REGULATION OF THE NITROGEN FIXATION GENES IN THE HETEROCYSTOUS CYANOBACTERIUM ANABAENA SP. STRAIN PCC 7120

A Dissertation

by

KRITHIKA KUMAR

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2011

Major Subject: Microbiology

REGULATION OF THE NITROGEN FIXATION GENES IN THE

HETEROCYSTOUS CYANOBACTERIUM ANABAENA SP. STRAIN PCC 7120

A Dissertation

by

KRITHIKA KUMAR

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	James Golden
	Michael Benedik
Committee Members,	Deborah Siegele
	Deborah Bell-Pedersen
	Daniel Ebbole
Head of Department,	U.J. McMahan

December 2011

Major Subject: Microbiology

ABSTRACT

Regulation of the Nitrogen Fixation Genes in the Heterocystous Cyanobacterium

Anabaena sp. strain PCC 7120. (December 2011)

Krithika Kumar, B.S., Bangalore University

Co- Chairs of Advisory Committee: Dr. James W. Golden
Dr. Michael Benedik

Many multicellular cyanobacteria produce specialized nitrogen-fixing heterocysts. During diazotrophic growth of *Anabaena* (*Nostoc*) sp. strain PCC 7120, a regulated developmental pattern of single heterocysts separated by about 10 to 20 photosynthetic vegetative cells is maintained along filaments. Heterocyst structure and metabolic activity function together to accommodate oxygen-sensitive nitrogen fixation, catalyzed by nitrogenase.

In this work, we show that the promoter of the *nifHDK* genes that encode nitrogenase, lies upstream from the intergenic region between *nifH* and *nifU*. Excision of the *fdxN* element is required for transcription of the *nifHDK* genes. Fluorescence microscopy of reporter strain P_{nifHD}-*gfp*, in the chromosomal *nif* locus indicated that expression of *nifHDK* is blocked in mutants that are unable to excise the *fdxN* element after nitrogen deprivation. We proposed that a promoter upstream of the element, likely P_{nifB}, is required for transcription of the *nifHDK* genes. Indeed, the P_{nifHD}-*gfp* reporter at an ectopic site did not show GFP fluorescence. A P_{nifB}-*gfp* reporter was expressed specifically in heterocysts indicating that a promoter for the *nifB* gene lies in the

intergenic region upstream of *nifB*. A stem loop structure located in the intergenic region between *nifH* and *nifU* may act as a processing site for production of *nifHDK* transcripts.

We also provide evidence that DevH, a transcriptional regulator, is involved in regulating the *nifB-fdxN-nifSUHDK* genes. DevH is a protein belonging to the cAMP receptor protein (CRP) family of proteins that are widespread in bacteria and regulate genes in response to a gamut of physiological conditions. We show that DevH binds specifically to the *nifB* upstream region but not to the immediate upstream region of *nifH*. We predict that DevH binds to an NtcA-like binding site upstream of *nifB* and functions as an activator of the *nifB-fdxN-nifSUHDK* genes.

Finally, we show that *sigE*, which is expressed at 16 hours after nitrogen deprivation, is required for normal expression of some heterocyst specific genes, including *nifHDK*. A *sigE* mutant shows delayed and reduced expression of *nifHDK* and some middle and late genes. We hypothesize that DevH in concert with SigE upregulates the expression of *nifHDK* in heterocysts after nitrogen deprivation.

DEDICATION

To my family, for everything.

ACKNOWLEDGEMENTS

Though this dissertation is credited to me, it would not have been possible without the support and guidance of many. I am indebted to them.

First and foremost, my deepest gratitude to my advisor Dr. James Golden for his guidance, mentoring, and unrelenting expectation for high standards. I owe much of my scientific learning to him. I am grateful to my co-advisor Dr. Michael Benedik for his support, particularly in the last two years of my graduate experience. His insights on my experiments have been productive and helpful. Many thanks to my committee member Dr. Daniel Ebbole for his advice and guidance. My heartfelt gratitude to my committee members Dr. Deborah Siegele and Dr. Deborah Bell-Pedersen for their faith and trust in me. I am indebted to them not only for their invaluable scientific advice but also for being my pillars of support. Dr. Siegele, Dr. Bell-Pedersen, and Dr. Susan Golden have all provided me with excellent female role models that have and will continue to improve the quality of my life.

A sincere thanks to all my friends and colleagues for making my graduate experience enjoyable. I would like to especially thank current and past members of the Golden labs, particularly Ramona Neunuebel for her friendship and mentoring, Thammajun Wood for being a close friend and confidant, Guogang Dong for his warm smile and utter humility, and Sushantha Saha for being my Indian buddy. Thanks also to Ashley Smallwood for helping me edit this dissertation.

I would like express my most heartfelt thanks to Nick Ratterman, my husband and life

companion, for being a constant source of love and encouragement, and for his genuine pride and joy in my accomplishments. Many thanks to Nick's family for loving me as their own.

Most importantly, I would like to express my deepest and most sincere gratitude to my parents R. T. Kumar and Srimathi Kumar for their unconditional love, unwavering faith, and encouragement. A very special thanks to my sister Krupa Kumar for her love, pride, and belief in me. My gratitude to my extended family for their love and support.

Last but not the least, thank you, India for molding me into the person I am today.

TABLE OF CONTENTS

Page

ABSTRACT iii

DEDICATION v

ACKNOWLEDGEMENTS vi

TABLE OF CONTENTS viii

LIST OF FIGURES x

LIST OF TABLES xii

CHAPTER I INTRODUCTION 1

- Cellular differentiation, multicellularity, and transport allow specialized functions 3
- Initiation and early stages of heterocyst development 11
- Heterocyst cellular differentiation produces an environment for nitrogen fixation 17
- RNA polymerase sigma factors are developmentally regulated 22
- Heterocyst pattern formation requires dynamic signaling 24
- Nitrogen fixation in heterocysts 29

CHAPTER II EXCISION OF THE FDXN ELEMENT IS REQUIRED FOR TRANSCRIPTION OF THE NIFHDK GENES 32

- Overview 32
- Introduction 33
- Materials and methods 36
- Results 41
- Discussion 53

CHAPTER III ROLE OF DEVH AS THE TRANSCRIPTIONAL REGULATOR OF THE NIFHDK GENES IN THE FILAMENTOUS CYANOBACTERIUM ANABAENA SP. STRAIN PCC 7120 57

- Overview 57
- Introduction 57
- Materials and methods 60
- Results 63
- Discussion 66

CHAPTER IV THE SIGE GENE IS REQUIRED FOR NORMAL EXPRESSION OF THE HETEROCYST-SPECIFIC GENES IN ANABAENA PCC 7120 69

Overview 69
Introduction 70
Materials and methods 75
Results 83
Discussion 100

CHAPTER V CONCLUSION 105

Overview 105
Cis-acting sequences required for transcription of the *nif* genes 105
Consensus binding sequences in late promoters 107
DevH as a transcriptional regulator of the *nifB-fdxN-nifSUHDK* operon 108
Secondary effector in DevH binding 110
Regulation of *nifHDK* expression by SigE 110
Regulatory targets of SigE 112
Future studies 112
Closing comments 115

REFERENCES 116

APPENDIX 147

VITA 161

LIST OF FIGURES

Page

FIG. 1. Heterocyst development in *Anabaena* PCC 7120..4

FIG. 2. Heterocyst development in *Anabaena* PCC 7120..5

FIG. 3. Transmission electron micrograph of the junction between two vegetative cells 8

FIG. 4. Developmentally regulated sigma factors in *Anabaena* PCC 7120....................10

FIG. 5. Model of regulatory interactions during heterocyst development. 13

FIG. 6. Schematic of the *nif* gene cluster in Anabaena sp. strain PCC 7120 35

FIG. 7. GFP fluorescence from GFP reporter expression from KKC7...........................43

FIG. 8. GFP fluorescence of AMC1774. 45

FIG. 9. GFP fluorescence of KKC3. 45

FIG. 10. PCR analysis of mutants. 46

FIG. 11. GFP fluorescence of KKC4. 47

FIG. 12. KKC1 at neutral site did not show GFP fluorescence after nitrogen deprivation...49

FIG. 13. P_{nifB}-*gfp* reporter shows GFP fluorescence after nitrogen deprivation 51

FIG. 14. Predicted stem-loop structure at the 5' end of the stable *nifH* transcript 52

FIG. 15. DevH binds to the *nifB* upstream region...64

FIG. 16. DevH does not bind to the *nifU* region 65

FIG. 17. Comparison of NtcA consensus binding sequence and DevH putative binding site present upstream of *nifB* 68

FIG. 18. Northern RNA blot analysis of *sigE* transcripts during heterocyst development 84

FIG. 19. NtcA binds to the *sigE* upstream intergenic region 87

FIG. 20. GFP reporter expression from different upstream regions of *sigE* 89

FIG. 21. Northern RNA blot analysis of *nifH* transcripts from the wild type and *sigE* mutant strain AMC649 during heterocyst development 91

FIG. 22. Time-lapse microscopy images of the P_{nifHD}-gfp reporter construct integrated into the nifH locus on the chromosome in wild-type (AMC1774), sigE overexpression (AMC1775), and sigE mutant (AMC1777) backgrounds.. 93

FIG. 23. Heterocyst pattern is influenced by *sigE* 95

FIG. 24. Quantitative reverse transcription PCR (qRT-PCR) measurements of mRNA abundance in wild-type and *sigE* mutant strains 98

LIST OF TABLES

Page

TABLE 1. Primers used in this study (Chapter II) 37

TABLE 2. Strains and plasmids used in this study (Chapter II) 38

TABLE 3. Primers used in this study (Chapter III) 61

TABLE 4. Strains and plasmids used in this study (Chapter IV) 76

TABLE 5. Primers used in this study (Chapter IV) 78

TABLE 6. Strains and plasmids used in this study (Appendix) 156

TABLE 7. Primers used in this study (Appendix) 159

CHAPTER I

INTRODUCTION*

Multicellularity has arisen several times during evolution and although common to eukaryotes, it is also found in prokaryotes (90). Organisms composed of multiple differentiated cell types can possess structures, functions, and behaviors that are more diverse and more efficient than what can be achieved in unicellular organisms (90). The generation of multicellularity and cell-type diversity enables an organism to acquire specialized functions and advantages in feeding, dispersion, and protection. Among multicellular prokaryotes, heterocystous cyanobacteria of the genera *Anabaena* and *Nostoc* offer a relatively simple model for the study of cellular differentiation and multicellularity.

Cyanobacteria are a large group of Gram-negative prokaryotes that perform oxygenic photosynthesis. They have evolved multiple specialized cell types, including nitrogen-fixing heterocysts, spore-like akinetes, and the cells of motile hormogonia filaments. Of these, the development of heterocysts in the filamentous cyanobacterium *Anabaena* (also *Nostoc*) sp. strain PCC 7120 (hereafter *Anabaena* PCC 7120) has been the best studied (Fig. 1).

This dissertation follows the style of Journal of Bacteriology.

* Reprinted with permission from "Cyanobacterial Heterocysts" in "Cold Spring Harbor Perspectives in Biology" 2: a000315 by Kumar *et al.*, CSHL Press, Copyright [2010] CSHL Press

Heterocyst development offers a striking example of cellular differentiation and developmental biology in a very simple form: filaments are composed of only two cell types and these are arrayed in a one-dimensional pattern similar to beads on a string (Figs. 1 and 2).

Many cyanobacterial species are capable of nitrogen fixation. However, oxygenic photosynthesis and nitrogen fixation are incompatible processes because nitrogenase is *inactivated by oxygen. Cyanobacteria mainly use two mechanisms to separate these activities: a biological circadian clock to separate them temporally, and multicellularity and cellular differentiation to separate them spatially. For example, the unicellular *Cyanothece* sp. strain ATCC 51142 stores glycogen during the day and fixes nitrogen at night (169). Heterocyst-forming cyanobacteria differentiate highly specialized cells to provide fixed nitrogen to the vegetative cells in a filament.

In the presence of a source of combined nitrogen such as nitrate or ammonium, *Anabaena* PCC 7120 grows as long filaments containing hundreds of photosynthetic vegetative cells. In the absence of combined nitrogen, it produces heterocysts, which are terminally differentiated nitrogen-fixing cells that form at semi-regular intervals between stretches of vegetative cells to produce a multicellular pattern of single heterocysts every ten to twenty vegetative cells (Fig. 1). Heterocyst development involves integration of multiple external and internal signals, communication between the cells in a filament, and temporal and spatial regulation of genes and cellular processes. The study of heterocyst development in *Anabaena PCC* 7120 has proven to be an excellent model for

the study of cell fate determination, pattern formation, and differential gene expression during prokaryotic multicellular development. Various aspects of heterocyst development, signaling, and regulation have been the subject of several recent reviews (1, 57, 79, 114, 198, 203).

It should be noted that cyanobacteria have recently attracted increased attention because of their important roles in environmental carbon and nitrogen fixation (118), and their potential for providing renewable chemicals and biofuels (40).

Cellular differentiation, multicellularity, and transport allow specialized functions

Cyanobacteria have a Gram-negative cell wall that includes two distinct membranes, the plasma membrane and an outer membrane, and a peptidoglycan layer, which is thicker than in other Gram-negative bacteria, sandwiched between these two membranes (80). External to the cell wall is a carbohydrate-enriched glycocalyx that can have different relative amounts of three recognizable layers: a closely associated sheath, a defined capsule, and loosely associated slime. These layers protect the cells from desiccation and presumably from phages and predators. Cyanobacteria contain extensive internal thylakoid membranes (Fig. 2), which are the site of photosynthetic reactions.

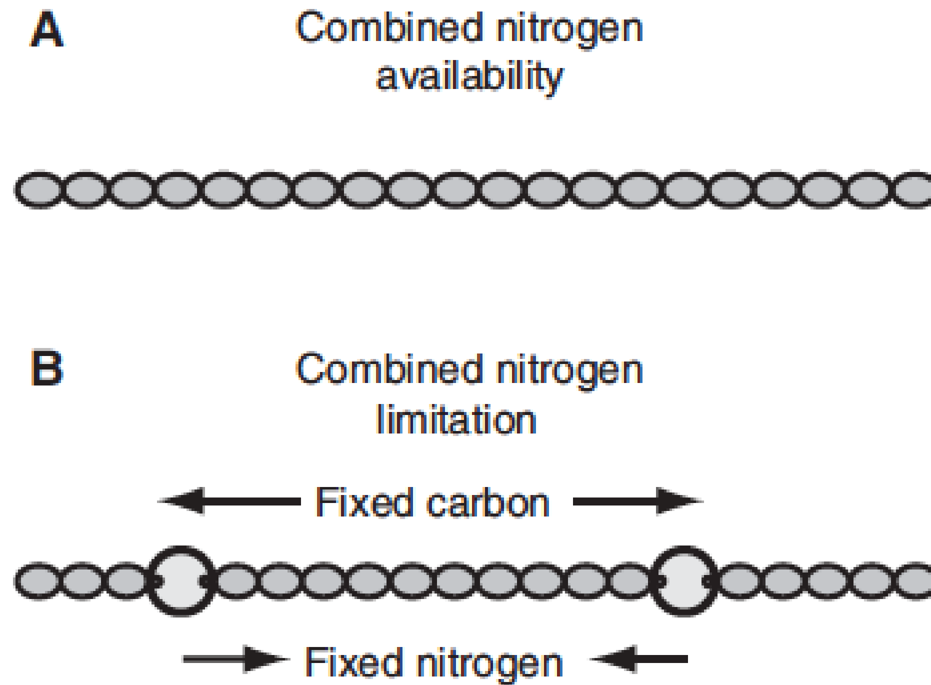

FIG. 1. Heterocyst development in *Anabaena* PCC 7120. (A) *Anabaena* PCC 7120 grown in medium containing a source of combined nitrogen grows as filaments of photosynthetic vegetative cells. (B) In the absence of combined nitrogen, heterocysts differentiate at semiregular intervals forming a developmental pattern of single heterocysts every 10 to 20 vegetative cells along the filaments. Heterocysts are often larger than vegetative cells, have a thicker multilayered envelope, and usually contain cyanophycin granules at the poles adjacent to their vegetative cell

.

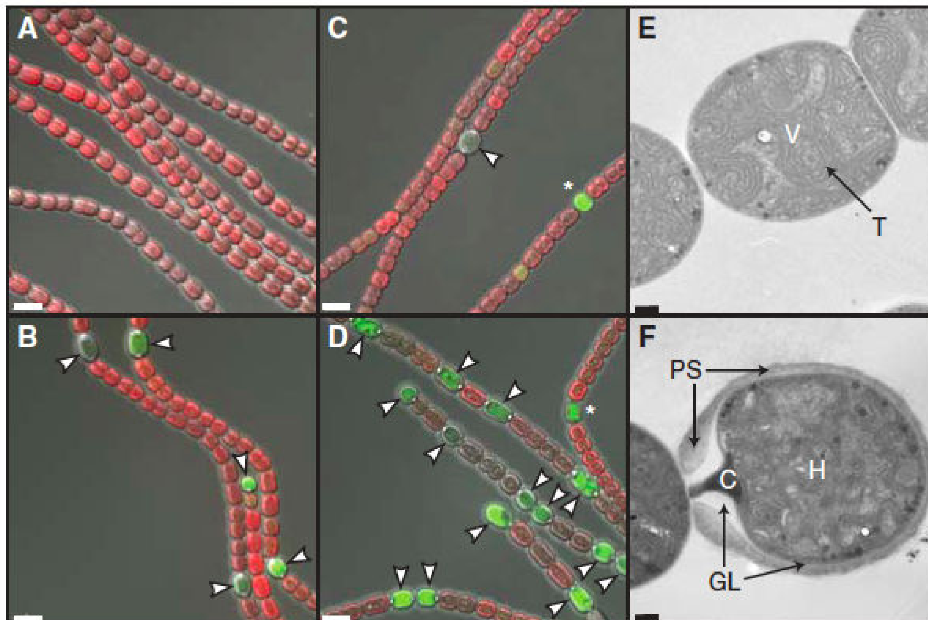

FIG. 2. Heterocyst development in *Anabaena* PCC 7120. Filaments of the wild type carrying a *patS-gfp* reporter grown in medium containing ammonium (A), and have undergone heterocyst development one day after transfer to medium without combined nitrogen (B). A *patS* mutant strain carrying the same *patS-gfp* reporter grown in media containing ammonium contains a small number of heterocysts (C), and one day after transfer to medium without combined nitrogen shows a higher than normal frequency of heterocysts and an abnormal developmental pattern (D). (a, b, c, d) Merged DIC (grayscale), autofluorescence (red), and *patS-gfp* reporter fluorescence (green) images; arrowheads indicate heterocysts; asterisks indicate proheterocysts; size bar, 5 µm. (e, f) Transmission electron micrographs of wild-type vegetative cells (V) and a heterocyst (H) at the end of a filament; CB, carboxysome; T, thylakoid membranes; PS, polysaccharide layer; GL, glycolipid layer; C, cyanophycin polar granule; size bar, 0.2 µm.

Heterocysts are typically distinguishable from vegetative cells by their somewhat larger and rounder shape, diminished pigmentation, thicker cell envelopes, and usually prominent cyanophycin granules at poles adjacent to vegetative cells (Fig. 2). The additional envelope layers surrounding heterocysts help to protect the enzyme nitrogenase from oxygen (50). For details of the heterocyst cell wall and envelope, readers are referred to recent reviews (4, 134, 142). Mature heterocysts provide the microoxic environment required for nitrogen fixation, spatially separating oxygen-evolving photosynthesis in vegetative cells from nitrogen fixation. Differentiating cells undergo many metabolic and morphological changes (67). Oxygen producing photosystem PSII is dismantled during differentiation and heterocysts show an increased rate of respiration (188). Morphological changes include the deposition of two additional envelope layers around the heterocyst: an inner "laminated" layer composed of two heterocyst-specific glycolipids (HGL) and an outer polysaccharide layer (HEP) (30, 31, 134). The heterocyst envelope polysaccharide layer is sometimes subdivided into a well-defined homogeneous inner layer and an external fibrous layer.

Heterocysts and vegetative cells are mutually interdependent. Because they lack photosystem II and carbon fixation, heterocysts are dependent on vegetative cells for a source of reductant and carbon, which is probably partially supplied as sucrose (38, 109). In *Anabaena* PCC 7120, vegetative cells must also supply glutamate to heterocysts, which convert it to glutamine and other amino acids (111). In return, vegetative cells obtain fixed nitrogen in the form of amino acids from the heterocysts (114). The temporal and spatial distributions of fixed carbon and nitrogen were studied

using high-resolution nanometer-scale secondary ion mass spectrometry (NanoSIMS), in *Anabaena oscillarioides* (144). Newly fixed nitrogen is rapidly exported from heterocysts and distributed to nearby vegetative cells.

The exchange of metabolites and intercellular signals that control the regulated spacing of the heterocysts require movement of molecules between cells along a filament, possibly through a continuous periplasm (55). According to this model, molecules exported from one cell would diffuse through the periplasm and then be taken up by other cells along a filament. It was recently shown that GFP expressed from the *patS* promoter, which is expressed at a low basal level in vegetative cells and strongly upregulated in differentiating heterocysts, and targeted to the periplasm by a cleavable twin-arginine signal peptide, could diffuse through the heterocyst periplasm to nearby vegetative cells (110). GFP attached to the cytoplasmic membrane was only seen in heterocysts and not in adjacent vegetative cells. However, another group found that GFP targeted to the periplasm of vegetative cells or heterocysts using different cell-type specific promoters (P-*hepA*, P-*patB*, or P-*rbcL*) and the signal sequence of the *Escherichia coli* TorA protein attached to GFP, showed no intercellular diffusion from one cell to the next (200). The exported GFP diffused around individual cells but not beyond the cell borders. Although it is unclear why these two research groups came to different conclusions, it may be due to the different signal sequences that were used. The data obtained by Zhang et al. could be the result of a failure to cleave the TorA signal peptide from the GFP reporter, which could result in the GFP being anchored to the membrane or localized to the space on the inner side of the peptidoglycan layer. Electron

micrographs of *Anabaena* PCC 7120 intercellular junctions appear to show an intact peptidoglycan layer around each cell and sometimes a distinct "junctional space" between these peptidoglycan layers (Fig. 3) is observed, which could be a barrier for secreted large molecules.

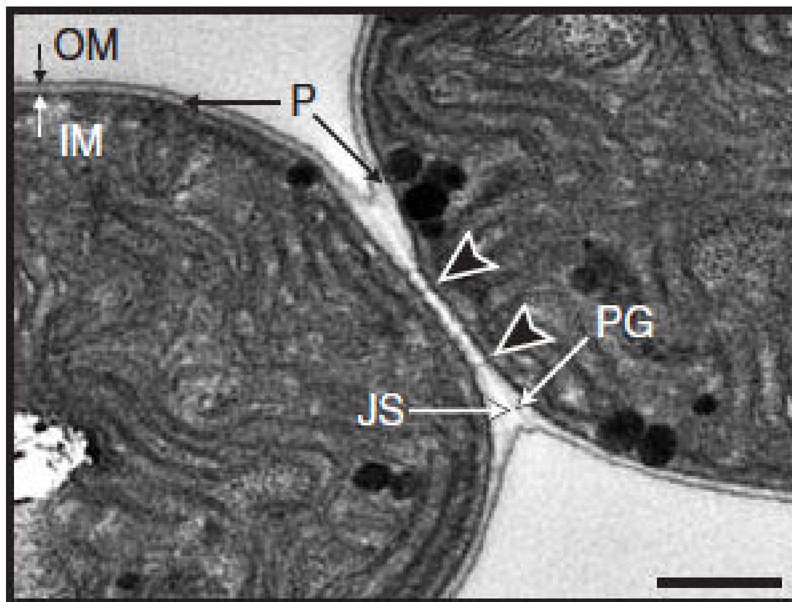

FIG. 3. Transmission electron micrograph of the junction between two vegetative cells. Arrowheads indicate microplasmodesmata, which are potential cell-to-cell channels. IM, inner membrane; OM, outer membrane; P, periplasm; PG, peptidoglycan; JS, junctional space. Size bar, 0.5 µm.

Recent data support the exchange of molecules through intercellular junctions or channels directly connecting the cytoplasm of adjacent cells (123). Calcein, a small fluorescent molecule loaded into the cytoplasm of cells was found to quickly diffuse between cells. Electron micrographs suggest the existence of connections between adjacent cells named microplasmodesmata (Fig. 3) (59, 60). The microplasmodesmata may be channels formed by protein oligomers, and FraG (SepJ) has been suggested as a candidate channel forming protein (56, 130). FraG has a large extracytoplasmic domain that may be involved in spanning the cell wall and bridging the gap between adjacent cells (56). Localization studies with a FraG-GFP reporter showed that it is expressed in both cell types and that it is localized at intercellular septa (56). *fraG* mRNA levels increase after nitrogen depletion and *fraG* null mutants are unable to differentiate heterocysts completely (56, 130). The diffusion of calcein required FraG, suggesting that FraG could be the channel forming protein connecting the cytoplasm of two adjacent cells (123).

The timeline of heterocyst development begins with sensing combined-nitrogen limitation and culminates with nitrogen fixation in the mature heterocyst. Heterocyst development is complete in about 20 hours and involves cellular differentiation of selected vegetative cells into heterocysts and less obvious changes in gene expression and metabolism in the remaining vegetative cells. Several studies employing DNA microarray methods show global changes in gene expression after nitrogen step-down (29, 43, 44, 156). The process of differentiation is reversible if a source of combined

nitrogen is added within 9-12 hours after nitrogen deprivation, after which point the cells are “committed” to forming heterocysts (Fig. 4) (168, 195).

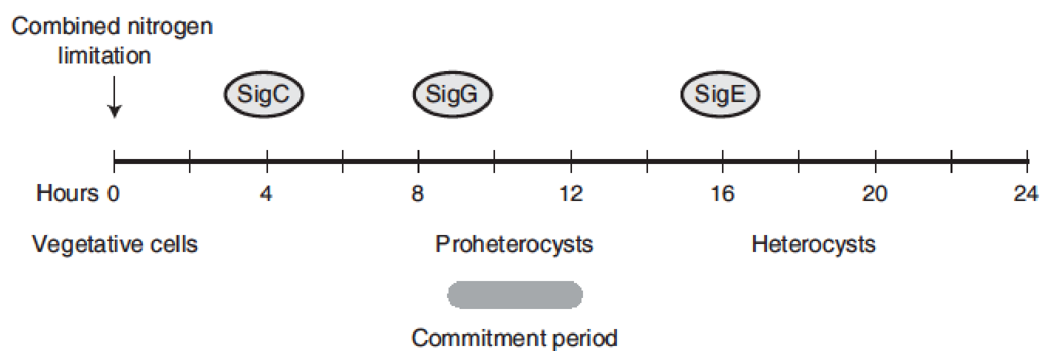

FIG. 4. Developmentally regulated sigma factors in *Anabaena* PCC 7120. The position of the sigma factors icons mark the times that GFP transcriptional reporters for the sigma factors are upregulated during heterocyst development. Filaments are composed of vegetative cells before nitrogen limitation. Immature proheterocysts are observed during approximately the same period of time when cells become committed to complete differentiation. Mature heterocysts are present by 20 hours.

Initiation and early stages of heterocyst development

The presence of a source of combined nitrogen such as ammonium or nitrate inhibits the differentiation of heterocysts. In cyanobacteria, 2-oxoglutarate, an intermediate in the Krebs cycle, constitutes the signal for nitrogen deprivation (103, 198). The Krebs cycle in cyanobacteria is incomplete because of the lack of 2-oxoglutarate dehydrogenase. As a result, 2-oxoglutarate's main function is to serve as a precursor in a variety of biosynthetic reactions. It is the primary carbon skeleton for incorporation of ammonium and is considered the metabolic junction between carbon and nitrogen balance in cyanobacteria (125, 129, 179). Nitrogen-limiting conditions result in an increase in the levels of 2-oxoglutarate. An artificial analog of 2-oxoglutarate, 2,2-difluoropentanoic acid, DFPA, added to medium resulted in heterocyst development even in the presence of ammonium, showing that 2-oxoglutarate level plays a key role in controlling heterocyst development (103).

NtcA, a transcriptional regulator belonging to the CRP (cyclic AMP receptor protein) family of proteins, senses 2-oxoglutarate levels (Fig. 5). The protein is conserved in all cyanobacteria and regulates a number of genes involved in carbon and nitrogen metabolism (79, 109). In *Anabaena* PCC 7120, NtcA is required for the expression of the genes in pathways for ammonium and nitrate assimilation, as well as heterocyst development. The *ntcA* gene is induced soon after nitrogen deprivation and is autoregulated (128, 147, 148). *ntcA* mutants are unable to use nitrate as the sole source of nitrogen and are blocked from initiating heterocyst development (58, 183). Genes that are activated by NtcA typically have the consensus binding site TGTA-(N_8)-TACA

centered at -41.5 nucleotides upstream of the transcription start point (TSP) (79). The DNA binding activity of NtcA is enhanced in the presence of 2-oxoglutarate and 2-oxoglutarate is necessary for transcriptional activation by NtcA (166, 178). Additionally, DFPA, the synthetic analogue of 2-oxoglutarate, stimulates DNA binding activity of NtcA *in vitro* (36, 103).

HetR is a master regulator of heterocyst development and plays a key role in differentiation and pattern formation (23). *hetR* is one of the earliest genes induced in differentiating cells and is positively autoregulated (Fig. 5). Transcription of *hetR* increases as early as 30 minutes after nitrogen deprivation and by 3.5 hours expression is confined to spaced foci arranged in a pattern similar to that of differentiating cells (10, 24). Null mutants of *hetR* fail to produce heterocysts, and overexpression of *hetR* (23, 24) and particular point mutants (94) result in increased heterocyst frequency.

HetR protein is a serine type protease that has autoprotease activity and DNA binding activity that requires formation of a HetR homodimer (83, 205). Mutations in either activities block heterocyst development at an early stage. The serine residues S152 and S179 are required for autoprotease activity and heterocyst differentiation. The heterocyst inhibitory peptide PatS interferes with HetR DNA-binding activity *in vitro* (83). A $hetR_{R223W}$ point mutant is insensitive to the main inhibitory signals of pattern formation, PatS and HetN (Fig. 5), and produces a conditionally lethal phenotype due to complete differentiation under nitrogen limiting conditions (94).

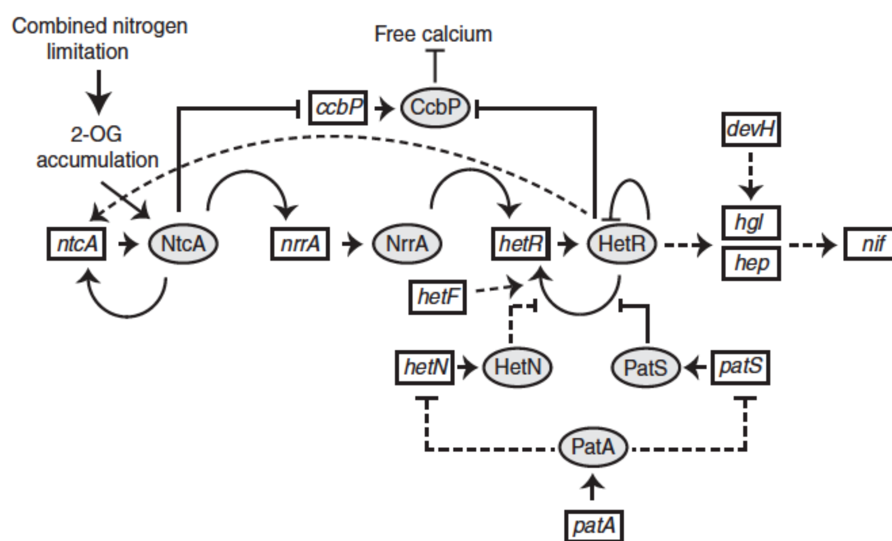

FIG. 5. Model of regulatory interactions during heterocyst development. For clarity, the figure shows only selected genes, proteins and events. Open box represent genes and gray ovals represent proteins. Lines ending in arrows and bars indicate positive and negative interactions, respectively. Dashed lines indicate indirect and/or unknown interactions or missing steps. Short arrows are between genes and their products.

Expression of *ntcA* and *hetR* show a mutual dependency during heterocyst development (128). *hetR* is not induced in an *ntcA* mutant and *ntcA* expression is transiently induced in a HetR-dependent manner (58, 128). Expression of some genes involved in the developmental process is dependent on both *ntcA* and *hetR*; however, in some cases the dependency is probably due to the interdependent upregulation of *hetR* and *ntcA* (52, 73, 177). NrrA, a response regulator (Fig. 5), has been identified as the regulatory link between NtcA and HetR (42). *nrrA* is transcribed in differentiating cells within 3 h after nitrogen deprivation and is directly dependent on NtcA (42, 126). Earlier work on the transposon reporter strain TLN14, now known to be a fusion to the *nrrA* gene, showed rapid induction one hour after nitrogen stepdown (25). An *nrrA* mutant strain shows a delay in heterocyst development caused by a delay in accumulation of HetR, and extra copies of *nrrA* result in increased expression of *hetR* and thus, increased heterocyst frequency (42, 43).

