Anabaena sp. Strain PCC 7120 *hetY* Gene Influences Heterocyst Development

Ho-Sung Yoon,† Martin H. Lee, Jin Xiong, and James W. Golden*

Department of Biology, Texas A&M University, College Station, Texas 77843-3258

The filamentous cyanobacterium *Anabaena* (*Nostoc*) sp. strain PCC 7120 responds to starvation for fixed nitrogen by producing a semiregular pattern of nitrogen-fixing cells called heterocysts. Overexpression of the *hetY* gene partially suppressed heterocyst formation, resulting in an abnormal heterocyst pattern. Inactivation of *hetY* increased the time required for heterocyst maturation and caused defects in heterocyst morphology. The 489-bp *hetY* gene (alr2300), which is adjacent to *patS* (asl2301), encodes a protein that belongs to a conserved family of bacterial hypothetical proteins that contain an ATP-binding motif.

Anabaena (*Nostoc*) sp. strain PCC 7120 is a filamentous cyanobacterium capable of both photosynthesis and nitrogen fixation. In an environment lacking a combined nitrogen source, approximately 10% of the photosynthetic vegetative cells terminally differentiate into nitrogen-fixing heterocysts at regular intervals along filaments (11). Differentiation requires approximately 18 to 24 h. Mature heterocysts exhibit a number of distinctive morphological, biochemical, and genetic characteristics that differ from their vegetative-cell progenitors. Heterocysts provide nitrogen fixation products to neighboring vegetative cells and in turn receive fixed carbon produced from photosynthesis. The regular spacing of heterocysts along the chain of vegetative cells presumably has evolved to permit efficient distribution of these metabolites within a long filament.

Multicellular development requires integration and coordination of complex environmental and internal regulatory signals. A number of genes have been identified that are involved in the regulation of heterocyst development and pattern formation (11, 18). Our laboratory has shown that the *patS* gene, which encodes a small peptide, is important for normal heterocyst pattern formation (19, 20). It is thought that PatS works by lateral inhibition, such that PatS peptide produced by a differentiating cell inhibits the differentiation of its neighbors to establish a pattern of single heterocysts along chains of vegetative cells.

We report here a genetic analysis of a gene, designated *hetY*, which is located downstream of and divergently transcribed to *patS*, and we show that *hetY* influences heterocyst development.

Identification of the *hetY* gene. Our investigators previously described a cosmid library clone, 8E11, that suppresses heterocyst development, and we localized this activity to a 3.3-kb region present in subclone pAM1035 (Fig. 1) (19). We identified the *patS* gene on subclone pAM1882 as responsible for the

strong heterocyst inhibition phenotype (Fig. 1) (19). However, additional analysis of the pAM1035 insert showed that a region adjacent to *patS* independently interfered with heterocyst formation and produced an abnormal pattern when introduced into *Anabaena* stain PCC 7120 cells on shuttle vectors (Fig. 1). Plasmid pAM1244 contains a *Sca*I-*Xba*I fragment, with an Ω Sp^r/Sm^r (spectinomycin and streptomycin resistance) cassette obtained from pAM684 (14) and inserted into the *Pac*I site to facilitate cloning, cloned into the shuttle plasmid pCCB110 (17) (Fig. 1). Exconjugants containing pAM1244 were capable of diazotrophic growth on BG-11$_0$ medium, which lacks combined nitrogen, but produced heterocysts at a reduced frequency compared to the wild type. Growth of *Anabaena* strain PCC 7120 and genetic techniques followed standard procedures (4, 5, 10).

The pAM1244 insert contains open reading frame (ORF) alr2300, which was named *hetY*. To determine if expression of the HetY product from extra-copy *hetY* was responsible for the heterocyst suppression phenotype, we disrupted the *hetY* ORF. An Ω Sp^r/Sm^r cassette from pAM684 was inserted into the unique *Acc*I site in pAM1035, interrupting *hetY*, to produce pAM1301. The fragment carrying the inactivated *hetY* was inserted into shuttle vector pCCB110 to make pAM1361. In contrast to pAM1244, pAM1361 had no effect on the production of heterocysts.

