

ORIGINAL PAPER

Elsie L. Campbell · Michael F. Cohen · John C. Meeks

A polyketide-synthase-like gene is involved in the synthesis of heterocyst glycolipids in *Nostoc punctiforme* strain ATCC 29133

Received: 4 November 1996 / Accepted: 6 January 1997

Abstract A Tn5-1063-derived mutant of *Nostoc punctiforme* strain ATCC 29133 was unable to fix N₂ in air although it reduced acetylene in the absence of O₂. Mutant strain UCD 307 formed cells morphologically similar to heterocysts, but it failed to synthesize the characteristic heterocyst glycolipids. Sequence analysis around the site of insertion revealed an ORF of 3,159 base pairs, termed *hglE*. *hglE* putatively encodes a 115.4-kDa protein containing two domains with conserved amino acid sequences identified with acyl transferase and the chain length factor variation of β -ketoacyl synthase active sites. These active sites are characteristic of polyketide and fatty acid synthases. The *N. punctiforme* strain 29133 *hglE* gene is transcribed only under nitrogen-limiting growth conditions. The *hglE* gene, or similar sequences, was found in all other heterocyst-forming cyanobacteria surveyed and was absent in unicellular *Synechococcus* sp. strain PCC 7942. Based on these results, we propose that the synthesis of heterocyst glycolipids follows a pathway characteristic of polyketide synthesis and involves similar large, multienzyme complexes.

Key words Fox⁻ mutant · Heterocyst glycolipid · *hglE* · *Nostoc punctiforme* · Polyketide synthase

Abbreviations *Ap* Ampicillin · *AT* Acyl transferase · *Chl a* Chlorophyll *a* · *CLF* Chain length factor · *Cm* Chloramphenicol · *Em* Erythromycin · *Hgl* Heterocyst glycolipid · *Km* Kanamycin · *KS* β -Ketoacyl synthase · *Nm* Neomycin, *PKS* Polyketide synthase · *FAS* Fatty acid synthase

Introduction

Cyanobacteria are the only oxygenic phototrophs with members that are capable of oxygen-sensitive nitrogen fixation. Some unicellular and filamentous cyanobacteria carry out N₂ fixation only when the oxygen tension in the environment is low, e.g., during the dark phase of a natural diel cycle when photosynthetic O₂ production ceases and respiratory O₂ consumption increases (Gallon 1992). Many filamentous cyanobacteria fix N₂ concurrently with photosynthetic O₂ production; in those strains that differentiate heterocysts, N₂ fixation in an oxic environment appears to be confined to these specialized cells (Elhai and Wolk 1990; Theil et al. 1995). Certain filamentous cyanobacteria such as *Trichodesmium* spp. do not appear to differentiate specialized cells, yet fix N₂ and evolve O₂ concurrently (Capone et al. 1990); the mechanisms by which the nitrogenase is protected from O₂ inhibition in these strains is unknown.

Heterocysts have three characteristics that collectively appear to yield an anoxic or micro-oxic cellular environment (Fay 1992): (1) they lack the photosystem II O₂-producing reaction; (2) they have an increased rate of respiratory O₂ uptake; and (3) they synthesize an outer envelope consisting of an alternating layer of polysaccharides (outermost) and glycolipids, which together impede the entry of O₂ and other gases (Walsby 1985; Murry and Wolk 1989). Heterocysts differentiate from vegetative cells only under conditions of combined nitrogen deprivation (Wolk et al. 1994). Formation of the outer polysaccharide layer is an early event in the differentiation process, and the thickening wall is diagnostic of a proheterocyst (Wilcox et al. 1973). Assembly of the glycolipid layer occurs during the commitment stage when a proheterocyst matures into a terminally differentiated heterocyst (Wilcox et al. 1973). Heterocysts that form in the absence of O₂ appear to lack the glycolipid layer and are unable to protect nitrogenase from O₂ inactivation under atmospheric conditions (Rippka and Stanier 1978). Mutants incapable of supplying reductant for respiration

E. L. Campbell · M. F. Cohen · J. C. Meeks (✉)
Section of Microbiology, Division of Biological Sciences,
University of California, Davis, CA 95616, USA
Tel. +1-916-752-3346; Fax +1-916-752-9014
e-mail: jcmeeks@ucdavis.edu

(Summers et al. 1995) or lacking the polysaccharide (Holland and Wolk 1990) or glycolipid layers (Ernst et al. 1992; Black et al. 1995) cannot fix nitrogen in air, but do so under anoxic conditions when supplied with reductant.