In *Anabaena* PCC 7120, the calcium-binding luminescent protein aequorin was used to detect an increase in intracellular calcium levels after nitrogen deprivation (170). The Ca^{2+} reporter obelin was used to show a 10-fold higher Ca^{2+} concentration in heterocysts compared to vegetative cells at 4 hours after nitrogen deprivation (204). The increase in Ca^{2+} concentrations is due to the decreased expression of *ccbP*, which encodes a calcium sequestering protein, CcbP (204). *ccbP* message is downregulated in differentiating cells and absent in mature heterocysts (Fig.5). Inactivation of *ccbP* causes a multiple-contiguous-heterocyst (Mch) phenotype and overexpression inhibits heterocyst development. It has been hypothesized that a regulatory pathway consisting of HetR,

CcbP, and NtcA controls intracellular free calcium; HetR specifically degraded CcbP in a calcium dependent manner and *ccbP* downregulation required 2-oxoglutarate-dependent binding of NtcA to its promoter (160). The increase in Ca^{2+} in differentiating cells is thought to be important for HetR's Ca^{2+}-dependent serine protease and/or other Ca^{2+}-dependent proteolytic activities.

HetF influences heterocyst development by a positive effect on *hetR* expression (Fig. 5) (153, 189). In mutants of *hetF*, expression of *hetR* was not localized to heterocysts and initiation of heterocyst development was not seen (189). Overexpression of *hetF* produces a multiple-contiguous-heterocyst phenotype but only in the absence of combined nitrogen (189). Like HetR, HetF is likely to be a protease (153). An elegant mosaic analysis was used to show that HetF is required specifically in cells that differentiate (153). However, it is also thought that HetF plays a role in restricting *hetR* expression and the accumulation of HetR protein to differentiating cells (189).

Like *hetF, patA* also influences heterocyst development via a positive effect on *hetR* expression, but its affect is related to pattern formation (Fig. 5) (105, 153). The *patA* gene encodes a response regulator similar to the CheY protein, which functions as a phosphorylation-activated switch (105). *patA* mutants form heterocysts almost exclusively at the ends of filaments. This mutant phenotype is maintained even when *hetR* is overexpressed in a *patA* mutant background, suggesting that *patA* acts downstream of *hetR* (24, 105). It is possible that PatA influences heterocyst development by attenuating the negative effects of the main inhibitory signals of heterocyst pattern formation, PatS and HetN (138).

The *hetC* gene, which encodes a member of the family of ATP-binding cassette type exporters, is required for an early step in the differentiation of heterocysts as observed by a P_{hetC}-*gfp* reporter, which showed an increase in expression in proheterocysts and heterocysts (93, 127). A *hetC* mutant carrying a P_{hetR}-*gfp* reporter shows a pattern of weakly fluorescent cells that are blocked from further stages of development (182, 192). It is possible that the *hetC* mutant fails to complete an early step in morphogenesis of the envelope that then triggers a developmental checkpoint that prevents further differentiation.

Two novel genes, *hetL* and asr1734, have been shown to be involved in regulating heterocyst development, but their exact roles and biochemical functions remain unclear. HetL is a pentapeptide-repeat protein composed almost entirely of 40 tandem pentapeptide repeats forming 10 complete coils (133). The *hetL* gene was isolated in a genetic screen designed to identify genes involved in PatS signaling (106). A *hetL*-null mutant showed normal heterocyst development and diazotrophic growth, indicating that *hetL* plays a non-essential role in heterocyst development. Overexpression of *hetL* in a *patS*-overexpression strain allows heterocysts to form by bypassing the PatS inhibitory signal. Overexpression of *hetL* in the wild type produces a multiple-contiguous-heterocyst phenotype. *hetL* overexpression even induces partial heterocyst development in an *ntcA*-null mutant.

The asr1734 gene is found in only heterocyst-forming cyanobacteria and a $P_{asr1734}$-*gfp* reporter showed localized expression in proheterocysts and heterocysts after nitrogen step-down. An asr1734 knockout mutant shows elevated levels of *ntcA* mRNA and

forms 15% heterocysts and a weak Mch phenotype in media without fixed nitrogen. Overexpression of asr1734 inhibits heterocyst development in the wild type and in two genetic backgrounds that stimulate heterocyst formation, a *patS* null and a $hetR_{R223W}$ mutant, suggesting that Asr1734 acts downstream of PatS and HetR (190).

Cyclic-di-GMP signaling appears to be involved in heterocyst development. The all2874 gene, which encodes a diguanylate cyclase, is required for normal heterocyst development during growth under conditions of high light intensity (131). An all2874 mutant shows a significant reduction in heterocyst frequency and reduced vegetative cell size. An all2874 mutant strain carrying a P_{patS}-*gfp* transcriptional reporter did not show normal upregulation of the reporter, indicating that the decrease in heterocyst frequency is due to an early block in differentiation.

Heterocyst cellular differentiation produces an environment for nitrogen fixation

The middle and later stages of heterocyst development are distinguished by structural and physiological changes. These changes begin with morphogenesis of the heterocyst envelope by the deposition of an outer polysaccharide layer and an inner glycolipid layer, which decrease the entry of oxygen into the heterocyst (Fig. 2) (50). Mutants that lack the envelope polysaccharide or the glycolipid layer are unable to grow diazotrophically in the presence of air (49, 82, 135, 186, 187).

Deposition of the external polysaccharide layer is one of the earliest morphological changes during heterocyst differentiation. DevR and HepK, which comprise a two component regulatory system, are involved in biosynthesis of the polysaccharide layer

(206). The *hep* genes *hepA*, *hepB*, and *hepC*, and a cluster of genes present around *hepA* are required for the deposition of the polysaccharide layer (82, 181, 208).

The heterocyst glycolipid layer is assembled beneath the polysaccharide layer and is composed of fatty alcohols glycosidically linked to sugar residues. The *hglB*, *hglC*, *hglD*, and *hglE* genes along with a cluster of nearby genes are required for the synthesis of these glycolipids (28, 49). DevH, a trans-acting regulatory protein, is required for the formation of the glycolipid layer, either by directly regulating the expression of the genes or indirectly through other gene products (Fig. 5) (150). The *hglK* gene is required for the localization of the glycolipids and may be directly involved in their deposition (9). ORF all5341, named *hglT*, is predicted to encode a glycosyl transferase and is required for the formation of the glycolipid layer (5).

Differentiation and maturation of heterocysts is dependent on DevBCA and HgdD, which are thought to be a glycolipid exporter and outer membrane efflux tunnel, respectively (51, 121), and autolysin HcwA, which is presumably involved in cell wall remodeling (207). The regulatory genes *hepK*, *hepN*, *henR*, and *hepS* are required for normal heterocyst maturation (104).

Recent proteomic studies of the heterocyst and vegetative cell wall have identified proteins in the inner and outer membranes as a step towards understanding the dynamics of the cell wall proteome during cell differentiation (119, 120). Analysis of proteins of the outer membrane of vegetative cells and of heterocysts revealed a high similarity, suggesting that the relative protein concentrations within the outer membrane may vary between the two cell types but not the composition (134).

The cell wall-related proteins can be categorized as proteins involved in (I) signal transduction, (II) synthesis, and (III) transport. Among the signal transduction proteins in *Anabaena* PCC 7120, histidine kinases and protein phosphatases have recently been shown to act downstream of NtcA and have been suggested to be involved in heterocyst cell wall formation (37, 180). In *Anabaena* PCC 7120, 131 genes encode putative histidine kinases, 52 genes code for serine/threonine and tyrosine kinases, and 13 genes code for a group of histidine and Ser/Thr hybrid (HSTKI) proteins (136, 143). Two genes, *pkn30* and *pkn44*, that encode HSTKI proteins are involved in regulation of HGL2 synthesis (159). A single mutant of either kinase has no obvious phenotype, but a double mutant has an immature HEP layer and is defective in synthesis of the HGL2. PrpJ, a Ser/Thr phosphatase encoded by all1731, is involved in the regulation of the HGL1 (86).

Among the proteins that define cell wall structure, OmpA, encoded by all3289 and alr5049, acts as a physical linker between the outer membrane and the peptidoglycan layer. The alr1278 gene encoding OstA/Imp is thought to be involved in remodeling of the lipid content of the outer membrane during heterocyst development (134). OstA/Imp is involved in transport of lipid A to the cell surface (15). Expression analysis showed that alr1278 was upregulated early following nitrogen deprivation (43).

The outer membrane of *Anabaena* PCC 7120 includes integral membrane proteins and lipoproteins attached to the outer membrane by N-terminal lipids (134). These proteins comprise porin-type proteins, proteins that define the cell wall structure, and proteins involved in transport and assembly (134). In *Anabaena* PCC 7120, all4499

encoded by ORF all4499 and all4550 encoded by ORF alr4550, which were identified based on homology to the outer membrane porin OprB of *Pseudomonas aeruginosa*, were identified in membranes of vegetative cells and heterocysts and were the most abundant outer membrane proteins in the heterocyst (119, 121). Genome wide expression analysis of genes after nitrogen deprivation showed an increase in expression of all4499 8 hours after nitrogen deprivation and a decrease in expression of all4550 24 hours after nitrogen depletion. In addition, ORF alr0834, classified as encoding a general porin, is upregulated during the early stages of heterocyst differentiation .

A large number of outer membrane proteins are involved in protein transport and assembly. all0406 encodes a protein with high similarity to autotransporters, which are gram-negative extracellular proteins that are characterized by the ability to translocate themselves across the outer membrane (134). alr3592 encodes a protein that belongs to the PapC family of proteins, which are involved in the excretion of pilus subunits in Gram-negative pathogens (39, 134). However, neither the presence of the protein in the outer membrane nor its developmental regulation has been studied. Two other proteins, all5116 and alr1659, show homology to the FhaC two-partner secretion system but their role in protein transport has not been confirmed. Omp85, a protein that belongs to a superfamily of polypeptide transporters that includes FhaC, has three homologues in *Anabaena* PCC 7120, alr0075, alr2269, and alr4893. alr2269 is the most abundant in vegetative cells and heterocysts as identified by mass spectrometry. It shows highest homology to the chloroplast-localized Toc75 protein involved in protein translocation (119, 121, 134). HgdD, a TolC like protein, is an outer membrane efflux channel and

functions in HGL-layer assembly. The *hgdD* gene is upregulated during heterocyst development and the protein is essential for HGL-layer formation. A *hgdD* knockout mutant is defective in deposition of the HGL layer and the phenotype is similar to the defect observed for a *devBCA* knockout. These data suggest the formation of a DevBCA-HgdD secretion complex that is essential for the formation of the HGL layer (121).

Heterocyst-specific uptake hydrogenase, encoded by the *hupSL* operon, recovers reductant from H_2 produced by nitrogenase (165). In *Nostoc punctiforme* the upstream region of the *hupSL* operon contains a putative NtcA binding site (81). Mobility shift assays showed that NtcA binds to the *hupSL* promoter. However, deletion of the NtcA binding site did not affect the expression of *hupSL* significantly. Study of cis-acting regulatory regions in the *hupSL* promoter in *N. punctiforme* using truncated sequences of the *hupSL* promoter fused to either *gfp* or *luxAB* reporter genes indicates that cis-acting regulatory sequences required for heterocyst-specific expression of the *hupSL* genes may be located in a short 316 –bp DNA sequence starting at 57-bp upstream of the translation start point.

Whereas NtcA is not required for transcriptional regulation of *hupSL* in *Nostoc punctiforme*, NtcA is required for the induction of *hupSL* in *Anabaena variabilis* after nitrogen step-down as shown by the reduced induction of *hupSL* in an *ntcA* mutant. A consensus NtcA-binding site is upstream of *hupSL*, and NtcA binds to this region (184).

Heterocyst development is accompanied by changes in the photosynthetic apparatus and carbon metabolism to provide ATP and low potential reductant for nitrogen fixation

(115, 188). Respiration by cytochrome c oxidase generates ATP for nitrogen fixation and helps create the microoxic environment. Whereas *cox1* encoding mitochondrial type cytochrome c oxidase is expressed in vegetative cells, *cox2* and *cox3* are expressed only in heterocysts and are required for growth during nitrogen fixation (89, 174).

RNA polymerase sigma factors are developmentally regulated

The control of gene expression in bacteria is primarily regulated at the level of transcription initiation. Bacteria contain an RNA polymerase (RNAP) enzyme that requires an additional σ factor for promoter recognition and transcription initiation. In many cases a bacterium utilizes specific sigma (σ) factors to activate a set of genes in response to environmental and intracellular signals (27).For example, in *Bacillus subtilis*, temporally and spatially regulated σ factors modulate the expression of genes during starvation-induced sporulation (27, 99). It was hypothesized that global changes in gene expression during heterocyst development in *Anabaena* PCC 7120 are regulated by σ factors (2, 95). The σ 70 family of sigma factors has been divided in four major groups: group 1 is responsible for transcription of housekeeping genes, group 2 is similar to group 1 but not essential under laboratory growth conditions, group 3 is involved in the expression of genes associated with sporulation, motility, stress response, and heat shock, and group 4 participates in multiple biological processes (27). *Anabaena* PCC 7120 has twelve putative sigma factors identified by sequence similarity (2, 196). Genetic analysis has not identified any sigma factor genes that are essential for

heterocyst differentiation or nitrogen fixation, suggesting that there is a level of functional redundancy between the sigma factors (95).

Genetic analysis using *gfp* as a reporter identified three sigma factor genes, *sigC*, *sigG*, and *sigE*, that are developmentally upregulated after nitrogen deprivation (2). The expression of *sigC* is upregulated 4 hours after nitrogen step-down and by 10 hours the expression is restricted mostly to single cells in a heterocyst-like pattern (Fig. 4). SigC may be involved in regulating early heterocyst-specific genes (2). The expression of *sigE* (*alr*4249) is low in vegetative cells grown in nitrate-containing media and is upregulated in individual cells at about 16 hours after nitrogen depletion. At 24 hours after nitrogen step-down, a *sigE* reporter is expressed in mature heterocysts suggesting that SigE may be involved in the expression of late heterocyst specific genes such as the *nif* and *hup* genes. Expression of a *sigG* reporter is seen in cells in media containing nitrate but decreases in all cells soon after nitrogen depletion. At around 9 hours after nitrogen step-down, increased expression is localized to single cells and by 16 hours, the expression is localized to morphologically distinguishable proheterocyst cells; at 24 hours, the heterocysts show decreased reporter expression. These data suggest that SigG is involved in the expression of genes during the middle stages of differentiation such as those involved in morphogenesis and the creation of a microoxic environment, and possibly genes that are necessary for commitment to complete differentiation(2).

Heterocyst pattern formation requires dynamic signaling

Pattern formation has been observed in many prokaryotic systems. *Myxococcus xanthus* forms a complex three-dimensional pattern when starving bacteria self-organize to form fruiting bodies (87), and *Pseudomonas aeruginosa* produces biofilms during infection (96). Temporary patterns have been studied in *Escherichia coli* and *Salmonella typhimurium* when these organisms swim in gradients of nutritional chemoattractant (11, 22). The developmentally regulated pattern of heterocysts in multicellular cyanobacteria offers a model to understand biological pattern formation, signals that control pattern formation, and downstream regulatory pathways.

A longstanding model of heterocyst pattern formation is that differentiating cells produce an inhibitory signal that diffuses along a filament to form a gradient that inhibits differentiation of neighboring cells in a concentration dependent manner (185, 188). The *patS* gene in *Anabaena* PCC 7120 is essential for normal pattern formation and is predicted to encode a 13- to 17- amino-acid peptide depending on the start codon used *in vivo* (Fig. 5) (194, 195). All filamentous cyanobacteria for which a genome sequence is available contain *hetR* and *patS* genes, and although the *patS* genes encode peptides from 13 to 90 amino acids, they all have an identical five amino acid sequence, RGSGR, at their C-terminal ends (199). This PatS pentapeptide sequence may be the essential element of the predicted diffusible inhibitor controlling, at least in part, heterocyst pattern.

The *patS* gene is transcribed early during heterocyst development. Expression analysis using northern RNA blots and a *patS*-*lacZ* reporter showed that *patS* is

upregulated by 6 hours after nitrogen deprivation (194). A *patS-gfp* reporter strain in nitrate-containing media showed low levels of expression in all cells. At 8 to 10 hours after nitrogen step-down increased fluorescence was localized to small groups of cells or individual cells (195). By 12 to 14 hours after nitrogen step-down, most of the GFP signal was localized in individual cells that were arranged in a pattern similar to that of mature heterocysts, and by 18 hours the fluorescence was localized almost entirely to heterocysts (also see Fig. 2).

A *patS* null mutant produces heterocysts even in the presence of nitrate, and forms multiple contiguous heterocysts after nitrogen deprivation with up to 30% of the cells differentiating to form heterocysts while extra copies or overexpression of *patS* results in a complete suppression of heterocysts (194, 195). Mutations in the last five carboxy-terminal residues of PatS result in a loss-of-function phenotype. A synthetic RGSGR pentapeptide, PatS-5, blocks heterocyst differentiation at nanomolar concentrations while a four amino acid synthetic peptide was unable to inhibit differentiation (194). Addition of PatS-5 to growth medium of a *patS* null mutant at a concentration that decreases the frequency of heterocysts to the wild-type level of about 10 percent does not restore a wild-type heterocyst pattern, but ectopic expression of *patS* from a heterocyst-specific promoter, *hepA*, in a *patS* null mutant background does restore the normal pattern. These data are consistent with PatS functioning as a diffusible inhibitor and that a gradient of PatS is required for the establishment of a normal pattern (66). However, other somewhat more complicated mechanisms are possible and direct support

for this hypothesis would require, for example, the development of methods to detect gradients of small molecules in cyanobacterial filaments.

The PatS signal must move from differentiating cells to neighboring cells to inhibit HetR, however, a *patS5* minigene expressed in heterocysts in a *patS* null background does not produce a normal pattern, indicating that PatS-5 produced by the minigene cannot function as a cell-to-cell signal and may remain in the cytoplasm of the differentiated cell, possibly because of the lack of a domain that is needed for transport to adjacent cells or export to the periplasmic space (191). The ability of large peptides containing an internal RGSGR sequence motif to inhibit heterocyst development support the idea that the receptor for PatS is localized in the cytoplasm (191). *In vitro* studies show that the DNA-binding activity of HetR is inhibited by PatS-5 pentapeptide in a dose-dependent manner, implying that the HetR-PatS ratio is important for controlling heterocyst differentiation and indicating that HetR is a PatS receptor (83). Upregulation of *hetR* expression is abolished when PatS-5 is added to the growth medium (83). $hetR_{\mathrm{R223W}}$ is insensitive to the PatS inhibitory signal and overexpression of the $hetR_{\mathrm{R223W}}$ allele in a $hetR_{\mathrm{R223W}}$ mutant background results in a conditionally lethal phenotype after nitrogen deprivation (94). Overexpression of *patS* and *hetR* in a synthetic operon inhibits heterocyst formation indicating that *patS* acts downstream of *hetR*, which is consistent with PatS inhibiting the activity of HetR (138).

The *hetN* gene encodes a protein similar to ketoacyl reductase. Like *patS*, overexpression of *hetN* gene results in complete suppression of heterocyst development (Fig. 5). A *hetN* null mutant shows a wild-type pattern at 24 hours after nitrogen

depletion but forms multiple contiguous heterocysts by 48 hours, suggesting that *hetN* is not necessary for *de novo* heterocyst pattern, and instead is required for the maintenance of the preexisting heterocyst pattern (26). A *patS*-*hetN* double mutant shows complete differentiation of heterocysts in the absence of combined nitrogen, producing a lethal phenotype. These data indicate that the two inhibitory signal pathways may be independent (14). Together, *patS*- and *hetN*-mediated inhibitory pathways are the primary mechanism for establishing heterocyst pattern.

However, other factors can influence pattern formation such as the cell cycle, the physiology of individual cells, signals or nutrients from vegetative cells, and the products of nitrogen fixation (1, 66, 198). It is known that single cells from fragmented filaments will not differentiate but instead always divide first followed by differentiation of one of the daughter cells, and it has been suggested that an activator, possibly 2-oxoglutarate, of differentiation is produced by vegetative cells and serves as a regulator of pattern formation (203). There is some evidence for the products of nitrogen fixation supplied from heterocysts contributing to the average spacing between heterocysts along filaments (1, 195). However, data from *Anabaena variabilis*, which can fix nitrogen in vegetative cells, have been interpreted as an argument against the products of nitrogen fixation having a significant effect on heterocyst pattern (168). This seemingly simple question has been difficult to answer because mutants in which heterocysts cannot supply nitrogen to the filament cannot grow diazotrophically, which makes determining the maintenance of heterocyst pattern impossible.

In *Nostoc punctiforme*, the novel genes *patU* and *patN* are required to inhibit excess heterocyst differentiation (115). In *Anabaena* PCC 7120, a cluster of genes that influence pattern formation and heterocyst differentiation, *hetZ*, *patU5*, and *patU3*, is conserved among heterocyst- and non-heterocyst-forming filamentous cyanobacteria (201). All three genes are upregulated after nitrogen deprivation and encode proteins with unknown function. A *hetZ* mutant does not differentiate heterocysts, unlike a *patU3* mutant, which forms multiple contiguous heterocysts. When *patU3* is inactivated in a *patA* mutant background, the *patA* mutant phenotype is abolished, restoring the formation of intercalary heterocysts (201). Although these genes are clearly involved in the regulation of heterocyst development, their biochemical functions remain to be determined.

In *Anabaena* PCC 7120, three DNA rearrangements take place during the late stages of heterocyst development that affect nitrogen fixation and uptake hydrogenase operons (32-34, 61-63). The rearrangements are the result of developmentally programmed site-specific recombination between direct repeats that flank the DNA elements that are deleted from the chromosome (61). One rearrangement excises an 11-kb element from the *nifD* gene (20, 63). The excision is catalyzed by *xisA* located on the 11-kb element (19, 62, 74, 102). A *xisA* null mutant forms heterocysts but is unable to excise the element or grow on media lacking a source of combined nitrogen (64, 65). A second rearrangement excises a 55-kb element from the *fdxN* gene (65) and requires the *xisF*, *xisH*, and *xisI* genes (35, 149). A third programmed DNA arrangement deletes a 10.5-kb element from the *hupL* gene and requires the *xisC* recombinase gene (33, 34). All three

DNA elements appear to be parasitic DNA sequences that are passed to daughter vegetative cells as they are carried innocuously in genes that are only required in heterocysts. And, all three DNA elements have acquired the ability to excise from the chromosome in terminally differentiated heterocysts so that the genes in which they reside can be expressed. This developmental regulation has apparently evolved independently for each element but the regulatory mechanism is not yet known.

Nitrogen fixation in heterocysts

Heterocyst development culminates with nitrogen fixation, which is the ATP-dependent process of reducing atmospheric nitrogen to ammonia by the enzyme nitrogenase. In many heterocyst-forming cyanobacteria, including *Anabaena* PCC 7120, nitrogenase is synthesized only in heterocysts, where it is protected from irreversible inactivation by oxygen (50). However, some heterocyst-forming cyanobacteria contain an additional set of *nif* genes that allows nitrogen fixation in vegetative cells under anoxic conditions (167). Nitrogenase is a well-conserved enzyme in all nitrogen-fixing organisms and consists of two components: dinitrogenase (Mo-Fe protein) and dinitrogenase reductase (Fe protein) (154). Dinitrogenase reductase serves to transfer electrons from electron donors such as ferredoxin or flavodoxin to dinitrogenase. In *Anabaena* sp. PCC 7120, the *fdxH* gene codes for a heterocyst-specific ferredoxin (112). In addition to the common molybdenum nitrogenase, some cyanobacteria contain an alternative nitrogenase that utilizes a vanadium cofactor (146).

In *Anabaena* PCC 7120, the nitrogen-fixation (*nif*) genes are expressed specifically in heterocysts late during development, between 18 and 24 hours after nitrogen deprivation (47, 64). The *nifHDK* genes encode the molybdenum-containing nitrogenase enzyme complex. Upstream of the *nifHDK* genes is another *nif* gene cluster, the *nifB-fdxN-nifS-nifU* (122). The *nifVZT* genes form a *nif* gene cluster separated from the main *nif* gene cluster (161). Other cyanobacterial *nif* genes include *nifE*, *nifN*, *nifX*, and *nifW* (167, 188). Little is known about the transcriptional regulators that control the expression of the *nif* genes in cyanobacteria. It has been recently shown that SigE is required for normal levels of *nifH* expression; in a *sigE* mutant strain, expression of *nifH* is delayed and reduced. Disruption of a cyanobacterial regulatory protein *PipX* led to impaired diazotrophic growth, reduced nitrogenase activity, and impaired activation of the nitrogenase structural genes. that nitrogenase activity of the mutant was low under oxic conditions, It is likely resulting from inefficient protection against oxygen (78).

In *Anabaena variabilis*, a promoter upstream of *nifB1* promoter contributes the majority of expression of the *nifHDK1* genes, while a second promoter within the *nifU* gene contributes to a lower proportion of expression (172). Characterization of the previously reported *nifH1* transcriptional start sites by 5'RACE revealed that these 5' ends resulted from processing of larger transcripts rather than by *de novo* transcription initiation. The 5' positions of the *nifH1* transcripts lie at the base of a stem-loop structure that may stabilize the *nifHDK1* transcripts. (172).

Although important progress has been made in our understanding the initial and middle stages of the development of heterocyst, the genetic, molecular and biochemical

events that govern the later stages of development, particularly the culminating event of nitrogen fixation, remain unknown. In case of the *nif* genes, sigma-54 and cognate activators, which play regulatory roles in proteobacterial diazotrophs, are absent from cyanobacteria, indicating that a different *nif* regulatory network is likely to control the transcription of the *nif* genes in these organisms. The identification of novel regulatory mechanisms, will further the understanding of prokaryotic gene regulation. Cyanobacteria are widely distributed and contribute significantly to the marine and global nitrogen cycle. They are of significance in the production of renewable chemical products, including biofuels, various alcohols, organic molecules such as proteins, carbohydrates, and lipids, and other chemicals. The identification of regulators of cyanobacterial late genes can increase the efficiency of production of such resources, particularly those that require a microoxic environment - products that are sensitive to oxygen or whose production is sensitive to oxygen.

CHAPTER II

EXCISION OF THE FDXN ELEMENT IS REQUIRED FOR TRANSCRIPTION OF THE NIFHDK GENES

Overview

In the heterocystous cyanobacterium *Anabaena* sp. strain PCC 7120, the nitrogen-fixation (*nif*) genes are expressed specifically in heterocysts during the late stages of development. We used *gfp* transcriptional fusions to study the regulation of the *nifHDK* genes in *Anabaena* sp. strain PCC 7120. Reporter fusions containing 500 bp of the upstream region of *nifH* did not show GFP fluorescence after nitrogen deprivation indicating that elements essential for transcription are farther upstream. Strain AMC1774 containing a *gfp* fusion to a region extending from the 5' end of *nifD* to the upstream region of *nifB* showed regulated *gfp* expression in differentiating cells at 14 hours after nitrogen deprivation. Mutants that did not excise the *fdxN* element within the *nifB-fdxN-nifS-nifU* gene cluster did not show GFP fluorescence in AMC1774 and were unable to grow under conditions of nitrogen deprivation. Absence of GFP fluorescence from a reporter construct at a neutral site showed that transcription of the *nifHDK* genes requires the *nifB* promoter. The previously identified transcription start site at 123-bp upstream of the *nifH* ATG is positioned at the base of a stem-loop structure that may contribute to increased stability of the *nifHDK* transcripts.

Introduction

The filamentous cyanobacterium *Anabaena* sp. strain PCC 7120 (hereafter *Anabaena* PCC 7120) responds to nitrogen deprivation by differentiating a developmental pattern of single heterocysts separated by about 10 to 20 photosynthetic vegetative cells. Mature heterocyst structure and metabolic activity serve to spatially separate the oxygen-sensitive process of nitrogen fixation from oxygen-yielding photosynthesis. The process of nitrogen fixation, in which atmospheric nitrogen is reduced to ammonia, is catalyzed by nitrogenase, a well-conserved enzyme in all nitrogen-fixing organisms (154). In *Anabaena* PCC 7120, the *nifHDK* genes encode the major structural components of the molybdenum-containing nitrogenase enzyme complex: *nifH* encodes dinitrogenase reductase and *nifD* and *nifK* encode the dinitrogenase alpha and beta subunits, respectively. Upstream of the *nifHDK* gene cluster is another set of *nif* genes, *nifB-fdxN-nifS-nifU* (122). The *nifB*, *nifS* and *nifU* genes encode co-factors required for nitrogenase and *fdxN* encodes a bacterial-type ferredoxin.

Although *nifH* was the first *nif* gene to be sequenced (117), little is known about the regulation of the *nifHDK* genes. In most diazotrophic organisms, expression of the *nif* genes is regulated by environmental cues such as oxygen tension and availability of combined nitrogen (41, 158). Heterocyst-specific *nif* genes are developmentally regulated and expressed specifically in heterocysts between 18 and 24 hours after nitrogen deprivation (186), when the environment within is microoxic, thus protecting the nitrogenase enzyme from irreversible inactivation by oxygen (47, 64). For *Anabaena* PCC 7120, heterocyst-specific putative transcription start sites have been mapped as

mRNA 5' ends upstream of the *nifH*, *nifB*, and *fdxH* genes (12, 72, 122). However, a defined *nif* or heterocyst-specific promoter has not been identified (12).

In *Anabaena* PCC 7120, two site-specific DNA rearrangements that affect nitrogen fixation take place in the *nif* locus during the late stages of heterocyst development (Fig 6). One rearrangement excises an 11-kb element from the *nifD* gene and is catalyzed by the *xisA* gene product located on the element (63). A second rearrangement, which excises a 59-kb element from the *fdxN* (bacterial-type ferredoxin) gene, requires the *xisF*, *xisH*, and *xisI* genes that are located on the *fdxN* element (8, 35, 62). Mutants that do not excise either element from heterocysts do not grow on media lacking a source of combined nitrogen (33).

In the current study, we used *gfp* reporter fusions to identify cis-acting sequences required for transcription of the *nifHDK* genes. We show that excision of the *fdxN* element within the *nifB-fdxN-nifSU* genes is required for the transcription of the *nifHDK* genes. This suggests that transcription of these genes initiates upstream of the *fdxN* element. In *Anabaena variabilis,* it was recently shown that the majority of *nifH1* transcripts originate from the *nifB1* promoter (173). Our data suggest that transcription of the *nifHDK* genes in *Anabaena* PCC 7120 also originates from the *nifB* promoter.

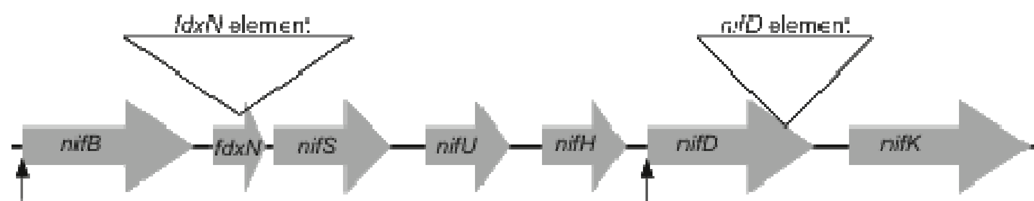

FIG. 6. Schematic of the *nif* gene cluster in Anabaena sp. strain PCC 7120. Arrows indicate sites of *gfp* fusion in AMC1774 and KKC4.

Materials and methods

Strains and conditions. Strains used in this study are listed in Table 1. *Anabaena* sp. strain PCC 7120 and its derivatives were grown in BG-11 or BG-11_0 (which lacks sodium nitrate) medium at 30°C as previously described (64). Cultures were grown under fluorescent white light illumination at 75-100 µmol photons m^{-2} s^{-1}. Plasmids were transferred to *Anabaena* PCC 7120 by conjugation from *Escherichia coli* following published protocols (46, 65). For *Anabaena* PCC 7120 strains carrying plasmids, the antibiotics neomycin (Nm; 25 µg/ml) and/or streptomycin (Sm) plus spectinomycin (Sp) (2 µg/ml each) were used for growth on BG-11 or BG-11_0 agar medium in plates; the antibiotic concentrations were reduced by half for liquid cultures. Heterocyst differentiation was induced by nitrogen deprivation as previously described (2).

