Journal of Bacteriology, July 2011, p. 3482–3489 Vol. 193, No. 14
0021-9193/11/$12.00 doi:10.1128/JB.05010-11
Copyright © 2011, American Society for Microbiology. All Rights Reserved.

Identification of Ten *Anabaena* sp. Genes That under Aerobic Conditions Are Required for Growth on Dinitrogen but Not for Growth on Fixed Nitrogen▽†

Sigal Lechno-Yossef,[1] Qing Fan,[1]‡ Elizabeth Wojciuch,[1] and C. Peter Wolk[1,2]*

MSU-DOE Plant Research Laboratory[1] and Department of Plant Biology,[2] Michigan State University, East Lansing, Michigan 48824

Received 5 April 2011/Accepted 9 May 2011

Heterocysts are specialized cells required for aerobic fixation of dinitrogen by certain filamentous cyanobacteria. Numerous genes involved in the differentiation and function of heterocysts in *Anabaena* sp. strain PCC 7120 have been identified by mutagenizing and screening for mutants that require fixed nitrogen for growth in the presence of oxygen. We have verified that 10 *Anabaena* sp. genes, *all1338*, *all1591*, *alr1728*, *all3278*, *all3520*, *all3582*, *all3850*, *all4019*, *alr4311*, and *all4388*, identified initially by transposon mutagenesis, are such genes by complementing or reconstructing the original mutation and by determining whether the mutant phenotype might be due to a polar effect of the transposon. Elucidation of the roles of these genes should enhance understanding of heterocyst biology.

Anabaena sp. strain PCC 7120 (hereinafter referred to as *Anabaena* sp.) is a filamentous cyanobacterium in which in the absence of fixed nitrogen, approximately every 10th cell differentiates into a specialized cell called a heterocyst. Heterocysts are the site of nitrogen fixation by the oxygen-sensitive enzyme nitrogenase. Within heterocysts, nitrogenase is protected from oxygen by a double-layered envelope comprising a laminated layer of heterocyst envelope glycolipids (HGL) that impedes entry of oxygen into the cell and a homogeneous layer of heterocyst envelope polysaccharide (HEP) that protects HGL layers from physical damage, an increased rate of respiration, and inactivation of at least the oxygen-generating complex of photosystem II (39). An estimate that ca. 15 to 25% of the *Anabaena* sp. DNA sense strand is transcribed only in heterocysts (21) could not distinguish how many of the genes involved are actually essential for development or diazotrophy.

Numerous genes involved in the differentiation and function of heterocysts in *Anabaena* sp. have been identified by mutagenizing and screening for mutants that require fixed nitrogen for growth in the presence of oxygen. Having identified such mutants that nonetheless reduced acetylene under anoxic conditions, Ernst et al. (9) wrote that "the inability to grow on molecular nitrogen in the presence of 0.2 atm … of O_2 does not preclude the possibility that nitrogenase may be expressed under other conditions, for example, in the absence of oxygen." Therefore, they coined the term "Fox" mutants, mutants that are "incapable of *f*ixation in the presence of *ox*ygen," to distinguish them from "Fix" mutants, unable to fix N_2 under all conditions. Fan et al. (11) identified *conR* (*all0187*) as a gene that requires fixed nitrogen for growth in the presence of oxygen. Their observation was confirmed by Mella-Herrera et al. (24), who observed that for at least a short period after N_2 fixation normally starts, a *conR* mutant shows aerobic nitrogenase activity. (They suggested that the inability of the cells to grow on N_2 was caused by the inability of heterocysts to transfer fixed nitrogen to neighboring cells.) What have been called Fox mutants have been tested for growth, or the lack of growth, in the presence of oxygen and absence of fixed nitrogen, but they have seldom been tested for aerobic nitrogenase activity. Because a *conR* mutant could express some nitrogenase activity aerobically, we propose that Fox mutants be defined slightly more broadly as those "requiring *f*ixed nitrogen for growth in the presence of *ox*ygen" rather than the original definition (9), those "incapable of *f*ixation in the presence of *ox*ygen," and will use the broader definition. By either definition, mutants are often easily distinguished because they soon yellow and cease growth under aerobic conditions in a medium lacking fixed nitrogen (9). Often, the mutation is evidently heterocyst related, e.g., *hep* and *hgl* genes that are required specifically for the biosynthesis of HEP and HGL, respectively, have heretofore been proven to be Fox genes. In other instances, a mutation may affect all cells of a filament, e.g., a mutation in a gene such as *sepJ* (*fraG*) (13, 26) that leads to extensive fragmentation of filaments. In that instance, heterocysts may soon lack nutrients with which to fix nitrogen or to remain internally anoxic. Similarly, a *hglK* mutant has a phenotype in vegetative cells as well as in heterocysts, but it is a Fox gene perhaps because of its effect on N_2 fixation in heterocysts (2).

Not all genes that regulate differentiation are Fox genes. For example, *hetR* is a Fox gene, because in its absence, no heterocysts differentiate and so no nitrogen fixation takes place (3), whereas *patA* and *patS* are not Fox genes. In a *patA* mutant, heterocysts form only at the ends of long filaments, but those cells that do differentiate fix nitrogen, so that the filaments grow (20). When a *patS* mutant is deprived of fixed nitrogen, the initial response is that contiguous vegetative cells differentiate into heterocysts, but within a few days, spaced heterocysts again become the norm (44). Genes required for

* Corresponding author. Mailing address: MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824. Phone: (517) 353-2049. Fax: (517) 353-9168. E-mail: wolk@msu.edu.

† Supplemental material for this article may be found at http://jb.asm.org/.

‡ Present address: Northwestern University, Feinberg School of Medicine, 303 East Chicago Avenue, Chicago, IL 60611-3008.

▽ Published ahead of print on 20 May 2011.

heterocyst-specific metabolism can be Fox genes, e.g., *nifH*, *nifD*, and *nifK*, that encode the subunits of nitrogenase, but *fdxH*, a ferredoxin that is heterocyst specific is—although important—not essential for aerobic N_2 fixation and so is not a Fox gene (23). Mutations in one of the two cytochrome oxidase operons, *coxABCII* and *coxABCIII*, allow continued diazotrophy under oxic conditions, so that these are not Fox genes, but a combination of mutations in *coxII* and *coxIII* does not allow aerobic nitrogen fixation (36), so that the dual mutation has a Fox^- phenotype. Similarly, mutation of two protein kinase genes, *pkn30* and *pkn44*, but neither alone, results in a Fox^- phenotype (31). In short, a gene need not be a Fox gene to be important for heterocyst formation or function, and a Fox gene—while it need not be expressed only in heterocysts—evidently has, in the presence of oxygen, some critical importance for heterocyst development or function but is not essential for growth on fixed nitrogen.