Nostoc punctiforme strain ATCC 29133 is a facultatively heterotrophic cyanobacterium whose vegetative cells have multiple developmental alternatives including the differentiation of heterocysts. Methods have been developed for transposon mutagenesis of *N. punctiforme* strain 29133 using a modified Tn5 constructed by Wolk et al. (1991) for recovery of the transposon and flanking genomic DNA from mutant strains and for molecular genetic characterization of the interrupted gene (Cohen et al. 1994). Approximately 0.4% of the antibiotic-resistant *N. punctiforme* strain 29133::Tn5-1063 mutants recovered under photoautotrophic growth conditions with NH_4^+ as the nitrogen source have a Fox^- (unable to fix nitrogen in air) phenotype; rather than failing to differentiate a heterocyst, the majority of these Fox^- mutants formed defective heterocysts where nitrogenase was active under anoxic incubation conditions. We have previously characterized one such Fox^- mutant that failed to develop a mature heterocyst wall; the Tn5-1063 insertion was in a gene (*devR*) encoding a response regulator protein that consists of the receiver domain only (Campbell et al. 1996), similar to Spo0F of *Bacillus subtilis* and CheY of *Escherichia coli* [see Parkinson and Kofoid (1992)]. The results indicated the operation of a signaling pathway in heterocyst development, apart from the initiation of differentiation. We report here the characteristics of a second Fox^- *N. punctiforme* strain 29133::Tn5-1063 mutant and identify the gene responsible for the phenotype as encoding a multidomain protein involved in the synthesis of heterocyst glycolipids.

Materials and methods

Bacterial strains, growth conditions, and physiological measurements

Nostoc punctiforme strain ATCC 29133 (PCC 73102, type strain of *N. punctiforme*; Rippka and Herdman 1992) and mutant strain UCD 307 were grown in liquid and agar-solidified medium as described (Enderlin and Meeks 1983). When necessary, cultures were supplemented with 2.5 mM NH_4Cl and buffered with 5 mM Mops (pH 7.8). Neomycin (Nm) at $10 \mu\text{g ml}^{-1}$ was used for selection and routine culture of recombinant *N. punctiforme* 29133 strains. Selection plates were incubated at 28°C under continuous illumination from cool-white fluorescent lamps at 8 W m^{-2} with 1% (v/v) CO_2 enrichment. Mutagenesis with the transposon Tn5-1063 (Wolk et al. 1991) and isolation of the Fox^- mutants and their preliminary characterization (Cohen et al. 1994) have been described previously. Oxic and anoxic acetylene reduction assays as a measure of nitrogenase activity were performed as described (Campbell and Meeks 1992). Heterocyst glycolipids were extracted from cultures with chloroform:methanol (2:1, v/v), concentrated under nitrogen gas, spotted on a silica gel G TLC plate (Analtech), separated by chromatography using chloroform:methanol:acetic acid:water (170:30:20:7.4, by vol.), and visualized by charring after spraying with 25% (v/v) H_2SO_4 as described (Murry and Wolk 1989). Luciferase activity was measured by adding $50 \mu\text{l}$ 2.65 mM *n*-decylaldehyde (Aldrich) to $50 \mu\text{l}$ of cells using a Model 6100 Pico-Lite luminometer (Packard Instruments); pho-

tons emanating from the sample were counted for 1 min following a 1-min delay and were normalized to chlorophyll *a* (Chl *a*) content, determined in methanolic extracts (Enderlin and Meeks 1983).

