JOURNAL OF BACTERIOLOGY, Nov. 1995, p. 6440–6448 Vol. 177, No. 22
0021-9193/95/$04.00+0

The *hglK* Gene Is Required for Localization of Heterocyst-Specific Glycolipids in the Cyanobacterium *Anabaena* sp. Strain PCC 7120

KRISTIN BLACK,† WILLIAM J. BUIKEMA, AND ROBERT HASELKORN*

Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, Illinois 60637

Received 19 June 1995/Accepted 9 August 1995

Mutant strain 543 of the cyanobacterium *Anabaena* sp. strain PCC 7120 was originally isolated as a Fox$^-$ mutant following chemical mutagenesis. Ultrastructural analysis shows that in nitrogen-replete media the vegetative cells of the mutant are more cylindrical and have thicker septa than those of the wild type, while in nitrogen-free media the mutant heterocysts lack the normal glycolipid layer external to the cell wall. Although this layer is absent, strain 543 heterocysts nevertheless contain heterocyst-specific glycolipids, as determined by thin-layer chromatography. The mutation in strain 543 is in a gene we have named *hglK*, encoding a protein of 727 amino acids. The wild-type HglK protein appears to contain four membrane-spanning regions followed by 36 repeats of a degenerate pentapeptide sequence, AXLXX. The mutation in strain 543 introduces a termination codon immediately upstream of the pentapeptide repeat region. A mutant constructed by insertion of an antibiotic resistance cassette near the beginning of the *hglK* gene has the same phenotype as strain 543. We propose that *hglK* encodes a protein necessary for the localization of heterocyst glycolipids and that this function requires the pentapeptide repeats of the HglK protein.

Cyanobacteria are an ancient and diverse group of prokaryotes related to plant and algal chloroplasts that often serve as primary producers in aquatic ecosystems (15, 24, 46). Many filamentous cyanobacteria follow developmental programs that lead to the formation of cell types specialized for particular environmental conditions. When fixed-nitrogen sources (e.g., nitrate or ammonia) are limiting in the environment, some filamentous cyanobacteria form specialized cells, called heterocysts, for the fixation of atmospheric dinitrogen. These highly specialized cells develop at semiregular intervals between the undifferentiated vegetative cells. These two cell types are mutually interdependent, with heterocysts reliant on vegetative cells for carbohydrate that provides the reductant for nitrogen fixation and vegetative cells reliant on heterocysts for fixed nitrogen (27).

The process of heterocyst development from vegetative cells can be described as a series of progressively more differentiated stages occurring over a period of approximately 24 h, culminating in a terminally differentiated microaerobic cell capable of nitrogen fixation (22, 53, 54). It has been estimated that 600 to 1,000 genes are expressed specifically in heterocysts (39). Several genes necessary for the differentiation process have been described. The developmentally regulated excision of three chromosomal elements allows the transcription of genes of the nitrogen fixation pathway (26, 41). The *hetR* gene product is necessary to form the earliest morphological markers of heterocyst differentiation (10), and its expression begins to increase within 30 min of fixed-nitrogen deprivation (9). The *hetP* gene also appears to be necessary for early stages in heterocyst differentiation (21). The *hetN* gene prevents heterocyst differentiation when present in multiple copies (5). Mutations in two other genes are associated with alterations in the heterocyst pattern. The *patA* gene is necessary for the formation of intercalary heterocysts (36), and the *patB* gene product may be indirectly necessary to suppress supernumerary heterocysts through its postulated function as a transcriptional regulator that responds to the redox state of the cell (37).

Heterocysts provide the microaerobic environment necessary for the function of nitrogenase (20). Thus, heterocyst differentiation involves the inactivation of the oxygen-evolving photosystem II complex, the induction of respiratory complexes to scavenge traces of oxygen, and the formation of specialized structures to limit the diffusion of oxygen into the cell. The heterocyst structures thought to be involved in limiting oxygen diffusion include a dense thylakoid web or honeycomb, rich in respiratory enzymes, formed near the junction between heterocysts and vegetative cells (42), and the two heterocyst-specific envelope layers formed outside the vegetative cell wall (43). The outer fibrous and homogeneous layers are composed of polysaccharides (13, 14), and the inner laminated layer is composed of unusual glycolipids. The glycolipids have been isolated from several cyanobacteria, and their structures have been determined (7, 35, 38, 50, 51). All of the structures described contain glucose in α-glycosidic linkage to a 26- or 28-carbon di- or trihydroxylated fatty acid. Measurement of cross sections of individual glycolipid lamellae suggests that each lamella may be a single glycolipid molecule thick (4, 55). No other components of the laminated layer have been identified, nor have the mechanisms of heterocyst glycolipid synthesis or transport been described.

Previously, we reported the isolation of a chemically induced mutant strain of *Anabaena* sp. strain PCC 7120, strain 541, that was unable to fix nitrogen in the presence of oxygen (11). In this study, we continue the characterization of a likely sibling strain, 543, that was isolated at the same time and from the same region of the original selection filter. Strain 543 has a phenotype and complementation properties identical to those of strain 541, which has been lost. Here, we show that the mutation in this strain affects a gene we have named *hglK*, which encodes a protein containing four potential membrane-spanning domains and unusual alanine- and leucine-rich pentapeptide repeats. We propose that the full-length *hglK* gene

* Corresponding author. Mailing address: Department of Molecular Genetics and Cell Biology, University of Chicago, 920 East 58th St., Chicago, IL 60637. Phone: (312) 702-1069. Fax: (312) 702-3172. Electronic mail address: r-haselkorn@uchicago.edu.

† Present address: Department of Biological Sciences, Stanford University, Palo Alto, CA 94305.

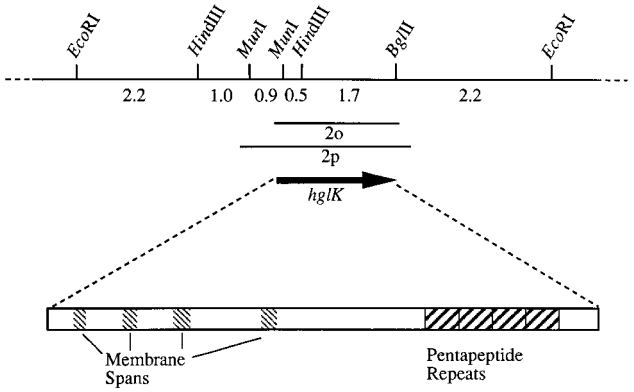

FIG. 1. Restriction map of the chromosomal region of *hglK* and an outline of the domain structure of the protein. The two fragments which were used for sequencing (2o and 2p), the position and direction of the *hglK* ORF, and an expanded view of the protein structure are shown. The membrane-spanning domains near the amino terminus and the pentapeptide repeats near the carboxy terminus are indicated. Map distances are indicated in kilobases.

product is involved in the transport or assembly of heterocyst glycolipids.