We have used transposon screening to identify Fox genes. However, a transposon mutation that confers a Fox^- phenotype can also result from a combination of a spontaneous mutation unrelated to the presence of the transposon and antibiotic resistance conferred by the transposon or by a polar effect of the transposon on downstream genes. That a transposon-intercepted gene is a Fox gene can be verified by reconstruction of the mutation or by complementation, provided that the possibility of a polar effect is carefully considered.

We present 10 genes whose mutants have a Fox^- phenotype and that have not previously been shown conclusively to be Fox genes in *Anabaena* sp. Mutations in some of these genes result in heterocysts that have structural abnormalities; some others have metabolic defects. Bioinformatic analysis provides hints as to what some of the latter defects may be. Considering that approximately 75 Fox genes have been published (see Table S1 in the supplemental material), the 10 genes reported in this paper represent more than 10% of the currently known Fox genes. Additionally, if the estimate is correct that *Anabaena* sp. has approximately 100 to 140 Fox genes (38), the majority of Fox genes have been discovered, many of them by our transposon mutagenesis screen (10, 11, 13, 17, 19, 37, 40).

MATERIALS AND METHODS

Cyanobacteria were grown, with shaking, in flask cultures of AA/8 liquid medium, with or without nitrate (16a), or in medium AA, with or without nitrate, solidified with 1.2% home-purified (Difco) Bacto agar (16a) and supplemented with antibiotics as appropriate, at 30°C. Cultures were illuminated initially as described previously (10, 17) and more recently with Philips F32T8/TL741Universal/Hi-Vision lamps (ca. 30 μmol s^{-1} m^{-2}) (LI-COR Bioscience light meter LI-250A). Mutagenesis with transposon Tn*5*-1063, screening of mutants, complementation, and insertional mutagenesis were performed as described previously (10, 17). Strains and plasmid constructions are listed in Table 1. Samples stained with an aqueous solution (17) of Alcian Blue (13a) were visualized with a Wild M20 microscope and photographed with a Nikon CoolPix 4300 digital camera.

The evolutionary history of *hisD* was inferred using the neighbor-joining method (29). The optimal tree with the sum of branch length of 8.38 is shown. Bootstrap value (100 replicates) larger than 75% are shown above the branches (12). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Dayhoff matrix-based method (30) and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the data set (complete deletion option). The final data set totaled 381 positions. Phylogenetic analyses were conducted by using MEGA4 (34).

RESULTS AND DISCUSSION

We present below information concerning 10 genes that we have identified as Fox genes on the basis of corresponding mutants that require fixed nitrogen to grow aerobically. Table 2 summarizes the information.

all1338. *all1338* was intercepted by a transposon twice in our mutagenesis study and is the last gene of a possible operon. FQ1470, a strain with a mutation in this gene, was complemented by pRL3007, which bears *all1338* as its only intact gene, implying that it is a Fox gene. The heterocyst envelopes of FQ1470 and of a second *all1338* mutant, FQ1265, are abnormal (Fig. 1a and b). In addition, internal vacuoles (Fig. 1a) are more frequently observed in heterocysts of FQ1265 than in FQ1470. All1338 is annotated as a hypothetical protein. It is unique to, and conserved among, all 42 genomically sequenced cyanobacteria from different taxonomic groups. It is rich in aspartic and glutamic acids and appears to be soluble. No function can now be assigned to Alr1338 on the basis of sequence similarity.

all1591. *all1591* was intercepted by a transposon 12 times in our mutagenesis study, one resulting mutant being FQ747. Its downstream gene, *all1590*, is similarly oriented. FQ747 was complemented by pRL3170, which carries *all1591*, and was not complemented by pRL2879, which carries *all1590*. PCR analysis of the complemented mutant, performed as described in reference 17, showed that pRL3170 has, in part, integrated into the genome of FQ747 upstream of the transposon and provided no evidence of integration downstream of the transposon. We conclude that only *all1591* and not *all1590* could have contributed to the complementation, implying that *all1591* is a Fox gene. An *all1591* insertional mutant, SR2813a, is also unable to grow in the absence of combined nitrogen. FQ747 and SR2813a, when deprived of fixed nitrogen, have the same phenotype: the heterocyst envelope, as seen with a microscope, ranges in shape from a slightly to widely open horseshoe (Fig. 1c). *all1591* is annotated as *hisD*. *hisD* codes for histidinol dehydrogenase, an enzyme that catalyzes the last step of histidine biosynthesis from 5-phosphoribosyl-α-pyrophosphate. Three of the four heterocyst-forming cyanobacteria with a fully sequenced genome have two copies of HisD, but *Nostoc punctiforme* has only a single copy, which is similar to All1591. Phylogenetic analysis of HisD from cyanobacteria, with the protein from *Escherichia coli* as a reference, shows that All1591 and its orthologs from other heterocyst-forming cyanobacteria cluster with HisD from *Lyngbya* sp. and *Trichodesmium erythraeum*, cyanobacteria that are also filamentous, whereas the other HisD copy, Alr3056 and its orthologs, cluster with the homologs from *E. coli* and unicellular cyanobacteria (Fig. 2). Interestingly, heterocyst-forming cyanobacteria but no other cyanobacteria also have two copies of *hisC* (corresponding to *all4966* and *alr2092* in *Anabaena* sp.), coding for histidinol phosphate aminotransferase which functions upstream of HisD in the metabolic pathway. Perhaps heterocysts employ specialized enzymes for histidine biosynthesis, although there is no evidence that either *hisC* gene is required for heterocyst function. Why a modification of histidine biosynthesis should affect heterocyst morphology (Fig. 1c) is unclear. Microarray data (data in reference 7 analyzed by the method in reference 41) showed that the expression of *all1591*