Escherichia coli strain DH5 α -MCR was used for all standard cloning experiments including subcloning of appropriate fragments following standard protocols (Sambrook et al. 1989). *E. coli* DH5 α -MCR carrying various recombinant plasmids was grown in Luria-Bertani broth (Sambrook et al. 1989) supplemented with antibiotics at the following concentrations: kanamycin (Km) at $50 \mu\text{g ml}^{-1}$, ampicillin (Ap) at $100 \mu\text{g ml}^{-1}$, and chloramphenicol (Cm) at $30 \mu\text{g ml}^{-1}$.

Molecular genetic techniques

Genomic DNA isolations from cyanobacteria were performed as described previously (Cohen et al. 1994). All DNA-modifying enzymes were purchased from New England Biolabs, Gibco BRL, or Boehringer Mannheim and were used according to the manufacturer's instructions. Recovery of the transposon with flanking genomic DNA in strain UCD 307 was performed as described (Cohen et al. 1994) by digesting mutant genomic DNA with *ScaI* followed by ligation and transformation into *E. coli* DH5 α -MCR, resulting in plasmid pSCR209. A Klenow-blunted 3.0-kb fragment from a *BamHI/EagI* double digest of pSCR209 (restriction sites in the transposon portion) containing all of the flanking genomic DNA was cloned into the *SmaI* site of pBluescript II KS+ (Stratagene), resulting in plasmid pSCR210; this plasmid served as template for sequencing using synthetic oligonucleotide primers and the dideoxy nucleotide method (kit purchased from United States Biochemical). All clones containing the insert from the ligation that gave rise to pSCR210 were such that the transcriptional reading frame was opposite that of *lacZ* from pBluescript. Following the method of interposon vector construction of Black and Wolk (1994), pSCR212 was made from pSCR209 for use in reconstruction of the mutant. Plasmid pSCR212 carries an inactivated transposase gene in the transposon interrupting *hglE* on a vector (pRL271; $\text{Em}^r \text{Cm}^r$) containing *SacB* to positively select for double recombinants in the genome (Cai and Wolk 1990; Cohen et al. 1994; Campbell et al. 1996). Again, only one orientation was obtained in all positive clones from this ligation.

Plasmid pSCR212 was mated into wild-type *N. punctiforme* strain 29133 as described (Campbell et al. 1996). Despite over 50 filter matings, no antibiotic-resistant colonies were recovered; however, only 112 bp of genomic DNA was available on one side of the interposon, which might have limited recombination. To obtain a larger genomic fragment for use in reconstruction of the mutation, the flanking DNA from pSCR210 was used as a probe (random priming kit purchased from Gibco-BRL; ^{32}P purchased from DuPont NEN) to a random sheared *N. punctiforme* strain 29133 genomic cosmid library (Cohen et al. 1994). Nine cosmids hybridized when probed with the flanking DNA from pSCR210. Upon subsequent isolation and characterization, three cosmids showed similar restriction patterns in Southern hybridizations. We attempted to clone several larger genomic fragments (5.5–12 kb) from cosmid pSCR213 into pBluescript and/or pACYC177 without success. Most clones analyzed from these ligations were either smaller or larger than predicted and did not have the expected pattern of restriction bands. Two pBluescript-based clones derived from ligation of a 12-kb *HpaI* genomic fragment contained two plasmids; one appeared to be of the predicted 12-kb size, while the other was larger. Efforts to stabilize the desired plasmid by growth on either rich or minimal medium or by lowered temperature (30°C) were not successful. We were, however, able to clone smaller fragments (2.2- and 1.5-kb *HindIII* fragments, a 2.5-kb *HincII*/partial *HindIII* fragment, a 1.7-kb *NaeI/ScaI* fragment, and a 2.16-kb *XmnI/ScaI* fragment) into pBluescript II KS+, and in all cases the cloned fragments were oriented such that the transcriptional direction was opposite that of *lacZ*. The two *HindIII* fragments, which together contained the sequence of the entire ORF, served as templates for sequencing of both strands and verified the previous sequence obtained from pSCR210. Additionally, a 1.0-kb

*Bsp*MI fragment completely internal to the ORF and containing the β -ketoacyl synthase active site could not be cloned into pBlue-script. None of the above listed fragments appeared to be stable in the conjugal vector pARO190 in either orientation. Thus, neither reconstruction of the mutation by a single internal recombination event nor complementation in *trans* could be attempted.