MATERIALS AND METHODS

Strains and culture conditions. *Anabaena* sp. strain PCC 7120 and mutant strain 543 have been described previously (11, 46). Liquid cultures were grown in BG11 medium (46) supplemented as appropriate with 17 mM $NaNO_3$ or 2.5 mM $(NH_4)_2SO_4$ and 10 mM HEPES (pH 8.0). Ammonia-grown cultures were monitored daily for pH changes, and 5 M NaOH was added to maintain the pH between 7.0 and 8.0. Cultures were typically grown in air supplemented with 2% CO_2. Solid media contained 1.3% purified agar (BBL Becton Dickinson Microbiology Systems). Neomycin was used at 30 μg/ml, spectinomycin and streptomycin were used at 2 μg/ml, and erythromycin was used at 5 μg/ml when required. *Escherichia coli* cultures were maintained in L broth supplemented with ampicillin (100 μg/ml), chloramphenicol (10 μg/ml), kanamycin (50 μg/ml), spectinomycin (100 μg/ml), or streptomycin (20 μg/ml) as necessary for the selection and maintenance of plasmids (47).

Electron microscopy. Cultures were prepared for electron microscopy by pelleting the cyanobacterial filaments in a clinical centrifuge and resuspending them in half-strength Karnovsky fixative (33)–2% formaldehyde–2.5% glutaraldehyde in 50 mM cacodylate buffer (pH 7.4)–3 mM $CaCl_2$. Fixation was done for 1 h at room temperature and then for 2 h at 4°C. The cells were washed three times in 50 mM cacodylate buffer and postfixed in 2% $KMnO_4$ overnight at 4°C. Fixed material was warmed to room temperature and washed with distilled water five times. Samples were typically embedded in 1% SeaKem low-gelling-temperature agarose. One- to two-millimeter cubic blocks were dehydrated through a graded ethanol series (two changes of 15 min each in 70, 85, and 95% ethanol and three changes of 20 min each in 100% ethanol) and two 30-min incubations in propylene oxide. The samples were infiltrated with a 1:1 mixture of propylene oxide-Epon 812 for 4 h and then with a 1:3 mixture of propylene oxide-Epon 812 overnight with gentle agitation. The samples were incubated in Epon 812 under a vacuum at room temperature for 4 h and then embedded in BEEM capsules with fresh degassed Epon 812 overnight at 55°C. Thin sections were collected on Parlodion-carbon-coated copper grids and stained with 3% uranyl acetate for 30 min in a moist chamber. The sections were examined in a Siemens 101 electron microscope operated at 80 kV.

Thin-layer chromatography. Cell extracts were prepared for thin-layer chromatography by the method of Nichols and Wood (44). Cultures grown in ammonia medium were divided. One part was washed three times and then resuspended in nitrogen-free medium; the other part was pelleted and frozen. The cultures were grown for 2 days and pelleted. The cell pellets were extracted with 5 volumes of 2:1 chloroform-methanol. The extracts were spotted on Silica Gel G plates and developed in 85:15:10:3.7 (vol/vol) chloroform-methanol-acetic acid-water. Separated glycolipids were visualized by spraying the chromatogram with 20% H_2SO_4 and heating it over a Bunsen burner.

Subcloning and sequencing. A cosmid complementing the mutation in strains 541 (11) and 543 was partially digested with *Sau*3AI and used to make 1- to 2-kb, 2- to 4-kb, and 4- to 8-kb mini-libraries by insertion into the *Bam*HI site of the conjugable shuttle vector pCCB110 (10). These libraries were conjugated into strain 543, and complemented exconjugants were selected by growth on nitrogen-free medium. Complementing plasmids were recovered and physically mapped. The smallest complementing fragment (2.2 kb) was subcloned as an *Eco*RI fragment into pBluescript SK^- to generate pKB2o. Nested deletions of the insert were made by using exonuclease III. Deletions from one side of the insert were generated with pKB2o; deletions from the other side of the insert were done on a derivative of pKB2o, pKB2o(−), from which the *Sal*I fragment of the composite pCCB110-pBluescript polylinker had been removed. Sequencing of pKB2o was performed by the dideoxynucleotide chain-termination method (48) using Sequenase reagents (U.S. Biochemicals) and double-stranded templates. A larger, complementing fragment (3.3 kb) from the *Sau*3AI library was subcloned

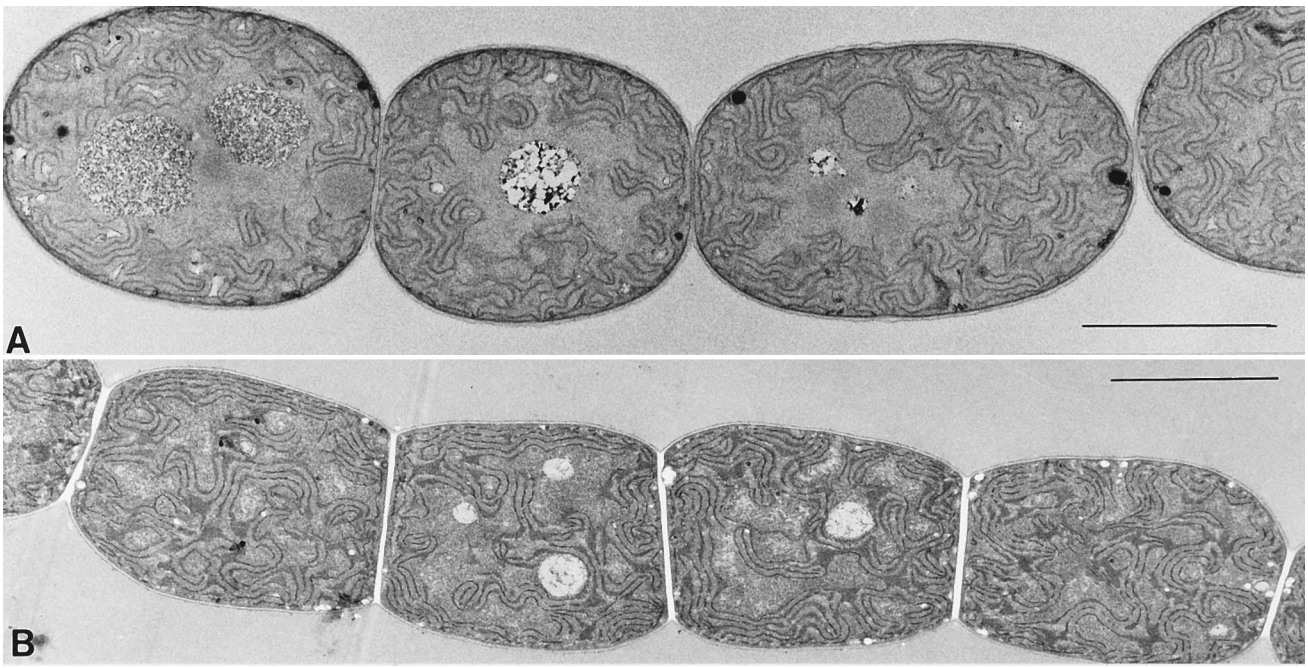

FIG. 2. Transmission electron micrographs of wild-type *Anabaena* sp. strain PCC 7120 and strain 543 grown in the presence of nitrate. (A) A filament of the wild type. (B) A filament of strain 543, showing the rectangular cell shape. Bars, 2 μm.