The complete *hetY* ORF is contained on a *Sca*I-*Taq*I fragment internal to the pAM1244 insert. This smaller fragment was cloned into shuttle vector pAM832 (a higher-copy-number variant of pCCB110 [J. Golden, unpublished data]) to make pAM1232. This construct, like pAM1244, caused a marked reduction in heterocyst frequency. To determine if higher expression of *hetY* would result in complete heterocyst suppression, we constructed a *hetY* overexpression vector by cloning the *Sca*I-*Taq*I fragment into pAM692 to make pAM1202. Plasmid pAM692 contains tandem *rbcL* and *glnA* promoters in shuttle vector pAM504 (17), and it produces high expression levels for cloned inserts (T. S. Ramasubramanian, unpublished data). After introduction into *Anabaena* strain PCC 7120, pAM1202 produced the same phenotype as pAM1232 (Fig. 2), indicating that expression of *hetY* was not limiting for the weak heterocyst suppression phenotype. A control construct con-

* Corresponding author. Mailing address: Department of Biology, Texas A&M University, 3258 TAMU, College Station, TX 77843-3258. Phone: (979) 845-9823. Fax: (979) 845-2891. E-mail: jgolden@tamu.edu.

† Present address: Institute of Agricultural Science and Technology, Kyungpook National University, Daegu, South Korea.
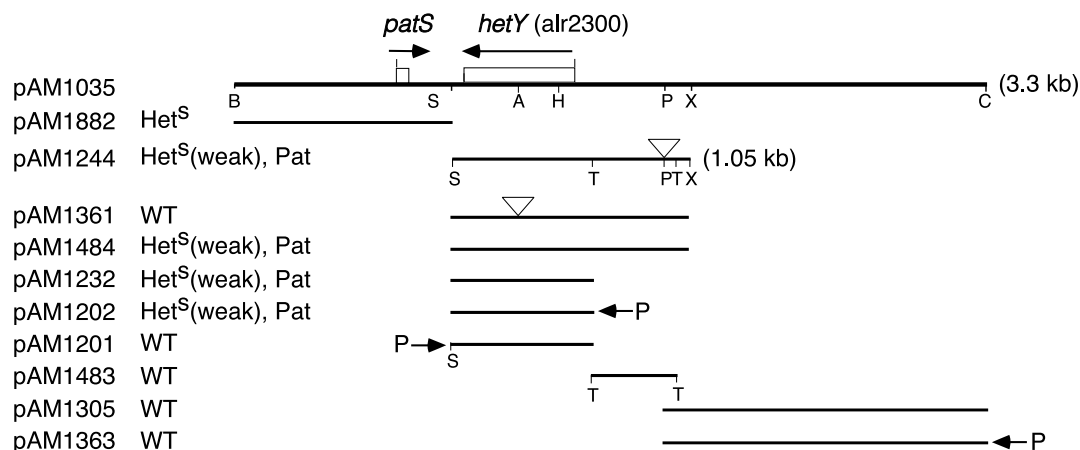
FIG. 1. Heterocyst suppression by DNA fragments subcloned from cosmid 8E11. Each shuttle plasmid was conjugated into *Anabaena* strain PCC 7120, and the heterocyst phenotype was determined. Phenotypes were scored as wild type (WT), heterocyst suppression (Het[s]), and abnormal heterocyst pattern (Pat). Only selected restriction sites are shown. Symbols: arrows over ORF, direction of transcription; P with arrow, external promoter; inverted triangle, insertion of Ω Sp[r]/Sm[r] cassette; A, *Acc*I; B, *Bam*HI; C, *Cla*I; H, *Hinc*II; P, *Pac*I; S, *Sca*I; T, *Taq*I; X, *Xba*I.

taining *hetY* in the reverse orientation relative to the external promoters, pAM1201, did not affect heterocyst development.

DNA sequences upstream of *hetY* on pAM1244 and pAM1035 did not affect heterocyst development (Fig. 1). The
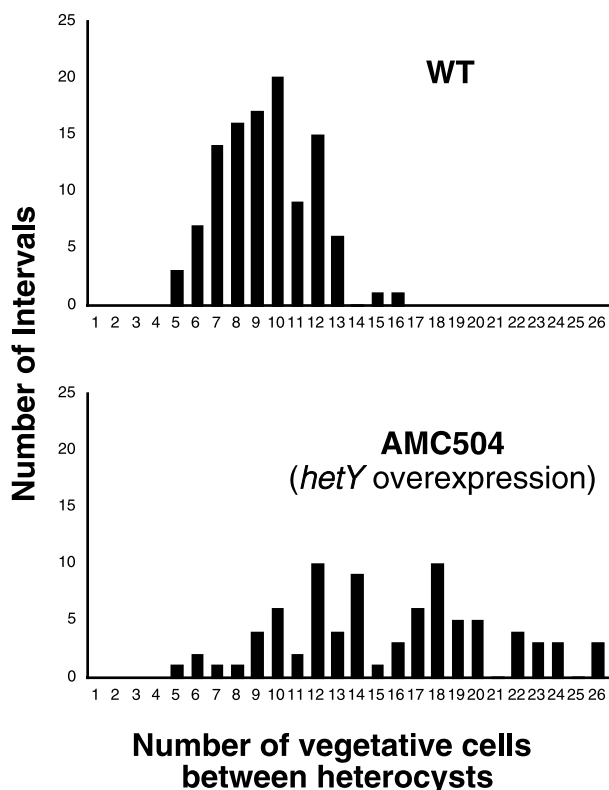