Sequence homology searches were performed at National Center for Biotechnology Information (NCBI) using the BLAST network service (Altschul et al. 1990). Alignment of *hglE* with other similar genes was done using the PILEUP program of Wisconsin GCG with default settings. The 3,701-bp sequence encompassing *hglE* is available from NCBI under accession no. L48681.

The 1.5-kb cloned *Hind*III fragment was used as the probe for Northern and Southern hybridizations. RNA and DNA blots were prepared as described (Summers et al. 1995).

Results

Phenotype of strain UCD 307

Mutant strain UCD 307 was generated by transposition of Tn5-1063 (Cohen et al. 1994). Strain UCD 307 grew at a photoautotrophic rate similar to that of the wild-type when supplemented with ammonium or nitrate as nitrogen sources, and the filaments typically contained more than 100 vegetative cells with no unusual morphology. However, within 72 h after deprivation of combined nitrogen (N-step-down), the filaments had fragmented into short, three- to four-cell filaments. Strain UCD 307 was unable to grow with N_2 as the sole nitrogen source or to reduce acetylene in air, but it had an anoxic acetylene reduction rate of 173 ± 9.6 (\pm SE, $n = 4$) nmol C_2H_4 formed min^{-1} [mg chlorophyll *a* (Chl *a*)] $^{-1}$ within 36 h after N-step-down. Wild-type *Nostoc punctiforme* strain 29133 cul-

tured with N_2 as the sole nitrogen source had a rate of 129 ± 33 (\pm SE, $n = 3$) nmol C_2H_4 formed min^{-1} (mg Chl *a*) $^{-1}$ under air. Epifluorescence microscopy using 580 nm excitation (primarily absorbed by phycoerythrin) and > 590 nm emission (primarily from allophycocyanin and Chl *a*) of strain UCD 307 24 h after N-step-down revealed a pattern of nonfluorescing cells in the filaments identical to the pattern of heterocyst (also nonfluorescing cells) spacing seen in wild-type cultures (data not shown). Under phase-contrast illumination, the nonfluorescing cells of strain UCD 307 had the enlarged shape and optical properties characteristic of heterocysts, but many of these cells also lysed under pressure of the cover slip; this fragility may, in part, explain the filament fragmentation. TLC profiles of lipid extracts of strain UCD 307 and of the wild-type 36 h following N-step-down are shown in Fig. 1. The unique heterocyst glycolipid III band is distinctly absent in strain UCD 307. This mutant has previously been shown to make visibly normal akinetes upon phosphate starvation (Cohen et al. 1994).

Molecular genetic characterization of strain UCD 307

The transposon and flanking genomic DNA from strain UCD 307 was recovered following a *Sca*I digest, ligation, and subsequent subcloning to yield a plasmid (pSCR210) that contained 112 bp of genomic DNA on the transposase side and approximately 2.7 kb of genomic DNA on the *luxAB* side of the transposon. Based on BLAST analysis, the amino acid sequence derived from double-stranded nucleotide sequence of a 3,159-bp ORF within this region and from the two *Hind*III subclones of cosmid clone pSCR213 yields a large, 1,053-amino-acid protein containing two active sites (Fig. 2a,b). Both active sites are characteristic of a well-defined group of enzymes termed polyketide synthases (PKS) and fatty acid (FAS) synthases. We have designated this ORF as *hglE* for heterocyst glycolipid. The putative acyl transferase (AT) active site of *hglE* shows the highly conserved motif GHSXG, where the serine is thought to be critical in the formation of the acyl-enzyme intermediate (Fig. 2a; Donadio et al. 1991). A putative β -ketoacyl:acyl carrier protein synthase (KS) domain of *hglE* was determined based on conserved regions as analyzed by comparisons of sequences encoding similar enzymes (Fig. 2b; Siggaard-Anderson 1993). The active site for KS enzymes has the consensus motif GPX₅TACSS (underlined region in Fig. 2b), where the cysteine (indicated by an asterisk in Fig. 2b) is important in thioester formation (Donadio et al. 1991). *HglE* contains 16 highly conserved amino acids identified within a region of approximately 200 amino acids that defines the KS domain outside of the active site.