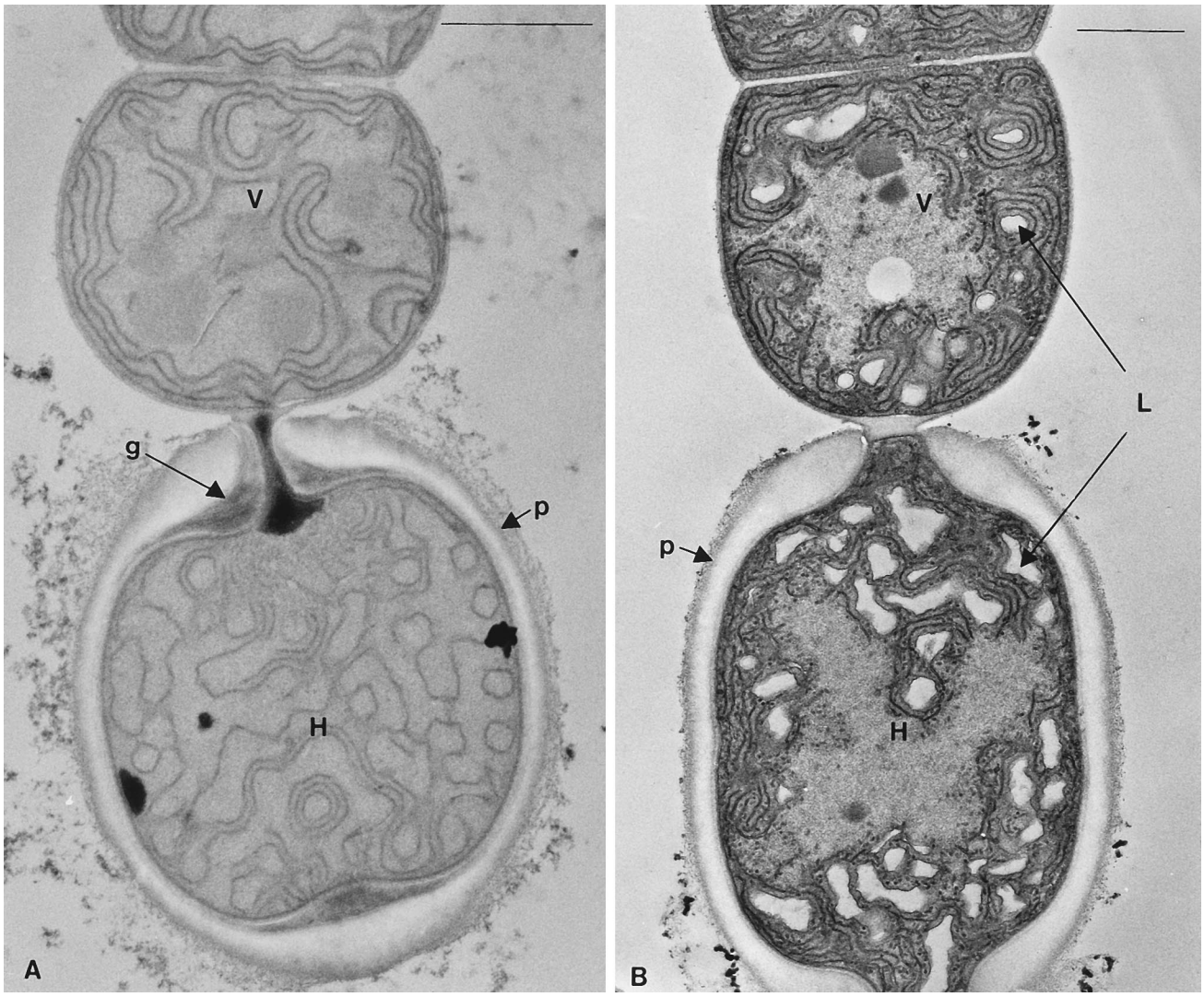

FIG. 3. Transmission electron micrographs of wild-type *Anabaena* sp. strain PCC 7120 and strain 543 grown in the absence of fixed nitrogen. (A) A filament of the wild type. Both vegetative cells (V) and a heterocyst (H) are shown. The fibrous and homogeneous polysaccharide layers (p) and glycolipid layer (g) are visible around the heterocyst. (B) A filament of strain 543. Both vegetative cells and the heterocyst have lacunae (L) between their thylakoid membranes. The heterocyst does not have an organized glycolipid layer, nor are its thylakoids arranged in the typical heterocyst conformation. Bars, 1 μm.

as an *Eco*RI fragment into pBluescript SK$^-$ to generate pKB2p and used to extend the sequence in both directions by using synthetic oligonucleotides as sequencing primers (Fig. 1).

Nucleotide sequence analysis used the University of Wisconsin Genetics Computer Group's sequence software package, as well as the following programs: PLSearch (developed by R. F. Smith and T. F. Smith, Molecular Biology Computer Research Resource [49]), MacPattern (23), BLAZE (6), BLAST (1), FastA (45), Mulfold (28, 56), DNA Strider (40), CodonUse (unpublished program written by C. Halling using the algorithm of Devereux, Haeberli, and Smithies [16] and a codon table compiled from all published *Anabaena* sp. strain PCC 7120 sequences and all sequenced but unpublished genes in the Haselkorn laboratory collection), TopPred (52), and DNAlysis (9). Hydropathy analysis with TopPred used the GES (Goldman-Engelman-Steitz) scale (18).

Characterization of the original mutation. Chromosomal DNA from strain 543 was used as the template in PCRs to generate three fragments of the mutated *hglK* gene. Primers corresponding to positions −487 (CCCTTCTGCTACCTGT TACCCTGC) and +1164 (CGGGATCCGTAATTCGGGGATGGCGG) relative to the ATG of the *hglK* open reading frame (ORF) were synthesized to generate fragment 1. Fragment 2 was generated by primers at positions +1102 (CGGGATCCTCATAGATACCATCCAACAG) and +1683 (CGGGATCCT GTACCCCGCAAATCAGCTCC). Fragment 3 was generated from primers at positions +1615 (CGGGATCCAATGCACGCTTAATTGGT) and +2184 (CG GGATCCCCAGATATATTACAG). PCR products were cloned into pCRII (Invitrogen). The two different orientations of fragment 1 in pCRII were cut with *Hin*dIII and recircularized. A clone of fragment 1 was cut with *Xba*I and *Mun*I to delete a 500-bp fragment and was also cut with *Kpn*I and *Mun*I to delete a 1,100-bp fragment. These seven constructs were used to sequence the mutant *hglK* gene. When base changes were evident from the sequence of the wild-type gene, additional independent PCR products were sequenced to distinguish between the actual mutation and PCR artifacts.