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristics and/or derivation[a]
Fox$^-$ derivatives of *Anabaena* sp. strain PCC 7120	
FQ202	Bm^r Nm^r Sm^r; *alr1728*::Tn*5*-1063
FQ204	Bm^r Nm^r Sm^r; *all1591*::Tn*5*-1063
FQ211	Bm^r Nm^r Sm^r; *all3520*::Tn*5*-1063
FQ228	Bm^r Nm^r Sm^r; *all3278*::Tn*5*-1063
FQ324	Bm^r Nm^r Sm^r; *alr1728*::Tn*5*-1063
FQ384	Bm^r Nm^r Sm^r; *all3850*::Tn*5*-1063
FQ406	Bm^r Nm^r Sm^r; *all4019*::Tn*5*-1063
FQ747	Bm^r Nm^r Sm^r; *all1591*::Tn*5*-1063
FQ807	Bm^r Nm^r Sm^r; *alr4311*::Tn*5*-1063
FQ1265	Bm^r Nm^r Sm^r; *all1338*::Tn*5*-1063
FQ1470	Bm^r Nm^r Sm^r; *all1338*::Tn*5*-1063
FQ1580	Bm^r Nm^r Sm^r; *all3582*::Tn*5*-1063
FQ1595	Bm^r Nm^r Sm^r; *all4388*::Tn*5*-1063
SR2816a	Cm^r Em^r; mutation in *all3850* resulting from single recombination with pRL2816a; see below
SR2821	Cm^r Em^r; mutation in *alr1728* resulting from single recombination with pRL2821; see below
Escherichia coli strains	
DH5αMCR	Invitrogen Corp., received courtesy of J. C. Meeks
DH10B	F$^-$ *mcr*A Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ*ΔM15 Δ*lacX74 recA1 endA1 araD139* Δ(*ara leu*)*7697 galU galK rpsL nupG* λ$^-$
HB101	F$^-$ *thi-1 hsdS20* (r_B^- m_B^-) *supE44 recA13 ara-14 leuB6 proA2 lacY1 galK2 rpsL20* (Sm^r) *xyl-5 mtl-1*
Plasmids	
anc0756	Cm^r Em^r; *Anabaena* sp. chromosomal DNA from bp 4233705 to 4250432 in the BamHI site of pRL838
anc1006	Cm^r Em^r; *Anabaena* sp. chromosomal DNA from bp 3952462 to 3970608 in the BamHI site of pRL838
anc1512	Cm^r Em^r; *Anabaena* sp. chromosomal DNA from bp 2070875 to 2084189 in the BamHI site of pRL838
anp00427	Ap^r; *Anabaena* sp. chromosomal DNA from bp 5161083 to 5168600 in the BamHI site of pUC18
anp00676	Ap^r; *Anabaena* sp. chromosomal DNA from bp 4325494 to 4333569 in the BamHI site of pUC18
anp00920	Ap^r; *Anabaena* sp. chromosomal DNA from bp 4640078 to 4648267 in the BamHI site of pUC18
anp02032	Ap^r; *Anabaena* sp. chromosomal DNA from bp 5253901 to 5262374 in the BamHI site of pUC18
anp02580	Ap^r; *Anabaena* sp. chromosomal DNA from bp 4321373 to 4329600 in the BamHI site of pUC18
anp03035	Ap^r; *Anabaena* sp. chromosomal DNA from bp 4241824 to 4248722 in the BamHI site of pUC18
anp03337	Ap^r; *Anabaena* sp. chromosomal DNA from bp 1859242 to 1867412 in the BamHI site of pUC18
anp04001	Ap^r; *Anabaena* sp. chromosomal DNA from bp 4833192 to 4840653 in the BamHI site of pUC18
anp04055	Ap^r; *Anabaena* sp. chromosomal DNA from bp 1510438 to 1518550 in the BamHI site of pUC18
anp04490	Ap^r; *Anabaena* sp. chromosomal DNA from bp 1584150 to 1591582 in the BamHI site of pUC18
pGEM-T Easy	Ap^r; cloning vector (Promega Corp.)
pK18	Km^r; cloning vector (27)
pRL443	Ap^r Tc^r; conjugative plasmid (10)
pRL838	Cm^r Em^r; BAC vector (18) (GenBank accession no. AF403425)
pRL2801	Ap^r; internal fragment of *all3850* amplified by PCR using genomic DNA from *Anabaena* sp. as template and primers IDT147 (5′-TAAATTCTCAGGCAGCAGATGA-3′) and IDT148 (5′-TGTAACACCGATGATTCTGGAG-3′), cloned in pGEM-T Easy
pRL2807	Ap^r; internal fragment of *alr1728* amplified by PCR using genomic DNA from *Anabaena* sp. as template and primers IDT161 (5′-AATCTCGCTACTCGCTTACCAG −3′) and IDT162 (5′-GTTTAGCATGACGGGGTTTAAG-3′), cloned in pGEM-T Easy
pRL2816a	Ap^r Cm^r Em^r; PstI fragment containing C.CE3-*oriT* cassette from pRL2665b (17) transferred to the PstI site of pRL2801
pRL2821	Ap^r Cm^r Em^r; PstI fragment containing C.CE3-*oriT* cassette from pRL2665b (17) transferred to the PstI site of pRL2807
pRL2831a	Sm^r Sp^r; P_{glnA}-containing RSF1010 derivative (17, 40)
pRL2831b	Sm^r Sp^r; same as pRL2831a, but P_{glnA} oppositely oriented (17)
pRL2833a	Cm^r Em^r; P_{glnA}-containing pDU1 derivative (10)
pRL2833b	Cm^r Em^r; same as pRL2833a, but P_{glnA} oppositely oriented (10)
pRL2879	Sm^r Sp^r; *all1590*, downstream of *all1591*, was excised as an EaeI-RsaI fragment of anp03337 and transferred between PspOMI and StuI of pRL2831a
pRL2888	Sm^r Sp^r; *all4018*-containing XmnI-NheI fragment of anp04001 transferred between XbaI and StuI of pRL2831a
pRL3006	Sm^r Sp^r; *all3520*-containing HpaI-Acc65I fragment of anp03035 transferred between StuI and BsiWI of pRL2831a
pRL3007	Sm^r Sp^r; *all1338*-containing HincII-EaeI fragment of anp04490 transferred between StuI and PspOMI of pRL2831a