Both FASs and PKSs are currently grouped into three major classes: Type I includes those in which the KS active site is part of a single, large, multifunctional protein with domains catalyzing all or part of the associated carrier protein, transferase, reductase(s), dehydrase, and thioesterase activities; in Type II synthases, each catalytic

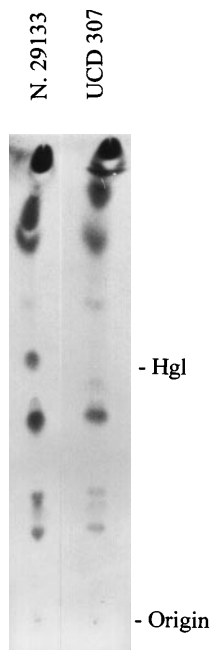


Fig. 1 Thin-layer chromatograms of chloroform-methanol extracts from cultures of wild-type *Nostoc punctiforme* strain 29133, grown with N_2 and mutant strain UCD 307 36 h following N-step-down. The position of heterocyst glycolipid III (Hgl) is indicated

Fig. 2 a Amino acid alignment of the acyl transferase (AT) domain of HglE and other proteins. Conserved residues in the AT active site motif (GHSXG) are shown in *bold type*. FabD (accession P25725) is from *Escherichia coli*, Fix23-2 (accession S18954) is from *Rhizobium meliloti*, and HglC (accession U13677) is from *Anabaena* strain 7120. Analysis of the HglE sequence starts at amino acid 701. A gap was inserted in the HglC sequence to improve its short-range alignment. **b** Amino acid alignment of the β -ketoacyl synthase (KS) domain of HglE and other proteins. Highly conserved residues in the KS domain are shown in *bold type*. The consensus active site motif of Type II KSs is GPX₂TACSS (*underlined*), with the location of the cysteine residue indicated by an *asterisk*. This cysteine is substituted in Type II chain length factors such as HglE (accession L48681) from *Nostoc punctiforme* strain 29133, HglC (accession U13677) from *Anabaena* strain 7120, and Strep. ORFIV (accession X55942) from *Streptomyces coelicolor*. Type II KSs are NodE (accession P06230) from *Rhizobium meliloti*, EryA (accession M63677) from *Saccharopolyspora erythraeus*, B.su.PKS (accession PN0637) from *Bacillus subtilis*, and HglD (accession U13677) from *Anabaena* strain 7120, of which only partial sequence is available. Analysis of the HglE sequence starts at amino acid 251