E. coli strains were maintained in LB (Lennox L) liquid or agar medium supplemented with the appropriate antibiotics.

TABLE 1. Primers used in this study (Chapter II)

Primer	Sequence	Source of reference
KKO 1	AGTAAAGGAGAAGAACTTT	This study
KKO 2	CCCAGCTTGCATGCCTGCAG	This study
KKO 3	ATTAGGGAATAGGAAGAAGC	This study
KKO 4	CATTATCTGCCT CAGTG	This study
KKO 5	ATTAGGGAATAGGAAGAAGC	This study
KKO 6	CATTATCTGCCTCAGTG	This study
KKO 7	CAATTACTAATTACCACATC	This study
KKO 8	TGAATTTCGCAAATGTCGGT	This study

Plasmid constructions. Plasmids and primers used in this study are listed in Table 2. Shuttle vectors pAMKK5, pAMKK6, pAMKK7, and pAMKK25 were constructed by amplifying different *nifH* upstream regions using primers KKO9 and KKO10 (221-bp fragment), KKO11 and KKO12 (350-bp fragment), KKO13 and KKO14 (700-bp fragment), and KKO15 and KKO16 (500-bp fragment), respectively. Amplified fragments were cloned into SalI and SacI sites of pAM1956. The *gfp* reporter plasmid was constructed by amplifying *gfpmut2* using primers KKO1 and KKO2 from pAM1956 (194) and cloning the fragment into SpeI and SacI restriction sites of the conjugal suicide plasmid pRL277 to generate pAMKK1. The P_{nifHD}-*gfp* reporter pAMKK2 was constructed by cloning a PCR-amplified fragment, using primers KKO3 and KK04, of the *nifD* upstream region from +3-bp to -2271-bp relative to the *nifD* ATG into pAMKK1 using XhoI and SpeI sites. For construction of a neutral site vector, a 1.5-kb *gvp* region was amplified using primers KKO5 and KKO6 and cloned into SacI and PstI sites on the P_{nifHD}-*gfp* plasmid to generate pAMKK3. The P_{nifB}-*gfp* transcriptional reporter was constructed by amplifying the *nifB* upstream region with primers KKO7 and

KKO8, containing SacI and Acc651 restriction sites, respectively, and cloning the amplified fragment into pAM1956 (194) to generate pAMKK4. The *xisF* mutant(35) and the conjugal expression library (106) were constructed as previously described. All plasmid constructs were verified by DNA sequencing.

TABLE 2. Strains and plasmids used in this study (Chapter II)

Strain	Characteristics	Source of reference
E. coli strains		
AM1358	*E. coli* strain containing conjugal helper plasmid	(106)
AM1359	*E. coli* strain containing conjugal helper and transfer plasmid	(46)
AM1460	Conjugal transfer plasmid pRK2013	(113)
AM1824	Overexpression library of random gene fragments expressed from the *rbcL* promoter	(106)
Anabaena sp. Strains		
AMC1774	*Anabaena* PCC 7120 carrying pAMKK2 recombined into the chromosome at the *nif* locus	This study
KKC1	*Anabaena* PCC 7120 carrying pAMKK3 recombined into the chromosome at the neutral site locus	This study
KKC2	*Anabaena* PCC 7120 carrying pAMKK4	This study
KKC3	Mutant of AMC1774 obtained by UV mutagenesis	This study

TABLE 2. continued…

Strain	Characteristics	Source of reference
KKC4	Mutant of AMC1774 carrying plasmid 1824-1	This study
KKC5	*Anabaena* PCC 7120 carrying pAMKK5	This study
KKC6	*Anabaena* PCC 7120 carrying pAMKK6	This study
KKC7	*Anabaena* PCC 7120 carrying pAMKK7	This study
KKC25	*Anabaena* PCC 7120 carrying pAMKK25	This study
AMC444	*Anabaena* PCC 7120 carrying pAM882 containing 834-bp internal fragment of *xisF*	(35)
Plasmids		
pAM1956	Shuttle vector pAM505 containing promoterless *gfpmut2* with unique upstream cloning sites - SalI, SacI, KpnI, and SmaI	(194)
pAMKK6	Shuttle vector pAM1956 containing 350-bp fragment upstream of *nifH* gene transcriptionally fused to *gfp*	This study
pAMKK7	Shuttle vector pAM1956 containing 700-bp fragment upstream of *nifH* gene transcriptionally fused to *gfp*	This study
pAMKK25	Shuttle vector pAM1956 containing a 500-bp fragment upstream of the *nifH* gene transcriptionally fused to *gfp*	This study
pRL277	Conjugal suicide vector; Sp^r, Sm^r	(10)

Isolation of mutants by UV mutagenesis. UV mutagenesis was performed according to published protocols (187) with minor modifications: following UV exposure, culture flasks were wrapped in yellow cellophane and were grown under fluorescent white light. Cultures were plated on BG-11_0 plates with streptomycin and spectinomycin to obtain individual colonies.

Complementation with the expression library. A library of plasmid DNA was transformed into *E. coli* strain AM1358 (106). The pooled library clones in strain AM1358 were transferred into *Anabaena* strain AMC1774 by triparental conjugation with *E. coli* strain AM1460 following standard protocols (106). Conjugation mixtures were plated on Millipore HATF filters laid on BG-11_0 plates with 5% LB and incubated at 30°C under 75-100 μmol photons m^{-2} s^{-1}. After 24 hours, the filters were moved onto BG-11_0 plates with streptomycin plus spectinomycin and neomycin.

Approximately 30,000 exconjugant colonies were screened for changes in GFP fluorescence and diazotrophic growth.

Microscopy. Fluorescence and bright-field images were captured using an Olympus IX70 inverted microscope with a Hamamatsu OrcaER C4742-95 charge-coupled-device camera and Simple PCI software version 6.1 as previously described (2). A Piston green fluorescent protein (GFP) band-pass filter set (no. 41025; Chroma Technology Corp.) was used for fluorescence images. Photomicrographs were minimally processed with Adobe Photoshop 8.0 to improve brightness and contrast.

RNA isolation. Cultures were induced for heterocyst development as previously described (2). Cells were collected at 0 and 24 h after nitrogen deprivation. Total RNA

was isolated using hot-phenol extraction: cells were collected from exponentially growing cultures and resuspended in resuspension buffer (0.3M sucrose, 10mM sodium acetate-pH 4.5, 500 mM EDTA- pH 8) and lysed in lysis buffer (2% w/v SDS, 10mM sodium acetate-pH 4.5). RNA was extracted with hot phenol, precipitated using 0.2M lithium chloride and ethanol, and resuspended in DEPC-treated water.

Northern RNA blot analysis. For each sample, 10 µg of total RNA was electrophoresed on a 1.0% agarose gel, transferred onto GeneScreenPlus filters and probed with DNA probes specific for *nifB* and *nifH*. Probes were amplified by PCR with previously published primers (116), labeled with [γ^{32}P]dATP by end-labeling, and purified with Wizard SV gel spin columns (Promega). Blots were blocked, hybridized with DNA probes, and washed using standard protocols (155). Membranes were exposed to a phosphorimager plate that was scanned with a Typhoon phosphorimager.

Results

P*nifH-gfp* reporter constructions on shuttle plasmids. To identify sequences upstream of the *nifH* gene that are required for transcription, we initially attempted to make transcriptional fusions of the *nifH* upstream region to *gfp* on shuttle plasmid pAM1956 (194). We first studied strain KKC5 containing the intergenic region (221-bp) between *nifU* and *nifH* fused to *gfp* on pAM1956. Because a putative *nifH* transcription start site had been mapped to position -123 relative to the *nifH* start codon (Fig. 6), we expected the *nifHDK* promoter to be in this region, but KKC5 did not show P_{nifH}-driven *gfp* expression in heterocysts when transferred to media without a source of combined

nitrogen (data not shown). Strain KKC6 and KKC25 containing 350-bp (data not shown) and 500-bp of the *nifH* upstream region fused to *gfp* on pAM1956 also did not show expression of P_{nifH}-*gfp* in heterocysts when deprived for nitrogen (Fig. 7). To confirm the absence of GFP in KKC5, KKC6, and KKC25, a western immunoblot experiment was performed. No GFP protein was detected (data not shown) indicating there was no *nifH* promoter activity in these constructs. These results suggest that elements necessary for transcription of the *nifHDK* genes are not present within the 500-bp region upstream of *nifH*.

We next constructed strain KKC7 containing 700-bp of the *nifH* upstream region transcriptionally fused to *gfp* on pAM1956. When transferred to media without a source of combined nitrogen, KKC7 showed strong expression of *gfp* in vegetative cells and heterocysts (data not shown). It is not clear if a true *nifH* promoter drives the expression of the plasmid-borne *gfp* reporter in KKC7 or if the transcription is an artifact of the construction. However, similar observations resulting from using a 700-bp fragment of the *nifH* upstream region had been made previously, leading to speculation that an upstream sequence could function as a promoter in vegetative cells (7).

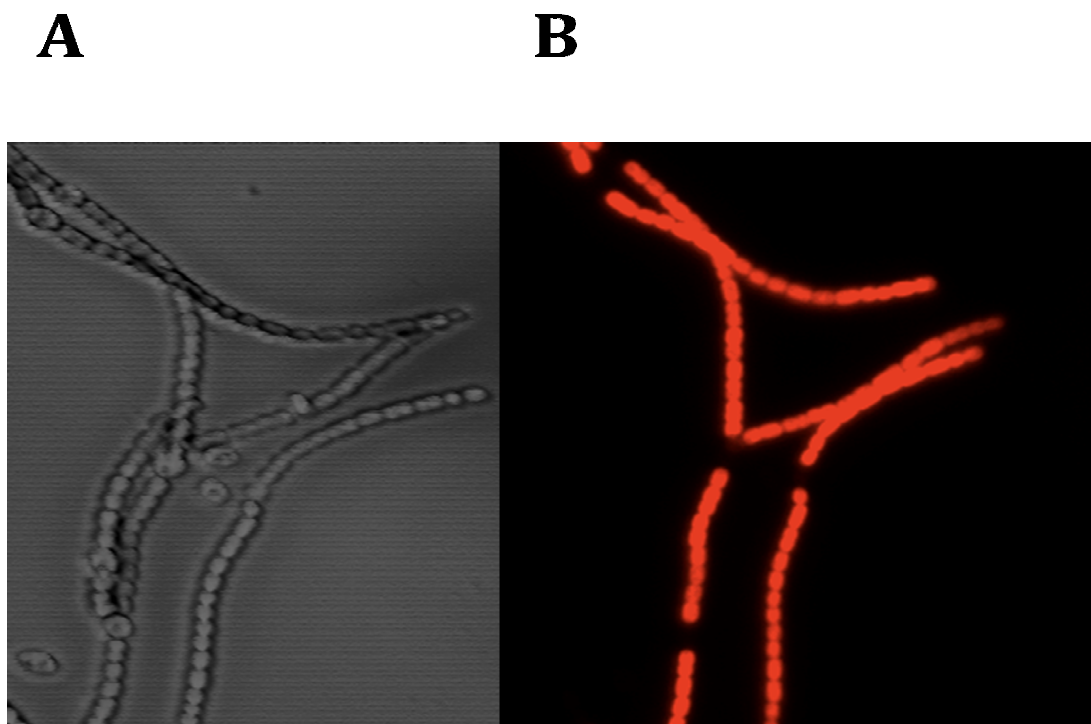

FIG. 7. GFP fluorescence from GFP reporter expression from KKC7. (A) DIC image (B) Merged autofluorescence and GFP fluorescence image of KKC7. Images were captured 22 hours after nitrogen deprivation. Scale bar, 10 μm.

Construction of P*nifHD-gfp* reporter in the chromosome. To mitigate potential problems with analyzing reporter fusions on a multicopy plasmid, we constructed a reporter fusion using an amplified DNA fragment extending from the start of the *nifD* ORF to the end of the *nifS* ORF on suicide plasmid pAMKK2 (Table 1). The suicide plasmid was integrated into the chromosome at the *nif* locus by a single crossover event, which would result in driving the expression of the *gfp* reporter gene from the entire chromosomal region upstream of *nifD*. In cultures grown without nitrate, strain AMC1774, carrying plasmid pAMKK2 in the *nif* locus, showed strong heterocyst-specific GFP fluorescence (Fig. 8). Time-lapse microscopy showed that GFP fluorescence was upregulated at approximately 14 hours after nitrogen deprivation (data not shown).

Excision of the *fdxN* element is required for *nifHDK* expression. UV mutagenesis was performed on AMC1774 to generate random mutants that showed abnormal regulation of the P_{nifHD}-*gfp* reporter. We screened 30,000 colonies to isolate mutants that did not show GFP in heterocysts. One mutant, KKC3, which differentiated heterocysts of normal frequency and morphology, did not show GFP reporter fluorescence in heterocysts and was unable grow on media lacking a source of combined nitrogen (Fig. 9). Complementation of KKC3 with an expression library restored the WT phenotype (data not shown). The complementing plasmid carried genes *xisF* and its flanking gene *all1460*. *xisF* and its neighboring genes *xisH* and *xisI* are involved in excision of the *fdxN* element in the *nifB-fdxN-nifS-nifU* locus (35, 149).

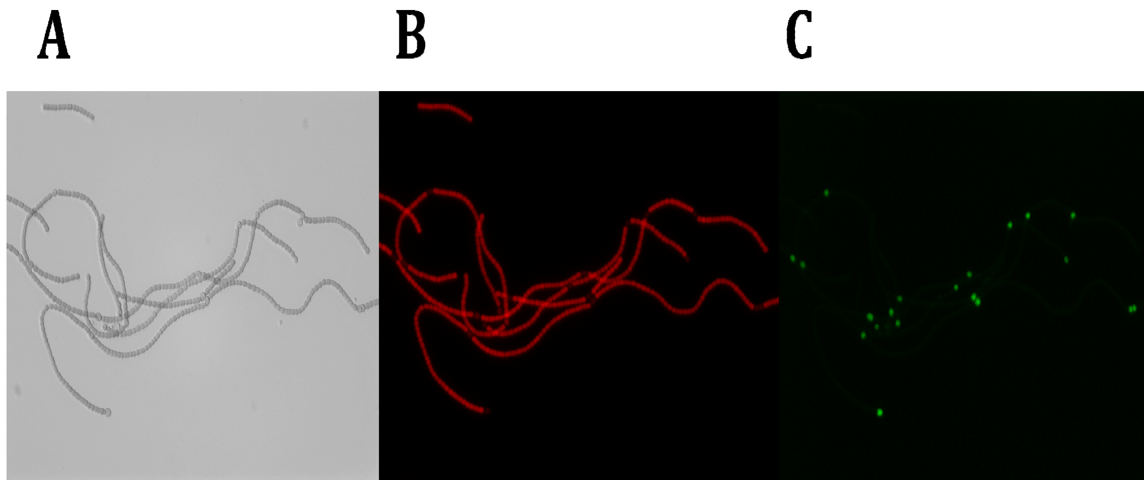

FIG. 8. GFP fluorescence of AMC1774. (A) Merged DIC image. (B) Autofluorescence image. (C) Corresponding GFP fluorescence image. Heterocysts were induced to differentiate by transfer of AMC1774 to BG-11_0 medium. Images were captured 22 hours after nitrogen deprivation. Scale bar, 10 µm.

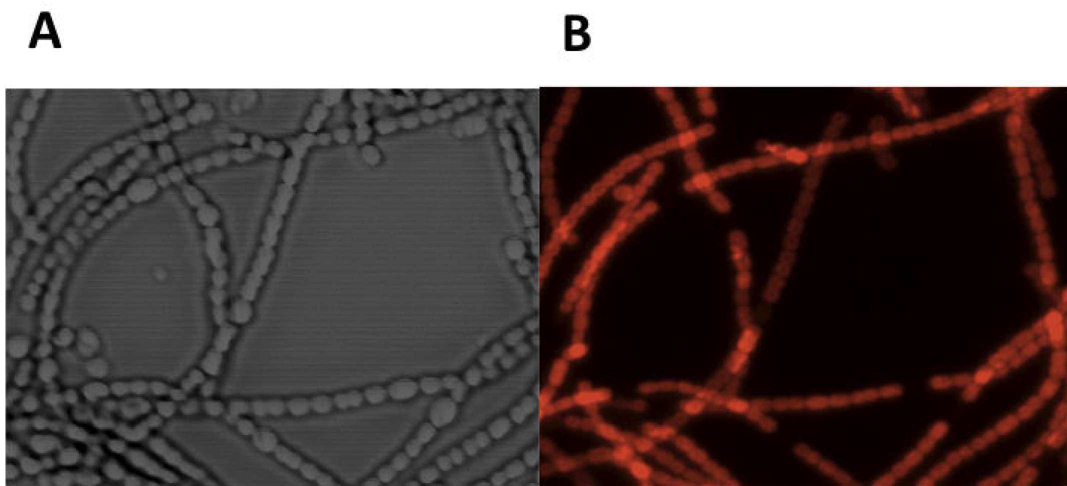

FIG. 9. GFP fluorescence of KKC3. (A) DIC images of mutants show heterocyst formation. (B) GFP fluorescence image shows that the heterocysts lack GFP expression. Images were collected at 22 hours after nitrogen deprivation. Scale bar, 10 µm.

A mutant lacking one or more of these genes does not excise the element or grow diazotrophically, indicating that the rearrangement of this element is required for nitrogen fixation (35). We hypothesized that KKC3 was unable to excise the *fdxN* element in heterocysts after nitrogen deprivation. PCR analysis showed

that mutant KKC3 did not excise the *fdxN* element but excised the 11-kb *nifD* element normally (Fig. 10). These results indicate that failure to excise the 59-kb *fdxN* element blocks transcription of *nifHDK* genes, which suggests that sequences essential for *nifHDK* transcription are upstream of the element

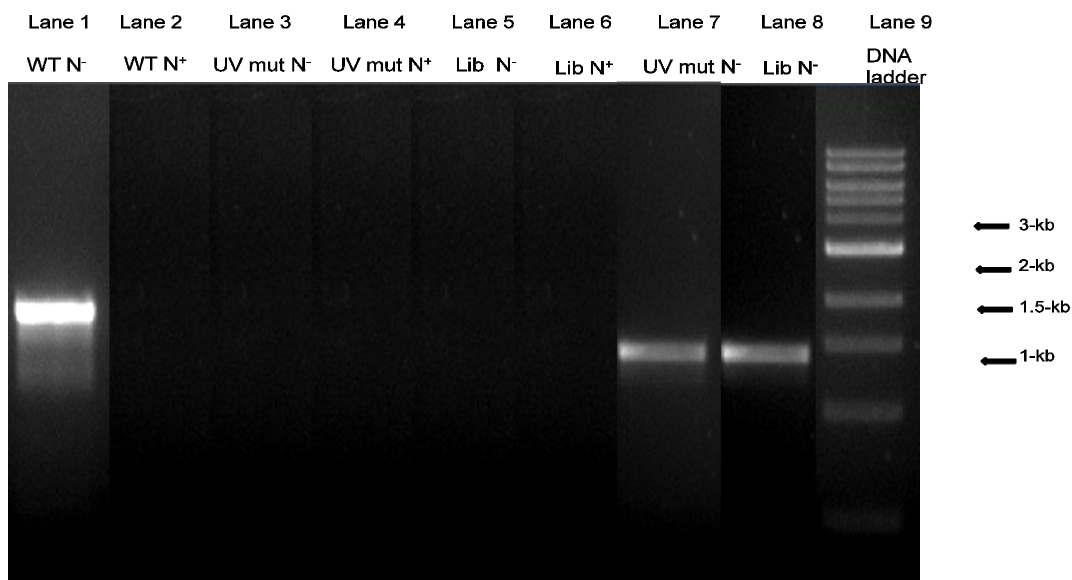

FIG. 10. PCR analysis of mutants. Lanes 1-6 indicate DNA fragments amplified by primers KKO7 and KKO17. Lanes 7 and 8 indicate DNA fragments amplified with primers KKO18 and KKO19.

We also screened for abnormal expression of the P_{nifHD}-*gfp* reporter by transferring a high-copy-number expression library of random gene fragments into AMC1774. Mutant phenotypes were expected to be a result of overexpression of a gene fragment, possibly a transcription repressor or activator, from the high copy number expression plasmid. We isolated one strain, KKC4, which was unable to grow under conditions of nitrogen deprivation and did not express P_{nifHD}-*gfp* (Fig. 11). KKC4 carried a plasmid with genes that localized to the 59-kb *fdxN* element on the chromosome. PCR analysis showed that KKC4 excised the *nifD* element but did not excise the *fdxN* element (Fig. 10).

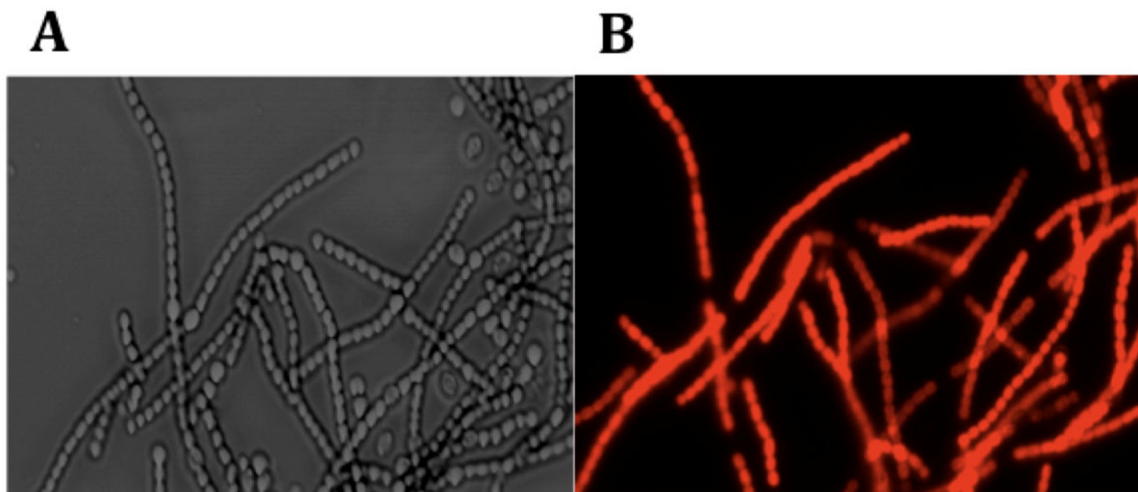

FIG. 11. GFP fluorescence of KKC4. (A) DIC images of mutant show heterocyst formation. (B) GFP fluorescence image shows that the heterocysts lack GFP expression. Images were collected at 22 hours after nitrogen deprivation. Scale bar, 10 µm.

thus corroborating our results from the mutant isolated by UV mutagenesis that excision of the *fdxN* element is required for *nifHDK* transcription. We hypothesized that transcription of *nifHDK* initiates at the distant upstream *nifB* promoter and that failure to excise the *fdxN* element interrupts transcription from P_{nifB} so that it cannot continue through to the *nifHDK* genes. Recently published RNA-seq data is consistent with our hypothesis; *nif* transcripts appear to begin at P_{nifB} and continue past *nifK* in RNA samples isolated 21 hours after nitrogen deprivation (53).

P_{nifB} is required for the normal expression of the *nifHDK* genes. To determine if transcription of the *nifHDK* genes requires the *nifB* promoter, we recombined the P_{nifHD}-*gfp* reporter construct on plasmid pAMKK3 at an ectopic site (neutral site (*gvp*) locus in the chromosome). This strain, KKC1, did not show GFP fluorescence in heterocysts after nitrogen deprivation indicating that transcription of *nifHDK* requires elements outside of this fragment (Fig. 12). In *A. variabilis*, the expression of *nifH1* requires the strong promoter upstream of *nifB1* and a weak promoter in the *nifU1* gene (173). Because our reporter construct in KKC1 at the neutral site contains the *nifU* gene and its immediate upstream region, the absence of GFP fluorescence in this strain suggests that *Anabaena* PCC 7120 *nifU* does not possess a second promoter that is strong enough to drive transcription of the P_{nifHD}-*gfp* reporter.

A promoter is present in the intergenic region upstream of *nifB*. To determine if a promoter is present immediately upstream of the previously identified *nifB* transcription

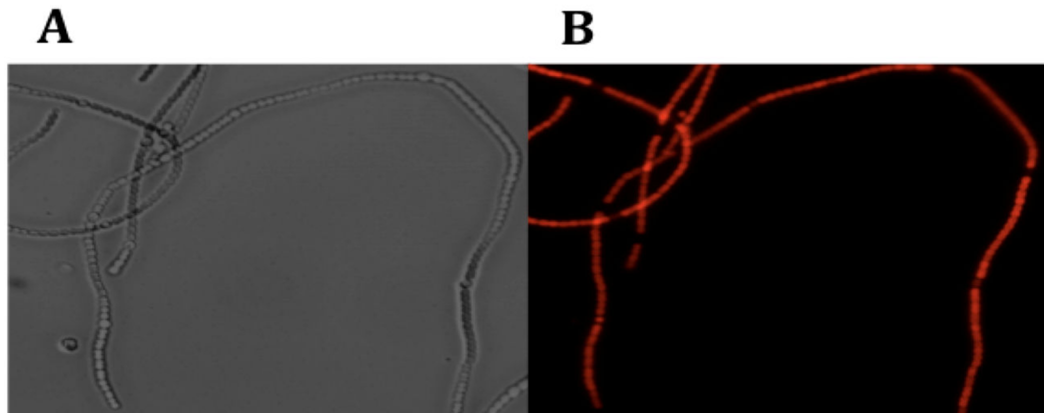

FIG. 12. KKC1 at neutral site did not show GFP fluorescence after nitrogen deprivation. (A) Merged DIC image. (B) Corresponding GFP fluorescence image. Images were collected at 22 hours after nitrogen deprivation. Scale bar, 10 µm.

start site, we constructed a P_{nifB}-*gfp* transcriptional fusion on pAM1956 using the intergenic region between *nifB* and its upstream gene asl1518. The resulting strain KKC2 showed heterocyst-specific GFP fluorescence when grown without a source of combined nitrogen (Fig. 13). A control strain containing the pAM1956 vector alone produced no detectable GFP fluorescence (data not shown). These results indicate that the *nifB* promoter lies within the intergenic region between *nifB* and asl1518. Recently published RNA-seq data supports our findings; a 5' end was identified at -283-bp relative to the *nifB* start codon (53), as was previously described (122).

Previous data have shown significantly lower levels of transcript from the *nifB-fdxN-nifSU* genes compared to the *nifHDK* genes (122). A 1.75-kb band thought to be the *nifB* transcript was detected, however, the levels were much lower than those of the *nifH* transcripts (122). The larger *nifB-fdxN-nifSU* transcript was not detectable (64, 122). We hypothesize that the difference in transcript levels must be due to higher stability of the *nifHDK* transcripts than the *nifB-fdxN-nifS-nifU* transcripts. The presence of stem-loop structures close to the 5' or 3' ends of the mRNA has been shown to increase stability of transcripts (92, 101, 132). We examined the 5' upstream region of *nifH* for secondary structures using MFold (209). This region showed a potential secondary structure with a stem-loop located at -123 bp relative to the *nifH* start codon that is likely to contribute to the stability of the *nifHDK* transcripts (Fig. 14).

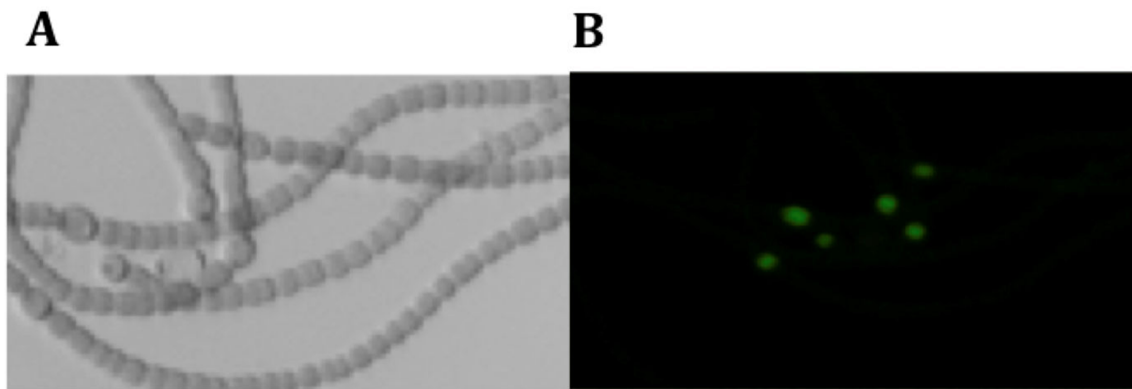

FIG. 13. P_{nifB}-*gfp* reporter shows GFP fluorescence after nitrogen deprivation. (A) DIC images of reporter shows heterocyst formation. (B) GFP fluorescence image shows that the heterocysts have GFP fluorescence. Images were collected at 22 hours after nitrogen deprivation. Scale bar, 10 µm.

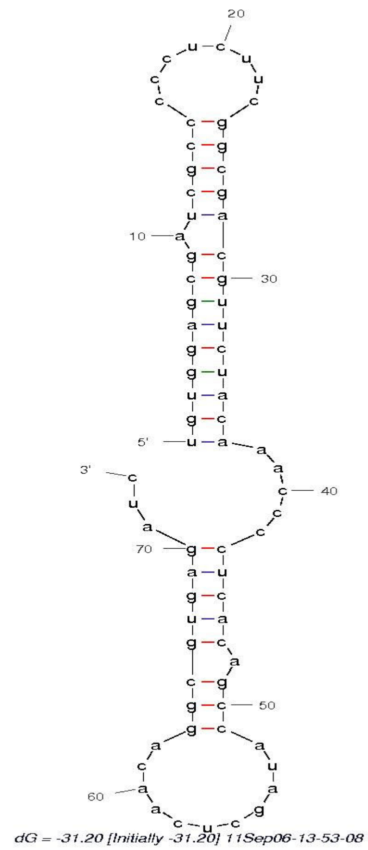

FIG. 14. Predicted stem-loop structure at the 5' end of the stable *nifH* transcript.

Discussion

Although the *nif* genes in *Anabaena* PCC 7120 were identified over 30 years ago, our understanding of the molecular and biochemical events that govern the regulation of these genes is limited. Using *gfp* as a reporter to study transcription from regions upstream of the *nifHDK* genes, we show that excision of the *fdxN* element within the *fdxN* gene is required for diazotrophic growth and expression of the *nifHDK* genes. This result shows that there is no functional promoter in the region immediately upstream of *nifH* and therefore suggests that transcription must originate at the far-upstream *nifB* promoter.

Although previous studies had identified a putative *nifH* transcription start site at -123 bp relative to the *nifH* start codon, we show that transcriptional fusions of the *nifU-nifH* intergenic region and sequences up to 500 bp upstream of *nifH* on shuttle plasmids did not produce expression of P_{nifH}-*gfp* reporter fusions. Attempts to identify a promoter within the *nifU* gene and in the upstream intergenic region that could drive normal transcription of *gfp* were not successful because of unregulated expression in vegetative cells. However, we succeeded in obtaining a regulated reporter construction by homologous recombination of a P_{nifHD}-*gfp* reporter at the *nif* locus in the chromosome to produce strain AMC1774, which showed strong heterocyst-specific expression of GFP fluorescence. A UV-induced mutant of strain AMC1774 was obtained that was unable to grow diazotrophically and did not show GFP fluorescence in heterocysts after nitrogen deprivation. This mutant, KKC3, was complemented by a high-copy-number plasmid carrying *xisF* and its flanking gene *all1*460, suggesting that the original mutant was

defective in excision of the *fdxN* element during heterocyst differentiation. Further analysis of the mutant by PCR confirmed that KKC3 did not excise the *fdxN* element. Screening for dark mutants of AMC1774 after transferring a high-copy-number expression library into the strain generated mutant KKC4 that carried a plasmid with genes that localized to the *fdxN* element. PCR analysis showed that KKC4 was not able to excise the element indicating that excision of the 59-kb element is required for *nifHDK* transcription. This data, together with results obtained from the UV mutagenesis, suggests that a promoter upstream of the *fdxN* element is required for transcription of the *nifHDK* genes.

In *A. variabilis*, two promoters contribute to the expression of the *nifHDK1* genes: one promoter upstream of *nifB1* contributes 75% of the transcripts and a second promoter within *nifU1* contributes 25% of the transcripts (172). To determine if a second promoter was present in the PCC 7120 *nifU* ORF, we monitored expression of the P_{nifHD}-*gfp* reporter construct integrated into an ectopic locus in the chromosome. This strain did not show GFP fluorescence in heterocysts after nitrogen deprivation (Fig. 12). Northern blot analysis of AMC444, a mutant that does not excise the *fdxN* element, did not show the presence of *nifH* transcripts (data not shown). Together, our data show that the *nifU* ORF does not contain a second promoter that contributes to *nifHDK* expression.

In *A. variabilis*, the putative transcription start site originally mapped to the intergenic region upstream of *nifH1* is a processed site; *nifHDK1* transcription is initiated at two transcription start sites, one site upstream of nifB and a second site within *nifU1* (173). In *Anabaena* PCC 7120, RNA-seq data shows continuous reads that extend from the

nifB promoter to the *nifK* gene (53). The transcription start site for the *nifB-fdxN-nifSUHDK* genes lies upstream of *nifB* at -283 bp from the *nifB* start codon, which is the same as the originally mapped site (122). We show that reporter fusions containing the intergenic region upstream of *nifB* transcriptionally fused to *gfp* on pAM1956 showed GFP fluorescence specifically in heterocysts.

Previous studies had identified a putative transcription start site for *nifH* at position -123 bp relative to the start codon (72, 85). The 5' end at this position is likely to be a processed end rather than a *de novo* transcription start, and the *nifHDK* transcript may be a cleavage product of a larger *nifB-fdxN-nifSUHDK* transcript. It has been shown that the presence of stem-loop structures close to the 5' or 3' ends contribute to transcript stability (92, 101, 132). We examined the upstream region of *nifH* for the presence of stem-loop structures that may contribute to the stability of the *nifHDK* transcripts. A stem loop structure, whose base is positioned at -123 bp of the *nifH* transcript, was identified by Mfold (209). This stem-loop closely resembles a stem-loop sequence near the 5' end of the *nifH1* transcript in *A. variabilis* (172), suggesting that processing of the *nifHDK* transcript at the stem-loop could contribute to higher stability of these transcripts when compared to the *nifB-fdxN-nifSU* transcripts.