Insertional mutants. Insertional mutants were generated by using spectinomycin-streptomycin Ω cassette (25) insertions at *Sau*3AI sites within the *hglK* gene. *Sau*3AI partial digests of pKB2p were ligated to *Bam*HI fragments of the Ω cassette, individual colonies were selected for resistance to spectinomycin and streptomycin, and the insertion site was located by restriction analysis. The position of the insertion site was confirmed by sequencing the junctions between the Ω cassette and the *hglK* fragment with unique primers near the ends of the Ω cassette (primer KGB5, CAGGGGAATTGATCCGGTG; primer KGB6, GGGCTTTACTAAGCTGAT). Three insertion strains were selected for further analysis. Two strains contained insertions at the first *Sau*3AI site 225 nucleotides within the *hglK* coding region. The other insertion was at the first *Sau*3AI site 178 nucleotides upstream of the first ATG of the *hglK* coding region. *hglK* fragments containing the Ω cassette were subcloned into the conjugable selection plasmid pRL271 (12, 17). These constructs were conjugated into wild-type *Anabaena* sp. strain PCC 7120, and spectinomycin- and streptomycin-resistant clones were selected. pRL271 contains the *sacB* gene, encoding levan sucrase and conferring sensitivity to 5% sucrose (12). Single colonies were picked, grown in liquid BG11 medium with spectinomycin and streptomycin, and then plated on BG11 medium

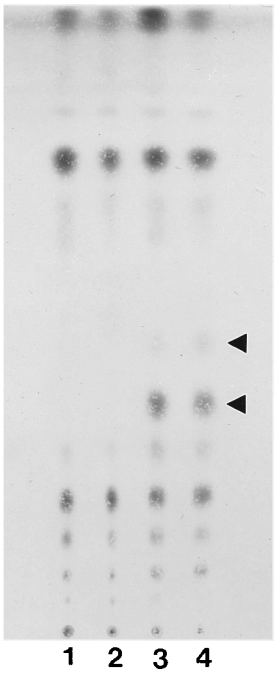

FIG. 4. Thin-layer chromatogram of lipid extracts of cultures of wild-type *Anabaena* sp. strain PCC 7120 (lanes 1 and 3) and strain 543 (lanes 2 and 4) grown under nitrogen-replete (lanes 1 and 2) and nitrogen-limited (lanes 3 and 4) conditions. Arrowheads indicate the positions of heterocyst-specific glycolipids (35).

with spectinomycin, streptomycin, and 5% sucrose to select colonies that had undergone a second recombination to delete the vector but retain the Ω cassette. Strains were further checked for sensitivity to erythromycin to confirm the second recombination. Purified strains were checked for morphology and the ability to grow on a nitrogen-free medium.

Survey blot. A Southern blot of representatives of the major cyanobacterial groups was created from chromosomal DNAs from *Oscillatoria* sp. strain PCC 7515, *Pseudanabaena* sp. strain PCC 7403, *Phormidium foveolarum*, *Nodularia* sp. strain PCC 73104, *Cylindrospermum* sp. strain PCC 7604, *Nostoc* sp. strain MAC R2, *Scytonema* sp. strain PCC 7110, *Calothrix* sp. strain PCC 7102, *Tolypothrix* sp. strain CCAP 1410/1, *Chlorogloeopsis* sp. strain PCC 6912, *Fischerella* sp. strain PCC 7414, *Synechococcus* sp. strain PCC 7942, a *Spirulina* sp., *Anabaena variabilis*, and *Anabaena* sp. strain PCC 7120. Individual DNA samples were digested with *Hin*dIII, and a small portion of each was loaded on a short agarose gel to determine approximate DNA concentrations and completion of restriction. A 0.8% Tris-acetate-EDTA agarose gel was used to separate the digests for the garden blot with approximately equally staining amounts of DNA. The blot was probed with the 2.0-kb *Xba*I fragment internal to *hglK*. The initial hybridization and washing were performed at 66°C. A second round of hybridization and washing using the same blot and probe was performed at 60°C.

Northern (RNA) analysis. RNA was isolated from wild-type filaments at the time of fixed-nitrogen removal and 3, 6, 12, and 24 h postinduction as described elsewhere (26). Some preparations included 1 mM aurintricarboxylic acid, an RNase inhibitor, in the extraction buffer. Northern blots were generated by using 5 to 40 μg of total glyoxal-treated RNA from each time point separated on 1% agarose gels as described previously (26). The blots were probed with nick-translated pKB2o plasmid or with an *Xba*I internal fragment of the ORF. The blots were subsequently hybridized with probes for *nifHDK* (a heterocyst-specific message) or *psbA* (a vegetative cell and heterocyst message) to confirm RNA quality and accurate loading.

Nucleotide sequence accession number. The nucleotide sequence of *hglK* in the pKB2p insert has been submitted to GenBank under accession no. U13768.

RESULTS

Ultrastructure of strain 543. Strain 543 was originally isolated on the basis of its inability to fix nitrogen aerobically (the Fox^- phenotype). At the resolution of the light microscope, the mutant is seen to form regularly spaced heterocysts in the absence of a combined-nitrogen source. After 2 or more days in N-free medium, the heterocysts appear abnormal, showing various degrees of plasmolysis and appearing more rectangular than the wild type. In the presence of a combined-nitrogen source, the filaments are unusually straight, suggesting greater stiffness and less flexibility of cell junctions. The mutant cells are more cylindrical and less spherical than the wild type (11).

In thin sections of strain 543 grown in the presence of nitrate, the cells are shaped more like *Anabaena cylindrica* cells than *Anabaena* sp. strain PCC 7120 cells, with their rectangular longitudinal section (Fig. 2). In thin sections of strain 543 grown in the absence of combined nitrogen, the vegetative cells appear more rounded, like those of wild-type *Anabaena* sp. strain PCC 7120 (Fig. 2A and 3). The heterocysts appear to be incompletely differentiated (Fig. 3B). The polysaccharide layer of the outer envelope has formed, but the inner laminated glycolipid layer is absent. The intercellular junction between heterocysts and vegetative cells is broader than that of mature heterocysts of the wild type. The thylakoid membranes of the heterocyst remain mostly in the vegetative configuration. In addition, thylakoid membranes of heterocysts and, to a lesser extent, vegetative cells form lacunae, electron-translucent distended areas between the thylakoid membranes. The vegetative cells with lacunae are often found adjacent to heterocysts. Lacunae are not present in vegetative cells or mature heterocysts of wild-type cultures. Both heterocysts and vegetative cells of strain 543 contain glycogen granules when grown in the absence of combined nitrogen, a common feature of cultures starved for an essential nutrient other than carbon (29, 30).

Glycolipid analysis. Because electron microscopy revealed that the laminated heterocyst glycolipid layer was missing from strain 543, lipid extracts of the wild type and strain 543 grown in the presence and absence of combined nitrogen were compared with respect to heterocyst glycolipid content (Fig. 4). Thin-layer chromatograms of the lipid extracts showed no differences between the wild type and the mutant. Thus, strain 543 is not defective in the synthesis of the heterocyst glycolipids, although it is unable to form a laminated glycolipid layer.

Subcloning. One cosmid that complemented the mutation of strain 543 for diazotrophic growth was used as the source of a smaller complementing fragment for DNA sequencing. Four complementing plasmids were isolated from the 2- to 4-kb subclass and mapped. The smallest plasmid, pKB2o, containing an insert of 2.2 kb, was selected for sequencing (Fig. 1). All complementing plasmids conferred wild-type growth characteristics on strain 543.

Genetic structure of *hglK*. The sequence of the 2.2-kb insert was determined and found to contain one major ORF, designated *hglK* for heterocyst glycolipid. The stop codon of *hglK* overlapped the *Sau*3AI site used to subclone the fragment. The sequence was extended by sequencing the ends of a 3.3-kb complementing fragment that completely overlaps the original 2.2-kb fragment (Fig. 5). The codon usage of *hglK* agrees well with the predictions for the codon usage of an expressed *Anabaena* gene throughout its length. The 600-bp region upstream of *hglK* contains two ORFs of more than 100 bases, but neither ORF has the codon usage typical of *Anabaena* genes. The *hglK* sequence does contain an AT-rich 9-bp inverted repeat 57 bases upstream of the stop codon. RNA structure prediction for the sequence downstream of the *hglK* coding region using Mulfold predicts a large stem and loop beginning 15 bases downstream of the stop codon with a ΔG of −16.9 kcal/mol (ca. −70.7 kJ/mol) indicating that transcription of *hglK* might use a rho-independent termination mechanism.