a						
HglE	...	ALQRTEY	AQPAIGVLSA	GMYSILQQAG	FK.SDFVAGH	SFGELTALWA
FabD	...	ELNKTWQ	TQPALLTASV	ALYRVWQQQG	GKAPAMMAGH	SLGEYSALVC
Fix23-2		LDRKLDATKV	AQPLLFAIQA	ALSDSLVAMG	IKPTAVF.GH	SVGETIAAAYA
HglC		KKLLSDSLAM	FEAEIFCTRL	LTTIIRDDFQ	VHPKYVF.GY	SLGETSMMVA
b						
				*		
HglE		GGMNCVVDA	WPSSFGALKM	AISELVEHRS	DMMLTGGVDT	DNTIMAYISF
HglC		TGPSFTLTAV	ESSAFKAVEV	AQMLLMTQEV	DAVVVGAVDL	AGGVENVLLR
Strep. ORF IV		KGPCGVVAAD	EAGGLDALAH	AALAVRNGT.	DTVVCGATEA	PLAPYSIVCQ
NodE		RGPFVGVVTS	CSSANHAIAS	AVDQIKCGRA	DVMLAGGSDA	PLVWIVLKAW
EryA		EGPSISVDTA	CSSSLVALHL	AVESLRKGES	SMAVVGGAAV	MATPGVVFVDF
HglD		SGTHFCIDAA	CSSAFYTIKL	ASHYLWTGKA	DVMLAGAISC	SEPLFLRMLF
B.su.PKS		HGPSIALDTM	CSSSLTAIHL	ACQSLQRGEC	EAAFAGGVNV	SIHPNKYMLL
HglE		SKTPAVSPSE	N....VKPFD	AKSDGMMLGE	GIGMIVLKRL	EDAE.RDNDK
HglC		SQAAPINTGI	N....TLYSD	QKANGWTVGE	GAGAVVLKRH	QQA.I.ENGDR
Strep. ORF IV		LYPELSRAT	EPDRAYRPFT	EAACGFAPAE	GGAVLVVEEE	AAARERG.AD
NodE		EAMRALAPDTCRPFS	AGRKGVLGE	GAGMAVLESY	EHA.TARG.AT
EryA		SRQRALAADG	R....SKAFG	AGADGFGFSE	GVTLVLLERL	SEAR.RNGHE
HglD		SGIQGYPPENG	I....SSPLD	KSSRGLVTSE	GIGMVMKRY	SDAV.RDGDK
B.su.PKS		GQNKFMSSKG	R....CESFG	QGGDGYVPGE	GVGAVLLKPL	SKAV.EDGDH
HglE		IYAVIKGIGT	SSD.GRYKSI	YASRKEGQVK	ALERAYEDAG	FSPATVGLME
HglC		IYAVIDAISI	QSHTKV DGE	SVNQACQQ..ALKMAV	YHLEQVKYVE
Strep. ORF IV		VRATVAGHAA	TFT...GAGR	WAESREGLAR	AIQGALAEAG	CRPEEVDVVF
NodE		ILAEVAVGVL	SAD...AFHI	TAPAVHGPES	AMRACLADAG	LNAEDVDYLN
EryA		VLAVVVRGSAL	NQD.GASNGL	SAPSGPAQRR	VIRQALESCG	LEPGDVEDAVE
HglD		ILATVCGNGL	SND.GKGKHL	LSPNTKGQVT	AFERAYKEAQ	LDPKAIDYME
B.su.PKS		IYGIKGTAI	NHG.GKTNGY	SVPNPNAQAD	VIKKAFFEAK	VDPRTVSYIE
HglE		AHGTGT MAGD	PT EFGSLKDF	FDVHDDKKQH	IALGSVKSQI	GHTKAAAGAA
HglC		VCASGIPQED	EAEIAGLVQA	YPSVGD.GLH	CAIGSVKANI	GHTFVASGIA
Strep. ORF IV		ADALGVPEAD	RAEALALADA	LGPH...AAR	VPVTAPKTGT	GRAYCAAPVL
NodE		AHGTGT KAND	QNETTAIKRV	FGDH...AYS	MSISSTKSTH	AHCIGAASAL
EryA		AHGTGT ALGD	PIEANALDFT	YGRDRDADRP	LWLGSVKSNI	GHTQAAAGVT
HglD		CHATGTL LGD	TTEFN SIETF	FGAHEAAP..	.LIGSTKANV	GHLLVAAGMV
B.su.PKS		AHGTGT SLGD	PIEITGLSKV	FTQETD.DKQ	FC.....

activity is present as a single polypeptide encoded by a single gene, and the KS proteins function as heterodimers; Type III synthases lack the acyl carrier protein polypeptide (Siggaard-Andersen 1993; Hutchinson and Fujii 1995). A subclass of the Type II synthases (synonym ORF2) that lack the active site cysteine in the KS domain is now referred to as the chain length factor (CLF) to reflect a putative role in determining the hydrocarbon chain length in polyketide synthesis (McDaniel et al. 1993; Hutchinson and Fujii 1995).

HglE shows some degeneracy in the KS active site motif: it has only three identical residues and contains a tryptophan in place of the cysteine (Fig. 2b). Thus, HglE is more characteristic of the CLF of Type II PKSs, which all lack the active site cysteine at this position (Fig. 2b; McDaniel et al. 1993; Siggaard-Andersen 1993).