Our data suggests that the *nifB* promoter carries important cis-acting elements required for the transcription of the *nifB-fdxN-nifSU* and *nifHDK* genes. It was recently shown that AnCrpA, a transcription factor, regulates the expression of the *nif* genes in *Anabaena* PCC 7120 in the presence of nitrate (164). AnCrpA bound to the 5' upstream region of *nifB* but not to the 5' upstream region of *nifH*, as observed by EMSA (164).

The *fdxN* element, like the other DNA elements in *Anabaena* PCC 7120 (20, 34), appears to be a parasitic DNA sequence that is carried innocuously in heterocyst-specific genes. The element is excised from within the *fdxN* gene late during heterocyst development when the *nif* genes are transcribed (33). Our data strongly indicate that the excision is necessary for the expression of the *nifHDK* structural genes because transcription of the *nifHDK* genes requires the *nifB* promoter. Our findings suggest that other developmentally regulated genes in *Anabaena* PCC 7120 may use similar mechanisms to regulate gene expression.

CHAPTER III

ROLE OF DEVH AS THE TRANSCRIPTIONAL REGULATOR OF THE NIFHDK GENES IN THE FILAMENTOUS CYANOBACTERIUM ANABAENA SP. STRAIN PCC 7120

Overview

We used electrophoretic mobility shift assays to examine the binding of DevH to the upstream region of the *nifB* gene of the heterocystous cyanobacterium *Anabaena* sp. strain PCC 7120. DevH showed specific binding to the *nifB* upstream region. An NtcA-like binding site present 37 bp upstream of the *nifB* transcriptional start site is likely to function as the DevH binding site.

Introduction

Bacterial gene expression is regulated in response to extracellular and intracellular signals. Regulation of gene expression by transcriptional initiation is controlled by sigma factors in association with transcriptional factors. Among the transcriptional factors, the cyclic AMP receptor protein (CRP; also known as CAP, catabolite gene activator protein) family is a group of transcriptional regulators that are widespread among bacteria (210). These proteins function as activators and repressors (151) in response to a range of physiological conditions, including carbon and nitrogen metabolism, redox state of the cell, and other stress responses (16, 97). Proteins of this family are classified into three groups, CRP, FNR (fumarate and nitrate reductase regulator), and Ntr (nitrogen

regulator), based on sequence and function, and are characterized by a helix-turn-helix DNA-binding domain (DBD) at the C-terminus (210). CRPs are activated by small effector molecules; the most well know example is that of CRP in *E. coli*, which is activated in response to levels of cAMP (97).

The filamentous cyanobacterium *Anabaena* sp. strain PCC 7120 has 6 genes predicted to encode proteins that belong to the CRP family. *Anabaena* PCC 7120 responds to deprivation of combined nitrogen by undergoing a developmental program in which heterocysts are differentiated at semi-regular intervals of every 10 to 15 photosynthetic vegetative cells. Heterocyst development culminates with nitrogen fixation, which is the ATP-dependent process of reducing atmospheric nitrogen to ammonia by the enzyme nitrogenase. A representative member of CRPs in cyanobacteria, NtcA, functions as a regulator of genes involved in nitrogen metabolism in response to the accumulation of an effector, 2-OG, after nitrogen deprivation (36, 103, 166, 178). AnCrpA, another protein in this family, responds to cAMP levels and regulates the expressions of gene clusters related to nitrogen fixation in the presence of nitrate (164).

In *Anabaena* PCC 7120, nitrogenase is synthesized only in heterocysts, where it is protected from irreversible inactivation by oxygen (122). The nitrogen fixation (*nif*) genes that encode the structural components of the nitrogenase enzyme are upregulated during the late stages of heterocyst development at 12-18 hours after nitrogen deprivation (45, 64). Although heterocyst development is a well-studied model of prokaryotic multicellularity and nitrogen fixation (1, 54, 100), transcriptional regulatory factors that govern the later stages of development, particularly the culminating event of

nitrogen fixation, are poorly understood. The *nif* genes in proteobacteria that fix nitrogen are regulated by Sigma-54 (21). However, Sigma-54 and cognate activators are absent from cyanobacteria. It has recently been shown that expression of the *sigE*, which encodes a cyanobacterial group 2 sigma factor, is required for normal levels of *nifH* expression in *Anabaena* PCC 7120 (116). Expression of *nifH* is delayed and reduced in a *sigE* mutant strain.

DevH, a transcriptional regulator belonging to the CRP family of proteins, shows sequence similarity at the amino acid level to NtcA. The *devH* gene is upregulated about five-fold at 24 hours after nitrogen deprivation (73). DevH has been shown to bind to its own promoter and is thought to have autoregulatory activity (73). A *devH* mutant strain forms heterocysts at semi-regular intervals but does not fix nitrogen in the presence of oxygen (73). Heterocyst- specific programmed DNA rearrangements within the *nifD* and *fdxN* genes occur normally in these mutants, however, *nifHDK* transcripts are undetectable (150). Because the *devH* mutant is able to fix nitrogen at low levels in the absence of oxygen, it was concluded that DevH does not function as a transcriptional regulator of the *nifHDK* genes (150). Rather, it was concluded that the *devH* mutant is unable to fix nitrogen because it lacks the laminated layer of envelope glycolipids, which results in the entry of oxygen into the heterocyst, and disappearance of the *nifHDK* transcripts because of lack of transcription, degradation, or both.

Because gene expression is not always completely inhibited in the absence of an activator (48, 69), we speculated that DevH could function as a regulator of the *nifHDK* genes. Previous data suggests that the *nifB* promoter drives the expression of the *nifHDK* genes. In this study we present evidence that DevH binds specifically to the region upstream of *nifB*. We propose that an NtcA-like sequence at 37-bp upstream of the *nifB* transcriptional start site is the DevH binding site.

Materials and methods

Strains and conditions. *E. coli* strain pAMKK26 for overexpression of recombinant DevH (rDevH) was grown in LB (Lennox L) liquid or agar medium supplemented with kanamycin (50 µg/ml) at 37°C.

Plasmid constructions. Primers used in this study are listed in Table 3. A plasmid for expression of rDevH in *E. coli* was constructed by cloning a PCR-amplified fragment of the *devH* open reading frame, using primers KKO47 and KKO48, into pET28a (Novagen) using NdeI and XhoI sites, to generate pAMKK26.

TABLE 3. Primers used in this study (Chapter III)

Primer	Sequence	Source of reference
KKO47	AGCGGCCATATG CAATCTCCATCCTCC	This study
KKO48	GTGCTCGAGTTAGCTAGCTCTGTTGATTGA	This study
KKO49	AGAAGCGTTAAGTGATTGGCG	This study
KKO50	CAACGTTCTCGAAATCAT	This study
KKO51	CTTGCTTCGGTGTGAG	This study
KKO52	CGGCTTTGTTTTCCT	This study
KKO53	CTACCCTATGGCAGATGA	This study
KKO54	CTGATGATAGTTAAAGCTATCCAC	This study
KKO55	GCTTTCTCTGGAGTG	This study
KKO56	TTTTGATAATGATGGAATAATTG	This study

Expression of DevH and preparation of cell-free extracts. 6-His tagged recombinant DevH (rDevH) protein was overexpressed from the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible plasmid pAMKK26 in *E. coli* BL21(DE3)pLysS. An overnight culture of *E. coli* BL21(DE3)pLysS containing pAMKK26 was inoculated into 10 ml of LB liquid supplemented with kanamycin (50 µg/ml) and incubated at 37°C with shaking (200 rpm). When the optical density at 600 nm (OD_{600}) of the culture reached 0.6 to 1.0, IPTG was added to a final concentration of 1 mM. Cells were incubated for 4 hours and harvested by centrifugation at 5,000 x *g* for 10 min. The pellet was resuspended in 1 ml lysis buffer (20 mM $Na_2H_2PO_4$, 0.5 M NaCl; pH 7.4), and the cells were lysed with a French press (American Instrument Co., division of Travenol Laboratories, Inc., Silver Spring, Md.). Soluble fractions were separated by centrifugation at 14,000 x *g* for 15 min. Total protein concentration in the supernatant

was determined by the Bradford assay (17) with bovine serum albumin as the standard. Soluble fractions were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) for presence of rDevH protein.

Electrophoretic mobility shift assays (EMSA). A 135-bp DNA fragment upstream from the *nifB* transcriptional start site at -283 bp relative to the *nifB* start codon, was amplified by PCR using primers KKO49 and KKO50. A 158-bp fragment within the *nifU* gene was amplified by PCR using primers KKO51 and KKO52. A 150-bp fragment upstream of the *devH* transcriptional start site was PCR-amplified using primers KKO53 and KKO54 and was used as a positive control. A 200-bp fragment of the *hfq* upstream region, PCR-amplified using primers KK055 and KKO56, was used as a negative control. The PCR products were labeled with [γ-^{32}P]dATP using polynucleotide kinase. Protein binding assays (15 µl reactions) contained 25 fm of labeled probe in EMSA binding buffer (2 µg poly (dI-dC), 250 ng/µl BSA, 12% (w/v) glycerol, 4 mM Tris-HCl, 12 mM HEPES, 60 mM KCl, 1 mM EDTA, 1.5 mM DTT, 5 mM $MgCl_2$; pH 8.0) and increasing amounts of rDevH cell-free extracts. Reactions were incubated for 15 min at room temperature and were then analyzed on a nondenaturing 5% polyacrylamide gel in Tris-glycine running buffer (380 mM glycine, 50 mM Tris-HCl [pH 8.5], 2 mM EDTA) at 4ºC. The gel was vacuum dried, exposed to a phosphorimager plate and scanned with a Typhoon 9410 phosphorimager.

Results

DevH interacts with the upstream region of the *nifB* promoter in vitro. It has been proposed that the *nifHDK* genes are likely to be a part of a large *nifB-fdxN-nifSUHDK* operon (Chapter II). Previous data has shown that some heterocyst-specific promoters possess an NtcA binding site with a consensus sequence of TGTAN$_8$TACA (78). Because DevH shows high sequence similarity to NtcA, we examined the upstream region of *nifB* for an NtcA-like binding site. We identified an NtcA-like binding site 37 bp upstream of the *nifB* transcriptional start site. To test if DevH binds specifically to a DNA fragment containing this site, we used a radioactively labeled 135-bp PCR fragment encompassing the region between -260 bp and -395 bp upstream of the *nifB* start codon, in the presence and absence of rDevH cell-free extracts. Figure 15 shows a shift in the mobility of the labeled DNA fragment in the presence of increasing amounts of rDevH cell-free extract, while cell-free extracts of the pET28a(+) vector alone had no effect. No shift was observed when rDevH cell-free extracts were tested with a fragment of P*hfq* (data not shown). The binding was significantly reduced in the presence of unlabeled competitor DNA, indicating that rDevH binds specifically to the upstream region of *nifB*.

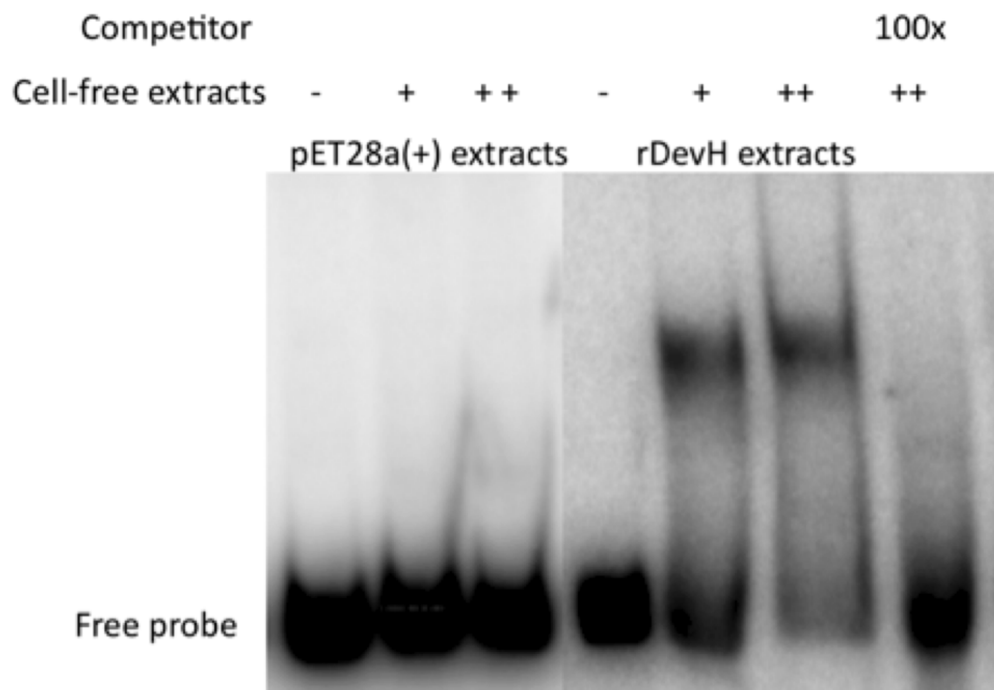

FIG. 15. DevH binds to the *nifB* upstream region. Binding of rDevH to a 135-bp fragment from the region upstream of *nifB* was examined by electrophoretic mobility shift assay (EMSA). Lanes 1, 2, and 3 are reaction mixtures contained 25 fmol of labeled DNA probe with increasing amounts of pET28a(+) cell-free extracts. Lanes 4, 5, and 6, are reaction mixtures contained 25 fmol of labeled DNA probe with rDevH cell-free extracts. The last lane contains a 100-fold molar excess of unlabeled probe as a specific competitor.

In *Anabaena variabilis*, two promoters control transcription of the *nifHDK* genes. One promoter upstream of *nifB1* contributes to 75% of the transcripts, and a second promoter in *nifU1* contributes to 25% of the transcripts (172). We investigated whether DevH binds to the *nifU* gene in *Anabaena* PCC 7120. We used a 164-bp PCR fragment encompassing the region between -300 bp and -464 bp from the 3′ end of *nifU*, which corresponds to the promoter region in *A. variabilis*; a 5′ end was found at -320-bp from the *nifU* start codon. No band shift was observed with rDevH cell-free extracts indicating that DevH does not bind to the potential promoter region within the *nifU* gene (Fig. 16).

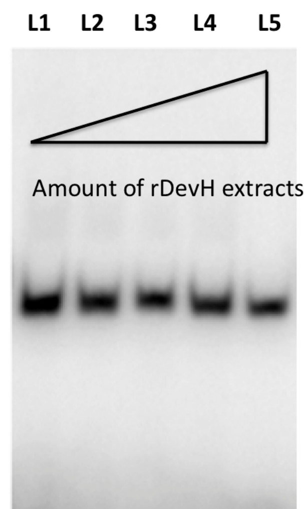

FIG. 16. DevH does not bind to the *nifU* region. Binding of rDevH to a 164-bp fragment from a region within the *nifU* ORF was examined by EMSA. Reaction mixtures contained 25 fmol of labeled DNA probe with rDevH cell-free extracts.

Discussion

The chromosome of *Anabaena* sp. strain PCC 7120 harbors several genes that encode proteins predicted to be members of the CRP family. In the current study, we investigated the role of one of these CRP family proteins, DevH, as a regulator of the *nifHDK* genes. Previous studies showed that the *Anabaena* PCC 7120 *devH* gene is required for nitrogen fixation under aerobic conditions (73). Because low levels of activity were observed in ethylene reductions assays under anaerobic conditions, it was concluded that DevH does not function as a regulator of the *nifHDK* genes. In work done later, DevH was characterized by Curtis and colleagues as being required for the formation of the heterocyst glycolipid layer and hence, for diazotrophic growth (150). However, mutations of genes encoding regulatory proteins do not entirely abolish expression of the target genes (48, 69). Thus, we examined the role of DevH as a potential regulator of the *nifHDK* genes.

We hypothesized that DevH could function as a regulator of the *nifHDK* genes by binding to the region upstream of *nifB*. EMSA results show that rDevH specifically bound to the upstream region of *nifB*. Our data suggest that absence of expression of the *nifHDK* genes in a *devH* mutant (150) results from absence of *nifB* promoter activity.

Bioinformatic analysis of the region upstream of the *nifB* transcriptional start site identified an NtcA-like binding site. Because NtcA and DevH belong to the same family of proteins, we speculate that this sequence could function as the DevH binding site. The putative DevH binding site is positioned at 37 bp upstream of the *nifB* transcriptional start site, at -320 bp relative to the *nifB* start codon. The putative DevH binding sequence

is an imperfect palindrome and differs by 1 bp from the NtcA consensus binding sequence (Fig. 17) (88). EMSA showed that DevH binds to a DNA fragment containing the NtcA-like binding sequence. DevH has previously been shown to bind to the *devH* promoter region (150), which contains two putative binding sites. One of these binding sites is the same imperfect palindrome as present upstream of the *nifB* transcriptional start site, adding support to our prediction that DevH binds to this sequence. NtcA binds the secondary metabolite 2-oxoglutarate (2-OG) (178), an intermediate of the Krebs cycle, as a secondary effector. It is possible that DevH also binds a secondary effector to enhance binding to its target promoters.

It is likely that DevH binds to the NtcA-like binding sequence upstream of the *nifB* transcriptional start site to upregulate the expression of the *nifB-fdxN-nifSUHDK* genes in heterocysts. The previous conclusion that DevH does not function as a transcriptional regulator of the *nifHDK* genes (73) can be explained by the idea that there may be some level of functional redundancy between DevH and another protein of the CRP family. For example, NtcA may weakly bind to the imperfect palindrome upstream of *nifB* resulting in low levels of *nifHDK* transcripts sufficient for the low level of nitrogenase activity under anaerobic conditions produced by the *devH* mutant.

The results presented in this work show that DevH binds specifically to the *nifB* upstream region. Because *devH* mutants are defective for *nifHDK* transcription and diazotrophic growth, our data strongly suggests that DevH functions as the transcriptional activator of the *nifHDK* genes.

TGTA(NNNNNNNNN)TACA

TCTATTCCAGAGATACA

FIG. 17. Comparison of NtcA consensus binding sequence and DevH putative binding site present upstream of *nifB*. The arrow indicates the base pair that differs between the two sites.

CHAPTER IV

THE SIGE GENE IS REQUIRED FOR NORMAL EXPRESSION OF THE HETEROCYST-SPECIFIC GENES IN ANABAENA PCC 7120 *

Overview

The filamentous cyanobacterium *Anabaena* (*Nostoc*) sp. strain PCC 7120 produces specialized cells for nitrogen fixation called heterocysts. Previous work showed that the group-two sigma factor *sigE* (alr4249, previously *sigF*) is upregulated in differentiating heterocysts 16 h after nitrogen step-down. We now show that the *sigE* gene is required for normal heterocyst development and normal expression levels of several heterocyst-specific genes. Mobility shift assays showed that the transcription factor NtcA binds to sites in the upstream region of *sigE,* and that this binding is enhanced by 2-oxoglutarate (2-OG). Deletions of the region containing the NtcA binding sites in P_{sigE}-*gfp* reporter plasmids showed that the sites contribute to normal developmental regulation but are not essential for upregulation in heterocysts. Northern RNA blot analysis of *nifH* mRNA revealed delayed and reduced transcript levels during heterocyst differentiation in a *sigE* mutant background. qRT-PCR analyses of the *sigE* mutant showed lower levels of transcripts for *nifH*, *fdxH*, and *hglE2*, but normal levels for *hupL*.

---

* Reprinted with permission from " The *sigE* gene is required for the normal expression of heterocyst-specific genes in *Anabaena* sp. strain PCC 7120" by Mella-Herrera *et al.*, Journal of Bacteriology, 193:1823-1832, Copyright [2011] by American Society of Microbiology Contributions to this work – Construction of strains AMC1774, AMC1775 and AMC1777, fluorescence microscopy of strains AMC1774, AMC1775 and AMC1777 and writing and editing of the publication

We developed a P_{nifHD}-*gfp* reporter construct that showed strong heterocyst-specific expression. Time-lapse microscopy of the P_{nifHD}-*gfp* reporter in a *sigE* mutant background had delayed development and undetectable GFP fluorescence. Overexpression of *sigE* caused accelerated heterocyst development, an increased heterocyst frequency, and premature expression of GFP fluorescence from the P_{nifHD}-*gfp* reporter.

Introduction

Bacteria often respond to environmental and intracellular conditions by regulation of gene expression, frequently via changes in transcriptional initiation of specific genes. Transcriptional regulation in bacteria is associated with two major families of sigma factors: sigma 70 (σ^{70}) which is subdivided into 4 groups (68) and sigma 54 (σ^{54}) which has no clear homologs in cyanobacteria or in gram-positive bacteria with high-GC content (162). Sigma factors are often key regulators in complex responses such as bacterial development and biofilm formation. For example, in *Bacillus subtilis*, temporally and spatially regulated σ factors control the expression of genes during starvation-induced sporulation (27, 99), and in the gram-negative bacterium *Pseudomonas aeruginosa*, sigma factors control the expression of antibiotic- and phagocyte-resistance genes during biofilm formation (145).

.

In cyanobacteria, sigma factors have been shown to be involved in a variety of regulatory responses (84, 139). In this paper we present data showing that the *Anabaena* (*Nostoc*) sp. strain PCC 7120 *sigE* gene (alr4249, previously *sigF*), which encodes a group 2 σ^{70} family sigma factor, is required for the normal expression of some genes involved in heterocyst development. The expression of *sigE* in the unicellular cyanobacterium *Synechocystis* PCC 6803 increases after nitrogen depletion and this response is dependent on the nitrogen regulator NtcA (139); and mobility shift assays using purified NtcA show an interaction with the *sigE* promoter region *in vitro*. Expression studies comparing a *sigE* null mutant with wild-type *Synechocystis* PCC 6803 showed that the expression of one of the nitrogen assimilation genes, *glnN*, was impaired in the mutant strain (125). DNA microarray experiments with *Synechocystis* PCC 6803 showed that the expression of genes involved in glycolysis, the oxidative pentose phosphate pathway, and glycogen catabolism were decreased in a *sigE* mutant strain, suggesting that *sigE* is involved in both nitrogen and sugar metabolism (141). Recent studies in *Synechocystis* PCC 6803 showed that *sigE* is posttranslationally activated by light-to-dark transition and that ChlH, the H subunit of Mg-chelatase, interacts with SigE *in vivo* and functions as an anti-sigma factor, transducing light signals to SigE in a process mediated by Mg^{2+} (140).

The multicellular cyanobacterium *Anabaena* PCC 7120 is a well-studied heterocyst-forming strain with tractable genetics and a fully sequenced genome. Heterocyst development is an established model for the study of bacterial differentiation, pattern formation, and nitrogen fixation (1, 100). *Anabaena* PCC 7120 grows as long chains of

photosynthetic vegetative cells in media containing combined nitrogen. In media lacking combined nitrogen, heterocysts differentiate from vegetative cells to form a semiregular pattern of single heterocysts every 10 to 20 vegetative cells along filaments. Heterocysts are capable of fixing nitrogen, which they supply to vegetative cells in the form of amino acids, and the vegetative cells supply fixed carbon to the heterocysts (1, 100). Heterocyst development can be synchronously induced by nitrogen step-down, the transfer of filaments to a medium lacking combined nitrogen.

Heterocysts are morphologically and physiologically differentiated from vegetative cells, and exhibit large-scale changes in gene expression. One of the earliest steps in heterocyst differentiation is the increased expression of *ntcA*, which is an auto-regulatory transcription factor belonging to the cAMP receptor protein (CRP) family (79). NtcA together with HetR, a master regulator for heterocyst differentiation, directly or indirectly regulate the expression of many genes involved in heterocyst differentiation and the assimilation of atmospheric nitrogen (100, 103, 125, 129, 179). NtcA activity is modulated by 2-oxoglutarate (2-OG), which is an intermediate in the incomplete Krebs pathway present in cyanobacteria, and its increase constitutes the signal for nitrogen limitation.

During heterocyst differentiation, a polysaccharide layer is deposited around the proheterocyst, followed by the deposition of a glycolipid layer between the outer membrane and the polysaccharide layer (4). Synthesis and deposition of the polysaccharide layer requires the expression of a cluster of genes that constitute the HEP island (82). The expression of these genes is upregulated during the early stages of

heterocyst differentiation. Synthesis of the glycolipid layer requires the expression of *hglB*, *hglC*, *hglD*, *hglE* (*hglE*$_A$), *hglK,* and *devBCA*, which are all upregulated early during heterocyst differentiation (4, 49, 86). A second cluster of *hgl* genes includes the *hglE2* gene (all1646), which shows *devH*-dependent upregulation after heterocyst induction (150). The morphologically differentiated heterocysts become micro-oxic partly because these layers limit oxygen diffusion into the heterocyst (50). In addition, heterocysts increase respiration to decrease oxygen levels (174).

Complete maturation of heterocysts involves the expression of genes in the late stages of heterocyst development that are necessary for nitrogen fixation and assimilation into amino acids. The *nifHDK* operon encodes the nitrogenase polypeptides and is upregulated 12 to 18 hours after nitrogen step-down in heterocysts (47, 64). Other late heterocyst-specific genes include *fdxH*, which encodes a [2Fe-2S] ferredoxin (112), and *hupL*, which encodes the large subunit of uptake hydrogenase (32, 34). These genes are regulated during heterocyst development and potentially by oxygen or metabolite levels in heterocysts, but the factors required for the regulation of these late genes are not yet known.

The role of sigma factors in transcriptional regulation during heterocyst development has been studied for many years. Northern RNA blot analysis showed that two sigma factor genes, *sigB* and *sigC*, are upregulated 12 hours after nitrogen step-down (18). However, single and double mutants of *sigB* and *sigC* showed that they are not essential for heterocyst differentiation or diazotrophic growth. In later work, insertional inactivation of *sigC*, *sigF*, *sigB2*, *sigD*, or *sigE*, found that the mutants form heterocysts

and can grow on nitrate or diazotrophically, but that *sigB2* and *sigD* mutants are significantly slow to establish diazotrophic growth (95). A double mutant of *sigB2 sigD* can form proheterocysts but was unable to grow diazotrophically, possibly due to extensive fragmentation of filaments upon nitrogen deprivation (95). The failure to find any single sigma factor that is essential for heterocyst development suggests that the cyanobacterial group 2 sigma factors have partially overlapping functions, which has hampered genetic analysis (2, 95, 196).

A more recent study in *Anabaena* PCC 7120 used *gfp* transcriptional reporter fusions to analyze the spatial and temporal pattern of expression for all sigma factor genes located on the chromosome except *sigA* (2). These studies found that three sigma factor genes, two group 2 sigma factors, *sigC* and *sigE*, and one representative of group 4, *sigG*, are upregulated in heterocysts after induction of heterocyst development (2). Expression of *sigC* is upregulated by 4 h after nitrogen step-down suggesting that it may be involved in the early stages of heterocyst development. The *sigG* gene is expressed in vegetative cells grown in nitrogen containing media but after nitrogen step-down its expression in vegetative cells decreases and by 10 h individual differentiating cells show increased expression. These results suggest that SigG may be involved in the expression of middle-stage heterocyst genes in addition to genes in vegetative cells grown on nitrate. The expression of *sigE* is localized to differentiating heterocysts around 16 h after nitrogen step-down, suggesting that SigE might be important for the regulation of genes that are expressed in the late stages of heterocyst development.

In this work, we present evidence that *sigE* is required for normal heterocyst development and the normal expression of the heterocyst-specific *nifH*, *fdxH*, and *hglE2* genes.

Materials and methods

Strains and culture conditions. Strains and plasmids used in this study are listed in Table 4. *Anabaena* PCC 7120 and its derivatives were grown in BG-11 or BG-11_0, which lacks sodium nitrate, medium at 30°C with approximately 75 µmol photons m^{-2} s^{-1} illumination from fluorescent lights as previously described (64). *Escherichia coli* strains were grown in LB (Lennox L) medium containing appropriate antibiotics at 37˚C. *E. coli* strain DH10B was used for all cloning experiments. Shuttle plasmids and suicide plasmids were transferred into *E. coli* donor strain AM1359 (194) by electroporation and transferred to *Anabaena* PCC 7120 strains by conjugation using standard protocols (46) with some modifications (95).

TABLE 4. Strains and plasmids used in this study (Chapter IV)

Strains	Relevant Characteristics	Reference
Anabaena		
PCC 7120	Wild-type *Anabaena* sp. strain PCC 7120	R. Haselkorn
AMC649	*Anabaena* PCC 7120 *sigE::luxAB*-Sp^r Sm^r cassette at ClaI site of *sigE*; Sm^r Sp^r	(95)
AMC1452	*Anabaena* PCC 7120 carrying P_{sigE}-*gfpmut2* on pAM3652; Km^r Nm^r	(2)
AMC1773	*Anabaena* PCC 7120 carrying P_{petE}-*sigE* on pAM3961; Km^r Nm^r	This study
AMC1774	*Anabaena* PCC 7120 with integrated pRL277-P_{nifHD}-*gfpmut2*; Sm^r Sp^r	This study
AMC1775	*Anabaena* PCC 7120 with integrated pRL277-P_{nifHD}-*gfpmut2* and P_{petE}-*sigE* on pAM3961; Sm^r Sp^r Km^r Nm^r	This study
AMC1776	*Anabaena* PCC 7120 with integrated pRL278-P_{nifHD}-*gfpmut2*; Km^r Nm^r	This study
AMC1777	AMC649 carrying pRL278-PnifHD-*gfpmut2*; Km^r Nm^r	This study
AMC1789	*Anabaena* PCC 7120 carrying $P_{sigE\text{-P1}}$-*gfpmut2* on pAM4469, Km^r Nm^r	This study
AMC1790	*Anabaena* PCC 7120 carrying $P_{sigE\text{-P2}}$-*gfpmut2* on pAM4468, Km^r Nm^r	This study
AMC1791	*Anabaena* PCC 7120 carrying $P_{sigE\text{-P3}}$-*gfpmut2* on pAM4467, Km^r Nm^r	This study
AMC1792	*Anabaena* PCC 7120 carrying $P_{sigE\text{-P4}}$-*gfpmut2* on pAM4466, Km^r Nm^r	This study
Plasmids	Characteristics	Reference
pAM1956	Shuttle vector pAM505 containing promoterless *gfpmut2* with unique upstream cloning sites SalI, SacI, KpnI, and SmaI; Km^r Nm^r	(194)
pAM2770	Shuttle vector containing XhoI-P_{petE}-NdeI-*lacZ*α-SapI(Cys)-6His *(*stop)-ClaI; Km^r Nm^r	(194)
pAM3652	pAM1956 containing 800-bp fragment upstream of *sigE* (alr4249), P_{sigE}-*gfpmut2*	(2)
pAM3929	*ntcA* was cloned into expression plasmid pET-28b+ in EcoRI and XhoI sites	This study

TABLE 4. continued…

pAM3961	*sigE* (ORF alr4249) cloned into pAM2770 containing P_{petE} in NdeI and XmaI sites; Km^r Nm^r	This study
pAM4466	pAM1956 containing 79-bp fragment upstream of *sigE* (alr4249), $P_{sigE\text{-}P4}$-*gfpmut2*; Km^r Nm^r	This study
pAM4467	pAM1956 containing 260-bp fragment upstream of *sigE* (alr4249), $P_{sigE\text{-}P3}$-*gfpmut2*; Km^r Nm^r	This study
pRL277	Conjugal suicide plasmid; Sm^r Sp^r	(46)
pRL278	Conjugal suicide plasmid; Km^r Nm^r	(46)

DNA manipulations and plasmid constructions. Standard protocols were used for cloning, *E. coli* transformation, and PCR. Primers used in this study are listed in Table 5. Total DNA from *Anabaena* PCC 7120 was isolated as previously described (95). DNA sequencing of plasmid inserts was performed by the Gene Technologies Laboratory (Texas A&M University) following the Big Dye sequencing protocol (Applied Biosystems), and by GENEWIZ (La Jolla, CA).

A *sigE* overexpression plasmid was constructed by PCR amplification of *sigE* (alr4249) with primers containing an NdeI site at the 5' end and a XmaI site at the 3' end (AMO-2222 and AMO-2223) and then ligation of the fragment into the same sites in pAM2770, which contains a copper-inducible *petE* promoter.

Site-directed mutagenesis of the consensus NtcA binding site in the *sigE* promoter was conducted using standard protocols (91). Primers AMO-2217 and AMO-2218 were used to mutate the site in pAM3652 (P_{sigE}-*gfp*).

Plasmids containing *gfpmut2* transcriptional fusions with *sigE* upstream fragments (pAM3652, pAM4466, pAM4467, pAM4468, and pAM3859) were constructed in shuttle plasmid pAM1956 (Table 4). The *sigE* upstream fragments P1, P2, P3, and P4 were PCR amplified from total genomic DNA with primers shown in Table 5. The resulting fragments contained engineered SalI and Acc65I sites and were inserted at the same sites in pAM1956.