Predicted protein structure of HglK. Hydropathy analysis (34) suggests that the amino-terminal part of the predicted

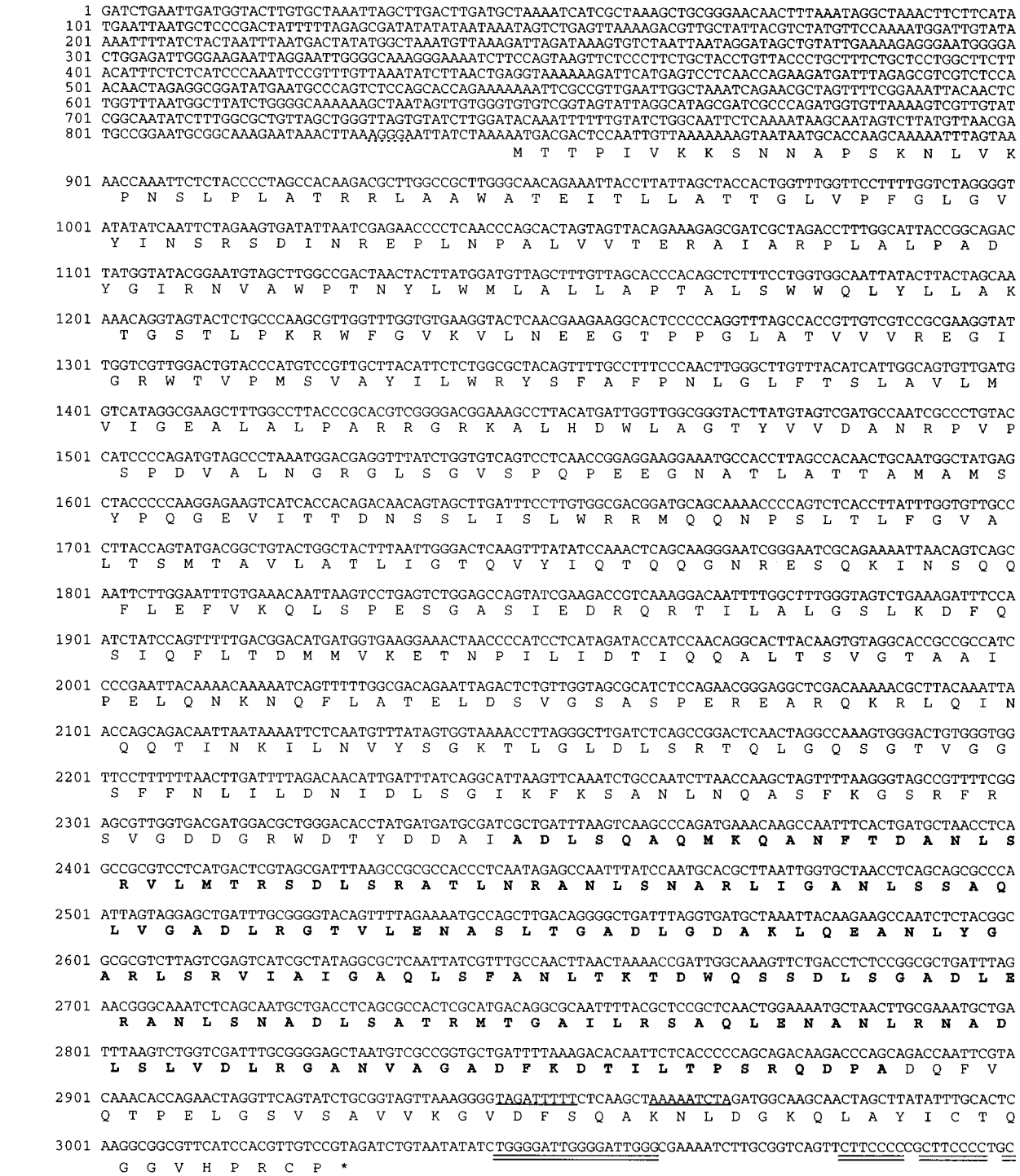

FIG. 5. Complete sequence of the pKB2p insert with translation of the *hglK* ORF. A potential ribosome binding site (dotted underline), a perfect inverted repeat (solid underline), heptamer repeats (double underlines), and the pentapeptide repeats (boldface) are shown.

protein is hydrophobic and contains four potential membrane-spanning domains. The carboxy terminus contains 36 alanine- and leucine-rich, degenerate pentapeptide repeats, starting at amino acid position 501. The nucleotide sequence does not show a repetitive structure in this region (Fig. 6). The consensus sequence of the repeats is ADLSG. The first and third positions are each 75% conserved; the three other positions are conserved between 27 and 30%. A larger repeating unit of

2346					501
GCT	GAT	TTA	AGT	CAA	ADLSQ
GCC	CAG	ATG	AAA	CAA	AQMKQ
GCC	AAT	TTC	ACT	GAT	ANFTD
GCT	AAC	CTC	AGC	CGC	ANLSR
GTC	CTC	ATG	ACT	CGT	VLMTR
AGC	GAT	TTA	AGC	CGC	SDLSR
GCC	ACC	CTC	AAT	AGA	ATLNR
GCC	AAT	TTA	TCC	AAT	ANLSN
GCA	CGC	TTA	ATT	GGT	ARLIG
GCT	AAC	CTC	AGC	AGC	ANLSS
GCC	CAA	TTA	GTA	GGA	AQLVG
GCT	GAT	TTG	CGG	GGT	ADLRG
ACA	GTT	TTA	GAA	AAT	TVLEN
GCC	AGC	TTG	ACA	GGG	ASLTG
GCT	GAT	TTA	GGT	GAT	ADLGD
GCT	AAA	TTA	CAA	GAA	AKLQE
GCC	AAT	CTC	TAC	GGC	ANLYG
GCG	CGT	CTT	AGT	CGA	ARLSR
GTC	ATC	GCT	ATA	GGC	VIAIG
GCT	CAA	TTA	TCG	TTT	AQLSF
GCC	AAC	TTA	ACT	AAA	ANLTK
ACC	GAT	TGG	CAA	AGT	TDWQS
TCT	GAC	CTC	TCC	GGC	SDLSG
GCT	GAT	TTA	GAA	CGG	ADLER
GCA	AAT	CTC	AGC	AAT	ANLSN
GCT	GAC	CTC	AGC	GCC	ADLSA
ACT	CGC	ATG	ACA	GGC	TRMTG
GCA	ATT	TTA	CGC	TCC	AILRS
GCT	CAA	CTG	GAA	AAT	AQLEN
GCT	AAC	TTG	CGA	AAT	ANLRN
GCT	GAT	TTA	AGT	CTG	ADLSL
GTC	GAT	TTG	CGG	GGA	VDLRG
GCT	AAT	GTC	GCC	GGT	ANVAG
GCT	GAT	TTT	AAA	GAC	ADFKD
ACA	ATT	CTC	ACC	CCC	TILTP
AGC	AGA	CAA	GAC	CCA	SRQDP

FIG. 6. Comparison of the nucleotide and amino acid sequences of the pentapeptide repeats. The alignment on the left is the contiguous nucleotide sequence starting at nucleotide 2346 (as numbered in Fig. 5), and the alignment on the right is the corresponding protein sequence starting at amino acid 501. Note that the nucleic acid sequence does not show conservation but that the first alanine and third leucine of the protein sequence are highly conserved. The entire pentapeptide repeat domain can be divided into four subdomains of nine pentapeptide repeats each (separated by horizontal lines) on the basis of evenly spaced glutamine and arginine residues at the second position of the pentapeptides.

nine pentapeptides divides the region into four subdomains. This larger unit is suggested by repeated glutamine and arginine residues in the second position of the pentapeptide repeat. Alternatively, the hydropathy analysis suggests five repeats in this region. Sequence comparisons using the GenBank database and several comparison algorithms did not show significant similarities to other known proteins.