FIG. 2. Overexpression of *hetY* partially suppressed heterocyst formation, resulting in an abnormal pattern. The wild type and strain AMC504, which contains *hetY* overexpressed from pAM1202, were grown in liquid BG-11 medium and induced to form heterocysts by transfer to BG-11$_0$ medium. The heterocyst pattern was determined by microscopic examination of filaments 26 h after induction as previously described (20). The percentage of heterocysts in filaments was 10.6% for the wild type and 6.4% for AMC504.

*Taq*I fragment from pAM1244 was cloned into pAM504 to make pAM1483. *Anabaena* strain PCC 7120 cells containing pAM1483 showed normal heterocyst development. Similarly, no heterocyst suppression activity was found in the distal third of the pAM1035 fragment, between the *Pac*I and *Cla*I sites. This was tested with two constructs. For pAM1305, all other sequences were deleted from the original cosmid, 8E11. First, all internal *Cla*I fragments were deleted from 8E11 to make pAM1007. A *Cla*I-*Sal*I fragment was then deleted from pAM1007 to make pAM1023, which eliminated sequences to the right of the map shown in Fig. 1. Finally, the *Bam*HI-*Pac*I fragment on the left was deleted from pAM1023 to make pAM1305 (Fig. 1). The second construct, pAM1363, contained the *Cla*I-*Pac*I fragment from pAM1035 cloned into pAM743, which contains a 270-bp *glnA* promoter fragment cloned into the shuttle plasmid pAM504. Both pAM1305 and pAM1363 had no effect on heterocyst development when introduced into *Anabaena* strain PCC 7120 (Fig. 1).

***hetY* overexpression phenotype.** *hetY* partially suppressed heterocyst development when it was present in extra copies on shuttle plasmids pAM1244 and pAM1232 and when it was overexpressed from an external promoter on pAM1202 (strain AMC504) (Fig. 1 and 2). The average number of vegetative cells between heterocysts increased, while the percentage of heterocysts in filaments was reduced from 10.6% in the wild type to 6.4% in the *hetY*-overexpressing strain AMC504 (Fig. 2). After several days of growth in BG-11$_0$ medium, strains overexpressing *hetY* displayed a lighter green color than wild-type cultures, which suggests nitrogen limitation. However, their growth rate was not obviously impaired.