TABLE 5. Primers used in this study (Chapter IV)

Gene and experiment	Forward Primer	Sequence (5'→3')	Reverse primer	Sequence (5'→3')
sigE probe	AMO-2222	CCTATTCCCGGGCTCAAATCAACC AACC	AMO-2223	AACAGTCATATGGAA ATTATGTACCAAAC
sigE upstream for EMSA	RAM-Fw	CTTGAATTCTAAAGATAAATATTC TTATTGCC	RAM-Rv	GTGGAATTCAAATAT TTATTGCACGTAT
P*sigE*-P1	AMO-2558	GTCGTCGACTTTTATACAAACAA TTTTGTG	AMO-2554	GGTGGTACCTTGAGG GATTCATGCTTT
P*sigE*-P2	AMO-2557	GTCGTCGACTTTATAAAAGTATA TAATTATT	AMO-2554	GGTGGTACCTTGAGG GATTCATGCTTT
P*sigE*-P3	AMO-2556	GTCGTCGACAGATTAAATGATTA ATAAA	AMO-2554	GGTGGTACCTTGAGG GATTCATGCTTT
P*sigE*-P4	AMO-2555	GTCGTCGACATGAATCTGAAAAT TGCT	AMO-2554	GGTGGTACCTTGAGG GATTCATGCTTT
nifH probe	AMO-622	TTCACGGTCAACCTTACGG	AMO-1038	CGGTAAAGGCGG
Site directed mutagenesis of P_{sigE}-gfp	AMO-2217	GTATTTTGAATGCAGATTTTATTC	AMO-2218	GAATAAAATCTGCATT CAAAATAC
nifH qRT PCR	AMO-2185	GCACAAGAAATCTACATC	AMO-2186	TACGAAGTGAATCATT TG
hupL qRT PCR	AMO-2189	TGCTTCTCACTTAACCTCTG	AMO-2190	GTCAATGGCGAACAAT CC
fdxH qRT PCR	AMO-2191	TACCAAGTTAGATTGATC	AMO-2192	AAGTAACACAAAGTA GAG
hglE2 qRT PCR	AMO-2197	TTGTTGAGGTGTATTGAG	AMO-2198	CAGGAACATCAGTGAT AC

TABLE 5. continued…

Gene and experiment	Forward Primer	Sequence (5'→3')	Reverse primer	Sequence (5'→3')
rpoA qRT PCR	AMO-2204	AGTTTGGACAAATGGTAG	AMO-2205	GCTTGAGACAGTTAT AGG
hglE2 qRT PCR	AMO-2197	TTGTTGAGGTGTATTGAG	AMO-2198	CAGGAACATCAGTGAT AC
rpoA qRT PCR	AMO-2204	AGTTTGGACAAATGGTAG	AMO-2205	GCTTGAGACAGTTAT AGG

Electrophoretic mobility shift assay (EMSA). A 6×His tagged NtcA protein was over-expressed from the IPTG-inducible plasmid pET-28b+ and purified using a HisTrap HP 1 ml Nickel column (GE Healthcare) using standard procedures. A 207-bp DNA fragment upstream the *sigE* gene containing a putative consensus NtcA binding site was amplified by PCR using oligonucleotides RAM-Fw and RAM-Rv, each containing an engineered EcoRI site at the 5' end. The PCR product was digested with EcoRI and labeled with [α-^{32}P]dATP (3000 Ci mmol^{-1}) by fill-in using the DNA polymerase Klenow fragment. Binding assays were carried out in a final volume of 10 µl containing 10 mM Tris-HCl (pH 7.5), 50 mM KCl, 1 mM dithiothreitol, 5 ng µl^{-1} bovine serum albumin, 12.5% (v/v) glycerol, and 10 fmol of the DNA fragment; the reaction was supplemented with 2-oxoglutarate (0.6 mM) where indicated. The labeled probe was incubated with 1 to 4 pmol of purified 6×His-NtcA protein for 30 minutes at room temperature. DNA fragments were separated by electrophoresis on a non-denaturing 5% polyacrylamide gel, vacuum-dried, and visualized by phosphorimaging (Fujifilm).

RNA isolation. A RiboPure-Bacteria kit (Ambion) was used to isolate total RNA from *Anabaena* PCC 7120 and its derivates. One flask was grown for each time-point sample for each strain. Filaments from growing BG-11 cultures were inoculated into 250 ml flasks containing 100 ml of BG-11(NH_4), which is BG-11_0 supplemented with 2.5 mM ammonium chloride and 5 mM MOPS (pH 8.0) to obtain a starting OD_{750} of 0.025 - 0.035. These cultures were grown on an orbital shaker under standard conditions overnight to OD_{750} 0.05 - 0.075, and then the cells were collected at room temperature by centrifugation in 50 ml conical tubes, washed twice with BG-11_0, and transferred to 100 ml of BG-11_0 in 250 ml flasks. The flasks were incubated under standard conditions for the appropriate time after nitrogen step-down. Then, each culture was poured into a conical 250 ml centrifuge tube containing 100 g of crushed ice to rapidly chill the sample, and the filaments were collected by centrifugation at 5,000 × g for 10 minutes at 4°C. The cell pellet was immediately frozen at -80°C and stored until RNA was isolated.

Northern RNA blot analysis. For each sample, ten µg of total RNA was separated on a 1.5% agarose denaturing formaldehyde gel in MOPS buffer (155), and then transferred by capillary action to a MAGNACHARGE nylon membrane (GE Osmonics) with 10× SSPE (64).DNA probes were amplified by PCR with appropriate primers (Table 5), labeled with [α-^{32}P]-dCTP by random-primer labeling, and purified on Micro Bio-Spin P-30 columns (Bio-Rad). Blots were blocked, hybridized with radioactively labeled DNA probes, and washed using standard protocols (155), and then exposed to a phosphorimager plate and scanned with a Phosphorimager BAS-5000 (Fujifilm).

Time-lapse microscopy. Time-lapse microscopy was used to record differential interference contrast (DIC), autofluorescence of photosynthetic pigments, and GFP-fluorescence images during heterocyst development. Filaments of each reporter strain were grown in nitrate-containing BG-11 medium to an OD_{750} of 0.2-0.3, washed with BG-11_0, and resuspended in a small volume of BG-11_0 for heterocyst induction. 10 μl of induced *Anabaena* filaments were placed into a single-chambered coverglass (Lab-Tek chamber slide system; Nalge Nunc International) and covered with a thin agarose pad supported by dialysis membrane. The agarose pad was made by slicing open a dialysis tube (1.5 cm by 2.5 cm VWR Scientific, MW cut off = 3,500 kDa) that had been equilibrated in BG-11_0, laying the dialysis membrane on a glass slide, and then covering the membrane with 300 μl of warm filtered BG-11_0 containing 0.7% agarose. After the agarose cooled and solidified, small square sections of the agarose pad were cut (about 0.4 by 0.4 cm), and used to cover the induced filaments in the single-chambered coverglass. Excess liquid medium was removed and the agarose pad was then carefully surrounded on all four sides by a total of about 2 ml of warm BG-11_0 containing 0.7% agarose in order to maintain moisture in the thin agarose pad. For time-lapse microscopy, the chamber was covered with its plastic lid.

Microscopy was performed with a DeltaVision Core system (Applied Precision) with a WeatherStation attached to an Olympus IX71 inverted microscope. The temperature was adjusted to 30ºC. SoftWoRx software was used to acquire time-lapse images for 24 to 42 h with a 15-minute time delay using a 40× objective. The time-lapse series was started approximately 10 to 20 minutes after nitrogen step-down. The *Anabaena*

filaments were illuminated to support their growth during time-lapse microscopy with approximately 75 µmol photons m^{-2} s^{-1} by a fiber optic LED white light source (Schott IFC 60825) attached to the DeltaVision Core system and controlled by softWoRx software, which controlled switching between the external white light, and the microscope white light and excitation light for taking images. For all images, a polychroic beam splitter was used. DIC exposure time was 0.05 seconds; for autofluorescence images, a TRITC filter set was used (EX555/EM617) with an exposure time of 0.05 seconds; and for GFP images, a Chroma GFP filter set was used (HQ EX470/EM515) with a 0.5 second exposure time. Time-lapse images were processed using softWoRx software, and individual images from specific time-points were exported as required.

Scoring of the heterocyst spacing pattern along filaments was performed essentially as previously described (195). Detached single heterocysts and aggregates of heterocysts were not scored. Because the strains AMC1774 and AMC1775 contained a *nifHD-gfp* reporter, GFP fluorescence was used as an additional character to identify proheterocysts. Strain AMC1777 was scored only by morphology because the *nifHD-gfp* reporter did not produce detectable GFP fluorescence in the *sigE* mutant background.

Quantitative Real-Time RT-PCR Analysis (qRT-PCR). qRT-PCR analysis was carried out with a StepOnePlus 96 well qRT-PCR system (Applied Biosystems) and *Power* SYBR Green PCR Master Mix (Applied Biosystems) mixed with 100 pmol/µl of each primer and 50 ng of total cDNA. All primers used in this study are listed in Table 5. Primers for *nifH*, *fdxH*, *hupL*, *hglE2*, and *rpoA* were designed with Vector NTI software

(Invitrogen) and each primer was analyzed using the BioBike web interface for sequence-similarity searches against the *Anabaena* PCC 7120 genome. cDNA synthesis from total RNA samples (grown on nitrate, or 12, 24, and 36 h after nitrogen step-down) was performed with an iScript cDNA kit (Bio-Rad). All samples were run in triplicate using the following amplification conditions: one cycle of 2 minutes at 52°C and 10 minutes at 95°C, and then 40 cycles at 95°C for 15 seconds and 60°C for 30 seconds. Ct (cycle threshold) values for each gene were averaged and normalized against *rpoA*.

Results

***sigE* transcripts.** Northern RNA blot analysis of *sigE* was performed to determine transcript size and abundance at different developmental times. The northern RNA blot results showed that *sigE* is developmentally regulated and that transcript levels were highest at 12 and 24 h after nitrogen step-down (Fig. 18). The *sigE* transcripts were present as a smear, suggesting a short half-life, with an upper edge around 2.5 to 3.0 kb, which is consistent with a monocistronic transcript. *sigE* transcript levels were decreased

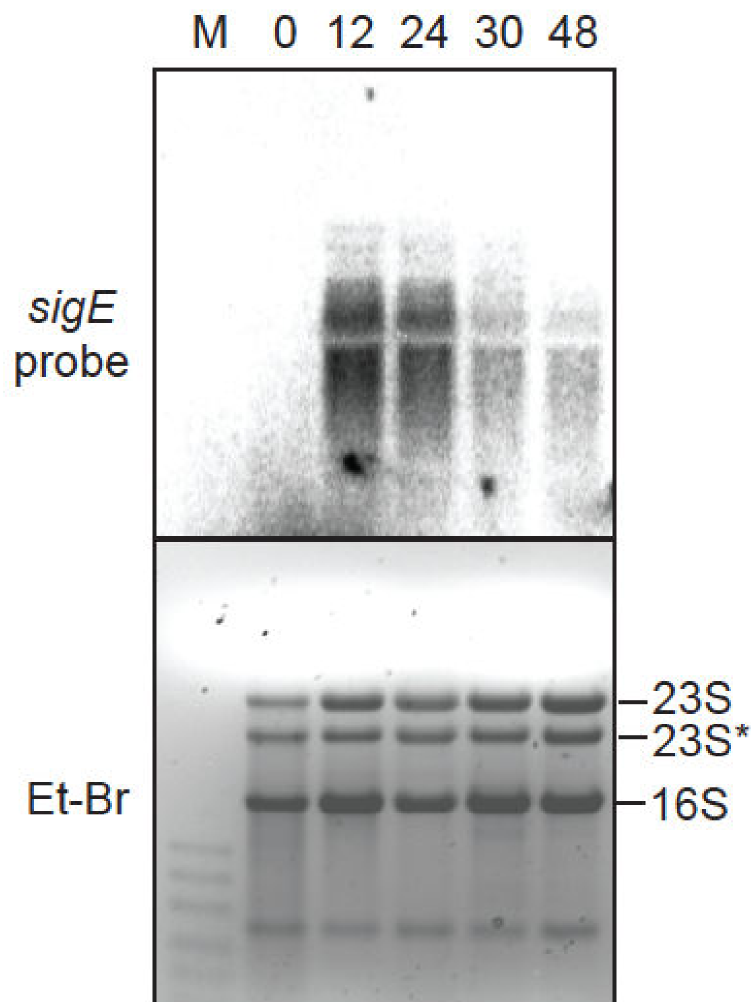

FIG. 18. Northern RNA blot analysis of *sigE* transcripts during heterocyst development. Lanes are labeled with hours post nitrogen step-down from BG-11 to BG-11_0 medium. The upper panel shows a northern RNA blot probed with a radioactive probe for *sigE* transcripts. The lower panel is the ethidium bromide stained gel to show sample loading. Lane M, Low Range ssRNA Ladder (NEB) size marker.

at 30 and 48 h after nitrogen step-down. Several attempts to map the 5' ends of *sigE* transcript using rapid amplification of 5' cDNA ends (RACE) were unsuccessful, possibly because the transcripts were unstable.

RNA-seq data from total RNA samples obtained at 0, 6, 12, and 21 hours after nitrogen step-down were consistent with the northern blot data, with RPKM (Reads Per Kilobase of CDS [coding sequence] model per Million mapped reads) values of 9.08, 8.90, 20.51, and 24.86, respectively (53). The RNA-seq data did not show a clear 5' end for *sigE* transcripts, which is consistent with our RACE results. At all time points, staggered overlapping reads extended to about -330 nucleotides upstream of the first base of the ORF with a gap in reads between about -60 and -140.

NtcA interacts with the upstream region of *sigE* in vitro. It has been shown that some heterocyst-specific promoters carry an NtcA-binding site with a consensus sequence of tGTAN$_8$TACa, which is often located approximately 22 nucleotides upstream from a -10 TAN$_3$T box (43, 78). Our bioinformatics analysis of the *sigE* promoter region identified a canonical NtcA binding site twenty-two nucleotides upstream of a TAN$_3$T sequence, from position -683 to -698 relative to the *sigE* translational start site, which is in agreement with an independent analysis of NtcA binding sites in the *Anabaena* PCC 7120 genome (163). A second potential NtcA binding site that differs at one of the highly conserved bases is present at position -632 to -647. To determine if these sites interacted with NtcA, we performed electrophoretic mobility shift assays (EMSA) using a radioactively labeled 207-bp PCR fragment containing both putative NtcA binding sites, purified 6×His-tagged NtcA protein, and

the NtcA effector metabolite 2-oxoglutarate (2-OG). Figure 19 shows NtcA, indicating NtcA binding to this region at two sites. The affinity of NtcA for this two band-shifts of the labeled DNA fragment in the presence of increasing amounts of promoter fragment increased in the presence of 2-OG and the binding was significantly reduced in the presence of unlabeled competitor DNA.

To determine if the highly conserved NtcA binding site is required for developmentally regulated expression of *sigE*, the binding site was mutated in plasmid pAM3652, which carries P_{sigE}-*gfp*. The NtcA binding site 5'-TGTA-N_8-TAC-3' in the *sigE* promoter region was replaced with 5'-GTGC-N_8-TAC-3', and the plasmid was introduced into wild-type *Anabaena* PCC 7120. The resulting strain, AMC1778, was analyzed by time-lapse microscopy after nitrogen step-down (data not shown). No significant differences in the pattern of GFP fluorescence was detected, showing that the conserved NtcA binding site is not essential for *sigE* expression *in vivo*, but does not eliminate the possibility that the site contributes to normal expression.

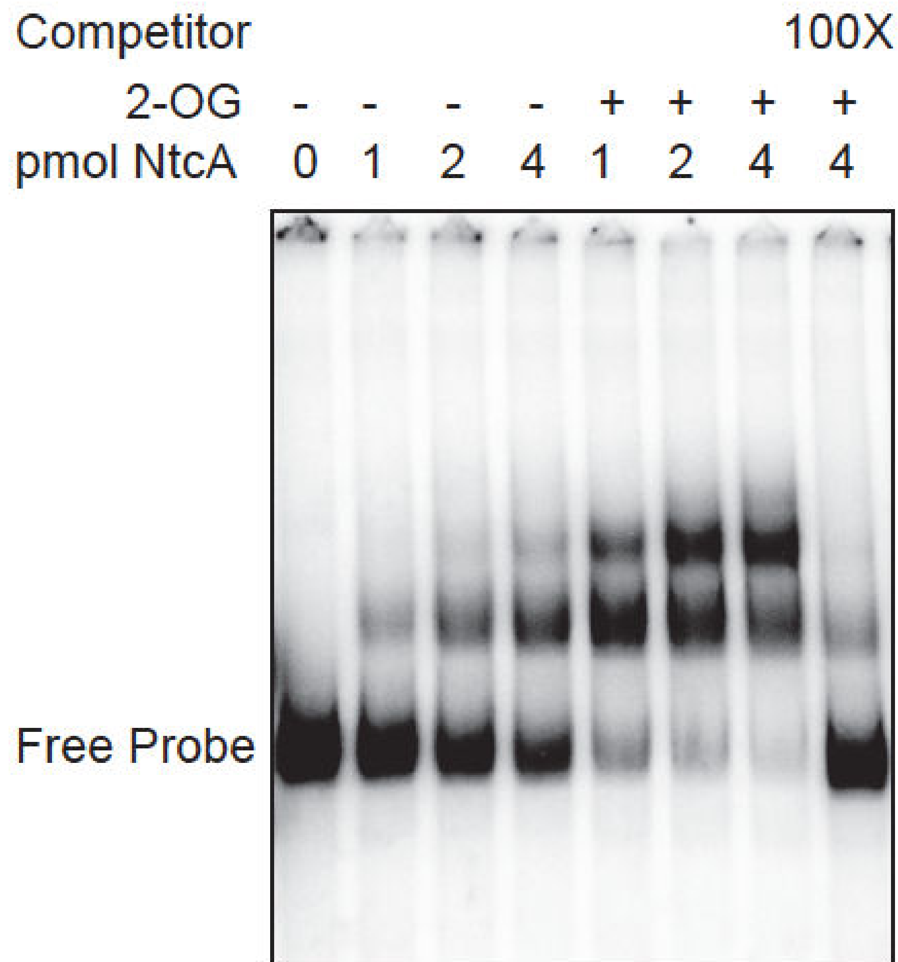

FIG. 19. NtcA binds to the *sigE* upstream intergenic region. Binding of purified 6×His-tagged NtcA to a 207-bp fragment from the intergenic region upstream of *sigE* was assayed by electrophoretic mobility shift assay (EMSA). Sample mixtures contained 0.1 fmol of labeled DNA fragment with or without the presence of 0.6 mM 2-oxoglutarate (2-OG). The last lane contains a 100-fold molar excess of the unlabeled probe as competitor.

To identify the minimal upstream region required for heterocyst-specific *sigE* expression, a deletion series of four fragments were used to drive *gfp* reporter expression (Fig. 20). Reporter plasmids containing $P_{sigE\text{-P}}$-*gfp* (-800 bp), $P_{sigE\text{-P1}}$-*gfp* (-500 bp), $P_{sigE\text{-P2}}$-*gfp* (-360 bp), $P_{sigE\text{-P3}}$-*gfp* (-260 bp) and $P_{sigE\text{-P4}}$-*gfp* (-79 bp) were transferred into wild-type *Anabaena* PCC 7120 to obtain reporter strains AMC1452, AMC1789, AMC1790, AMC1791, and AMC1792, respectively (Table 1). Observation of GFP fluorescence 24 h after nitrogen step-down revealed that the strains containing $P_{sigE\text{-P}}$-*gfp*, $P_{sigE\text{-P1}}$-*gfp*, $P_{sigE\text{-P2}}$-*gfp*, and $P_{sigE\text{-P3}}$-*gfp* all showed higher levels of GFP fluorescence in heterocysts relative to vegetative cells, but that $P_{sigE\text{-P}}$-*gfp*, which contained the entire upstream intergenic region including the two NtcA binding sites, produced the strongest differential upregulation in heterocysts. The strain containing $P_{sigE\text{-P4}}$-*gfp* produced low-level expression in all cells. A strain carrying promoterless *gfp* on pAM1956 was dark. These data show that a 260-bp upstream region was sufficient to provide developmental regulation but that the entire upstream intergenic region, which includes the two NtcA binding sites, was required for robust regulation.

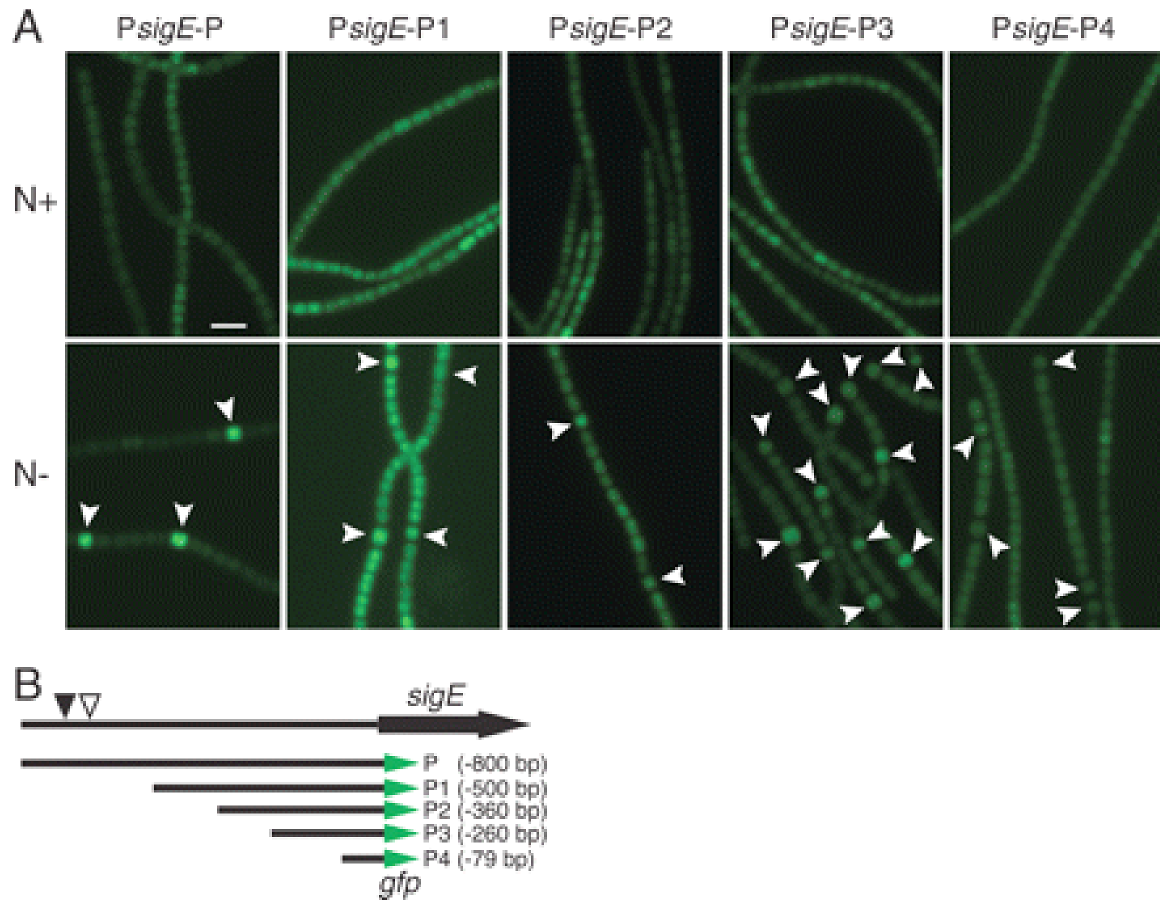

FIG. 20. GFP reporter expression from different upstream regions of *sigE*. (A) GFP fluorescence micrographs of filaments for reporter strains grown with (N+, upper panel) and without (N-, lower panel) combined nitrogen for 24 h. (B) Map showing upstream regions of *sigE* driving the *gfpmut2* reporter. Region P contains the entire 800-bp upstream intergenic region. The position of a conserved NtcA binding site is marked with a filled triangle and a second potential binding site is marked with an open triangle. Heterocysts are indicated by arrowheads. Scale bar, 10 µm.

***nifH* expression is delayed and reduced in the *sigE* mutant strain.** To determine if SigE is required for expression of genes in the later stages of heterocyst development, we analyzed the expression of the *nifH* gene, which encodes the nitrogenase iron protein, in wild-type and *sigE* mutant backgrounds. Total RNA from the wild type and the *sigE* mutant strain AMC649 was isolated at different time points after nitrogen step-down and then analyzed by northern RNA blot. Figure 21 shows the expression of *nifH* in the wild type and *sigE* mutant when grown in medium containing nitrate, labeled 0, and at 12, 24, 30, 36, and 48 h after nitrogen step-down. The *nifH* probe hybridizes with three transcripts of approximately 1.1, 2.8, and 4.7 kb, which contain the *nifH*, *nifHD*, and *nifHDK* open reading frames, respectively (64). Transcripts containing *nifH* in the wild-type background were detected at low levels at 12 h after nitrogen step-down and reached their highest level by 24 h. In the *sigE* mutant strain AMC649, *nifH* transcripts were detected only weakly at 24 h and 30 h after nitrogen step-down, increased at 36 h, but at lower levels than the wild-type strain, and significantly decreased at 48 h. These results show that expression of *nifH* is delayed and reduced in the *sigE* mutant, and together with data showing that *sigE* is upregulated at later times in proheterocysts, suggest that SigE is required for normal levels of *nifH* expression.

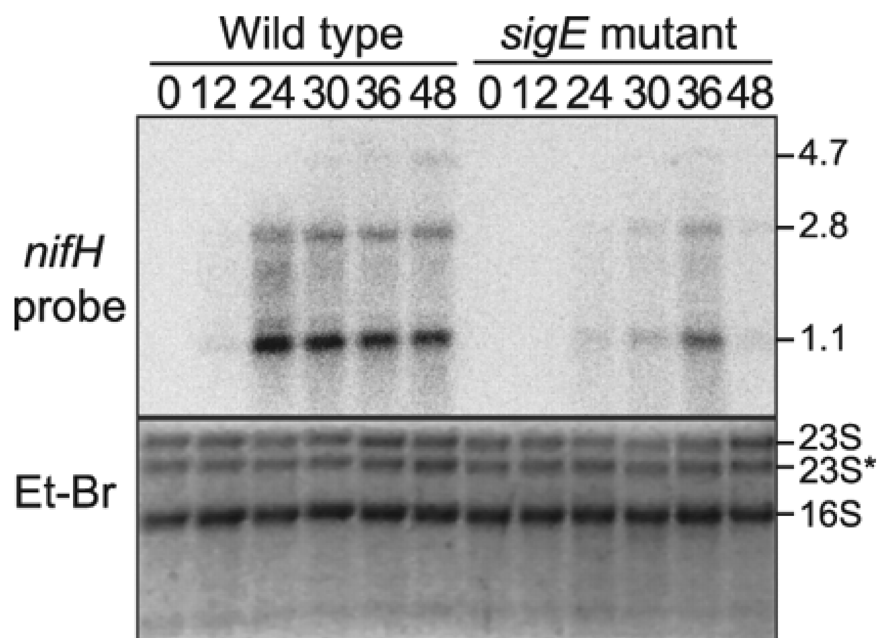

FIG. 21. Northern RNA blot analysis of *nifH* transcripts from the wild type and *sigE* mutant strain AMC649 during heterocyst development. Lanes are labeled with hours post nitrogen step-down from BG-11 to BG-11_0 medium. The upper panel shows a northern RNA blot probed with a radioactive probe for the *nifH* transcript. The lower panel is an ethidium bromide stained gel to show sample loading. Three transcripts containing *nifH* were detected at approximately 1.1 kb, 2.8 kb, and 4.7 kb. The *sigE* mutant strain AMC649 had delayed and reduced levels of *nifH* transcripts.

***sigE* influences the expression of a $P_{nifHD\text{-}gfp}$ reporter during heterocyst development.** Time-lapse microscopy of heterocyst development was used to determine if *sigE* affects expression of a P_{nifHD}-*gfp* reporter and/or heterocyst pattern *in vivo*. Strains containing a P_{nifHD}-*gfp* reporter integrated into the *Anabaena* PCC 7120 chromosome at the *nifH* locus were used for these studies. The P_{nifHD}-*gfp* reporter was placed into three genetic backgrounds: wild type (AMC1774), a *sigE* knockout mutant (AMC1777), and a *sigE* overexpression strain (AMC1775). Heterocyst development was induced and DIC, autofluorescence, and GFP fluorescence images were acquired every 15 minutes for 24 h for strains AMC1774 and AMC1775, and for 42 h for strain AMC1777. Figure 22 contains images corresponding to the indicated time-points that were extracted from the time-lapse series.

At 1 hour after nitrogen step-down, no GFP fluorescence or heterocyst morphological characteristics were evident in any of the three strains (AMC1774, AMC1775, and AMC1777).

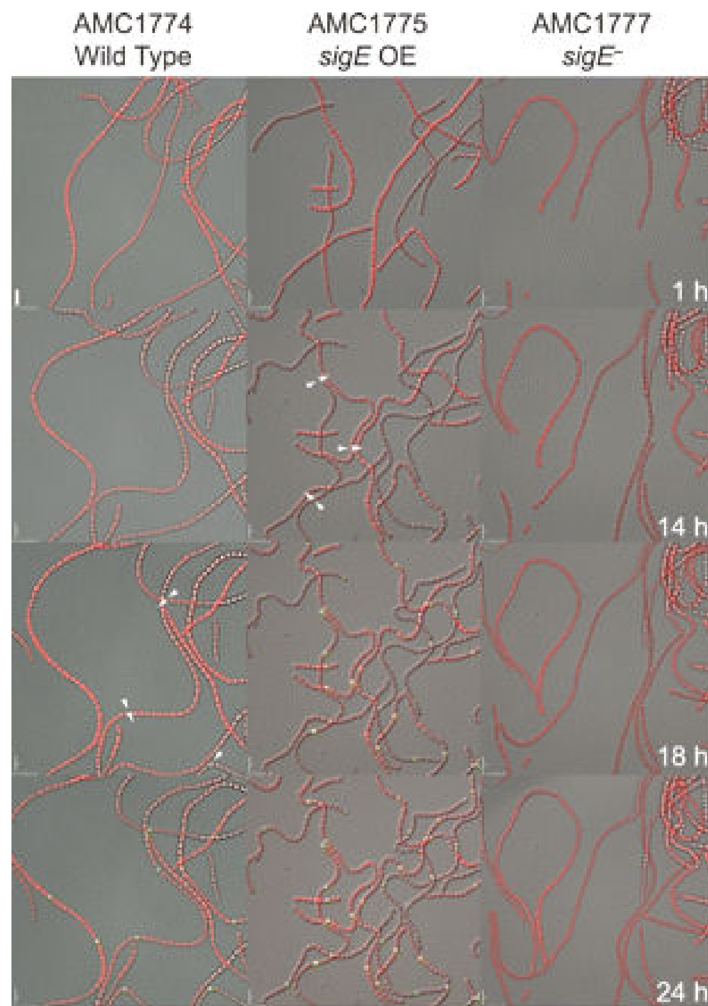

FIG. 22. Time-lapse microscopy images of the P_{nifHD}-gfp reporter construct integrated into the nifH locus on the chromosome in wild-type (AMC1774), sigE overexpression (AMC1775), and sigE mutant (AMC1777) backgrounds. Images are shown from 1, 14, 18, and 24 hours after nitrogen step-down. Images are merged DIC (grayscale), autofluorescence (red), and GFP reporter fluorescence (green). Arrows indicate weak GFP fluorescence from proheterocysts. Scale bars, 10 µm.

By 14 h, when morphological differentiation of proheterocysts was not yet obvious in the wild type, the *sigE* overexpression strain (AMC1775) showed GFP fluorescence in individual cells that were morphologically distinguishable as proheterocysts. AMC1774 and AMC1777 filaments did not show any GFP fluorescence or proheterocysts at 14 h after nitrogen step-down.

At 18 h after nitrogen step-down, the wild type strain (AMC1774) showed GFP fluorescence in some individual cells, which is somewhat delayed compared to expression of *nifH* determined by northern RNA blot experiments (Fig. 21); this is likely due to differences in growth conditions between batch cultures and samples on the microscope and the time required to accumulate sufficient GFP to be detectable under our time-lapse microscopy conditions, which uses low excitation intensity to avoid photobleaching or cell damage. For the wild-type strain at 18 h, slight morphological changes in the cells that showed GFP fluorescence were apparent and the average number of vegetative cells in the interval between heterocysts or GFP-bright proheterocysts was 26.6 (Fig. 22). The *sigE* overexpression strain showed strong GFP fluorescence at 18 h and showed clear heterocyst morphological differentiation with an average interval of 12.4. In the *sigE* mutant strain (AMC1777), no GFP fluorescence or morphological differentiation were detected 18 h after heterocyst induction.