Confirmation of the mutant phenotype. The original mutation in strain 543 was identified by cloning and sequencing PCR-generated fragments of 543 genomic DNA. A single base change that creates a stop codon at amino acid residue 496, five residues before the pentapeptide repeats, was found.

To confirm the mutant phenotype, two strains were constructed by using antibiotic resistance cassette insertions within the *hglK* gene. In one strain, the cassette was placed near the beginning of the gene, 225 bases downstream of the first ATG. This strain is unable to grow on media lacking a nitrogen source. Moreover, this strain has cylindrical vegetative cells with thick septa, like strain 543, when grown on nitrogen-replete medium. The second cassette insertion was 178 nucleotides upstream from the translation start. This strain was indistinguishable from the wild type, suggesting that the start of *hglK* transcription is close to the coding region.

Cyanobacterial survey. To study the distribution of *hglK* among the cyanobacteria, a Southern blot was made using DNA from a variety of cyanobacterial strains representing the major taxonomic groups of filamentous cyanobacteria (sections II, IV, and V) and a representative unicellular cyanobacterium, *Synechococcus* sp. strain PCC 7942 (section I) (46). Under stringent hybridization conditions (66°C hybridization and washes), DNAs from all heterocystous cyanobacteria hybridized to the probe, but DNA from nonheterocystous cyanobacteria did not (Fig. 7). When less stringent conditions (60°C hybridization and washes) were used, a few very faint new bands appeared in *Anabaena* sp. strain PCC 7120, *Cylindrospermum*, *Calothrix*, and *Oscillatoria* DNAs (data not shown). *Oscillatoria* sp. strain PCC 7515 is a nonheterocystous filamentous cyanobacterium, and its lane on the Southern blot showed no signal following the high-stringency hybridization. Thus, it appears that *hglK* is conserved among, and limited to, the heterocystous cyanobacteria, but it is possible that there may be a gene with similarity to *hglK* in some nonheterocystous cyanobacteria.

Northern analysis. Northern analysis of total RNA isolated from developing cells did not reveal discrete bands corresponding to the *hglK* gene message (data not shown). Rather, smears starting at approximately 2.5 kb were detected in RNA from the vegetative cell culture and at all times during development (3). The exact position of the RNA is obscured by rRNA bands that overlap the 2.5-kb region. The intensity of the smear increased approximately fivefold between 0 and 6 h and remained the same during subsequent development. Hybridization of the same blots with a *nifHDK* probe revealed discrete bands of the expected sizes, 4.5, 2.8, and 2.2 kb, in addition to smears probably due to degraded mRNA ranging over the region from 2.5 down to 1.4 kb (3). Primer extension experiments using primers within the *hglK* coding region or 125 bases upstream of the coding region did not reveal a discrete 5′ end.

DISCUSSION

We have described a mutant of *Anabaena* sp. strain PCC 7120, strain 543, that is unable to fix nitrogen aerobically. However, strain 543 can fix nitrogen microaerobically, which demonstrates that the strain is capable of rearranging the *nif* genes and that the nitrogen fixation pathway is intact. Upon nitrogen starvation, this strain develops heterocysts that lack the glycolipid layer of the heterocyst envelope and appear in some aspects to be arrested at an early stage of development. In undifferentiated mutant filaments, the vegetative cells are more cylindrical than wild-type vegetative cells; however, vegetative cells in nitrogen-starved filaments have the spherical wild-type shape. Normal wild-type vegetative cells in both nitrogen-starved and nitrogen-replete filaments have the same spherical shape. Both heterocysts and vegetative cells of the nitrogen-starved mutant develop thylakoid distensions that are electron translucent (lacunae). The mutation can be complemented by a plasmid carrying the *hglK* gene, and complemented strains have the wild-type vegetative cell shape and are capable of nitrogen fixation. From these morphological data, we initially believed the mutant to have a defect in glycolipid synthesis, transport, assembly, or scaffolding or a combination of these functions.

The possibility of a heterocyst glycolipid synthesis defect was tested by isolating lipids from both the mutant strain and the wild-type strain. The profiles of glycolipids were similar in the two strains. The two strains make heterocyst-specific glycolipids in similar quantities with indistinguishable R_f values (Fig. 4). Thus, the synthesis of heterocyst-specific glycolipids appears to be unaffected in strain 543. Several mutants that have glycolipid synthesis defects that produce detectable R_f differences have been described previously (19). Several recently characterized genes, *hglB*, *hglC*, and *hglD*, encode components of a fatty acid synthase. Mutants defective in these genes do