Continued on following page

TABLE 1—*Continued*

Strain or plasmid	Relevant characteristics and/or derivation[a]
pRL3008	Sm^r Sp^r; *all3582*-containing HindIII-HincII fragment of anp00676 transferred between HindIII and Ecl136II of pRL2831b
pRL3009	Sm^r Sp^r; *all4388*-containing *Nla*III-HincII fragment of anp02032 transferred between SphI and Ecl136II of pRL2831b
pRL3026	Sm^r Sp^r; *alr4311*-containing Eco47III-NlaIII fragment of anp00427 transferred between Ecl136II and SphI of pRL2831b
pRL3038	Sm^r Sp^r; *all4019*-containing SphI-HindIII fragment of anp04001 transferred between the same sites in pRL2831b
pRL3050	Sm^r Sp^r; *all3278*-containing NheI fragment of anc1006 transferred into BlnI of pRL2831a
pRL3106	Km^r; genes downstream of *all3582* were excised as an XcmI (blunted)-PstI fragment of anp02580 and cloned between SmaI and PstI of pK18
pRL3118	Cm^r Em^r; genes downstream of *all3582* were excised as a SacI-SphI fragment of pRL3106 and transferred between the same sites in pRL2833b
pRL3132	Km^r; *all4387*-containing SpeI-Cac8I fragment of anp02032 cloned between XbaI and SmaI of pK18
pRL3134	Cm^r Em^r; genes downstream of *all3520* were excised as a SapI (blunt)-BsrGI fragment of anc0756 and transferred between StuI and BsiWI of pRL2833a
pRL3142	Cm^r Em^r; *all4387*-containing PstI-Acc65I fragment of pRL3132 transferred between NsiI and BsiWI of pRL2833a
pRL3155	Ap^r Cm^r Em^r; genomic DNA of SR2821 was digested with HindIII, religated, and transformed to *E. coli*. Plasmids recovered on plates containing Cm contain pRL2821 and its neighboring genomic sequences
pRL3170	Cm^r Em^r; *all1591*-containing EaeI-PvuII fragment of anp03337 transferred between PspOMI and StuI of pRL2833a
pRL3171	Cm^r Em^r; *all3850*-containing PstI-DraI fragment of anp00920 transferred between NsiI and StuI of pRL2833a. This fragment extends 41 bp 3′ and 396 bp 5′ from all3850 (88 bp 3′ from the next upstream orf, asl3851)
pRL3172	Cm^r Em^r; *all4018*-containing NsiI-XhoI fragment of pRL2888 transferred between the same sites of pRL2833a
pRL3180	Sm^r Sp^r; *asl3849*-containing ApaI-XmnI fragment, which overlaps the 3′ ends of neighboring ORFs *asr3848* and *all3850*, of anp00920 transferred between ApaI and StuI of pRL2831a
pRL3810	Km^r Nm^r; the coliphage T7 terminator as a BamHI-BglII fragment of pET-3 (28) inserted into the BamHI site at the end of the polylinker of pRL3040a (40)
pRL3811	Km^r Nm^r; SacI-XhoI fragment of pRL3171 transferred between the SacI and XhoI sites of pRL3810
pUC18	Ap^r; cloning vector (43)

[a] Drug resistance is indicated by a superscript r. The drugs are abbreviated as follows: Ap, ampicillin; Bm, bleomycin; Cm, chloramphenicol; Em, erythromycin; Km, kanamycin; Nm, neomycin; Sm, streptomycin; Sp, spectinomycin; Tc, tetracycline. *Anabaena* sp., *Anabaena* sp. strain PCC 7120; BAC, bacterial artificial chromosome.

increased significantly after 3 and 8 h of nitrogen deprivation, whereas the expression of *alr3056* did not. Similarly, the expression of *hisC* (*alr2092*) increased after 8 h of N deprivation, and that of *all4966* did not change significantly.

alr1728. *alr1728* was intercepted by a transposon four times in our mutagenesis study and so it is unlikely to be a false-positive result. Because the gene downstream from *alr1728* is encoded on the opposite strand of DNA, if the phenotype-causing mutation was within *alr1728*, no polar effect of the transposon on the downstream gene would need to be considered. Strains FQ202 and FQ324 were complemented by plasmid anc1512 (Table 1), implying that if the mutation causing the phenotype is not in *alr1728* (bp 2077564 to 2078901), it is within 7 kb of that gene. Rather than continue with complementation, we took the alternative approach of reconstructing the mutation, in this instance by homologous recombination with an internal fragment of the gene. We thereby obtained mutant strain SR2821 (Table 1). To determine whether the mutational construction was actually within *alr1728*, genomic DNA of the mutant was cut with HindIII, diluted, recircularized by ligation, and transferred to *E. coli* by electroporation. Because pRL2821 includes a replication origin and antibiotic resistance determinants that function in *E. coli* and has sites for HindIII only within its internal fragment of *alr1728*, we could recover (as pRL3155) a portion of pRL2821 fused to a portion of the flanking region of insertion. We thereby determined that the insertion was within *alr1728*. SR2821 proved to be Fox^-, showing that *alr1728* is a Fox gene. Moreover, SR2821 shared with transposon mutant FQ202 (Fig. 1d) the structural phenotype that some but not all of its heterocysts were visibly open at one end.

Alr1728 is annotated as a hypothetical protein. The C-terminal 2/3 of the protein belongs to COG2339 (35), which includes membrane proteinases that regulate anti-sigma factors. Alr1728 may belong to a group of regulated intramembrane proteolytic (RIP) proteins (8) that posttranslationally activate membrane-bound transcription factors (MTF) (16). Ruanbao Zhou (S. Dakota State University) (personal communication) is currently testing whether Alr1728 may interact with Alr4305, a putative membrane-bound transcription factor. The presence of such a Fox gene in *Anabaena* sp. suggests that one or more RIP-MTF interactions may be important in heterocyst differentiation. (Due to an annotational error, Fox gene All0187 [11] might, but should not, be interpreted as an MTF [24].) The N-terminal 1/3 of Alr1728 contains a fork-

TABLE 2. Ten genes newly identified and/or validated as Fox genes and their known properties