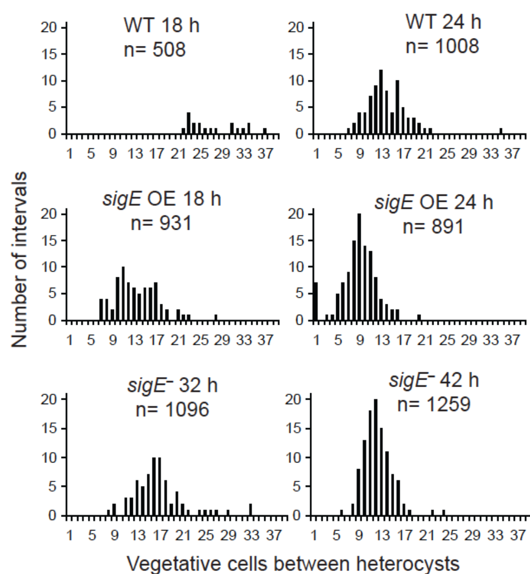

FIG. 23. Heterocyst pattern is influenced by *sigE*. Heterocyst pattern was determined for the strains shown in Fig. 22, which have a *nifHD-gfp* reporter in wild-type (AMC1774), *sigE* overexpression (AMC1775), and *sigE* mutant (AMC1777) backgrounds. For the wild-type (WT) and *sigE* overexpression (OE) backgrounds, heterocyst pattern was scored by observing cell morphology and GFP fluorescence at 18 and 24 h after nitrogen step-down. For the *sigE* mutant background, which had delayed development and very low levels of *nifHD-gfp* reporter expression, the heterocyst pattern was scored by cell morphology at 32 and 42 h after nitrogen step-down. For each sample, the number of vegetative cells between heterocysts was scored, with adjacent heterocysts scored as an interval of zero. The x-axis shows the number of intervals scored for each interval length. The average number of vegetative cells between heterocysts was 26.6 at 18 h and 13.1 at 24 h for wild type; 12.4 at 18 h and 7.9 at 24 h for the *sigE* overexpression strain; and 15.8 at 32 h and 11.3 at 42 h for the *sigE* mutant.

At 24 h after nitrogen step-down, the P_{nifHD}-*gfp* reporter in a wild-type background (AMC1774) showed GFP fluorescence in individual differentiated cells (Fig. 22) with an interval length of 13.1 vegetative cells (7.1 percent heterocysts) (Fig. 23). The *sigE* overexpression strain (AMC1775) showed a higher frequency of GFP-fluorescent heterocysts with an interval length of only 7.9 (11.2 percent heterocysts). These data show that extra copies of *sigE* caused accelerated heterocyst formation and expression of *nifH*, and an increase in heterocyst frequency. At 24 h after nitrogen step-down, the *sigE* knockout mutant strain (AMC1777) did not show detectable GFP fluorescence or heterocyst morphological changes. By 32 h after nitrogen step-down, the *sigE* mutant showed morphological differentiation of some individual cells and an interval length of 15.8 (6.0 percent heterocysts), but no GFP fluorescence was evident. At 42 h, the *sigE* mutant had a pattern of mature heterocysts similar to the wild type with an average interval length of 11.3 (8.1 percent heterocysts) (Fig. 23), but no GFP fluorescence was detectable even at this time.

The *sigE* gene is required for expression of other heterocyst-specific genes. Previous observations of unsynchronized cultures of a *sigE* mutant strain (AMC649) found little effect on heterocyst development or diazotrophic growth (95), and in our current experiments, we observed no significant difference in the growth rate of *sigE* mutant strains compared to the wild type once diazotrophic growth was established. However, the *sigE* expression data (Fig. 18), and the delay in heterocyst development (Figs. 22 and 23) and the delayed and reduced *nifH* expression (Fig. 21) observed in the *sigE* mutant suggest that the SigE sigma factor may be required for transcription of

genes expressed in the middle to late stages of heterocyst differentiation. To determine if SigE is required for the expression of other heterocyst-specific genes, we used qRT-PCR to examine the expression levels of genes that are upregulated during middle and late stages of heterocyst development. We selected *nifH*, *fdxH* (encoding a heterocyst-specific ferredoxin) and *hupL* (encoding the large subunit of uptake hydrogenase) as genes that are expressed during late stages of heterocyst development (32, 175). For a gene expressed during the middle stages of heterocyst development, we selected *hglE2* (all1646), which encodes a type I polyketide synthase potentially involved in heterocyst glycolipid synthesis (28). For a baseline control mRNA, the expression of the housekeeping gene *rpoA*, which encodes the RNA polymerase alpha subunit, was analyzed by qRT-PCR during heterocyst development. *rpoA* message levels remained essentially invariant and at relatively low levels at all time points (Fig. 24); therefore, we used *rpoA* message as an internal baseline control for our qRT-PCR experiments.

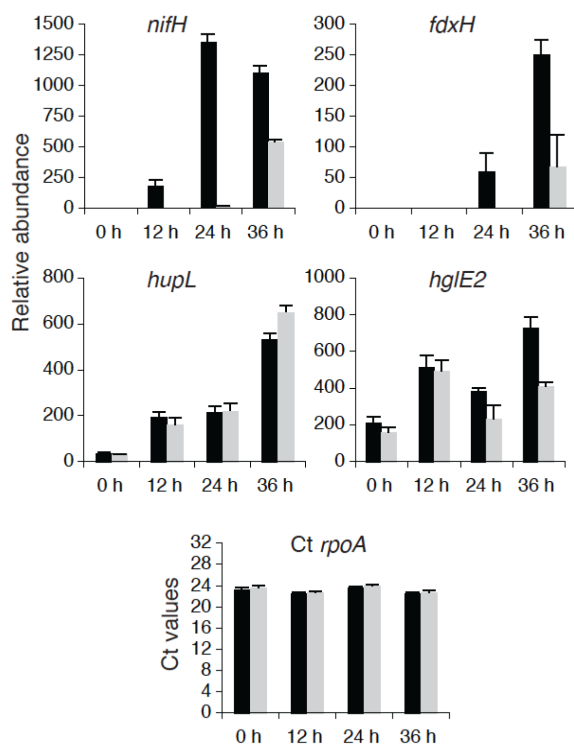

FIG. 24. Quantitative reverse transcription PCR (qRT-PCR) measurements of mRNA abundance in wild-type and *sigE* mutant strains. Relative mRNA abundance was determined for *nifH, fdxH, hupL,* and *hglE2* using total RNA from wild-type *Anabaena* PCC 7120 (black bars) and the *sigE* mutant AMC649 (grey bars) at four time points after nitrogen step-down. The data represent average values and standard error measurements from three technical replicates, shown as percent relative abundance normalized to the RNA polymerase alpha subunit (*rpoA*) gene. The Ct *rpoA* graph shows the average Ct values for *rpoA* normalized against total RNA for wild-type *Anabaena* PCC 7120 and the *sigE* mutant AMC649.

Figure 24 shows the transcript levels of *nifH*, *fdxH*, *hupL, and hglE2* relative to *rpoA* in wild-type and *sigE* mutant backgrounds during heterocyst induction. In media containing nitrogen (0 h), *nifH* transcripts were not detected in either the wild type or the *sigE* mutant. At 12 h after nitrogen step-down, small amounts of *nifH* transcript were detected, but only in the wild-type strain. At 24 h a peak of *nifH* transcript was observed in the wild-type strain, while only trace amounts were detected in the *sigE* mutant strain. By 36h, the levels of *nifH* transcript in the *sigE* mutant strain had increased to about half the wild-type transcript level. These results are in accordance with the data obtained by northern RNA blot analysis (Fig. 21) and with the P_{nifHD}-*gfp* reporter, which supports the hypothesis that SigE regulates the expression of *nifH*.

Expression of the *fdxH* gene was similar to that of *nifH*, although somewhat delayed (Fig. 24). Message was first detected for wild-type filaments at 24 h and was abundant by 36 h, but for the *sigE* mutant strain, *fdxH* message was detectable at low levels in only the 36 h sample. These data suggest that SigE regulates the expression of *fdxH* during heterocyst development.

Although previous studies in *Anabaena* PCC 7120 showed expression of *hupL* only during the late stages of heterocyst development (32), *hupL* transcripts have been detected in ammonium-grown vegetative cells in *Anabaena variabilis* (13). Our present data show very low levels of *hupL* transcripts in samples from vegetative cells in nitrate-containing medium (0 h) and increasing levels at later stages of heterocyst development (Fig. 24). The *hupL* transcript levels were not significantly different between the *sigE*

mutant strain and the wild type, indicating that SigE is not involved in the regulation of the *hupSL* operon.

For the *hglE2* gene, transcripts were present in all samples, with the highest levels at 12 h and 36 h after heterocyst induction and a decrease at 24 h (Fig. 24). The *hglE2* transcripts in the *sigE* mutant were significantly lower than the wild type at 24 and 36 h, with the levels at 36 h about half the wild-type levels. These data suggest that SigE may regulate *hglE2* expression at late stages of heterocyst development but not in vegetative cells or at early stages of development. Together, these data indicate that SigE is involved with the regulation of some but not all genes at late stages of heterocyst development.

Discussion

Previous studies had found that none of the *Anabaena* PCC 7120 group 2 sigma factors *sigB*, *sigB2*, *sigC*, *sigD*, or *sigE* is individually essential for heterocyst development or diazotrophic growth (18, 95). However, more recent results showed that *sigC* and *sigE* are upregulated in differentiating heterocysts (2). In this work, using a more detailed examination of gene expression and heterocyst development, we show that the *sigE* gene is required for normal heterocyst development and for normal expression of certain heterocyst genes.

Expression analysis of *sigE* by northern RNA blot showed developmental upregulation of transcripts beginning 12 h after nitrogen step-down, which is consistent with our previous observations using a P_{sigE}-*gfp* reporter (2). The *sigE* transcripts

showed significant degradation and attempts to identify a specific *sigE* transcriptional start site were not successful, indicating a relatively short half-life, which is normal for most bacterial mRNAs.

Bioinformatic analysis identified a putative consensus NtcA binding site in the *sigE* upstream region, which could contribute to its developmental regulation. EMSA showed that NtcA binds to a DNA fragment containing this site and produced two shifted bands, indicating a second binding site. Binding was enhanced by the presence of 2-OG. Further analysis of the fragment sequence allowed the identification of another potential binding site that differs from the consensus at one position. Deletion of both binding sites resulted in a partial loss of developmental expression indicating that NtcA plays a partial role in the regulation of *sigE*. Because *ntcA* mutants are defective for heterocyst development and highly pleiotropic, it is not possible to use *ntcA* mutants to establish a direct role for NtcA in the regulation of heterocyst-specific genes, especially those upregulated at the middle to late stages of development. Many *Anabaena* genes show complex regulation by multiple promoters, and it is likely that *sigE* expression is regulated by multiple transcription factors to provide for expression in vegetative cells and upregulation in differentiating cells (124). Further biochemical studies will be required to determine the function of the two NtcA binding sites and if they interact with each other or alter DNA topology to influence the expression of *sigE* or the adjacent gene all4248.

Because *sigE* is expressed during the later stages of heterocyst development, we hypothesized that *sigE* may be involved in the expression of the nitrogen-fixation (*nif*)

genes. The regulation of the *Anabaena* PCC 7120 *nifHDK* operon has been studied for over twenty-five years (47, 64, 147, 149, 172, 176), however little is known about the transcription components that regulate this operon. Our results suggest that *sigE* is involved in *nifH* expression. However, regulated *nifH* expression was not completely abolished, which means that other transcription factors must be involved, and that sigma factors other than SigE must be able to provide transcription initiation. It is not clear if the residual expression is part of normal regulation or a consequence of crosstalk among the multiple group 2 sigma factors in *Anabaena* PCC 7120. Demonstrating direct regulation of *nifH* expression by SigE will require *in vitro* transcription assays using purified components. However, these experiments are especially challenging because the cis-acting sequences that constitute the *nifH* promoter have not been identified in *Anabaena* PCC 7120, and recent data with *Anabaena variabilis* suggest that multiple distant promoters and RNA processing events may be involved in expression of the *nifHDK* operon (172).

In addition to the delayed and reduced expression of *nifH*, the *sigE* mutant showed delayed heterocyst morphological differentiation; mature heterocysts were formed, but not until 32 h after nitrogen step-down (about 14 h longer than the wild-type strain). This indicates that the SigE regulon includes genes required for normal heterocyst differentiation and morphogenesis, but also that there must be some redundancy in the regulation. This redundancy may be a selective advantage because it would provide robustness to the developmental process.

Overexpression of *sigE* caused premature expression of P_{nifHD}-*gfp* and a higher heterocyst frequency than normal. The expression of P_{nifHD}-*gfp* in the *sigE* overexpression strain is first detectable 14 h after heterocyst induction (four hours earlier than the wild-type strain). Therefore, *sigE* is both necessary for normal *nifHDK* expression and sufficient to cause precocious expression of the operon. The simplest scenario is that SigE interacts directly with a promoter for the *nifHDK* operon, however, it is possible that the regulation is indirect. In addition, heterocyst morphogenesis and maturation occurred earlier and the final heterocyst frequency was higher in the *sigE* overexpression strain, indicating that the SigE regulon includes genes that influence the progression and timing of heterocyst differentiation.

Because *sigE* influenced the expression of *nifH*, we hypothesized that *sigE* might also regulate the expression of other late stage heterocyst-specific genes such as *fdxH* and *hupL* (32, 175). qRT-PCR revealed that the expression levels of *nifH* and *fdxH* were both reduced in the *sigE* mutant strain. On the other hand, the expression of *hupL* was upregulated normally in the *sigE* mutant after nitrogen step-down. These results show that *sigE* is required for the expression of only a subset of late-stage heterocyst-specific genes and indicate that at least some factors required for *hupL* expression are different from those required for *nifH* and *fdxH* expression.

The *hglE2* gene is necessary for the heterocyst glycolipid layer and is developmentally upregulated during the middle stage of heterocyst differentiation (49, 181). In the *sigE* mutant strain, the *hglE2* expression levels at 12 h were similar to the wild-type strain, however, by 36 h after nitrogen step-down the transcript levels of *hglE2*

were only half that of wild-type levels. These data suggest that *hglE2* is expressed in at least two ways: expression at early stages of heterocyst development that is SigE independent, and SigE-dependent expression at a later stage.

The focus of this work was on gene regulation related to heterocyst development and nitrogen fixation, but it seems likely that SigE in *Anabaena* PCC 7120 could also be involved in the regulation of sugar catabolic genes as has been demonstrated in other cyanobacteria (84, 141). Although *gfp* reporter experiments showed that the *sigE* gene was upregulated in differentiating cells after nitrogen stepdown, it was also expressed in uninduced filaments and in the vegetative cells of induced filaments, which is consistent with SigE being involved in gene expression in those cells.

The results presented in this work indicate that SigE is involved in the expression of heterocyst-specific genes that are upregulated during the later stages of development. A *sigE* mutant, similar to other cyanobacterial sigma factor mutants, showed incomplete genetic penetrance and expressivity. This has made the study of sigma factor function in cyanobacteria difficult. However, observations of quantitative changes in cellular and molecular phenotypes are beginning to tease out the roles played by sigma factors in cyanobacterial gene expression.

CHAPTER V

CONCLUSION

Overview

The filamentous cyanobacterium *Anabaena* sp. has evolved heterocysts, which are specialized cells that perform nitrogen fixation. Heterocyst development in *Anabaena* PCC 7120 offers a simple model to study the integration of extracellular and intracellular signals that control cellular differentiation and regulation of gene expression. Gene expression is developmentally regulated with distinct sets of genes being expressed at different stages during the formation of heterocysts. The culminating event of heterocyst development is the process of nitrogen fixation, which is catalyzed by the enzyme nitrogenase. The *nif* genes that encode the structural components of nitrogenase are strongly expressed in a fully mature heterocyst (122). Despite their early identification and abundant transcription, little was known about the factors that regulate these genes. The work presented here has advanced our understanding of heterocyst development by elucidating the regulatory mechanisms of the *nif* genes. It is likely that the same regulators and similar mechanisms control other late genes in *Anabaena* PCC 7120 and other heterocystous cyanobacteria.

Cis-acting sequences required for transcription of the *nif* genes

Fluorescence microscopy of GFP reporter strains has been used in the study of early stages of heterocyst development (24, 137). However, a concern with using GFP during

late stages of development, when the environment within the heterocyst is microoxic, is that maturation of the GFP fluorophore may not fold properly in the absence of sufficient oxygen (70). Our application of GFP as a reporter for expression of the *nif* genes has shown that GFP can be successfully used in the study of late genes of heterocyst development (Chapter II). A *nif-gfp* reporter was used to identify cis-acting sequences important for developmental regulation of the *nifHDK* genes in *Anabaena* PCC 7120 (Chapter II). Our results indicate that the excision of a 59-kb element from within the *fdxN* gene is required for transcription of the downstream *nifHDK* genes. In most bacteria, including cyanobacteria, promoters and other regulatory sequences lie within the intergenic region between the gene of interest and its flanking gene. Our study suggests that the promoter of the *nifHDK* genes lies several kilobases upstream in the *nifB* promoter. This was a surprising result because the *nifHDK* transcription start site was mapped, by S1 mapping, to -123 bp upstream of the *nifH* ATG (72, 85). Additionally, *nifB-fdxN-nifS-nifU* transcripts are present at significantly lower levels than the *nifHDK* transcripts, which was presumed to be due to differences in promoter activity (122). Our results suggest that the previously identified TSS is a processed site, and the difference in transcript levels is because of differential stability of the messages (Chapter II). S1 mapping cannot differentiate between true 5’ ends and processed 5’ ends, a caveat which we considered unlikely when we were designing experiments to identify the promoter of the *nifHDK* genes. Our results clearly indicate that future analyses of transcript 5’ ends in *Anabaena* PCC 7120 should employ techniques such as 5’ RACE with TAP (202) that can distinguish between transcription start sites and

mRNA processing sites.

Our data are supported by research on the closely related species *Anabaena variabilis*, in which it was recently shown that the *nifH1* gene is controlled by the *nifB1* promoter several kilobases upstream (172). This promoter accounts for the majority of the transcripts of the *nifH1* gene, while a second promoter within the *nifU1* gene accounts for a smaller proportion of the transcripts. Our data suggests that a single promoter upstream of *nifB* controls the *nifHDK* genes in *Anabaena* PCC 7120.

There are several examples of promoters located thousands of base pairs from the genes they control (171, 193). Our work with the *nif* genes suggests that other such complex promoters may be present in *Anabaena* PCC 7120 and possibly other cyanobacteria.

Consensus binding sequences in late promoters

A consensus promoter for heterocyst-specific genes has not been identified. Master regulators such as HetR and NtcA control the expression of multiple early genes of heterocyst development, and genes controlled by NtcA have a consensus binding sequence within 40-bp upstream of their TSS (88). NtcA binds to the *sigE* promoter, which is expressed during the later stage of heterocyst development (116). We identified an NtcA-like binding site at 37-bp upstream of the *nifB* promoter (Chapter III). It is likely that a protein belonging to the same family as NtcA acts as a master regulator of genes that are expressed during the middle and/or late stages of heterocyst differentiation. A search and comparison of NtcA-like binding sequences in the

promoters of middle and late genes such as *hepA, fdxH*, *hupSL*, and *nifH2* may reveal a late gene-specific consensus binding sequence. Biochemical and genetic experiments can be used to validate bioinformatics information and determine the specificity of such a late gene-specific binding sequence.

DevH as a transcriptional regulator of the *nifB-fdxN-nifSUHDK* operon

Heterocysts undergo morphological and physiological changes to protect nitrogenase against the cell's microoxic environment. For example, heterocysts synthesize a glycolipid envelope and express nitrogen-regulated cytochrome c oxidase-type respiratory terminal oxidases Cox2 and Cox3 (174). Cells become committed to differentiation into heterocysts at a time corresponding to these morphological and physiological changes. It is possible that the concentration of oxygen may act as a signal for commitment of a cell to differentiate, as well as for the expression of genes whose products require a microoxic environment to function. This suggests that the regulator of the *nif* genes may respond to O_2 availability.

We hypothesized that the regulator of the *nif* genes is a protein that would respond to the lack of combined nitrogen, as well as the microoxic environment of the heterocyst. The CRP family of transcriptional regulators consists of well-conserved proteins that regulate genes in response to various physiological states of the cell, e.g., carbon status, nitrogen status, and oxygen state of the cell, etc. (98). The *Anabaena* PCC 7120 genome has several genes that are predicted to encode proteins of the CRP family. NtcA, a master regulator of the response to nitrogen deprivation and of heterocyst differentiation,

is a representative member of the CRP family in *Anabaena* PCC 7120 (58).

DevH, a member of the CRP family with high sequence similarity to NtcA, is required for transcription of the *nif* genes and nitrogen fixation under oxic conditions (150). We identified an NtcA-like binding site at -37 upstream of the TSS in the PCC 7120 *nifB* promoter (Chapter III). To investigate if DevH binds specifically to this site, we performed EMSA using recombinant DevH protein extracts produced in *E. coli.* Our results show specific binding of DevH to the *nifB* promoter fragment (Chapter III) suggesting that DevH may act as a transcriptional activator of the *nifB-fdxN-nifSUHDK* operon. This is in contrast with previous data, which concluded that DevH does not act as a direct regulator of the *nif* genes because a *devH* mutant is able to fix nitrogen at low levels under anoxic conditions (150). However, it is known that mutations of regulatory genes do not always completely abolish expression of the genes that they control (48, 69). Our results make sense in light of this fact. It stands to reason that a small level of functional redundancy may exist between proteins that belong in the same family. For example, NtcA may weakly bind to the *nifB* promoter and activate transcription at levels sufficient to fix nitrogen under anoxic conditions. DevH is present in *Anabaena variabilis*, and proteins with similar sequences are present in other heterocystous cyanobacteria (Chapter III), suggesting that DevH may function as regulators of late gene expression in nitrogen fixing cyanobacteria.

Secondary effector in DevH binding

Secondary effectors are small molecules that bind to the effector domain of regulated transcription factor proteins and activate or enhance their binding to their target promoter. In many bacteria, the secondary messenger cAMP forms a complex with CRP proteins to activate binding to DNA and upregulate the expression of a large number of genes that respond to a range of physiological conditions within the cell. In *Anabaena* PCC 7120, the secondary metabolite 2-oxoglutarate (2-OG), an intermediate of the Krebs cycle (which is incomplete in cyanobacteria), has been shown to be a nitrogen status signal, which accumulates within the cells after nitrogen deprivation. 2-OG functions as an effector of NtcA, enhancing its DNA-binding activity. We speculate that a similar mechanism regulates DevH binding to its target DNA. Future experiments can identify a secondary effector that binds to DevH. It is also possible that DevH does not bind a secondary effector. We show here that in *Anabaena* PCC 7120, DevH binds to the *nifB* promoter in response to nitrogen starvation, but it remains to be determined what, if any, other factors are employed to control this activity during heterocyst development (Chapter III). One interesting possibility is that DevH and NtcA might function as a heterodimer at some promoters.

Regulation of *nifHDK* expression by SigE

In *Anabaena* PCC 7120, some sigma factors are expressed specifically in heterocysts. Among these, SigE is expressed late during heterocyst development, about 12 hours following nitrogen deprivation. The timing of *sigE* expression suggests that SigE is

involved in the expression of middle and late genes of heterocyst development, particularly the *nif* genes whose timing of expression is synchronized with that of *sigE*. Inactivation of *sigE* results in delayed and reduced levels of *nifH* transcripts in heteorocysts, indicating that SigE is important for the expression of *nif* genes (Chapter IV) (116). Time-lapse fluorescence microscopy of the P*nifHD-gfp* reporter in the wild-type background shows developmentally regulated expression of the *gfp* reporter 12 hours after nitrogen deprivation (chapter II). In the *sigE* mutant background, the *PnifHD-gfp* reporter does not express *gfp*, even at 42 hours after nitrogen deprivation (Chapter IV) (116). Overexpression of *sigE* results in early heterocyst differentiation, a higher frequency of heterocysts, and premature expression of *PnifHD-gfp* (Chapter IV) (116). These data clearly indicate that *sigE* plays a role in regulating the expression of the *nifH* gene. However, SigE is not essential for *nifHDK* expression, probably because some level of functional redundancy exists between sigma factors or because other specific transcription factors provide partial regulation of the promoter or mRNA stability.

The levels of the *nifB-fdxN-nifS-nifU* transcripts in the *sigE* inactivation mutant have not been determined. Because our data suggests that the *nifB* promoter drives the transcription of *nifHDK*, SigE most likely interacts with the *nifB* promoter. Based on these data, we predict that the expression of the *nifB-fdxN-nifS-nifU* genes, like the *nifHDK* genes, will be delayed and reduced in a *sigE* mutant. Future research that involves RNA-seq of a sigE mutant can be used to define the SigE regulon and in vitro transcription assays could be used to establish if SigE interacts directly with the *nifB*

promoter.

Regulatory targets of SigE

Because the expression of *sigE* is heterocyst-specific, we hypothesize that *sigE* controls heterocyst-specific promoters. The timing of *sigE* expression suggests that SigE plays a role in regulating late genes and possibly middle genes of heterocyst development. In addition to the *nif* genes, our results show that *sigE* is important in the normal expression of a middle and a late gene (*hepB* and *fdxH,* respectively). We predict that SigE regulates only a subset of middle and late genes because the expression of other middle and late genes (*hglE* and *hupL*, respectively) was not significantly different when compared to the wild-type (Chapter IV) (116). This suggests other pathways must exist that do not involve or are able to bypass the requirement for *sigE*. Future studies can be directed at determining if SigE interacts directly with the promoter of the *nif* genes by in vitro transcription assays. To identify the repertoire of targets that belong to the SigE regulon, RNA deep sequencing experiments could be performed to compare transcript levels of heterocyst-specific middle and late genes between wild-type and *sigE* mutant backgrounds.

Future studies

Heterocyst development in *Anabaena* PCC 7120 involves differential regulation of gene expression during cellular differentiation and morphogenesis. The use of new techniques such as RNA Deep Sequencing (53), along with classic genetic and

biochemical methods, has led to important discoveries of genes required for the differentiation of a vegetative cell to a heterocyst. However, our understanding of heterocyst development is far from complete, particularly for the middle and later stages of heterocyst development.

So far, a heterocyst-specific promoter has not been specifically identified. The results from our experiments with the *nifB* and *sigE* promoter show the presence of NtcA-like binding sites (Chapter II and IV) (116). Analysis of the promoter regions of other late genes may reveal NtcA-like binding sites. Coupled with expression profiling of a regulatory mutant, bioinformatic analysis of promoter sequences of target genes could identify consensus middle and late promoter sequence.

Our lab has determined that heterocyst development is reversible until approximately 9-12 h after nitrogen deprivation, when cells become committed to differentiate (195). RNA-seq data from cells 9-12 hours after nitrogen deprivation can be used to identify genes that are upregulated during this time period and further work could validate potential transcriptional regulators of these genes (53). I think that future work should be aimed at the identification and characterization of middle genes to understand what commits a cell to differentiate into a heterocyst.

A topic that has not been given much attention is the study of middle genes of heterocyst development in *Anabaena* PCC 7120. There is little knowledge of what genes and signals are expressed during the middle stages of heterocyst differentiation. Our data from *hgl-gfp* fusions confirms that the *hgl* genes are middle genes that are expressed between 9-12 hours after nitrogen deprivation (Appendix). It has been shown that sigma

factor SigG is also expressed during this stage of differentiation (2). Based on its timing of expression, it seems likely that SigG regulates a subset of middle genes.

Other physiological factors besides nitrogen deprivation may regulate the expression of middle and late genes, though our knowledge of this is limited. Could oxygen tension within the cell act as a checkpoint for the expression of the late genes? In other words, does the environment within the heterocyst reach a minimum oxygen concentration before some late genes, specifically the *nif* genes, are expressed? The concentration of oxygen within the heterocyst has not been precisely measured, and there is no knowledge of the levels of oxygen that would disrupt the physiological signals and events that control late gene expression. If oxygen concentration acts as a signal for *nif* gene expression, it makes sense that the cell would expend the energy needed to synthesize the structural components of oxygen-sensitive nitrogenase only after the internal environment is microoxic. The phenotype of a *devH* mutant, which is defective in the formation of the glycolipid layer, suggests that blocking the entry of oxygen may be required for transcription of the *nif* genes (150).

Our experiments with the *sigE* mutant demonstrate that SigE plays a role in the regulation of some middle and late genes, including *nifH* (116). And, our data shows that SigE does not affect expression of all middle and late genes, indicating that additional levels of control are involved in regulating these genes. Future analyses involving RNA-seq expression profiling experiments will help identify other regulators of middle and late genes. In vitro transcriptional studies could be used to validate the genomic data and to determine if interactions between putative regulators and their target genes are direct

or indirect.

Closing comments

Over the last few years, cyanobacteria have gained importance due to their potential use for producing renewable chemical products, including biofuels, various alcohols, and organic molecules such as proteins, carbohydrates, and lipids. The identification of the mechanisms controlling heterocyst middle and late gene regulation could be used to increase the efficiency of production of certain products, particularly those that require a microoxic or anaerobic environment for their production.

The identification of novel regulatory mechanisms in *Anabaena* PCC 7120 will extend the understanding of gene regulation in other cyanobacteria and prokaryotes. It is clear that *Anabaena* PCC 7120 exhibits complex multicellularity through a regulatory network that responds to external and internal cues. The study of multicellular developmental biology in *Anabaena* PCC 7120, as well as other bacterial developmental systems, has helped to uproot the common misconception of bacterial organisms as simple autonomous single cells. The research presented in this dissertation has increased our knowledge of the mechanisms by which a biological system accomplishes such amazing feats.

REFERENCES

1. **Aldea, M. R., K. Kumar, and J. W. Golden.** 2008. Heterocyst development and pattern formation, p. 75-90. *In* W. W. C. and B. L. Bassler (ed.), Chemical Communication Among Bacteria. ASM Press, Washington, DC.
2. **Aldea, M. R., R. A. Mella-Herrera, and J. W. Golden.** 2007. Sigma factor genes *sigC*, *sigE*, and *sigG* are upregulated in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **189:**8392-8396.
3. **Andrews, S. C., A. K. Robinson, and F. Rodriguez-Quinones.** 2003. Bacterial iron homeostasis. FEMS Microbiol Rev **27:**215-237.
4. **Awai, K., S. Lechno-Yossef, and C. P. Wolk.** 2009. Heterocyst envelope glycolipids, p. 179-202. *In* H. Wada and N. Murata (ed.), Lipids in photosynthesis: essential and regulatory functions. Springer, New York.
5. **Awai, K., and C. P. Wolk.** 2007. Identification of the glycosyl transferase required for synthesis of the principal glycolipid characteristic of heterocysts of *Anabaena* sp. strain PCC 7120. FEMS Microbiol. Lett. **266:**98-102.
6. **Barondeau, D. P., C. D. Putnam, C. J. Kassmann, J. A. Tainer, and E. D. Getzoff.** 2003. Mechanism and energetics of green fluorescent protein chromophore synthesis revealed by trapped intermediate structures. Proc Natl Acad Sci U S A **100:**12111-12116.
7. **Bauer, C. C., and R. Haselkorn.** 1995. Vectors for determining the differential expression of genes in heterocysts and vegetative cells of Anabaena sp. strain PCC 7120. J Bacteriol **177:**3332-3336.

8. **Bauer, C. C., K. S. Ramaswamy, S. Endley, L. A. Scappino, J. W. Golden, and R. Haselkorn.** 1997. Suppression of heterocyst differentiation in Anabaena PCC 7120 by a cosmid carrying wild-type genes encoding enzymes for fatty acid synthesis. FEMS Microbiol Lett **151:**23-30.

9. **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.

10. **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.

11. **Blair, D. F.** 1995. How bacteria sense and swim. Annu. Rev. Microbiol. **49:**489-522.

12. **Böhme, H., and R. Haselkorn.** 1988. Molecular cloning and nucleotide sequence analysis of the gene coding for heterocyst ferredoxin from the cyanobacterium Anabaena sp. strain PCC 7120. Mol Gen Genet **214:**278-285.

13. **Boison, G., H. Bothe, and O. Schmitz.** 2000. Transcriptional analysis of hydrogenase genes in the Cyanobacteria Anacystis nidulans and Anabaena variabilis monitored by RT-PCR. Curr Microbiol **40:**315-321.

14. **Borthakur, P. B., C. C. Orozco, S. S. Young-Robbins, R. Haselkorn, and S. M. Callahan.** 2005. Inactivation of *patS* and *hetN* causes lethal levels of heterocyst differentiation in the filamentous cyanobacterium *Anabaena* sp. PCC 7120. Mol. Microbiol. **57:**111-123.

15. **Bos, M. P., B. Tefsen, J. Geurtsen, and J. Tommassen.** 2004. Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. Proc. Natl. Acad. Sci. USA **101:**9417-9422.

16. **Botsford, J. L., and J. G. Harman.** 1992. Cyclic-Amp in prokaryotes. Microbiological Reviews **56:**100-122.

17. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **72:**248-254.

18. **Brahamsha, B., and R. Haselkorn.** 1992. Identification of multiple RNA polymerase sigma factor homologs in the cyanobacterium Anabaena sp. strain PCC 7120: cloning, expression, and inactivation of the sigB and sigC genes. J Bacteriol **174:**7273-7282.