***hetY* inactivation.** We interrupted the *hetY* ORF to determine the heterocyst phenotype of a loss-of-function mutant. The *Bam*HI-*Cla*I insert from pAM1301, which carries an Ω Sp[r]/Sm[r] cassette inserted into *hetY* (see above), was inserted into the *sacB*-containing suicide vector pRL271 (1) to make pAM1302. Plasmid pAM1302 was transferred into *Anabaena* strain PCC 7120 by conjugation, and one of the resulting single recombinants (AMC279) was selected for the isolation of double recombinants. Standard procedures were used to obtain
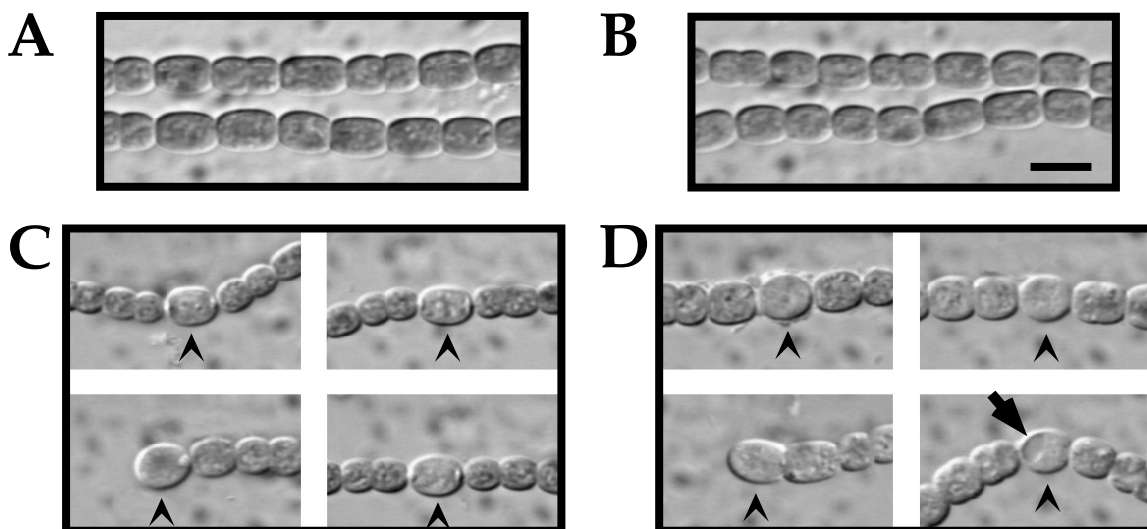
FIG. 3. Inactivation of *hetY* produced immature and abnormal heterocysts, as shown in photomicrographs of wild-type *Anabaena* strain PCC 7120 (A and C) and *hetY*-inactivated strain AMC642 (B and D) taken before (A and B) and 24 h after (C and D) heterocyst induction. Arrowheads mark heterocysts that are mature in the wild type (C) but immature (lacking polar cyanophycin granules) and abnormal in the *hetY* mutant (D). The arrow indicates defective vacuolated heterocyst morphology occasionally produced by strain AMC642. Bar, 5 μm.

double recombinants (3, 17). Two *hetY* mutant strains, AMC642 and AMC643, were confirmed by Southern blot analysis.

Strains AMC642 and AMC643 were morphologically and phenotypically indistinguishable. Both strains grew normally and had normal cell morphology when grown on BG-11 medium, which contains nitrate. However, after transfer to BG-11_0 medium, they showed a delay in heterocyst development and occasional defects in morphology. In contrast to the wild type, the *hetY* mutant strains had produced only immature proheterocysts by 24 h after nitrogen step-down (Fig. 3). Proheterocysts have a thickened cell envelope, lighter yellow-green color, and their cytoplasm appears less granular than vegetative cells. Mature heterocysts are identified microscopically by the presence of refractive cyanophycin granules at their poles that contact a neighboring vegetative cell. In the *hetY* mutants at 24 h, the polar cyanophycin granules were either absent or diminished in size. Additionally, the *hetY* mutant heterocysts occasionally showed a protoplast defect (Fig. 3). At 48 h after induction, some heterocysts had matured but many were still lacking normal polar granules. The cultures also displayed signs of nitrogen starvation, including less growth and a yellowish color compared to the wild type. We also observed that filaments of the mutant strains became more fragmented than those of the wild type in older cultures. At 24 h after nitrogen step-down, the pattern and frequency of immature heterocysts in the *hetY* mutant AMC642 were similar to those of mature heterocysts in the wild-type strain (Fig. 4).

We attempted to complement the *hetY* inactivation strains AMC642 and AMC643 with plasmid pAM1484 (Fig. 1). This plasmid contains a *Sca*I-*Xba*I fragment with *hetY* and its upstream region in conjugal vector pAM504. Plasmid pAM1484 was transferred into strains AMC642 and AMC643 by conjugation, and exconjugants were analyzed for their heterocyst development phenotype. On plates lacking fixed nitrogen, the frequency of heterocysts was less than that of the wild type,

which is the dominant *hetY* overexpression phenotype. Apparently, the elevated copy number of plasmid-borne *hetY* in this new construct leads to the same *hetY* overexpression phenotype produced by our other *hetY* constructs, which makes the complementation results ambiguous.

The growth defects observed for the *hetY* mutants were further examined by following the growth of cultures after nitrogen step-down. We compared the growth of three independently derived *hetY* strains against the wild type. The *hetY* mutant strains showed a growth lag that was approximately 2 days longer than the wild type, which is consistent with their slower differentiation of mature heterocysts (Fig. 5). However, once the mutant strains started growing, they showed approximately the same normal polynomial growth rate as the wild type. Batch cultures of cyanobacteria grown photoautotrophically do not show an exponential growth phase, presumably due to self-shading. HetY is apparently required for the normal time course of de novo heterocyst development, but once formed, the heterocysts function sufficiently well to support nearly normal growth.