Gene	Fig. 1 panel(s)	Annotation[a]	Morphological phenotype	Microarray result[b]	Comment(s)[c]
all1338	a and b	Hypothetical protein	Sometimes a gap between protoplast and envelope	No change	Protein conserved among cyanobacteria, no defined domains
all1591	c	Histidinol dehydrogenase	Heterocysts often open at end	Significant increase at 3 and 8 h	*hisD*; *alr3056* also annotated *hisD*: do heterocysts make their own His?
alr1728	d	Hypothetical protein with FHA domain	Some heterocysts are open at one end	Significant increase at 8 h	Regulatory, COG2339: *prsW* membrane proteinase, regulator of anti-sigma factor
all3278	e	Hypothetical protein	Open polar region in at least one end of the heterocyst	No significant change, very low expression	Ykud domain, enzymes that may be involved in cell wall synthesis
all3520	f	Unknown protein	Some heterocysts divide internally	Significant increase at 3 h	Only in heterocyst-forming cyanobacteria; signal peptide/transmembrane domain
all3582	g	Hypothetical protein	Heterocysts vacuolate	Very low expression	Signal peptide/two transmembrane domains; entire protein highly conserved in cyanobacteria, ca. 70 aa similar to cell division proteins in other bacteria
all3850	h	Unknown protein	Thick heterocyst envelope and sometimes polar bodies enlarged	Significant increase at 3 and 8 h	Peroxidase heme-binding signature at positions 339 to 349 appears only in cyanobacteria
all4019	i	G6P dehydrogenase	No abnormal structural phenotype	Significant increase at 3, 8, and 24 h	One of two G6PDs
alr4311	j	ATP-binding protein of ABC transporter	Some heterocysts are vacuolated	Very low expression	ABC transporter with no transmembrane domain. Likely an importer, functioning with Alr4309, Alr4310, or both
all4388	k	Hypothetical protein	Cell envelope not stained by Alcian Blue	Significant increase at 3, 8, and 24 h	Polysaccharide transporter

[a] FHA domain, forkhead-associated domain; G6P, glucose-6-phosphate; ABC, ATP-binding cassette.
[b] Microarray result for data in reference 7 analyzed by the method in reference 41. The strains were deprived of fixed nitrogen and studied after 3, 8, and 24 hours of nitrogen deprivation.
[c] aa, amino acids; G6PD, glucose-6-phosphate dehydrogenase.

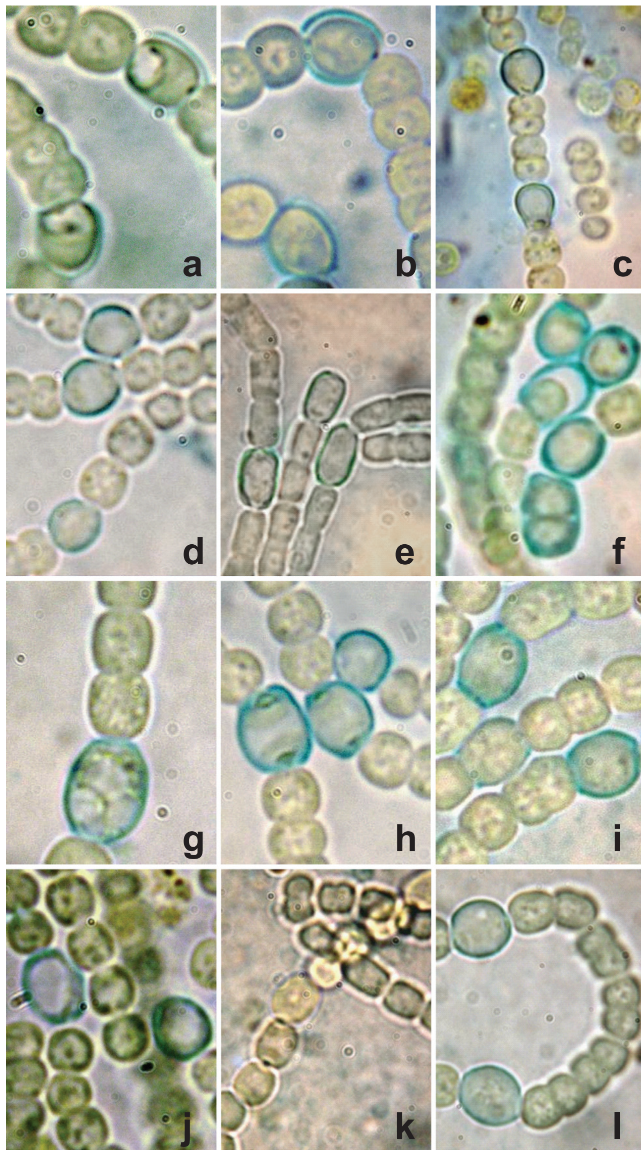

FIG. 1. Light micrographs of mutant filaments deprived of nitrogen on agar plates for 3 to 7 days and stained with Alcian Blue before microscopic observation. (a) Strain FQ1265 (*all1338*); (b) FQ1470 (*all1338*); (c) SR2813a (*all1591*); (d) FQ202 (*alr1728*); (e) FQ228 (*all3278*); (f) FQ211 (*all3520*); (g) FQ1580 (*all3582*); (h) SR2816a (*all3850*); (i) FQ406 (*all4019*); (j) FQ807 (*all4311*); (k) FQ1595 (*all4388*); (l) wild-type *Anabaena* sp. Aberrant overall coloration resulting from differing intensities of illumination was partially corrected with Adobe Photoshop CS5.

head-associated domain that may bind phosphoserine, phosphothreonine, or sometimes phosphotyrosine residues.

all3278. *all3278* was intercepted by a transposon 7 times in our mutagenesis study; its downstream gene is encoded on the opposite strand of DNA. Strain FQ228, mutated in this open reading frame (ORF), was complemented by pRL3050, implying that it is a Fox gene. When deprived of fixed nitrogen, heterocysts of FQ228 appear to have an open polar region in at least one end (Fig. 1e). All3278 is annotated as a hypothetical protein. Its product resembles Ykud domain enzymes (pfam03734), and its orthologs in *N. punctiforme* and