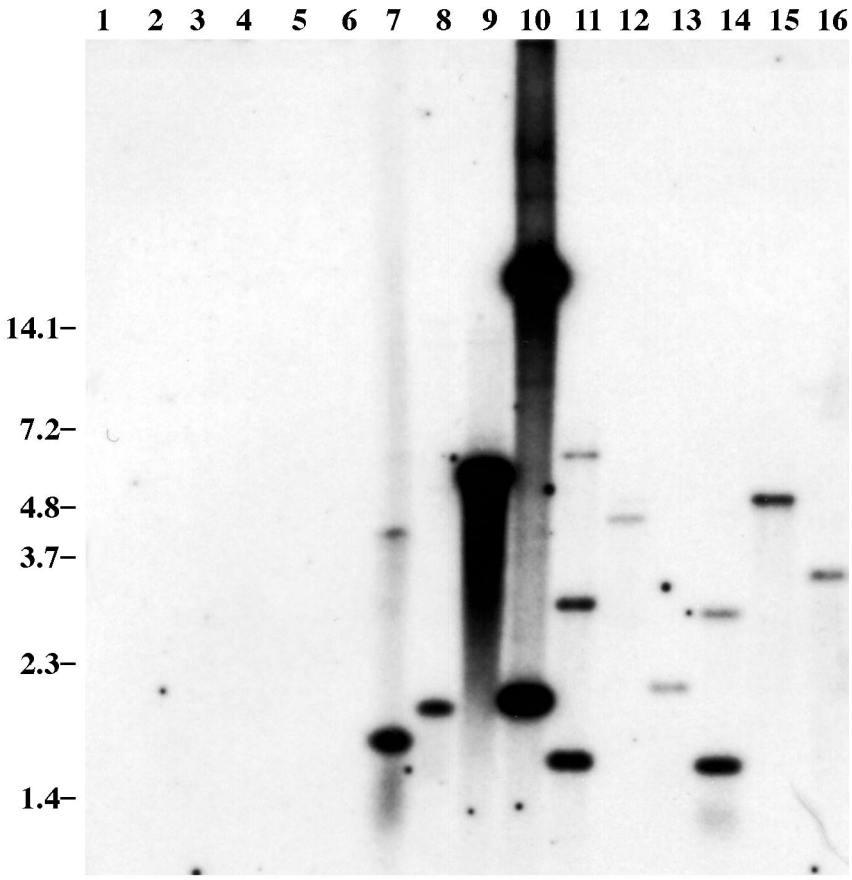

FIG. 7. Southern blot of *Hin*dIII digests of cyanobacterial DNAs probed with the internal 2.0-kb *Xba*I fragment of *hglK* to find similar genes in other cyanobacteria. Lanes: 1, *Bst*EII digest of phage lambda DNA; 2, *Synechococcus* sp. strain PCC 7942; 3, a *Spirulina* sp.; 4, *Oscillatoria* sp. strain PCC 7515; 5, *Pseudanabaena* sp. strain PCC 7403; 6, *P. faveolarum*; 7, *Nodularia* sp. strain PCC 73104; 8, *Cylindrospermum* sp. strain PCC 7604; 9, *A. variabilis*; 10, *Anabaena* sp. strain PCC 7120; 11, *Nostoc* sp. strain MAC R2; 12, *Scytonema* sp. strain PCC 7110; 13, *Calothrix* sp. strain PCC 7102; 14, *Tolypothrix* sp. strain CCAP 1410/1; 15, *Chlorogloeopsis* sp. strain PCC 6912; 16, *Fischerella* sp. strain PCC 7414. Hybridization was done at 66°C. The blot shows that the *hglK* gene is present in all section IV and section V cyanobacteria but not in the nonheterocystous cyanobacteria tested. Adobe Photoshop version 3.0 for the Macintosh was used to generate this figure. Sizes (in kilobases) are indicated on the left.

not synthesize heterocyst glycolipids (2). *hglK* does not fall into this class.

We propose that HglK is not involved in glycolipid synthesis per se but in the transport or assembly of glycolipids or both. Our data do not indicate how direct a role HglK plays in glycolipid localization, and it remains a possibility that HglK affects other gene products that interact more specifically with glycolipid. While a role for HglK in the formation of the heterocyst glycolipid layer is suggested by our data, the morphological data indicate that HglK may have a different role in vegetative cells. Potential target molecules include glycolipids, glycoproteins, and polysaccharides.

The cylindrical shape of cells in undifferentiated filaments and the presence of thylakoid lacunae in the nitrogen-starved filaments could both result from defects in transport or assembly. Misplaced or misaligned envelope components could alter cell shape, as is seen in the vegetative cells of nitrogen-replete filaments. Overloading of general lipid transport due to the absence of a differentiation-induced lipid transport mechanism could result in the formation of thylakoid lacunae composed of heterocyst glycolipids. Also, accumulation of unassembled heterocyst glycolipid could cause a feedback inhibition of general lipid transport that might result in the formation of thylakoid lacunae composed of a combination of heterocyst glycolipids and other lipids. The presence of fewer and smaller lacunae in the vegetative cells adjacent to heterocysts may result from their incomplete differentiation during the first few hours following nitrogen stepdown.

Two characteristics of strain 543 suggest that the mutation affects the outer membrane specifically. First, filaments grown in the presence of a nitrogen source form unusually broad intracellular junctions in which are deposited electron-translucent material that might be composed of peptidoglycan and possibly polysaccharides (Fig. 2) (8, 31, 32). Second, the heterocysts of strain 543 are missing the glycolipid layer which is deposited outside the outer membrane. Thus, two of the major alterations in the mutant flank the outer membrane, suggesting that the defect specifically affects this component.

Hydropathy predictions suggest that HglK is a membrane protein. Most proteins that are involved in lipid transport, assembly, or scaffolding are membrane proteins. The protein is probably well anchored in the membrane with four spanning domains in the N-terminal half of the protein.

An antiserum raised against HglK did not display a specific reaction either in Western blots (immunoblots) of total protein extracted from differentiated cultures or in immunolocalization experiments using electron microscopy of thin sections of *Anabaena* sp. strain PCC 7120 (3). These results indicate that

the HglK protein may be a low-abundance protein, suggesting that the protein does not play a structural role. In any case, the conservation of the *hglK* gene among the heterocystous cyanobacteria is suggestive of a role in heterocystous growth.

The role of the C-terminal alanine- and leucine-rich pentapeptide repeats is unknown. Since the loss of these repeats results in a phenotype indistinguishable from that of a null mutant of *hglK*, it is likely that they are essential for the function of this gene. It is tempting to hypothesize that the repeats interact with the heterocyst glycolipids. This interaction may be transient, resulting in transport of the glycolipids out of the cell or assembly of the glycolipids into the laminated structure typical of the heterocyst envelope.

ACKNOWLEDGMENTS

We thank Hewson Swift for guidance and training in electron microscopy, Jihong Liang for advice on RNA isolation, Martin Mulligan for DNA samples, Michael Fath and Mark Woelfle for critical reading of the manuscript, Le Phung for helpful discussions, and Sagami Paul and Lori Scappino for technical assistance.

This work was supported by grants from the National Institutes of Health (GM21823 and GM40685) and by a predoctoral traineeship to K.B. (GM 07183). Electron microscopy was done in a facility supervised by Hewson Swift and supported by a grant from the Center for Digestive Diseases (DK42086).

