent to genes without a clear connection to heterocyst function but which may be involved in vegetative cell metabolism or heterocyst morphogenesis. Metabolism under nitrogen-limiting conditions is inherently different from nitrogen-replete growth in that fixed nitrogen is provided exclusively from heterocysts rather than imported from the environment. A role for HetR as a general regulator of metabolism or heterocyst morphogenesis in the absence of nitrogen is consistent with the overexpression of hetP in a hetR deletion mutant resulting in the formation of heterocysts that were only capable of nitrogen fixation under anoxic conditions (18). It is possible that, while morphologically distinct heterocysts cells can form in the absence of hetR with overexpression of hetP, certain functional components are lacking that would normally be controlled by HetR. This finding is consistent with the arl1000 mutant described in this work that, although it was able to produce morphologically distinct heterocysts, only fixed nitrogen under anoxic conditions.

Category III HetR binding sites have inverted repeat sequences with a low Kd and high binding affinity in vitro. Like category II sites, mutation of nucleotides within the 17-bp inverted repeat abrogates HetR binding (18). The only two category III binding sites, those present in the hetP and hetZ promoters, are shifted at 1:1 and 6:1 HetR dimer-to-DNA ratios, respectively (22; P. Videau and S. M. Callahan, unpublished results). The hetP and hetZ genes have been reported to have overlapping function, and overexpression of hetP can partially bypass hetR (18, 51). It has been proposed that HetP acts as a slave to HetR such that HetP, and perhaps HetZ, executes the patterning decision made by HetR and PatS. Consistent with this hypothesis, the functions of hetP and hetZ and the timing of their upregulation coincide with the commitment phase of differentiation (Fig. 4) (39).

Recent work in Anabaena identified TSSs in the wild type and a ΔhetR mutant during growth with a source of nitrogen and 8 h following nitrogen removal (15). A set of TSSs, termed DIF+ (differentiation-related changes in transcription from TSSs in the wild type), were upregulated ∼8-fold in wild-type Anabaena but not in a ΔhetR mutant. Of the category II HetR binding sites identified in this work, only 4 promoters with category II sites contain TSSs in the DIF+ category; these TSSs are within the promoter regions of all0284, alr2835, alr3554, and alr4919. The TSSs nearest to the remaining category II sites did not show an obvious correlation with regulation by HetR. The distances of category II sites to

FIG4 A model depicting the genetic interactions involved in the development of heterocyst differentiation. The time during which the components of each phase function following the removal of combined nitrogen is indicated. Arrows indicate positive interactions, while bars indicate negative interactions. C1, C2, and C3 correspond to the category I, II, or III HetR binding sites, respectively, thought to contribute to the phase of differentiation. Non-diazotrophic processes refers to processes that are downregulated during diazotrophic growth.
the closest TSS were fairly evenly distributed, but this could be an artifact of an incomplete data set. The binding sites presented here are part of an initial characterization of the direct HetR regulon and probably do not encompass all HetR binding sites.

ACKNOWLEDGMENTS

We thank Samantha O’Hanlon for help with statistics and Lizbeth Videau for critical reading of the manuscript.

This work was supported by grant MCB-1121346 from the National Science Foundation.

REFERENCES


Expanding the HetR Regulon in *Anabaena*