19. **Brusca, J. S., C. J. Chastain, and J. W. Golden.** 1990. Expression of the *Anabaena* sp. strain PCC 7120 *xisA* gene from a heterologous promoter results in excision of the *nifD* element. J Bacteriol **172:**3925-3931.

20. **Brusca, J. S., M. A. Hale, C. D. Carrasco, and J. W. Golden.** 1989. Excision of an 11-kilobase-pair DNA element from within the *nifD* gene in *Anabaena variabilis* heterocysts. J. Bacteriol. **171:**4138-4145.

21. **Buck, M., and W. Cannon.** 1992. Activator-independent formation of a closed complex between sigma 54-holoenzyme and nifH and nifU promoters of Klebsiella pneumoniae. Mol Microbiol **6:**1625-1630.

22. **Budrene, E. O., and H. C. Berg.** 1995. Dynamics of formation of symmetrical patterns by chemotactic bacteria. Nature **376:**49-53.

23. **Buikema, W. J., and R. Haselkorn.** 1991. Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. Genes Dev. **5:**321-330.

24. **Buikema, W. J., and R. Haselkorn.** 2001. Expression of the *Anabaena hetR* gene from a copper-regulated promoter leads to heterocyst differentiation under repressing conditions. Proc. Natl. Acad. Sci. USA **98:**2729-2734.

25. **Cai, Y.** 1991. Characterization of insertion sequence IS892 and related elements from the cyanobacterium Anabaena sp. strain PCC 7120. J Bacteriol **173:**5771-5777.

26. **Callahan, S. M., and W. J. Buikema.** 2001. The role of HetN in maintenance of the heterocyst pattern in *Anabaena* sp. PCC 7120. Mol. Microbiol. **40:**941-950.

27. **Campbell, E. A., L. F. Westblade, and S. A. Darst.** 2008. Regulation of bacterial RNA polymerase sigma factor activity: a structural perspective. Curr. Opin. Microbiol. **11:**121-127.

28. **Campbell, E. L., M. F. Cohen, and J. C. Meeks.** 1997. A polyketide-synthase-like gene is involved in the synthesis of heterocyst glycolipids in *Nostoc punctiforme* strain ATCC 29133. Arch. Microbiol. **167:**251-258.

29. **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.

30. **Cardemil, L., and C. P. Wolk.** 1981. Isolated heterocysts of *Anabaena variabilis* synthesize envelope polysaccharide. Biochim. Biophys. Acta. **674:**265-276.

31. **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.

32. **Carrasco, C. D., J. A. Buettner, and J. W. Golden.** 1995. Programmed DNA rearrangement of a cyanobacterial *hupL* gene in heterocysts. Proc. Natl. Acad. Sci. USA **92:**791-795.

33. **Carrasco, C. D., and J. W. Golden.** 1995. Two heterocyst-specific DNA rearrangements of *nif* operons in *Anabaena cylindrica* and *Nostoc* sp. strain Mac. Microbiology **141 (Pt 10):**2479-2487.

34. **Carrasco, C. D., S. D. Holliday, A. Hansel, P. Lindblad, and J. W. Golden.** 2005. Heterocyst-specific excision of the *Anabaena* sp. strain PCC 7120 *hupL* element requires *xisC*. J. Bacteriol. **187:**6031-6038.

35. **Carrasco, C. D., K. S. Ramaswamy, T. S. Ramasubramanian, and J. W. Golden.** 1994. *Anabaena xisF* gene encodes a developmentally regulated site-specific recombinase. Genes Dev. **8:**74-83.

36. **Chen, H., S. Laurent, S. Bédu, F. Ziarelli, H. L. Chen, Y. Cheng, C. C. Zhang, and L. Peng.** 2006. Studying the signaling role of 2-oxoglutaric acid

using analogs that mimic the ketone and ketal forms of 2-oxoglutaric acid. Chem. Biol. **13:**849-856.

37. **Cheng, Y., J. H. Li, L. Shi, L. Wang, A. Latifi, and C. C. Zhang.** 2006. A pair of iron-responsive genes encoding protein kinases with a Ser/Thr kinase domain and a His kinase domain are regulated by NtcA in the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **188:**4822-4829.

38. **Cumino, A. C., C. Marcozzi, R. Barreiro, and G. L. Salerno.** 2007. Carbon cycling in *Anabaena* sp. PCC 7120. Sucrose synthesis in the heterocysts and possible role in nitrogen fixation. Plant Physiol. **143:**1385-1397.

39. **Daniels, R., and S. Normark.** 2008. Twin ushers guide pili across the bacterial outer membrane. Cell **133:**574-576.

40. **Dismukes, G. C., D. Carrieri, N. Bennette, G. M. Ananyev, and M. C. Posewitz.** 2008. Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. Curr. Opin. Biotechnol. **19:**235-240.

41. **Dixon, R.** 1998. The oxygen-responsive NIFL-NIFA complex: a novel two-component regulatory system controlling nitrogenase synthesis in gamma-proteobacteria. Arch Microbiol **169:**371-380.

42. **Ehira, S., and M. Ohmori.** 2006. NrrA directly regulates the expression of *hetR* during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **188:**8520-8525.

43. **Ehira, S., and M. Ohmori.** 2006. NrrA, a nitrogen-responsive response regulator facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. Mol. Microbiol. **59:**1692-1703.

44. **Ehira, S., M. Ohmori, and N. Sato.** 2003. Genome-wide expression analysis of the responses to nitrogen deprivation in the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. DNA Res. **10:**97-113.

45. **Elhai, J.** 1993. Strong and regulated promoters in the cyanobacterium *Anabaena* PCC 7120. FEMS Microbiol. Lett. **114:**179-184.

46. **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.

47. **Elhai, J., and C. P. Wolk.** 1990. Developmental regulation and spatial pattern of expression of the structural genes for nitrogenase in the cyanobacterium *Anabaena*. EMBO J. **9:**3379-3388.

48. **Eron, L., and R. Block.** 1971. Mechanism of initiation and repression of in vitro transcription of the lac operon of Escherichia coli. Proc Natl Acad Sci U S A **68:**1828-1832.

49. **Fan, Q., G. Huang, S. Lechno-Yossef, C. P. Wolk, T. Kaneko, and S. Tabata.** 2005. Clustered genes required for synthesis and deposition of envelope glycolipids in *Anabaena* sp. strain PCC 7120. Mol. Microbiol. **58:**227-243.

50. **Fay, P.** 1992. Oxygen relations of nitrogen fixation in cyanobacteria. Microbiol Rev **56:**340-373.

51. **Fiedler, G., M. Arnold, S. Hannus, and I. Maldener.** 1998. The DevBCA exporter is essential for envelope formation in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC 7120. Mol. Microbiol. **27:**1193-1202.

52. **Fiedler, G., A. M. Muro-Pastor, E. Flores, and I. Maldener.** 2001. NtcA-dependent expression of the *devBCA* operon, encoding a heterocyst-specific ATP-binding cassette transporter in *Anabaena* spp. J. Bacteriol. **183:**3795-3799.

53. **Flaherty, B. L., F. Van Nieuwerburgh, S. R. Head, and J. W. Golden.** 2011. Directional RNA deep sequencing sheds new light on the transcriptional response of Anabaena sp. strain PCC 7120 to combined-nitrogen deprivation. BMC Genomics **12:**332.

54. **Flores, E., and A. Herrero.** 2010. Compartmentalized function through cell differentiation in filamentous cyanobacteria. Nat Rev Microbiol **8:**39-50.

55. **Flores, E., A. Herrero, C. P. Wolk, and I. Maldener.** 2006. Is the periplasm continuous in filamentous multicellular cyanobacteria? Trends Microbiol. **14:**439-443.

56. **Flores, E., R. Pernil, A. M. Muro-Pastor, V. Mariscal, I. Maldener, S. Lechno-Yossef, Q. Fan, C. P. Wolk, and A. Herrero.** 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.

57. **Forchhammer, K.** 2004. Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signals to targets. FEMS Microbiol Rev **28:**319-333.

58. **Frias, J. E., E. Flores, and A. Herrero.** 1994. Requirement of the regulatory protein NtcA for the expression of nitrogen assimilation and heterocyst development genes in the cyanobacterium *Anabaena* sp. PCC 7120. Mol. Microbiol. **14:**823-832.

59. **Giddings, T. H., and L. A. Staehelin.** 1981. Observation of microplasmodesmata in both heterocyst-forming and non-heterocyst forming filamentous cyanobacteria by freeze-fracture electron microscopy. Arch. Microbiol. **129:**295-298.

60. **Giddings, T. H., and L. A. Staehelin.** 1978. Plasma membrane architecture of *Anabaena cylindrica*: occurrence of microplasmodesmata and changes associated with heterocyst development and the cell cycle. Eur. J. Cell. Biol. **16:**235-249.

61. **Golden, J. W.** 1997. Programmed DNA rearrangements in cyanobacteria, p. 162-173. *In* F. J. de Bruijn, J. R. Lupski, and G. Weinstock (ed.), Bacterial genomes: physical structure and analysis. Chapman & Hall, New York.

62. **Golden, J. W., C. D. Carrasco, M. E. Mulligan, G. J. Schneider, and R. Haselkorn.** 1988. Deletion of a 55-kilobase-pair DNA element from the chromosome during heterocyst differentiation of *Anabaena* sp. strain PCC 7120. J. Bacteriol. **170:**5034-5041.

63. **Golden, J. W., S. J. Robinson, and R. Haselkorn.** 1985. Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium *Anabaena*. Nature **314:**419-423.

64. **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.

65. **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.

66. **Golden, J. W., and H. S. Yoon.** 2003. Heterocyst development in *Anabaena*. Current opinion in microbiology **6:**557-563.

67. **Golden, J. W., and H. S. Yoon.** 1998. Heterocyst formation in *Anabaena*. Current opinion in microbiology **1:**623-629.

68. **Gruber, T. M., and C. A. Gross.** 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. Annu Rev Microbiol **57:**441-466.

69. **Guidi-Rontani, C., and A. Spassky.** 1985. RNA polymerase mutant able to express in vivo and in vitro the lactose operon in the absence of the cAMP-CRP complex. J Mol Biol **186:**527-532.

70. **Hansen, M. C., R. J. Palmer, Jr., C. Udsen, D. C. White, and S. Molin.** 2001. Assessment of GFP fluorescence in cells of Streptococcus gordonii under conditions of low pH and low oxygen concentration. Microbiology **147:**1383-1391.

71. **Hantke, K.** 2001. Iron and metal regulation in bacteria. Current opinion in microbiology **4:**172-177.

72. **Haselkorn, R., D. Rice, S. E. Curtis, and S. J. Robinson.** 1983. Organization and transcription of genes important in *Anabaena* heterocyst differentiation. Annales de microbiologie **134B:**181-193.

73. **Hebbar, P. B., and S. E. Curtis.** 2000. Characterization of *devH*, a gene encoding a putative DNA binding protein required for heterocyst function in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **182:**3572-3581.

74. **Henson, B. J., L. E. Pennington, L. E. Watson, and S. R. Barnum.** 2008. Excision of the *nifD* element in the heterocystous cyanobacteria. Arch. Microbiol. **189:**357-366.

75. **Hernandez, J. A., S. Lopez-Gomollon, M. T. Bes, M. F. Fillat, and M. L. Peleato.** 2004. Three fur homologues from Anabaena sp. PCC7120: exploring reciprocal protein-promoter recognition. FEMS Microbiol Lett **236:**275-282.

76. **Hernández, J. A., S. Lopez-Gomollón, A. Muro-Pastor, A. Valladares, M. T. Bes, M. L. Peleato, and M. F. Fillat.** 2006. Interaction of FurA from Anabaena sp. PCC 7120 with DNA: a reducing environment and the presence of Mn(2+) are positive effectors in the binding to isiB and furA promoters. Biometals **19:**259-268.

77. **Hernández, J. A., A. M. Muro-Pastor, E. Flores, M. T. Bes, M. L. Peleato, and M. F. Fillat.** 2006. Identification of a *furA* cis antisense RNA in the cyanobacterium Anabaena sp. PCC 7120. J Mol Biol **355:**325-334.

78. **Herrero, A., A. M. Muro-Pastor, and E. Flores.** 2001. Nitrogen control in cyanobacteria. J Bacteriol **183:**411-425.

79. **Herrero, A., A. M. Muro-Pastor, A. Valladares, and E. Flores.** 2004. Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. FEMS Microbiol Rev **28:**469-487.

80. **Hoiczyk, E., and A. Hansel.** 2000. Cyanobacterial cell walls: news from an unusual prokaryotic envelope. J. Bacteriol. **182:**1191-1199.

81. **Holmqvist, M., K. Stensjö, P. Oliveira, P. Lindberg, and P. Lindblad.** 2009. Characterization of the *hupSL* promoter activity in *Nostoc punctiforme* ATCC 29133. BMC Microbiol. **9:**54.

82. **Huang, G., Q. Fan, S. Lechno-Yossef, E. Wojciuch, C. P. Wolk, T. Kaneko, and S. Tabata.** 2005. Clustered genes required for the synthesis of heterocyst envelope polysaccharide in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **187:**1114-1123.

83. **Huang, X., Y. Dong, and J. Zhao.** 2004. HetR homodimer is a DNA-binding protein required for heterocyst differentiation, and the DNA-binding activity is inhibited by PatS. Proc. Natl. Acad. Sci. USA **101:**4848-4853.

84. **Imamura, S., and M. Asayama.** 2009. Sigma factors for cyanobacterial transcription. Gene Regul Syst Bio **3:**65-87.

85. **Jackman, D. M., and M. E. Mulligan.** 1995. Characterization of a nitrogen-fixation (nif) gene cluster from Anabaena azollae 1a shows that closely related cyanobacteria have highly variable but structured intergenic regions. Microbiology **141 (Pt 9):**2235-2244.

86. **Jang, J., L. Wang, R. Jeanjean, and C. C. Zhang.** 2007. PrpJ, a PP2C-type protein phosphatase located on the plasma membrane, is involved in heterocyst maturation in the cyanobacterium *Anabaena* sp. PCC 7120. Mol. Microbiol. **64:**347-358.

87. **Jelsbak, L., and L. Søgaard-Andersen.** 2002. Pattern formation by a cell surface-associated morphogen in *Myxococcus xanthus*. Proc. Natl. Acad. Sci. USA **99:**2032-2037.

88. **Jiang, F., S. Wisén, M. Widersten, B. Bergman, and B. Mannervik.** 2000. Examination of the transcription factor NtcA-binding motif by in vitro selection of DNA sequences from a random library. J Mol Biol **301:**783-793.

89. **Jones, K. M., and R. Haselkorn.** 2002. Newly identified cytochrome c oxidase operon in the nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120 specifically induced in heterocysts. J. Bacteriol. **184:**2491-2499.

90. **Kaiser, D.** 2001. Building a multicellular organism. Annu Rev Genet **35:**103-123.

91. **Ke, S. H., and E. L. Madison.** 1997. Rapid and efficient site-directed mutagenesis by single-tube 'megaprimer' PCR method. Nucleic Acids Res **25:**3371-3372.

92. **Kennell, D.** 2002. Processing endoribonucleases and mRNA degradation in bacteria. J Bacteriol **184:**4645-4657; discussion 4665.

93. **Khudyakov, I., and C. P. Wolk.** 1997. *hetC*, a gene coding for a protein similar to bacterial ABC protein exporters, is involved in early regulation of heterocyst differentiation in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **179:**6971-6978.

94. **Khudyakov, I. Y., and J. W. Golden.** 2004. Different functions of HetR, a master regulator of heterocyst differentiation in *Anabaena* sp. PCC 7120 can be separated by mutation. Proc. Natl. Acad. Sci. USA **101:**16040-16045.

95. **Khudyakov, I. Y., and J. W. Golden.** 2001. Identification and inactivation of three group 2 sigma factor genes in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **183:**6667-6675.

96. **Klausen, M., A. Heydorn, P. Ragas, L. Lambertsen, A. Aaes-Jørgensen, S. Molin, and T. Tolker-Nielsen.** 2003. Biofilm formation by *Pseudomonas aeruginosa* wild type, flagella and type IV pili mutants. Mol. Microbiol. **48:**1511-1524.

97. **Kolb, A., S. Busby, H. Buc, S. Garges, and S. Adhya.** 1993. Transcriptional regulation by Camp and its receptor protein. Annual Review of Biochemistry **62:**749-795.

98. **Korner, H., H. J. Sofia, and W. G. Zumft.** 2003. Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. FEMS Microbiol Rev **27:**559-592.

99. **Kroos, L.** 2007. The *Bacillus* and *Myxococcus* developmental networks and their transcriptional regulators. Annu. Rev. Genet. **41:**13-39.

100. **Kumar, K., R. A. Mella-Herrera, and J. W. Golden.** 2010. Cyanobacterial heterocysts. Cold Spring Harb Perspect Biol **2:**a000315.

101. **Kushner, S. R.** 2002. mRNA decay in Escherichia coli comes of age. J Bacteriol **184:**4658-4665; discussion 4657.

102. **Lammers, P. J., S. McLaughlin, S. Papin, C. Trujillo-Provencio, and A. J. Ryncarz, 2nd.** 1990. Developmental rearrangement of cyanobacterial *nif* genes: nucleotide sequence, open reading frames, and cytochrome P-450 homology of the *Anabaena* sp. strain PCC 7120 *nifD* element. J. Bacteriol. **172:**6981-6990.

103. **Laurent, S., H. Chen, S. Bedu, F. Ziarelli, L. Peng, and C. C. Zhang.** 2005. Nonmetabolizable analogue of 2-oxoglutarate elicits heterocyst differentiation under repressive conditions in *Anabaena* sp. PCC 7120. Proc. Natl. Acad. Sci. USA **102:**9907-9912.

104. **Lechno-Yossef, S., Q. Fan, S. Ehira, N. Sato, and C. P. Wolk.** 2006. Mutations in four regulatory genes have interrelated effects on heterocyst maturation in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **188:**7387-7395.

105. **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.

106. **Liu, D., and J. W. Golden.** 2002. *hetL* overexpression stimulates heterocyst formation in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **184:**6873-6881.

107. **Lopez-Gomollon, S., J. A. Hernandez, S. Pellicer, V. E. Angarica, M. L. Peleato, and M. F. Fillat.** 2007. Cross-talk between iron and nitrogen regulatory networks in anabaena (Nostoc) sp. PCC 7120: identification of overlapping genes in FurA and NtcA regulons. J Mol Biol **374:**267-281.

108. **Lopez-Gomollon, S., J. A. Hernandez, C. P. Wolk, M. L. Peleato, and M. F. Fillat.** 2007. Expression of furA is modulated by NtcA and strongly enhanced in heterocysts of Anabaena sp. PCC 7120. Microbiology **153:**42-50.

109. **Marcozzi, C., A. C. Cumino, and G. L. Salerno.** 2009. Role of NtcA, a cyanobacterial global nitrogen regulator, in the regulation of sucrose metabolism gene expression in *Anabaena* sp. PCC 7120. Arch. Microbiol. **191:**255-263.

110. **Mariscal, V., A. Herrero, and E. Flores.** 2007. Continuous periplasm in a filamentous, heterocyst-forming cyanobacterium. Mol Microbiol **65:**1139-1145.

111. **Martin-Figueroa, E., F. Navarro, and F. J. Florencio.** 2000. The GS-GOGAT pathway is not operative in the heterocysts. Cloning and expression of *glsF* gene from the cyanobacterium *Anabaena* sp. PCC 7120. FEBS Lett. **476:**282-286.

112. **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.

113. **Meeks, J. C., E. L. Campbell, and P. S. Bisen.** 1994. Elements interrupting nitrogen-fixation genes incCyanobacteria - presence and absence of a *nifD* element in clones of Nostoc sp strain. Mac. Microbiol-Uk **140:**3225-3232.

114. **Meeks, J. C., E. L. Campbell, M. L. Summers, and F. C. Wong.** 2002. Cellular differentiation in the cyanobacterium *Nostoc punctiforme*. Arch. Microbiol. **178:**395-403.

115. **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.

116. **Mella-Herrera, R. A., M. R. Neunuebel, K. Kumar, S. K. Saha, and J. W. Golden.** 2011. The sigE gene is required for normal expression of heterocyst-specific genes in Anabaena sp. strain PCC 7120. J Bacteriol **193:**1823-1832.

117. **Mevarech, M., D. Rice, and R. Haselkorn.** 1980. Nucleotide sequence of a cyanobacterial nifH gene coding for nitrogenase reductase. Proc Natl Acad Sci U S A **77:**6476-6480.

118. **Montoya, J. P., C. M. Holl, J. P. Zehr, A. Hansen, T. A. Villareal, and D. G. Capone.** 2004. High rates of N2 fixation by unicellular diazotrophs in the oligotrophic Pacific Ocean. Nature **430:**1027-1032.

119. **Moslavac, S., R. Bredemeier, O. Mirus, B. Granvogl, L. A. Eichacker, and E. Schleiff.** 2005. Proteomic analysis of the outer membrane of *Anabaena* sp. strain PCC 7120. J. Proteome Res. **4:**1330-1338.

120. **Moslavac, S., K. Nicolaisen, O. Mirus, F. Al Dehni, R. Pernil, E. Flores, I. Maldener, and E. Schleiff.** 2007. A TolC-like protein is required for heterocyst development in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **189:**7887-7895.

121. **Moslavac, S., V. Reisinger, M. Berg, O. Mirus, O. Vosyka, M. Ploscher, E. Flores, L. A. Eichacker, and E. Schleiff.** 2007. The proteome of the heterocyst cell wall in *Anabaena* sp. PCC 7120. Biological chemistry **388:**823-829.

122. **Mulligan, M. E., and R. Haselkorn.** 1989. Nitrogen fixation (*nif*) genes of the cyanobacterium *Anabaena* species strain PCC 7120. The *nifB-fdxN-nifS-nifU* operon. J. Biol. Chem. **264:**19200-19207.

123. **Mullineaux, C. W., V. Mariscal, A. Nenninger, H. Khanum, A. Herrero, E. Flores, and D. G. Adams.** 2008. Mechanism of intercellular molecular exchange in heterocyst-forming cyanobacteria. EMBO J. **27:**1299-1308.

124. **Muro-Pastor, A. M., E. Flores, and A. Herrero.** 2009. NtcA-regulated heterocyst differentiation genes hetC and devB from Anabaena sp. strain PCC 7120 exhibit a similar tandem promoter arrangement. J Bacteriol **191:**5765-5774.

125. **Muro-Pastor, A. M., A. Herrero, and E. Flores.** 2001. Nitrogen-regulated group 2 sigma factor from Synechocystis sp. strain PCC 6803 involved in survival under nitrogen stress. J Bacteriol **183:**1090-1095.

126. **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.

127. **Muro-Pastor, A. M., A. Valladares, E. Flores, and A. Herrero.** 1999. The *hetC* gene is a direct target of the NtcA transcriptional regulator in cyanobacterial heterocyst development. J. Bacteriol. **181:**6664-6669.

128. **Muro-Pastor, A. M., A. Valladares, E. Flores, and A. Herrero.** 2002. Mutual dependence of the expression of the cell differentiation regulatory protein HetR and the global nitrogen regulator NtcA during heterocyst development. Mol Microbiol **44:**1377-1385.

129. **Muro-Pastor, M. I., J. C. Reyes, and F. J. Florencio.** 2005. Ammonium assimilation in cyanobacteria. Photosynth. Res. **83:**135-150.

130. **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.

131. **Neunuebel, M. R., and J. W. Golden.** 2008. The *Anabaena* sp. strain PCC 7120 gene all2874 encodes a diguanylate cyclase and is required for normal heterocyst development under high-light growth conditions. J. Bacteriol. **190:**6829-6836.

132. **Newbury, S. F., N. H. Smith, and C. F. Higgins.** 1987. Differential mRNA stability controls relative gene expression within a polycistronic operon. Cell **51:**1131-1143.

133. **Ni, S., G. M. Sheldrick, M. M. Benning, and M. A. Kennedy.** 2009. The 2Å resolution crystal structure of HetL, a pentapeptide repeat protein involved in regulation of heterocyst differentiation in the cyanobacterium *Nostoc* sp. strain PCC 7120. J. Struct. Biol. **165:**47-52.

134. **Nicolaisen, K., A. Hahn, and E. Schleiff.** 2009. The cell wall in heterocyst formation by *Anabaena* sp. PCC 7120. Journal of basic microbiology **49:**5-24.

135. **Nicolaisen, K., S. Moslavac, A. Samborski, M. Valdebenito, K. Hantke, I. Maldener, A. M. Muro-Pastor, E. Flores, and E. Schleiff.** 2008. Alr0397 is an outer membrane transporter for the siderophore schizokinen in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **190:**7500-7507.

136. **Ohmori, M., M. Ikeuchi, N. Sato, P. Wolk, T. Kaneko, T. Ogawa, M. Kanehisa, S. Goto, S. Kawashima, S. Okamoto, H. Yoshimura, H. Katoh, T. Fujisawa, S. Ehira, A. Kamei, S. Yoshihara, R. Narikawa, and S. Tabat.** 2001. Characterization of genes encoding multi-domain proteins in the genome of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120. DNA Res. **8:**271-284.

137. **Olmedo-Verd, E., A. M. Muro-Pastor, E. Flores, and A. Herrero.** 2006. Localized induction of the ntcA regulatory gene in developing heterocysts of Anabaena sp. strain PCC 7120. J Bacteriol **188:**6694-6699.

138. **Orozco, C. C., D. D. Risser, and S. M. Callahan.** 2006. Epistasis analysis of four genes from *Anabaena* sp. strain PCC 7120 suggests a connection between PatA and PatS in heterocyst pattern formation. J. Bacteriol. **188:**1808-1816.

139. **Osanai, T., M. Ikeuchi, and K. Tanaka.** 2008. Group 2 sigma factors in cyanobacteria. Physiologia plantarum **133:**490-506.

140. **Osanai, T., M. Imashimizu, A. Seki, S. Sato, S. Tabata, S. Imamura, M. Asayama, M. Ikeuchi, and K. Tanaka.** 2009. ChlH, the H subunit of the Mg-chelatase, is an anti-sigma factor for SigE in Synechocystis sp. PCC 6803. Proc Natl Acad Sci U S A **106:**6860-6865.

141. **Osanai, T., Y. Kanesaki, T. Nakano, H. Takahashi, M. Asayama, M. Shirai, M. Kanehisa, I. Suzuki, N. Murata, and K. Tanaka.** 2005. Positive regulation of sugar catabolic pathways in the cyanobacterium Synechocystis sp. PCC 6803 by the group 2 sigma factor sigE. J Biol Chem **280:**30653-30659.

142. **Pereira, S., A. Zille, E. Micheletti, P. Moradas-Ferreira, R. De Philippis, and P. Tamagnini.** 2009. Complexity of cyanobacterial exopolysaccharides: composition, structures, inducing factors and putative genes involved in their biosynthesis and assembly. FEMS Microbiol. Rev. **33:**917-941.

143. **Phalip, V., J. H. Li, and C. C. Zhang.** 2001. HstK, a cyanobacterial protein with both a serine/threonine kinase domain and a histidine kinase domain: implication for the mechanism of signal transduction. Biochem J **360:**639-644.

144. **Popa, R., P. K. Weber, J. Pett-Ridge, J. A. Finzi, S. J. Fallon, I. D. Hutcheon, K. H. Nealson, and D. G. Capone.** 2007. Carbon and nitrogen fixation and metabolite exchange in and between individual cells of *Anabaena oscillarioides*. ISME J. **1:**354-360.

145. **Potvin, E., F. Sanschagrin, and R. C. Levesque.** 2008. Sigma factors in Pseudomonas aeruginosa. FEMS Microbiol Rev **32:**38-55.

146. **Pratte, B. S., K. Eplin, and T. Thiel.** 2006. Cross-functionality of nitrogenase components NifH1 and VnfH in *Anabaena variabilis*. J. Bacteriol. **188:**5806-5811.

147. **Ramasubramanian, T. S., T. F. Wei, and J. W. Golden.** 1994. Two *Anabaena* sp. strain PCC 7120 DNA-binding factors interact with vegetative cell- and heterocyst-specific genes. J. Bacteriol. **176:**1214-1223.

148. **Ramasubramanian, T. S., T. F. Wei, A. K. Oldham, and J. W. Golden.** 1996. Transcription of the *Anabaena* sp. strain PCC 7120 *ntcA* gene: multiple transcripts and NtcA binding. J. Bacteriol. **178:**922-926.

149. **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:**1241-1249.

150. **Ramirez, M. E., P. B. Hebbar, R. Zhou, C. P. Wolk, and S. E. Curtis.** 2005. *Anabaena* sp. strain PCC 7120 gene *devH* is required for synthesis of the heterocyst glycolipid layer. J. Bacteriol. **187:**2326-2331.

151. **Rasmussen, P. B., B. Holst, and P. Valentin-Hansen.** 1996. Dual-function regulators: the cAMP receptor protein and the CytR regulator can act either to repress or to activate transcription depending on the context. Proc Natl Acad Sci U S A **93:**10151-10155.

152. **Razquin, P., S. Schmitz, M. F. Fillat, M. L. Peleato, and H. Böhme.** 1994. Transcriptional and translational analysis of ferredoxin and flavodoxin under iron and nitrogen stress in Anabaena sp. strain PCC 7120. J Bacteriol **176:**7409-7411.

153. **Risser, D. D., and S. M. Callahan.** 2008. HetF and PatA control levels of HetR in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **190:**7645-7654.

154. **Rubio, L. M., and P. W. Ludden.** 2008. Biosynthesis of the iron-molybdenum cofactor of nitrogenase. Annu. Rev. Microbiol. **62:**93-111.

155. **Sambrook J. and Russell, D. W.** 2006. Northern hybridization

. Cold Spring Harb Protoc. **2006 doi:10.1101/pdb.prot3723**.

156. **Sato, N., M. Ohmori, M. Ikeuchi, K. Tashiro, C. P. Wolk, T. Kaneko, K. Okada, M. Tsuzuki, S. Ehira, H. Katoh, S. Okamoto, H. Yoshimura, T. Fujisawa, A. Kamei, S. Yoshihara, R. Narikawa, T. Hamano, S. Tabata, and S. Kuhara.** 2004. Use of segment-based microarray in the analysis of global gene expression in response to various environmental stresses in the cyanobacterium *Anabaena* sp. PCC 7120. J. Gen. Appl. Microbiol. **50:**1-8.

157. **Scarlato, V., I. Delany, and R. Rappuoli.** 2004. Fur functions as an activator and as a repressor of putative virulence genes in Neisseria meningitidis. Mol Microbiol **52:**1081-1090.

158. **Schmitz, R. A., K. Klopprogge, and R. Grabbe.** 2002. Regulation of nitrogen fixation in Klebsiella pneumoniae and Azotobacter vinelandii: NifL, transducing two environmental signals to the nif transcriptional activator NifA. J Mol Microbiol Biotechnol **4:**235-242.

159. **Shi, L., J. H. Li, Y. Cheng, L. Wang, W. L. Chen, and C. C. Zhang.** 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.

160. **Shi, Y., W. Zhao, W. Zhang, Z. Ye, and J. Zhao.** 2006. Regulation of intracellular free calcium concentration during heterocyst differentiation by HetR and NtcA in *Anabaena* sp. PCC 7120. Proc. Natl. Acad. Sci. USA **103:**11334-11339.

161. **Stricker, O., B. Masepohl, W. Klipp, and H. Böhme.** 1997. Identification and characterization of the *nifV-nifZ-nifT* gene region from the filamentous cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **179:**2930-2937.

162. **Studholme, D. J., M. Buck, and T. Nixon.** 2000. Identification of potential sigma(N)-dependent promoters in bacterial genomes. Microbiology **146 Pt 12:**3021-3023.

163. **Su, Z., V. Olman, F. Mao, and Y. Xu.** 2005. Comparative genomics analysis of NtcA regulons in cyanobacteria: regulation of nitrogen assimilation and its coupling to photosynthesis. Nucleic Acids Res. **33:**5156-5171.

164. **Suzuki, T., H. Yoshimura, S. Ehira, M. Ikeuchi, and M. Ohmori.** 2007. AnCrpA, a cAMP receptor protein, regulates nif-related gene expression in the cyanobacterium Anabaena sp. strain PCC 7120 grown with nitrate. FEBS Lett **581:**21-28.

165. **Tamagnini, P., E. Leitão, P. Oliveira, D. Ferreira, F. Pinto, D. J. Harris, T. Heidorn, and P. Lindblad.** 2007. Cyanobacterial hydrogenases: diversity, regulation and applications. FEMS Microbiol. Rev. **31:**692-720.

166. **Tanigawa, R., M. Shirokane, S. Maeda Si, T. Omata, K. Tanaka, and H. Takahashi.** 2002. Transcriptional activation of NtcA-dependent promoters of

Synechococcus sp. PCC 7942 by 2-oxoglutarate in vitro. Proc. Natl. Acad. Sci. USA **99:**4251-4255.

167. **Thiel, T., E. M. Lyons, J. C. Erker, and A. Ernst.** 1995. A second nitrogenase in vegetative cells of a heterocyst-forming cyanobacterium. Proc. Natl. Acad. Sci. USA **92:**9358-9362.