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myer, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Bauer, C. C.** 1994. Ph.D. thesis. University of Chicago, Chicago.
3. **Black, K.** 1995. Ph.D. thesis. University of Chicago, Chicago.
4. **Black, K., and R. Haselkorn.** Unpublished data.
5. **Black, T. A., and C. P. Wolk.** 1994. Analysis of a Het$^-$ mutation in *Anabaena* sp. strain PCC 7120 implicates a secondary metabolite in the regulation of heterocyst spacing. J. Bacteriol. **176:**2282–2292.
6. **Brutlag, D. L., J. Dautricourt, R. Diaz, J. Fier, B. Moxon, and R. Stamm.** 1993. Blaze^{TM}: an implementation of the Smith-Waterman sequence comparison algorithm on a massively parallel computer. Comput. Chem. **17:**203–207.
7. **Bryce, T. A., D. Welti, A. E. Walsby, and B. W. Nichols.** 1972. Monohexoside derivatives of long-chain polyhydroxy alcohols: a novel class of glycolipid specific to heterocystous algae. Phytochemistry **11:**295–302.
8. **Budel, B., and E. Rhiel.** 1985. A new cell wall structure in a symbiotic and a free-living strain of the blue-green alga genus *Chroococcidioopsis* (Pleurocapsales). Arch. Microbiol. **143:**117–121.
9. **Buikema, W. J.** Unpublished data.
10. **Buikema, W. J., and R. Haselkorn.** 1991. Characterization of a gene controlling heterocyst development in the cyanobacterium *Anabaena* 7120. Genes Dev. **5:**321–330.
11. **Buikema, W. J., and R. Haselkorn.** 1991. Isolation and complementation of nitrogen fixation mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **173:**1879–1885.
12. **Cai, Y., and C. 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.
13. **Cardemil, L., and C. P. Wolk.** 1976. The polysaccharides from heterocyst and spore envelopes of a blue-green alga: methylation analysis and structure of the backbones. J. Biol. Chem. **251:**2967–2975.
14. **Cardemil, L., and C. P. Wolk.** 1979. The polysaccharides from heterocyst and spore envelopes of a blue-green alga: structure of the basic repeating unit. J. Biol. Chem. **254:**736–741.
15. **Carr, N. G., and N. H. Mann.** 1994. The oceanic cyanobacterial picoplankton, p. 27–48. *In* D. A. Bryant (ed.), The molecular biology of cyanobacteria. Kluwer Academic Publishers, Dordrecht, The Netherlands.
16. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **10:**141–156.
17. **Elhai, J., and C. P. Wolk.** 1988. Conjugal transfer of DNA to cyanobacteria. Methods Enzymol. **167:**747–754.
18. **Engelman, D. M., T. A. Steitz, and A. Goldman.** 1986. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu. Rev. Biophys. Biophys. Chem. **15:**321–352.
19. **Ernst, A., T. Black, Y. Cai, J.-M. Panoff, D. N. Tiwari, and C. P. Wolk.** 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.
20. **Fay, P.** 1992. Oxygen relations of nitrogen fixation in cyanobacteria. Microbiol. Rev. **56:**340–373.
21. **Fernández-Pinas, F., F. Leganés, and C. P. Wolk.** 1994. A third genetic locus required for the formation of heterocysts in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **176:**5277–5283.
22. **Fleming, H., and R. Haselkorn.** 1974. The program of protein synthesis during heterocyst differentiation in nitrogen-fixing blue-green algae. Cell **3:**159–170.
23. **Fuchs, R.** 1991. MacPattern: protein pattern searching on the Apple Macintosh. Comput. Appl. Biosci. **7:**105–106.
24. **Giovannoni, S. J., S. Turner, G. J. Olsen, S. Barns, D. J. Lane, and N. R. Pace.** 1988. Evolutionary relationships among cyanobacteria and green chloroplasts. J. Bacteriol. **170:**3584–3592.
25. **Golden, J. W., and D. R. Wiest.** 1988. Genome rearrangement and nitrogen fixation in *Anabaena* blocked by inactivation of *xisA* gene. Science **242:**1421–1423.
26. **Haselkorn, R.** 1992. Developmentally regulated gene rearrangements in prokaryotes. Annu. Rev. Genet. **26:**111–128.
27. **Haselkorn, R., and W. J. Buikema.** 1992. Nitrogen fixation in cyanobacteria, p. 166–190. *In* G. Stacey, R. H. Burris, and H. J. Evans (ed.), Biological nitrogen fixation. Chapman & Hall, New York.
28. **Jaeger, J. A., D. H. Turner, and M. Zuker.** 1989. Predicting optimal and suboptimal secondary structure for RNA. Methods Enzymol. **183:**281–306.
29. **Jensen, T. E.** 1984. Cyanobacterial cell inclusions of irregular occurrence: systematic and evolutionary implications. Cytobios **39:**35–62.
30. **Jensen, T. E.** 1985. Cell inclusions in the cyanobacteria. Archiv. Hydrobiol. **71:**33–73.
31. **Jurgens, U. J., and J. Weckesser.** 1985. The fine structure and chemical composition of the cell wall and sheath layers of cyanobacteria. Ann. Inst. Pasteur Microbiol. **136A:**41–44.
32. **Jurgens, U. J., and J. Weckesser.** 1986. Polysaccharide covalently linked to the peptidoglycan of the cyanobacterium *Synechocystis* sp. strain PCC 6714. J. Bacteriol. **168:**568–573.
33. **Karnovsky, M. J.** 1965. A formaldehyde-glutaraldehyde fixative of high osmolarity for use in electron microscopy. J. Cell Biol. **27:**137A–138A.
34. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
35. **Lambein, F., and C. P. Wolk.** 1973. Structural studies on the glycolipids from the envelope of the heterocyst of *Anabaena cylindrica*. Biochemistry **12:**791–798.
36. **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. USA **89:**5655–5659.
37. **Liang, J., L. Scappino, and R. Haselkorn.** 1993. The *patB* gene product, required for growth of the cyanobacterium *Anabaena* sp. strain PCC 7120 under nitrogen-limiting conditions, contains ferredoxin and helix-turn-helix domains. J. Bacteriol. **175:**1697–1704.
38. **Lorch, S. K., and C. P. Wolk.** 1974. Application of gas-liquid chromatography to study of the envelope lipids of heterocysts. J. Phycol. **10:**352–355.
39. **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.
40. **Marck, C.** 1988. 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res. **16:**1829–1835.
41. **Matveyev, A. V., E. Rutgers, E. Soderback, and B. Bergman.** 1994. A novel genome rearrangement involved in heterocyst differentiation of the cyanobacterium *Anabaena* sp. PCC 7120. FEMS Microbiol. Lett. **116:**201–208.
42. **Murry, M. A., A. G. Olafsen, and J. R. Benemann.** 1981. Oxidation of diaminobenzidine in the heterocysts of *Anabaena cylindrica*. Curr. Microbiol. **6:**201–206.
43. **Murry, M. A., and C. P. Wolk.** 1989. Evidence that the barrier to the penetration of oxygen into heterocysts depends upon two layers of the cell envelope. Arch. Microbiol. **151:**469–474.
44. **Nichols, B. W., and B. J. B. Wood.** 1968. New glycolipid specific to nitrogen-fixing blue-green algae. Nature (London) **217:**767–768.
45. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
46. **Rippka, R., J. Deruelles, J. B. Waterbury, M. Herdman, and R. Y. Stanier.** 1979. Genetic assignments, strain histories and properties of pure cultures of cyanobacteria. J. Gen. Microbiol. **111:**1–61.
47. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
48. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
49. **Smith, R. F., and T. F. Smith.** 1990. Automatic generation of primary sequence patterns from sets of related protein sequences. Proc. Natl. Acad. Sci. USA **87:**118–122.
50. **Soriente, A., A. Gambacorta, A. Trincone, C. Sili, M. Vincenzini, and G.**

Sodano. 1993. Heterocyst glycolipids of the cyanobacterium *Cyanospira rippkae*. Phytochemistry **33:**393–396.
51. **Soriente, A., G. Sodano, A. Gambacorta, and A. Trincone.** 1992. Structure of the "heterocyst glycolipids" of the marine cyanobacterium *Nodularia harveyana*. Tetrahedron **48:**5375–5384.
52. **von Heijne, G.** 1992. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. J. Mol. Biol. **255:**487–494.
53. **Wilcox, M., G. J. Mitchison, and R. J. Smith.** 1973. Pattern formation in the blue-green alga, *Anabaena*. I. Basic mechanisms. J. Cell Sci. **12:**707–723.
54. **Wilcox, M., G. J. Mitchison, and R. J. Smith.** 1973. Pattern formation in the blue-green alga, *Anabaena*. II. Controlled proheterocyst regression. J. Cell Sci. **13:**637–649.
55. **Winkenbach, F., C. P. Wolk, and M. Jost.** 1972. Lipids of membranes and of the cell envelope in heterocysts of a blue-green alga. Planta **107:**69–80.
56. **Zuker, M.** 1989. On finding all suboptimal foldings of an RNA molecule. Science **244:**48–52.