We conclude that normal heterocyst development requires appropriate levels of *hetY* expression. HetY is not essential for heterocyst development or normal pattern formation, but its absence results in slow and sometimes abnormal differentiation. Overexpression of *hetY* caused fewer cells to differentiate, apparently by a dominant negative effect. This could be because of increased HetY activity itself or from interactions between excess HetY protein and other components in the regulatory pathway. However, the *hetY* overexpression phenotype could result from an aberrant response unrelated to HetY's normal function.

***hetY* RNA blot analysis.** The expression pattern of *hetY* upon heterocyst induction was examined by northern RNA blot analysis. RNA samples (20 μg) from differentiating filaments were prepared, separated by electrophoresis, and blotted as described previously (17). The blot was hybridized with a
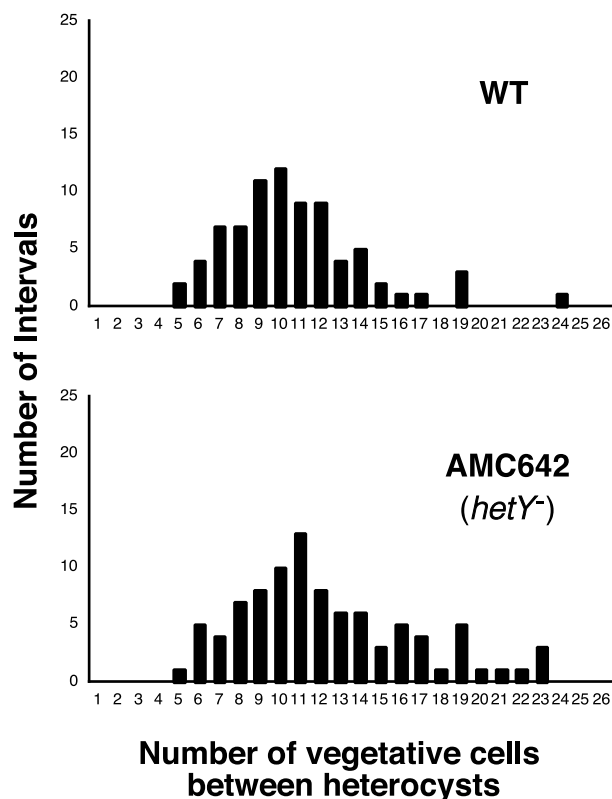
FIG. 4. Inactivation of *hetY* did not affect heterocyst pattern, as shown here by the heterocyst pattern in *hetY* inactivation strain AMC642. The wild type and *hetY* mutant AMC642 were grown, induced, and scored 24 h after induction as described in the legend for Fig. 2. At that time, all heterocysts of AMC642 were immature and lacked distinct polar cyanophycin granules at the cell poles.

strand-specific probe generated with a MAXIscript kit (Ambion). *hetY* message was detected in total RNA extracted from vegetative cells grown with nitrate and at 6 and 12 h after nitrogen step-down, but it was absent in samples from 18, 24, 30, and 36 h (data not shown). *hetY* transcripts were between 700 and 1,000 bp in vegetative cell RNA. However, the *hetY* message detected at 6 and 12 h after nitrogen step-down was

substantially degraded. The temporal expression pattern of *hetY* is different from that of *patS*, which shows increasing levels of expression during heterocyst induction (19). It also differed from that of another flanking gene, alr2298, which showed a low amount of message in nitrate-grown cells, none at 6 h, and then increasing amounts between 12 and 30 h before decreasing again at 36 h (data not shown). The *hetY* expression profile is consistent with HetY playing a role in vegetative cells as well as being involved in early heterocyst development.

***hetY* bioinformatics.** GenBank searches showed that the predicted product of the 489-bp *hetY* gene had extensive sequence similarity to a conserved family of small nucleotide-binding proteins of unknown function from diverse bacteria. HetY shows 30 to 55% deduced amino acid sequence identity and similar length to this family of proteins from various gram-positive and gram-negative bacteria.

Protein sequence motifs were identified by the PPSearch program (http://www.ebi.ac.uk/ppsearch/) (2). The protein sequence was also searched against a database of Clusters of Orthologous Groups (COG) using the program COGnitor (http://www.ncbi.nlm.nih.gov/COG/xognitor.html). HetY sequence was further used in an exhaustive search against a database for protein structures (PDB) using the Smith-Waterman search algorithm (15). Multiple sequence alignment was carried out using the T-Coffee program (http://www.ch.embnet .org/software/TCoffee.html) (12).