168. **Thiel, T., and B. Pratte.** 2001. Effect on heterocyst differentiation of nitrogen fixation in vegetative cells of the cyanobacterium *Anabaena variabilis* ATCC 29413. J. Bacteriol. **183:**280-286.

169. **Toepel, J., E. Welsh, T. C. Summerfield, H. B. Pakrasi, and L. A. Sherman.** 2008. Differential transcriptional analysis of the cyanobacterium *Cyanothece* sp. strain ATCC 51142 during light-dark and continuous-light growth. J. Bacteriol. **190:**3904-3913.

170. **Torrecilla, I., F. Leganés, I. Bonilla, and F. Fernández-Piñas.** 2004. A calcium signal is involved in heterocyst differentiation in the cyanobacterium *Anabaena* sp. PCC7120. Microbiology **150:**3731-3739.

171. **Truffa-Bachi, P., and G. N. Cohen.** 1973. Amino acid metabolism. Annu Rev Biochem **42:**113-134.

172. **Ungerer, J. L., B. S. Pratte, and T. Thiel.** 2008. Regulation of fructose transport and its effect on fructose toxicity in Anabaena spp. J Bacteriol **190:**8115-8125.

173. **Ungerer, J. L., B. S. Pratte, and T. Thiel.** 2010. RNA processing of nitrogenase transcripts in the cyanobacterium Anabaena variabilis. J Bacteriol **192:**3311-3320.

174. **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.

175. **Valladares, A., I. Maldener, A. M. Muro-Pastor, E. Flores, and A. Herrero.** 2007. Heterocyst development and diazotrophic metabolism in terminal respiratory oxidase mutants of the cyanobacterium Anabaena sp. strain PCC 7120. J Bacteriol **189:**4425-4430.

176. **Valladares, A., A. M. Muro-Pastor, M. F. Fillat, A. Herrero, and E. Flores.** 1999. Constitutive and nitrogen-regulated promoters of the petH gene encoding ferredoxin:NADP+ reductase in the heterocyst-forming cyanobacterium Anabaena sp. FEBS Lett **449:**159-164.

177. **Valladares, A., A. M. Muro-Pastor, A. Herrero, and E. Flores.** 2004. The NtcA-dependent P1 promoter is utilized for *glnA* expression in N_2-fixing heterocysts of *Anabaena* sp. strain PCC 7120. J. Bacteriol. **186:**7337-7343.

178. **Vazquez-Bermudez, M. F., A. Herrero, and E. Flores.** 2002. 2-Oxoglutarate increases the binding affinity of the NtcA (nitrogen control) transcription factor for the Synechococcus glnA promoter. FEBS Lett **512:**71-74.

179. **Vazquez-Bermúdez, M. F., A. Herrero, and E. Flores.** 2003. Carbon supply and 2-oxoglutarate effects on expression of nitrate reductase and nitrogen-

regulated genes in *Synechococcus* sp. strain PCC 7942. FEMS Microbiol. Lett. **221:**155-159.

180. **Wang, L., Y. P. Sun, W. L. Chen, J. H. Li, and C. C. Zhang.** 2002. Genomic analysis of protein kinases, protein phosphatases and two-component regulatory systems of the cyanobacterium *Anabaena* sp. strain PCC 7120. FEMS Microbiol. Lett. **217:**155-165.

181. **Wang, Y., S. Lechno-Yossef, Y. Gong, Q. Fan, C. P. Wolk, and X. Xu.** 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.

182. **Wang, Y., and X. Xu.** 2005. Regulation by *hetC* of genes required for heterocyst differentiation and cell division in *Anabaena* sp. strain PCC 7120. J. Bacteriol. **187:**8489-8493.

183. **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.

184. **Weyman, P. D., B. Pratte, and T. Thiel.** 2008. Transcription of *hupSL* in *Anabaena variabilis* ATCC 29413 is regulated by NtcA and not by hydrogen. Appl. Environ. Microbiol. **74:**2103-2110.

185. **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.

186. **Wolk, C. P.** 1996. Heterocyst formation. Annu. Rev. Genet. **30:**59-78.

187. **Wolk, C. P., Y. Cai, L. Cardemil, E. Flores, B. Hohn, M. Murry, G. Schmetterer, B. Schrautemeier, and R. Wilson.** 1988. Isolation and complementation of mutants of *Anabaena* sp. strain PCC 7120 unable to grow aerobically on dinitrogen. J. Bacteriol. **170:**1239-1244.

188. **Wolk, C. P., A. Ernst, and J. Elhai.** 1994. Heterocyst metabolism and development, p. 769-863. *In* D. A. Bryant (ed.), The molecular biology of cyanobacteria, vol. 1. Kluwer Academic Publishers, Dordrecht.

189. **Wong, F. C., and J. C. Meeks.** 2001. The *hetF* gene product is essential to heterocyst differentiation and affects HetR function in the cyanobacterium *Nostoc punctiforme*. J. Bacteriol. **183:**2654-2661.

190. **Wu, X., D. W. Lee, R. A. Mella, and J. W. Golden.** 2007. The *Anabaena* sp. strain PCC 7120 asr1734 gene encodes a negative regulator of heterocyst development. Mol. Microbiol. **64:**782-794.

191. **Wu, X., D. Liu, M. H. Lee, and J. W. Golden.** 2004. *patS* minigenes inhibit heterocyst development of *Anabaena* sp. strain PCC 7120. J. Bacteriol. **186:**6422-6429.

192. **Xu, X., and C. P. Wolk.** 2001. Role for *hetC* in the transition to a nondividing state during heterocyst differentiation in *Anabaena* sp. J. Bacteriol. **183:**393-396.

193. **Yanofsky, C., T. Platt, I. P. Crawford, B. P. Nichols, G. E. Christie, H. Horowitz, M. VanCleemput, and A. M. Wu.** 1981. The complete nucleotide

sequence of the tryptophan operon of Escherichia coli. Nucleic Acids Res **9:**6647-6668.

194. **Yoon, H. S., and J. W. Golden.** 1998. Heterocyst pattern formation controlled by a diffusible peptide. Science **282:**935-938.

195. **Yoon, H. S., and J. W. Golden.** 2001. PatS and products of nitrogen fixation control heterocyst pattern. J. Bacteriol. **183:**2605-2613.

196. **Yoshimura, H., S. Okamoto, Y. Tsumuraya, and M. Ohmori.** 2007. Group 3 sigma factor gene, *sigJ*, a key regulator of desiccation tolerance, regulates the synthesis of extracellular polysaccharide in cyanobacterium *Anabaena* sp. strain PCC 7120. DNA Res. **14:**13-24.

197. **Zhang, C., X. H. Xing, and K. Lou.** 2005. Rapid detection of a gfp-marked Enterobacter aerogenes under anaerobic conditions by aerobic fluorescence recovery. FEMS Microbiol Lett **249:**211-218.

198. **Zhang, C. C., S. Laurent, S. Sakr, L. Peng, and S. Bédu.** 2006. Heterocyst differentiation and pattern formation in cyanobacteria: a chorus of signals. Mol. Microbiol. **59:**367-375.

199. **Zhang, J. Y., W. L. Chen, and C. C. Zhang.** 2009. *hetR* and *patS*, two genes necessary for heterocyst pattern formation, are widespread in filamentous nonheterocyst-forming cyanobacteria. Microbiology **155:**1418-1426.

200. **Zhang, L. C., Y. F. Chen, W. L. Chen, and C. C. Zhang.** 2008. Existence of periplasmic barriers preventing green fluorescent protein diffusion from cell to

cell in the cyanobacterium *Anabaena* sp. strain PCC 7120. Mol. Microbiol. **70:**814-823.

201. **Zhang, W., Y. Du, I. Khudyakov, Q. Fan, H. Gao, D. Ning, C. P. Wolk, and X. Xu.** 2007. A gene cluster that regulates both heterocyst differentiation and pattern formation in *Anabaena* sp. strain PCC 7120. Mol. Microbiol. **66:**1429-1443.

202. **Zhang, Y., and M. A. Frohman.** 2000. Using Rapid Amplification of cDNA Ends (RACE) to Obtain Full-Length cDNAs. The Nucleic Acid Protocols Handbook **V:**267-288.

203. **Zhao, J., and C. P. Wolk.** 2008. Developmental Biology of Heterocysts, 2006, p. 397-418. *In* D. E. Whitworth (ed.), Myxobacteria: Multicellularity and Differentiation. ASM Press, Washington, D. C.

204. **Zhao, Y., Y. Shi, W. Zhao, X. Huang, D. Wang, N. Brown, J. Brand, and J. Zhao.** 2005. CcbP, a calcium-binding protein from *Anabaena* sp. PCC 7120, provides evidence that calcium ions regulate heterocyst differentiation. Proc. Natl. Acad. Sci. USA **102:**5744-5748.

205. **Zhou, R., X. Wei, N. Jiang, H. Li, Y. Dong, K. L. Hsi, and J. Zhao.** 1998. Evidence that HetR protein is an unusual serine-type protease. Proc. Natl. Acad. Sci USA **95:**4959-4963.

206. **Zhou, R., and C. P. Wolk.** 2003. A two-component system mediates developmental regulation of biosynthesis of a heterocyst polysaccharide. J. Biol. Chem. **278:**19939-19946.

207. **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.

208. **Zhu, J., R. Kong, and C. P. Wolk.** 1998. Regulation of hepA of Anabaena sp. strain PCC 7120 by elements 5' from the gene and by hepK. J Bacteriol **180:**4233-4242.

209. **Zuker, M.** 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res **31:**3406-3415.

210. **Zumft, W. G., H. Korner, and H. J. Sofia.** 2003. Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. Fems Microbiology Reviews **27:**559-592.

APPENDIX

Section 1

***nifH-gfp* reporter constructions on shuttle plasmids.** Heterocyst differentiation culminates with the expression of nif genes and the resulting nitrogen-fixation activity. As previously stated, little is known about the control of *nif* gene expression in cyanobacteria. The absence of *rpoN*, encoding sigma-54, and *nifA* means that heterocysts must use a novel mechanism for regulating the nif genes compared to other well-studied bacteria.

To understand the transcriptional regulation of the *nifHDK* genes, we initially attempted to make transcriptional fusions of the upstream region of *nifH* to *gfp* on a shuttle plasmid pAM1956. All of our shuttle-vector constructions were confirmed by DNA sequencing. Our first strain, KKC5, contained the intergenic region (221 bp) between *nifU* and *nifH* fused to *gfp* on pAM1956 (pAMKK5). We expected the *nifHDK* promoter to be in this region, but KKC5 did not show P*nifH*-driven *gfp* expression in heterocysts when transferred to media without a source of combined nitrogen. Strains KKC6 and KKC25, which contained 350 bp and 500 bp of *nifH* upstream sequence fused to *gfp* on pAM1956 (pAMKK6 and pAMKK25), also showed no expression of P*nifH*-*gfp* in heterocysts when deprived for nitrogen.

Because GFP fluorescence requires activation by oxygen (6), and the *nif* promoter is activated late during heterocyst development when the environment within the heterocyst is microoxic, we hypothesized that the absence of GFP in the above strains

was attributable to low oxygen concentration within the cell. To test this hypothesis, we grew strains KKC5, KKC6 and KKC25 under anaerobic conditions; under such conditions the *hgl* layer of heterocysts is not well formed and oxygen can enter the cell. Following anaerobic growth, we performed aerobic fluorescence recovery(197) by exposing the cells to air for 30 minutes. Strains KKC5, KKC6 and KKC25 still did not show GFP fluorescence following exposure to air, indicating that oxygen was not the limiting factor in observing GFP fluorescence in these strains. Rather, no GFP protein is made in these strains. Our data show that *gfp* can successfully be used as a reporter for the study of heterocyst-specific promoters: several *gfp* reporter strains show fluorescence in mature heterocysts, including constructs from our lab containing the promoters: *patS* (195), *hetR* (189), *hglE* (below), and the 1000-bp *nifH* fragment (below), among others. To confirm the absence of GFP in KKC5, KKC6 and KKC25, a western immunoblot experiment was performed. No GFP protein was detected (data not shown), which is consistent with no promoter activity from these constructs. Our results suggest that the *nif* promoter does not lie within the 500-bp upstream region of the *nifH* gene.

To identify the upstream region presumably containing the *nifH* promoter, we constructed strain KKC7 containing shuttle plasmid pAMKK7, in which 1-kb of the *nifH* upstream region was transcriptionally fused to *gfp* on pAM1956. Strain KKC7 showed strong expression of *gfp* in vegetative cells and heterocysts in media with and without a source of combined nitrogen. It is not clear if a true *nifH* promoter drives the expression of *gfp* in KKC7. However, this observation raises the possibility that the *nifH* gene is kept off in vegetative cells by the action of a specific repressor, which gets titrated out

by multiple copies of the promoter on the high-copy-number shuttle vector. To test this possibility, we cloned the 700-bp upstream region of *nifH* into a shuttle plasmid without *gfp*, and transferred this plasmid to AMC1774 (Chapter II). We hypothesized that multiple copies of the promoter would lead to titration of a repressor, and would allow P*nifHD*-*gfp* expression from the chromosome in vegetative cells. However, this strain showed normal heterocyst-specific GFP expression, suggesting that the constitutive expression from our plasmid construct was not the result of repressor titration. It remains to be determined if differences in DNA topology, or chromatin structure between the chromosome and the plasmid, might account for the constitutive expression of *gfp* in KKC7.

Translational fusions within the *nifH* gene. To determine if sequences within the *nifH* gene might influence stability of the *nifH* transcripts and affect *gfp* reporter expression, we made several reporter fusions of *gfp* fused to the *nifH* ATG, or downstream sequences within the *nifH* gene, in the chromosome. All fusions carried an upstream region containing sequences until the 3' end of the *nifS* gene. Fusions at the *nifH* ATG, +54, +99, +250, +450, and +750 showed no *gfp* expression after nitrogen deprivation. However, AMC1774, in which *gfp* is fused to the *nifD* ATG in the chromosome, showed developmentally-regulated GFP in heterocysts. It is not clear why the *nifH* ORF fusions did not express *gfp*.

Section II

***fdxH–gfp* reporter fusion**. The *fdxH* gene of *Anabaena* PCC 7120 encodes a heterocyst-specific [2Fe-2S] bacterial ferredoxin (FdxH) that is thought to serve as the electron donor to nitrogenase under iron-replete conditions. *fdxH* is upregulated during late stages of heterocyst development following nitrogen deprivation (53). Although FdxH is not necessary for diazotrophic growth, it is necessary for optimal growth and nitrogenase activity under conditions of nitrogen starvation (112). Because our initial efforts to express *gfp* from the *nifH* upstream region on a shuttle plasmid were unsuccessful, we attempted to use *fdxH* as a candidate gene in an effort to study heterocyst-specific late-gene regulation in *Anabaena* PCC 7120. We constructed reporter fusions of the upstream region between *fdxH* and its neighboring gene transcriptionally fused to *gfp* on pAM1956. KKC20 carrying this *gfp* reporter (pAMKK20) did not express *gfp*. It is possible that the promoter of *fdxH*, like *nifHDK,* lies far upstream of the gene itself.

Section III

Role of FurA in expression of the *nif* genes. The ferric uptake regulator (Fur) protein is a global transcriptional regulator in many bacteria. It controls genes involved in iron homeostasis, oxidative stress responses and central metabolism (3, 71). Fur functions mostly as a repressor that, under iron-rich conditions, uses iron as a co-repressor to bind to target DNA sequences known as iron boxes (76). In some cases, Fur has been shown to function as an activator of gene expression (157).

The genome of *Anabaena* PCC 7120 encodes three Fur homologues: FurA (encoded by all1691) is an essential protein in cyanobacteria and is upregulated about 15 hours after nitrogen deprivation in proheterocysts and mature heterocysts; FurB (encoded by all2473) and FurC (encoded by alr0957) do not show differential expression in response to nitrogen starvation (75). The amount of active FurA in *Anabaena* sp. is governed by iron availability, redox status of the cell, and *α-furA* antisense RNA, among other factors (3, 76). Regulation of *furA* occurs at two levels: transcriptional and post-transcriptional level. At the transcriptional level, *furA* is autoregulated, and heme binds to FurA protein, impairing its ability to bind to its DNA targets (76). NtcA, the master regulator of nitrogen metabolism, represses expression of FurA in nitrogen-replete cultures (108). At the post transcriptional level, *α-furA* antisense RNA controls the levels of FurA in the cell by interfering with furA transcript translation (77)

A clear link exists between the reserves of iron and nitrogen metabolism. Transcription of the *nifHDK* genes and excision of the 11-kb DNA fragment takes place in iron-starved *Anabaena* 7120 even in the presence of combined nitrogen (152). EMSA showed that FurA binds to a putative binding site (iron box) (GTAGAAGCAGTTTAGTTAACAGTAACA) upstream of *nifH* (107).

Prior to the identification of P*nifB* as the promoter that drives the expression of the *nifHDK* genes, we hypothesized that FurA may act as a transcriptional regulator, which binds to the iron box upstream of *nifH*. We predicted that transcription of *furA* increases as a result of low iron levels in heterocysts, and that then FurA, which senses oxidative stress caused by iron starvation would bind to the promoter of the *nifHDK* genes to

activate transcription.

To test the role of FurA in regulation of the *nif* genes, we overexpressed FurA alone, and FurA together with the late-gene specific sigma factor SigE in the AMC1774 background. Upon maximum induction of the *petE* promoter with 2 μM copper, there was no early or constitutive expression of *gfp* showing that FurA is not sufficient to cause expression of the nif genes. However, this result does not exclude the involvement of FurA in the regulation of the *nif* genes: it is possible that the 59-kb *fdxN* element must be excised to allow constitutive *nif* expression from overexpressing FurA alone, and/or FurA and SigE. This could be addressed with an experiment that would overexpress FurA, and FurA together with SigE in AMC445, a strain that has lost the *fdxN* element. Additionally, in vitro transcription assays with FurA and SigE could determine if FurA plays a direct role in regulating the *nif* genes.

Section IV

Transcriptional fusions of P_{hglD} and P_{hglE} to *gfp*. The heterocyst glycolipid layer is deposited to limit the entry of oxygen into the heterocyst. Synthesis of the glycolipid layer involves the *hglE* gene and the *hglD*, *hglC*, and *hglB* gene cluster (49). The *hglK* (9), *devBCA* gene cluster (51), and *devH* genes (150) are implicated in glycolipid transport and/or assembly. The *hglD* and *hglE* genes encode heterocyst glycolipid synthases, and are upregulated about 8 hours following nitrogen starvation. The *hglE* gene, or similar sequences, is found in all other heterocyst-forming cyanobacteria (28). Based on their timing of expression, the *hgl* genes are classified as middle-stage genes of

heterocyst development. Factors that regulate genes during this stage of heterocyst development remain poorly studied (150). To identify the regulatory factors and mechanisms involved in expression of the middle genes, we used P*hglD* and P*hglE* to create *gfp* reporter fusions that can be used to characterize the regulation of *hgl* and other middle genes.

To create the reporter construct, we constructed transcriptional fusions of the amplified upstream intergenic regions of *hglD* and *hglE* fused to *gfp* on pAM1956. Strain KKC23 and KKC24 carrying these plasmids showed heterocyst-specific *gfp* expression approximately 8-10 hours after nitrogen deprivation. The reporter strains were constructed with the intent of using them to isolate mutants in trans-acting factors that regulate the middle genes. However, we chose to focus our efforts on identifying trans-acting factors and cis-acting sequences of the *nif* genes as candidate late genes. In the future, an approach similar to the one described for the *nif* genes (Chapter II and III) could be used to identify trans-acting factors that regulate the middle genes. KKC23 and KKC24 can also serve as reporter strains to study the role of SigG, a middle-stage sigma factor, in regulating the middle genes of heterocyst development.

Section V

ORFs alr0340-all0341-all0342. In an effort to identify the trans-acting regulatory proteins that are required to transcribe the *nif* genes, we overexpressed a high-copy-number expression library of random gene fragments in AMC1774 (Chapter II). We screened for mutants that did not express *gfp*, and were unable to grow diazotrophically,

which would include mutants overexpressing a repressor on the expression plasmid. We also screened for constitutively bright mutants, which would include those that overexpress an activator from the expression plasmid.

We isolated a dark mutant (KKC14) that was unable to grow on media lacking a source of combined nitrogen. Sequencing of the library plasmid from KKC14 showed that this plasmid carried ORFs alr0340, all0341, both of which are proteins of unknown function, and all0342, a sodium glutamate symport protein. alr0340 encodes a protein that has 31% identity to the LuxR regulators.

ORFs alr0340 and all0342 were present as truncated versions on the library plasmid pAMKK14; alr0340 had an N-terminal truncation of 51 codons and alr0342 had a C-terminal truncation of 50 codons. alr0340 is in reverse orientation to the *rbcL* promoter on the library plasmid. However, an in-frame translational start site on the plasmid, present upstream of the alr0340 sequence, which could result in the production of a truncated protein carrying this extra N-terminal sequence.

To identify the gene responsible for the phenotype of KKC14, we cloned a wild-type copy of each gene, alr0340, all0341 or all0342, independently into pAM2770 (191), a shuttle plasmid carrying the copper-inducible *petE* promoter. The genes were overexpressed in AMC1774 by adding 2 μM copper for complete induction of P*petE*. Strains KKC16 and KKC17, overexpressing genes all0341 (pAMKK16) and all0342 (pAMKK17), respectively, showed normal growth and heterocyst-specific expression of P*nifHD-gfp* under nitrogen deprivation conditions. Strain KKC15 overexpressing alr0340 from plasmid pAMKK15 was unable to grow diazotrophically, but showed

heterocyst-specific expression of P*nifHD-gfp*.

The difference in phenotypes between the original library plasmid and pAMKK15 could result from differences in truncated versus full-length proteins produced by these plasmids, respectively: the truncated protein may produce a dominant negative effect in the mutant. To determine if the truncated version of alr0340 was responsible for the mutant phenotype, we cloned truncated alr0340 and a wild type copy of the gene independently in pAM1824 in reverse orientation relative to the *rbcL* promoter. Strains KKC18 and KKC19 showed heterocyst-specific GFP but did not grow diazotrophically. It is possible that the mutant phenotype of KKC14 is specific to the library plasmid: sequences upstream of alr0340 on the plasmid may contribute to the phenotype, and/or the mutant phenotype is a result of a synergistic interaction between two or more genes on the library plasmid. Our attempts to use the original library plasmid to disrupt the ORFs individually were unsuccessful because of limitations in availability of restriction sites and the large size of the plasmid.

Our data suggest that no ORF on the library plasmid (pAMKK14) is a direct regulator of the *nif* genes. Rather, it is likely that one or more genes indirectly affects nitrogen fixation by disrupting other physiological processes such as glutamate transport, or by interacting with the nitrogenase itself and rendering it nonfunctional.

Materials and methods

Strains and conditions. Strains and plasmids used in this study are shown in Table 6. *Anabaena* sp. strain PCC 7120 and its derivatives were grown in BG-11 or $BG\text{-}11_0$

medium at 30°C as previously described (64). Cultures were grown under fluorescent white light illumination at 75-100 μmol photons m^{-2} s^{-1}. Plasmids were transferred to *Anabaena* 7120 by conjugation from *Escherichia coli* following published protocols (46, 65). For *Anabaena* PCC 7120 strains carrying plasmids, the antibiotics neomycin (Nm; 25 μg/ml) and/or streptomycin (Sm) and spectinomycin (Sp) (2 μg/ml each) were used for growth on BG-11 or BG-11_0 agar medium in plates; the antibiotic concentrations were reduced by half for liquid cultures. Heterocyst differentiation was induced by nitrogen deprivation as previously described (2).

E. coli strains were maintained in LB (Lennox L) liquid or agar medium supplemented with appropriate antibiotics.

TABLE 6. Strains and plasmids used in this study (Appendix)

Strain	Characteristic	Source or reference
E. coli strains		
AM1359	*E. coli* strain containing conjugal transfer and helper plasmid	(46)
Anabaena sp. strains		
KKC5	*Anabaena* PCC 7120 carrying pAMKK5	This study
KKC6	*Anabaena* PCC 7120 carrying pAMKK6	This study
KKC7	*Anabaena* PCC 7120 carrying pAMKK7	This study
KKC8	*Anabaena* PCC 7120 carrying pAMKK8 recombined into the chromosome at the *nif* locus	This study
KKC9	*Anabaena* PCC 7120 carrying pAMKK9 recombined into the chromosome at the *nif* locus	This study
KKC10	*Anabaena* PCC 7120 carrying pAMKK10 recombined into the chromosome at the *nif* locus	This study
KKC11	*Anabaena* PCC 7120 carrying pAMKK11 recombined into the chromosome at the *nif* locus	This study

Table 6. Continued…

Strain	Characteristic	Reference
KKC12	*Anabaena* PCC 7120 carrying pAMKK12 recombined into the chromosome at the *nif* locus	This study
KKC13	*Anabaena* PCC 7120 carrying pAMKK13 recombined into the chromosome at the *nif* locus	This study
KKC14	*Anabaena* PCC 7120 carrying pAMKK14	This study
KKC15	*Anabaena* PCC 7120 carrying pAMKK15	This study
KKC16	*Anabaena* PCC 7120 carrying pAMKK16	This study
KKC17	*Anabaena* PCC 7120 carrying pAMKK17	This study
KKC18	*Anabaena* PCC 7120 carrying pAMKK18	This study
KKC19	*Anabaena* PCC 7120 carrying pAMKK19	This study
KKC20	*Anabaena* PCC 7120 carrying pAMKK20	This study
KKC21	*Anabaena* PCC 7120 carrying pAMKK21	This study
KKC22	*Anabaena* PCC 7120 carrying pAMKK22	This study
KKC23	*Anabaena* PCC 7120 carrying pAMKK23	This study
KKC24	*Anabaena* PCC 7120 carrying pAMKK24	This study
KKC25	*Anabaena* PCC 7120 carrying pAMKK25	This study
Plasmids		
pAMKK5	Shuttle vector pAM1956 containing 221-bp fragment upstream of *nifH* gene transcriptionally fused to *gfp*	This study
pAMKK6	Shuttle vector pAM1956 containing 350-bp fragment upstream of *nifH* gene transcriptionally fused to *gfp*	This study
pAMKK7	Shuttle vector pAM1956 containing 700-bp fragment upstream of *nifH* gene transcriptionally fused to *gfp*	This study
pAMKK8	Conjugal suicide vector, pAMKK1 containing the region between the *nifS* and *nifH* gene; *gfp* is fused to the *nifH* ATG	This study
pAMKK9	Conjugal suicide vector pAMKK1 containing the fragment between the *nifS* and *nifH* gene; *gfp* is fused to the +54 bp of the *nifH* ORF	This study
pAMKK10	pAMKK1 containing the region between the *nifS* and *nifH* gene; *gfp* is fused to the +99 bp of the *nifH* ORF	This study
pAMKK11	pAMKK1 containing the region between the *nifS* and *nifH* gene; *gfp* is fused to the +250 bp of the *nifH* ORF	This study

Plasmid constructions. Primers used in this study are shown in Table 7. Shuttle vectors pAMKK5 (primers KKO9, KKO10), pAMKK6 (primers KKO11, KKO12), and

pAMKK7 (primers KKO13, KKO14) were constructed by cloning PCR-amplified fragments of the *nifH* upstream regions into SalI and SacI sites of pAM1956. pAMKK8, in which *gfp* is translationally fused to the *nifH* ATG on pAMKK1, was constructed by cloning a PCR-amplified fragment of the *nifH* upstream region using primers KKO15 and KKO16 into XhoI and SpeI sites. The *nifH-gfp* translational reporters were constructed by amplifying regions of 54 bp (primers KKO17, KKO18), 99 bp (primers KKO19, KKO20), 250 bp (primers KKO21, KKO22), 450 bp (primers KKO23, KKO24), 750 bp (primers KKO25, KKO26), and cloning each into XhoI and SpeI sites of pAMKK1 to generate pAMKK9, pAMKK10, pAMKK11, pAMKK12, and pAMKK13, respectively. Plasmids pAMKK16, pAMKK17, and pAMKK18, for overexpression of ORFs alr0340, all0341 and all0342 independently from the *petE* promoter, were constructed by cloning PCR-amplified fragments of the genes, using primers KKO29 and KKO30 (alr0340), KKO31 and KKO32 (all0341), and KKO33 and KKO34 (all0342), into pAM2770 using SacI and XmaI sites. Plasmids pAMKK18 and pAMKK19 for overexpression of the complete and truncated version of the alr0340 gene, respectively, from the *rbcL* promoter, were constructed by cloning PCR-amplified fragments of the complete or truncated version of the gene, using primers KKO35 and KKO 36 (alr0340), and KKO37 and KKO38 (truncated alr0340) into pAM1824 using SalI and Sac*I* sites. The plasmids for constructing the overexpressing FurA (pAMKK21), and FurA and SigE (pAMKK22) were constructed by amplifying the genes and cloning the fragment using SalI and SacI restriction sites into the shuttle plasmid pAM2770. The P*fdxH-gfp* (primers KKO39, KKO40), P*hglD-gfp* (primers KKO43, KKO44) and P*hglE*-

gfp (primers KKO45, KKO46) transcriptional reporters were constructed by amplifying the entire upstream UTR of each gene and cloning into pAM1956 to generate pAMKK20 (P*fdxH-gfp*), pAMKK23 (P*hglD*-*gfp*), and pAMKK24 (P*hglE*-*gfp*).

Microscopy. Fluorescence and bright-field images were captured using an Olympus IX70 inverted microscope with a Hamamatsu OrcaER C4742-95 charge-coupled-device camera and Simple PCI software version 6.1 as previously described (131). A Piston green fluorescent protein (GFP) band-pass filter set (no. 41025; Chroma Technology Corp.) was used for fluorescence images. Photomicrographs were minimally processed with Adobe Photoshop 8.0 to improve brightness and contrast.

TABLE 7. Primers used in this study (Appendix)

Primer	Sequence	Primer
KKO9	AAGGTCGACGTAGAATCCGATAAGAT	This study
KKO10	CCAGAGCTCCAAGGGTGTTTTGGG	This study
KKO11	GATGTCGACGGCGATATTGTCAAAGTA	This study
KKO12	GAGCTCTCTGTCTAATGTTTTCGT	This study
KKO13	AGAGGTCGAC GCGGTGATGCGCTGA	This study
KKO14	ATTGAGCTCAGCTATCTGTCTAAT	This study
KKO15	TAGCTCGAGATTAGGGAATAGGAAGAAGC	This study
KKO16	ATTCTTAAGGTATGTTCTCTTTTCCTGCAATTG	This study
KKO17	TAGCTCGAGATTAGGGAATAGGAAGAAGC	This study
KKO18	ATTACTAGTACCGATACCGCCTTTACC	This study
KKO19	TAGCTCGAGATTAGGGAATAGGAAGAAGC	This study
KKO20	ACCACTAGTATTTCTGCCATAGCTGCAAGGGT	This study
KKO21	TAGCTCGAGATTAGGGAATAGGAAGAAGC	This study
KKO22	ATTACTAGTGAAACCGGTCAACAT	This study
KKO23	TAGCTCGAGATTAGGGAATAGGAAGAAGC	This study
KKO24	ATACTTAAGACCACCGGAGTGAGCATATTT	This study
KKO25	TAGCTCGAGATTAGGGAATAGGAAGAAGC	This study
KKO26	ATTACTAGTCTATTTGGTAGCTTCTGCGGG	This study
KKO29	TAAGAGCTC	This study
KKO30	TAACCCGGG	This study
KKO31	TAAGAGCTCATGATTCAAACTATGTGG	This study
KKO32	TAACCCGGGTTAACCACCACCACCGCC	This study

Table 7. Continued…

Primer	Sequence	Reference
KKO33	TAAGAGCTCATGAGTGACTCCAAAGGC	This study
KKO34	TAACCCGGGTCATTCGCTTTGCACCTC	This study
KKO35	TAAGAGCTCATGCCTAGCCTCAAACTC	This study
KKO36	TAAGTCGACTTAAAATAAACGCTGCAC	This study
KKO37	ATTGAGCTCATGCCTGCAGGTCGACGGTAT	This study
KKO38	TAAGTCGACTTAAAATAAACGCTGCAC	This study
KKO39	GTAGTCGACTCTCTGTCAAAGATGGTGGTGG	This study
KKO40	GTTGAGCTCTCTAACTTGGTAGCTAGCCATTTG	This study
KKO41	ATTGAGCTCATGACTGTCTACACAAATAC	This study
KKO42	ATTCTCGAGCTAAAGTGGCATGAG	This study
KKO43	GGTGTCGACGCCGATATCCTCCATGAC	This study
KKO44	GATGAGCTCGAATCTTTCTGCTCGGTG	This study
KKO45	ATTGTCGACAGTAATGCAGAATTAGTTAGTC	This study
KKO46	ATAGAGCTCAACAAAGCCTTTTCGCAG	This study

VITA

Name: Krithika Kumar

Address: Texas A&M University
TAMU 3258, College Station, TX 77843-3258

Education: B.S., Biotechnology, Bangalore University, 2004

Ph.D., Microbiology, Texas A&M University, 2011