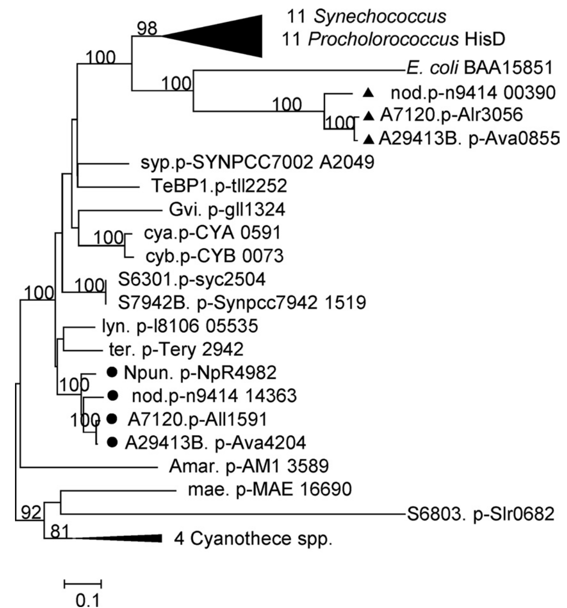

FIG. 2. Phylogenetic analysis of 46 HisD sequences from cyanobacteria. The small black circles represent orthologs of All1591 in heterocyst-forming cyanobacteria, and the small black triangles represent their paralogs in the same organisms. HisD from *E. coli* (GenBank accession no. BAA15851) was used as a reference. Strain abbreviations: A7120, *Anabaena* sp.; A29413B, *Anabaena variabilis* ATCC 29413; Amar, *Acaryochloris marina* MBIC11017; cya, *Synechococcus* sp. OS type a′; cyb, *Synechococcus* sp. OS type b′; Gvi, *Gloeobacter violaceus* PCC 7421; lyn, *Lyngbya* sp. strain PCC 8106; mae, *Microcystis aeruginosa* NIES843; nod, *Nodularia spumigena*; Npun, *Nostoc punctiforme* ATCC 29133; S6803, *Synechocystis* sp. strain PCC 6803; S7942B, *Synechococcus elongatus* PCC 7942; syp, *Synechococcus* sp. strain PCC 7002; TeBP1, *Thermosynechococcus elongatus* bp1; ter, *Trichodesmium erythraeum* IMS101. The bar indicates 0.1 substitution per site.

Anabaena variabilis are annotated as ErfK/YbiS/YcfS/YnhG proteins. ErfK-like proteins are proteins that are involved in cell wall synthesis; hence, this gene product may be required for the cell wall reorganization that is thought to take place during heterocyst differentiation (19, 42, 45).

all3520. *all3520* was intercepted 3 times in our experiments. It encodes an unknown protein that is unique to heterocyst-forming cyanobacteria and contains a possible transmembrane helix or signal peptide domain. The second in a sequence of 7 ORFs that have the same orientation, it begins only 20 bp 3′ from *all3521*, and so is presumably cotranscribed with that gene. *all3519*, 224 bp 3′ from *all3520*, may have its own promoter. pRL3006, whose only intact ORF is *all3520*, complemented FQ211, whereas pRL3134, which bears the downstream genes, failed to complement FQ211. Therefore, *all3520* appears to be a Fox gene. Remarkably, some heterocysts of nitrogen-deprived FQ211 appeared to divide internally (Fig. 1f, lower right). In microarray experiments, the gene was upregulated at 3 h of nitrogen deprivation (data in reference 7 analyzed by the method in reference 41). No functional domains could be identified using INTERPRO or Pfam searches, but a low similarity to bacterial surface proteins is seen in an iterative PSI-BLAST search (1). The relationship between the unusual mutant phenotype of *all3520* and the protein it encodes is unclear.

all3582. *all3582*, intercepted 5 times in our experiments, encodes a hypothetical protein that is highly conserved in many cyanobacteria. It contains a signal peptide or a transmembrane domain at its N terminus. The C terminus of the protein has similarity to the glyoxalase superfamily of proteins (cl00411) that contains some signal transduction and cell division proteins. The gene is the first in a sequence of 5 similarly oriented genes. FQ1580, mutated in this gene, was complemented by pRL3008, a replicating plasmid whose only intact ORF is *all3582*, and was not complemented by pRL3118, a plasmid that bears the four downstream genes, implying that *all3582* is a Fox gene. Heterocysts of FQ1580 appear vacuolated and sometimes granular (Fig. 1g).

all3850. Intercepted by a transposon 5 times in our experiments, *all3850* encodes an unknown protein that is conserved in cyanobacteria but is otherwise unique. Plasmids pRL3171 and pRL3811 bear *all3850*, and pRL3180 bears the downstream gene, *asl3849*. Because *all3850* mutant FQ384 was not complemented by pRL3171, pRL3180, or both plasmids, the mutation was reconstructed by insertional mutation with pRL2816a. In the resulting Fox^- mutant, SR2816a, the heterocyst envelope appears unusually thick, and the polar bodies are sometimes quite enlarged (Fig. 1h). SR2816a was also not complemented by *all3850* clone pRL3811, pRL3180, or both plasmids. It is unclear why neither *all3850* mutation was complemented. All3850 has a peroxidase heme-binding signature at positions 339 to 349 (ProSite PS00435) that includes a conserved His residue that is involved in binding the ligand. After nitrogen deprivation, expression of *all3850* is upregulated at 3 and 8 h, but not at 24 h, compared to a culture maintained with nitrate (data in reference 7 analyzed by the method in reference 41).

all4019. Intercepted 11 times in our transposon mutagenesis experiments, *all4019* is one of two genes in *Anabaena* sp. that encodes a glucose-6-phosphate dehydrogenase (G6PD). Plasmid pRL3038 bears *all4019* as its only gene, and compatible plasmid pRL3172 bears *all4018*, the next gene downstream from *all4019*, as its only intact gene. Strain FQ406, mutated in *all4019*, was complemented by pRL3038 with or without pRL3172, whereas pRL3172 alone failed to complement FQ406. We conclude that if *all4018* is a Fox gene, it is not the sole Fox gene affected by the FQ406 mutation, implying that *all4019* is a Fox gene. *all4018* is orthologous to *opcA*, a gene that in *Nostoc punctiforme* is cotranscribed along with *zwf*. In *N. punctiforme*, both genes are required for a Fox^+ phenotype, and *opcA* is required for the activity of G6PD, of which it is an allosteric effector (14, 33). It is likely, then, that in *Anabaena* sp. both genes, *all4019* and *all4018*, are required for heterocyst function. Although the transcript level of *all4019* was significantly higher 3, 8, and 24 h after nitrogen deprivation compared to a nitrate-grown culture (data in reference 7 analyzed by the method in reference 41), the mutant shows no unusual morphological phenotype (Fig. 1i). Interestingly, not only is *zwf*, an ortholog of *all4019*, required for heterocyst differentiation in *N. punctiforme* (32, 33), but akinete-like cells can also differentiate more abundantly in *zwf* mutants than in wild-type *N. punctiforme* (4).