Analysis of HetY using the COG database, which compares consensus phylogenetic sequence patterns, suggested it could be an ATPase or kinase. A search against PDB resulted in a significant hit (E value, 8e−13) with YjeE, an ATPase from *Haemophilus influenzae*, which has been implicated in cell wall biosynthesis (16). HetY and YjeE showed 31.6% identity and 62.4% similarity. To assess the statistical significance of sequence similarities between HetY and YjeE, we used a randomization test program, PRSS (13), through the Biology Workbench web server (http://workbench.sdsc.edu/). The PRSS statistical test, without influence of the alignment, further verified that the two sequences are significantly similar, with a P value of 9.43e−13. Both proteins have an identical
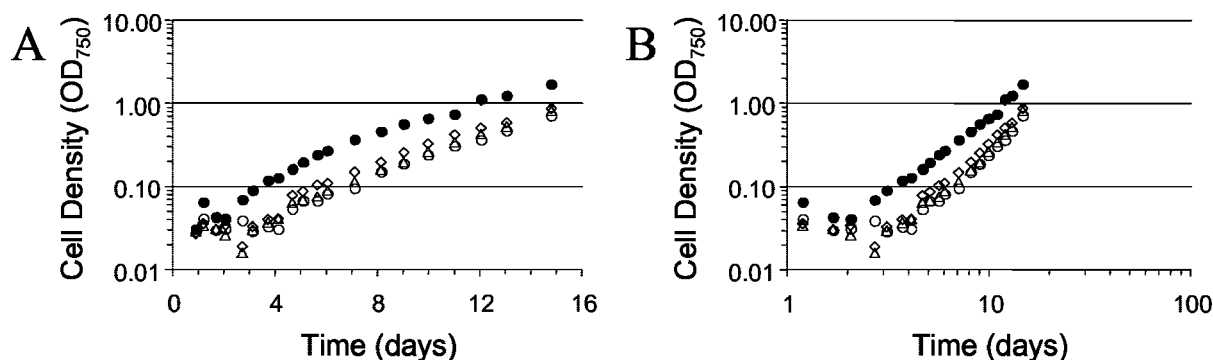