alr4311. *alr4311*, intercepted twice in our transposon mutagenesis experiments, encodes a putative ATP-binding cassette (ABC) protein. The gene is the last in a presumptive operon, downstream from three ORFs that encode hypothetical proteins. It is also immediately downstream from *nrrA* (*all4312*), a gene that on the opposite strand of DNA encodes for a nitrogen-responsive response regulator that is important in heterocyst differentiation (6, 7, 25). Strain FQ807, mutated in *alr4311*, was complemented by pRL3026, a plasmid that carries *alr4311* as its only intact ORF, implying that *alr4311* is a Fox gene. Heterocysts of FQ807 appear normal or (as in Fig. 1j) vacuolated. Although Alr4311 itself contains no predicted transmembrane domains, analysis of the predicted products, Alr4309 and Alr4310, of its two upstream genes show that they contain 4 and 3 transmembrane helices, respectively, suggesting that these three proteins may jointly compose a transporter (15). Because exporters normally contain the transmembrane domain and the ABC on a single polypeptide and ABC proteins that are not involved in transport normally contain two fused ABC domains, it is likely that Alr4311 with Alr4309 and/or Alr4310 function as an importer (5). Even if this were true, what they may transport is obscure.

all4388. Intercepted by a transposon as described earlier (α21 [9, 22]) and intercepted 5 times in our more recent transposon mutagenesis experiments, *all4388* encodes a putative polysaccharide transporter that was described as a Fox gene without the corresponding mutant having been complemented or reconstructed (22). As described in reference 22 for mutant α21 (Fig. 1l), the envelopes of heterocysts of FQ1595, which appear to retain a glycolipid layer, lack a polysaccharide layer according to light microscopy and staining with Alcian Blue. FQ1595 was complemented by pRL3009, whose only intact *Anabaena* sp. ORF is *all4388*. It was not complemented by pRL3142, which bears the downstream gene, *all4387*. These results support the proposition (22) that *all4388* is a Fox gene. Expression of *all4388* was upregulated 3, 8, and 24 h after nitrogen deprivation, whereas many of the genes required for deposition of HEP (17, 37) are first significantly upregulated at 8 h (data in reference 7 analyzed by the method in reference 41).

In summary, we have documented that 10 genes are required for heterocyst differentiation or function in *Anabaena* sp. These genes may be classified into three functional groups. (i) Presumptive regulatory genes include *alr1728*, whose sequence is similar to the sequences of known intramembrane proteolytic proteins, and *all3520*, which normally prevents internal division of heterocysts. (ii) Genes that are involved in the formation of structural elements of heterocysts include *all3278*, which may be involved in cell wall reconstruction; *all4388*, which is required for HEP deposition; and *all1338*, which may affect deposition of the heterocyst envelope. (iii) Genes that are involved in specialized heterocyst functions include *all1591*, presumably required for histidine biosynthesis in heterocysts; *all3850*, whose presumptive peroxidase heme-binding signature may be involved in sensing of reactive oxygen species or of O_2 or in other redox reactions required for heterocyst function; *all4019*, encoding glucose-6-phosphate dehydrogenase, thought to be required in heterocysts for the oxidative pentose phosphate cycle and the transfer of electrons from sugar to respiration or to assimilation of N_2 (33); and *alr4311*, which may encode an importer of an unidentified ligand. The role of *all3582* remains obscure.

ACKNOWLEDGMENT

This work was supported by the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy grant DOE FG02-91ER20021.

REFERENCES

1. **Altschul, S. F., et al.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Black, K., W. J. Buikema, and R. Haselkorn.** 1995. The *hglK* gene is required for localization of heterocyst-specific glycolipids in the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **177:**6440–6448.
3. **Buikema, W. J., and R. Haselkorn.** 1991. Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. Genes Dev. **5:**321–330.
4. **Campbell, E. L., M. L. Summers, H. Christman, M. E. Martin, and J. C. Meeks.** 2007. Global gene expression patterns of *Nostoc punctiforme* in steady-state dinitrogen-grown heterocyst-containing cultures and at single time points during the differentiation of akinetes and hormogonia. J. Bacteriol. **189:**5247–5256.
5. **Davidson, A. L., E. Dassa, C. Orelle, and J. Chen.** 2008. Structure, function, and evolution of bacterial ATP-binding cassette systems. Microbiol. Mol. Biol. Rev. **72:**317–364.
6. **Ehira, S., and M. Ohmori.** 2006. NrrA directly regulates expression of *hetR* during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **188:**8520–8525.
7. **Ehira, S., and M. Ohmori.** 2006. NrrA, a novel nitrogen-responsive response regulator facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. Mol. Microbiol. **59:**1692–1703.
8. **Ellermeier, C. D., and R. Losick.** 2006. Evidence for a novel protease governing regulated intramembrane proteolysis and resistance to antimicrobial peptides in *Bacillus subtilis*. Genes Dev. **20:**1911–1922.
9. **Ernst, A., et al.** 1992. Synthesis of nitrogenase in mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120 affected in heterocyst development or metabolism. J. Bacteriol. **174:**6025–6032.
10. **Fan, Q., et al.** 2005. Clustered genes required for synthesis and deposition of envelope glycolipids in *Anabaena* sp. strain PCC 7120. Mol. Microbiol. **58:** 227–243.
11. **Fan, Q., et al.** 2006. Signal transduction genes required for heterocyst maturation in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **188:**6688–6693.
12. **Felsenstein, J.** 1992. Estimating effective population size from samples of sequences: a bootstrap Monte Carlo integration method. Genet. Res. **60:** 209–220.
13. **Flores, E., et al.** 2007. Septum-localized protein required for filament integrity and diazotrophy in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **189:**3884–3890.

13a. **Gantar, M., J. Elhai, J. Jia, and M. Ow.** 1995. Abstr. 5th Cyanobacterial Mol. Biol. Workshop, p. 25.

14. **Hagen, K. D., and J. C. Meeks.** 2001. The unique cyanobacterial protein OpcA is an allosteric effector of glucose-6-phosphate dehydrogenase in *Nostoc punctiforme* ATCC 29133. J. Biol. Chem. **276:**11477–11486.
15. **Higgins, C. F., et al.** 1986. A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. Nature **323:**448–450.
16. **Hoppe, T., M. Rape, and S. Jentsch.** 2001. Membrane-bound transcription factors: regulated release by RIP or RUP. Curr. Opin. Cell Biol. **13:**344–348.

16a. **Hu, N. T., T. Thiel, T. H. Giddings, and C. P. Wolk.** 1981. New *Anabaena* and *Nostoc* cyanophages from sewage settling ponds. Virology **114:**236–246.

17. **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.
18. **Kaneko, T., et al.** 2001. Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. DNA Res. **8:**205–213.
19. **Leganés, F., et al.** 2005. Wide variation in the cyanobacterial complement of presumptive penicillin-binding proteins. Arch. Microbiol. **184:**234–248.
20. **Liang, J., L. Scappino, and R. Haselkorn.** 1992. The *patA* gene product, which contains a region similar to CheY of *Escherichia coli*, controls heterocyst pattern formation in the cyanobacterium *Anabaena* 7120. Proc. Natl. Acad. Sci. U. S. A. **89:**5655–5659.
21. **Lynn, M. E., J. A. Bantle, and J. D. Ownby.** 1986. Estimation of gene expression in heterocysts of *Anabaena variabilis* by using DNA-RNA hybridization. J. Bacteriol. **167:**940–946.
22. **Maldener, I., S. Hannus, and M. Kammerer.** 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.
23. **Masepohl, B., K. Schölisch, K. Görlitz, C. Kutzki, and H. Böhme.** 1997. The heterocyst-specific *fdxH* gene product of the cyanobacterium *Anabaena* sp. PCC 7120 is important but not essential for nitrogen fixation. Mol. Gen. Genet. **253:**770–776.
24. **Mella-Herrera, R. A., M. R. Neunuebel, and J. W. Golden.** 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.
25. **Muro-Pastor, A. M., E. Olmedo-Verd, and E. Flores.** 2006. All4312, an NtcA-regulated two-component response regulator in *Anabaena* sp. strain PCC 7120. FEMS Microbiol. Lett. **256:**171–177.
26. **Nayar, A. S., H. Yamaura, R. Rajagopalan, D. D. Risser, and S. M. Callahan.** 2007. FraG is necessary for filament integrity and heterocyst maturation in the cyanobacterium *Anabaena* sp. strain PCC 7120. Microbiology **153:**601–607.
27. **Pridmore, R. D.** 1987. New and versatile cloning vectors with kanamycin-resistance marker. Gene **56:**309–312.
28. **Rosenberg, A., et al.** 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56:**125–135.
29. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406–425.
30. **Schwarz, R., and M. Dayhoff.** 1979. Matrices for detecting distant relationships, p. 353–358. *In* M. Dayhoff (ed.), Atlas of protein sequences. National Biomedical Research Foundation, New York, NY.
31. **Shi, L., et al.** 2007. Two genes encoding protein kinases of the HstK family are involved in synthesis of the minor heterocyst-specific glycolipid in the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **189:**5075–5081.
32. **Summers, M. L., and J. C. Meeks.** 1996. Transcriptional regulation of *zwf*, encoding glucose-6-phosphate dehydrogenase, from the cyanobacterium Nostoc punctiforme strain ATCC 29133. Mol. Microbiol. **22:**473–480.
33. **Summers, M. L., J. G. Wallis, E. L. Campbell, and J. C. Meeks.** 1995. Genetic evidence of a major role for glucose-6-phosphate dehydrogenase in nitrogen fixation and dark growth of the cyanobacterium *Nostoc* sp. strain ATCC 29133. J. Bacteriol. **177:**6184–6194.
34. **Tamura, K., J. Dudley, M. Nei, and S. Kumar.** 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. **24:**1596–1599.
35. **Tatusov, R. L., et al.** 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. **29:**22–28.
36. **Valladares, A., A. Herrero, D. Pils, G. Schmetterer, and E. Flores.** 2003. Cytochrome c oxidase genes required for nitrogenase activity and diazotrophic growth in *Anabaena* sp. PCC 7120. Mol. Microbiol. **47:**1239–1249.
37. **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.
38. **Wolk, C. P.** 2000. Heterocyst formation in *Anabaena*, p. 83–104. *In* Y. V. Brun and L. J. Shimkets (ed.), Prokaryotic development. ASM Press, Washington, DC.
39. **Wolk, C. P., A. Ernst, and J. Elhai.** 1994. Heterocyst metabolism and development, p. 769–823. *In* D. A. Bryant (ed.), The molecular biology of cyanobacteria. Kluwer Academic Publishers, Dordrecht, Netherlands.
40. **Wolk, C. P., et al.** 2007. Paired cloning vectors for complementation of mutations in the cyanobacterium *Anabaena* sp. strain PCC 7120. Arch. Microbiol. **188:**551–563.
41. **Xu, X., J. Elhai, and C. P. Wolk.** 2008. Transcriptional and developmental responses by *Anabaena* to deprivation of fixed nitrogen, p. 383–422. *In* A. Herrero and E. Flores (ed.), Genomics and molecular biology of cyanobacteria. Horizon Scientific Press, Norwich, United Kingdom.
42. **Xu, X., I. Khudyakov, and C. P. Wolk.** 1997. Lipopolysaccharide dependence of cyanophage sensitivity and aerobic nitrogen fixation in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **179:**2884–2891.
43. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
44. **Yoon, H.-S., and J. W. Golden.** 2001. PatS and products of nitrogen fixation control heterocyst pattern. J. Bacteriol. **183:**2605–2613.
45. **Zhu, J., et al.** 2001. HcwA, an autolysin, is required for heterocyst maturation in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **183:**6841–6851.