FIG. 5. Growth of *hetY* insertion mutants showed an extended growth lag after nitrogen step-down. Wild type (closed circles) and three independently obtained *hetY* double recombinant inactivation strains, DX-12 (open triangles), AMC642 (open circles), and AMC643 (open diamonds), were grown in the presence of 17.6 mM nitrate (BG-11) and then transferred to $BG-11_0$ at time zero. Growth was followed for 15 days by measuring the optical density at 750 nm (OD_{750}). (A) Growth plotted as a semilog graph. (B) The same data plotted on a log-log graph to show more clearly the polynomial growth phase and the lag-to-growth transition.
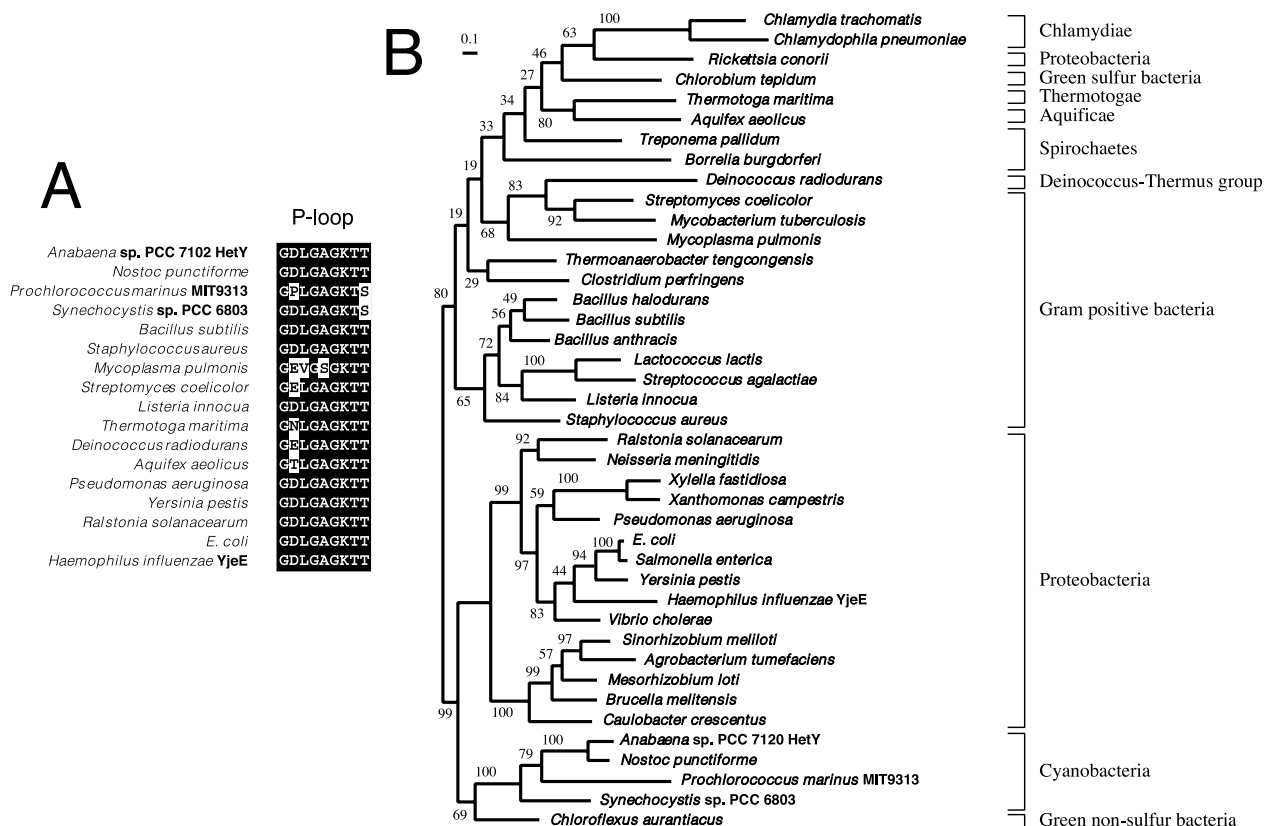
FIG. 6. (A) Multiple sequence alignment of the glycine-rich P-loop region of HetY with the corresponding region of proteins in the YjeE family. YjeE is an ATPase with known X-ray crystal structure. The nucleotide-binding P-loop sequences were identified with PPSearch. The aligned sequences have a consensus GX(L/V)G(A/S)GKT(T/S). (B) Phylogenetic tree for the HetY/YjeE family. The phylogenetic tree was constructed using the Bayesian inference method (7) with the MrBayes program (6), during which four simultaneous Markov chains were run for 1,000,000 iterations under the Jones, Taylor, and Thornton amino acid substitution model (8). The starting tree was estimated using the distance neighbor-joining method of CLUSTAL X with 1,000 bootstrap replications. The consensus Bayesian inference tree was obtained from 10,000 sampled trees after exclusion of the initial 90,000 trees as the "burn-in." The final tree has a maximum likelihood value −lnL = 10,733.8. The Bayesian maximum likelihood nonparametric bootstrap values are indicated near the base of each node. Branch lengths represent mean values of the sampled trees, with the scale bar corresponding to 0.1 amino acid substitutions per site.

glycine-rich P-loop region (Fig. 6A) that is characteristic of ATPases (16).

A phylogenetic analysis encompassing diverse bacterial lineages has shown that HetY and YjeE homologs are widely distributed in the bacterial domain (Fig. 6B). YjeE homologs from proteobacteria form a distinct clade that is a sister group with the cyanobacteria-green sulfur bacteria clade containing HetY. These two groups together form a monophyletic separation from gram-positive bacteria and other groups. The overall phylogeny roughly correlates with the known 16S rRNA phylogeny, suggesting that the HetY/YjeE family may be of ancient origin with mainly vertical transmission during evolution. The only exception appears to be *Rickettsia* sp., for which later gene transfer is suspected.

The HetY/YjeE family is found in almost all bacteria but is absent in *Archaea* and eukaryotic genomes, suggesting that it is present only in phylogenetic groups that synthesize peptidoglycans. The possibility that this family of proteins may be involved in regulating cell wall biosynthesis is supported by the fact that YjeE homologs in many genomes are adjacent to *amiB*, which is involved in recycling of peptidoglycans (16).

The YjeE ATPase is suggested to have properties of a molecular switch (16), which may offer clues to HetY function.

Cell wall defects can influence aerobic nitrogen fixation and heterocyst maturation. *Anabaena* sp. strain PCC 7120 strains with a mutation in *pbpB*, which encodes a penicillin-binding protein, have distorted vegetative cell and heterocyst morphology and cannot fix nitrogen aerobically (9). Strains with a mutation in *hcwA*, which encodes an autolysin, fail to induce the *hepA* gene, which is required for synthesis of the heterocyst envelope polysaccharide, and they are also defective for heterocyst maturation (21). It was proposed that remodeling of the peptidoglycan layer is required for heterocyst maturation, and disruption of this process affects a morphological regulatory checkpoint in heterocyst maturation (21). If *hetY* is involved in cell wall synthesis, it could affect cell communication and the exchange of metabolites along filaments, which could influence heterocyst pattern, as well as heterocyst maturation.

REFERENCES

1. **Black, T. A., Y. Cai, and C. P. Wolk.** 1993. Spatial expression and autoregulation of *hetR*, a gene involved in the control of heterocyst development in *Anabaena*. Mol. Microbiol. **9:**77–84. (Erratum, **10:**1153, 1993.)
2. **Bucher, P., and A. Bairoch.** 1994. A generalized profile syntax for biomolecular sequence motifs and its function in automatic sequence interpretation. Proc. Int. Conf. Intell. Syst. Mol. Biol. **2:**53–61.
3. **Cai, Y., and P. Wolk.** 1990. Use of a conditionally lethal gene in *Anabaena* sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. J. Bacteriol. **172:**3138–3145.
4. **Elhai, J., A. Vepritskiy, A. M. Muro-Pastor, E. Flores, and C. P. Wolk.** 1997. Reduction of conjugal transfer efficiency by three restriction activities of *Anabaena* sp. strain PCC 7120. J. Bacteriol. **179:**1998–2005.
5. **Golden, J. W., L. L. Whorff, and D. R. Wiest.** 1991. Independent regulation of *nifHDK* operon transcription and DNA rearrangement during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **173:**7098–7105.
6. **Huelsenbeck, J. P., and F. Ronquist.** 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:**754–755.
7. **Huelsenbeck, J. P., F. Ronquist, R. Nielsen, and J. P. Bollback.** 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. Science **294:**2310–2314.
8. **Jones, D. T., W. R. Taylor, and J. M. Thornton.** 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. **8:**275–282.
9. **Lazaro, S., F. Fernandez-Pinas, E. Fernandez-Valiente, A. Blanco-Rivero, and F. Leganes.** 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.
10. **Liu, D., and J. W. Golden.** 2002. *hetL* overexpression stimulates heterocyst formation in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **184:**6873–6881.
11. **Meeks, J. C., and J. Elhai.** 2002. Regulation of cellular differentiation in filamentous cyanobacteria in free-living and plant-associated symbiotic growth states. Microbiol. Mol. Biol. Rev. **66:**94–121.
12. **Notredame, C., D. G. Higgins, and J. Heringa.** 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J. Mol. Biol. **302:**205–217.
13. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
14. **Ramaswamy, K. S., C. D. Carrasco, T. Fatma, and J. W. Golden.** 1997. Cell-type specificity of the *Anabaena fdxN*-element rearrangement requires *xisH* and *xisI*. Mol. Microbiol. **23:**1241–1250.
15. **Smith, T. F., and M. S. Waterman.** 1981. Identification of common molecular subsequences. J. Mol. Biol. **147:**195–197.
16. **Teplyakov, A., G. Obmolova, M. Tordova, N. Thanki, N. Bonander, E. Eisenstein, A. J. Howard, and G. L. Gilliland.** 2002. Crystal structure of the YjeE protein from *Haemophilus influenzae*: a putative ATPase involved in cell wall synthesis. Proteins **48:**220–226.
17. **Wei, T.-F., T. S. Ramasubramanian, and J. W. Golden.** 1994. *Anabaena* sp. strain PCC 7120 *ntcA* gene required for growth on nitrate and heterocyst development. J. Bacteriol. **176:**4473–4482.
18. **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, vol. 1. Kluwer Academic Publishers, Dordrecht, The Netherlands.
19. **Yoon, H. S., and J. W. Golden.** 1998. Heterocyst pattern formation controlled by a diffusible peptide. Science **282:**935–938.
20. **Yoon, H. S., and J. W. Golden.** 2001. PatS and products of nitrogen fixation control heterocyst pattern. J. Bacteriol. **183:**2605–2613.
21. **Zhu, J., K. Jager, T. Black, K. Zarka, O. Koksharova, and C. P. Wolk.** 2001. HcwA, an autolysin, is required for heterocyst maturation in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **183:**6841–6851.