The site of Tn5-1063 transposition is 312 bp downstream from the AT active site beginning at position 2946 in the nucleotide sequence (illustrated in Fig. 3). A characteristic 9-bp repeat formed at sites of transpositions by

Tn5 (Berg and Berg 1987) was confirmed by sequence analysis of pSCR209. The orientation of the transposon insertion was such that the *luxAB* reporter genes of Tn5-1063 were in the same direction as *hglE*, thereby forming a transcriptional fusion (Fig. 3). The expression of LuxAB in the mutant showed negligible light emission in ammonium-grown cultures, with an increase beginning approximately 10 h following N-step-down and continuing to a steady-state level by 20 h (Fig. 4a). This induction pattern is comparable to DNA:RNA (Northern) blot analysis using a 1.5-kb *HindIII* fragment (Fig. 3) from the sequenced region containing the AT active site as a hybridization probe to RNA isolated following N-step-down in wild-type cells. The Northern blot (Fig. 4b) shows clearly that no message is present in ammonium-grown cells; unstable mRNA appears within 6 h following N-step-down, with an apparent maximal size of about 6.2 kb. The rRNA bands in this gel were distinct and of equal intensity following ethidium bromide staining (data not shown); therefore, the mRNA smearing was not due to nonspecific

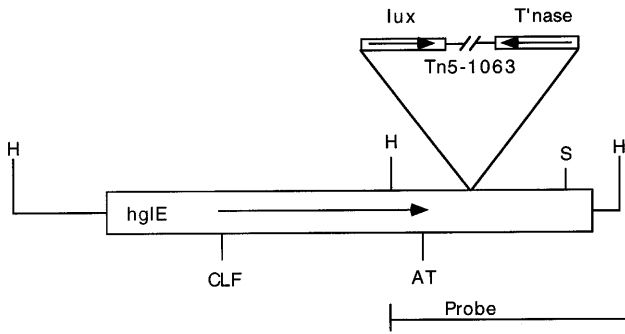


Fig. 3 Restriction and schematic map of the *hglE* region in *Nostoc punctiforme* strain 29133 showing the approximate site and orientation of the Tn5-1063 transposition. The size of the Tn5-1063 insertion is not drawn to scale, as indicated by the line breaks. Boxes indicate approximate location of genes and arrows indicate the direction of transcription of those genes. Locations of restriction sites are indicated above the *hglE* gene, and putative active sites are indicated below the gene. Probe is the 1.5-kb *Hind*III fragment used as a probe in Southern and Northern hybridizations (*lux* the bacterial luciferase reporter genes of Tn5-1063, *T'nase* the transposase gene of Tn5-1063, *H* *Hind*III restriction site, *S* *Spe*I restriction site, *CLF* chain length factor active site, *AT* acyl transferase active site)

degradation. Moreover, hybridization of this blot with a different probe [see Campbell et al. (1996)] did not show the smearing seen with the *hglE* probe.

The same 1.5-kb *Hind*III fragment was used as a probe to genomic DNA isolated from other cyanobacteria (Fig. 5). All heterocyst-forming strains show hybridizing bands at varying intensities, while no band was present in the lane of unicellular *Synechococcus* sp. strain PCC 7942 DNA. Hybridization of this blot with other probes verified that DNA is present in the *Synechococcus* strain 7942 lane (data not shown). The hybridization pattern between *Nostoc punctiforme* strains ATCC 29133 and PCC 73102, presumed to be identical organisms, is identical and distinctly different from the pattern in other *Nostoc* strains and *Anabaena* sp. strain PCC 7120. There are also other faint but noticeable hybridizing fragments in the two *N. punctiforme* samples.

Discussion

Strain UCD 307 is unable to fix nitrogen in air, but does express nitrogenase under anoxic or micro-oxic incubation conditions (Fox⁻). Fox⁻ rather than Fix⁻ (unable to fix N₂ under any condition; *sensu* Ernst et al. 1992) mutants appear to be the most common phenotype obtained – at least from transposon mutagenesis of *Nostoc punctiforme* strain 29133 (Cohen et al. 1994) and *Anabaena* strain 7120 (Ernst et al. 1992) – when survivors are screened for colony bleaching following N-step-down. The majority of these mutants have defects in the synthesis or assembly of components of the heterocyst envelope. For example, mutants of *Anabaena* strain 7120 that form an aberrant heterocyst envelope polysaccharide but synthesize glycolipids are Fox⁻ (Holland and Wolk 1990; Ernst et al. 1992); the

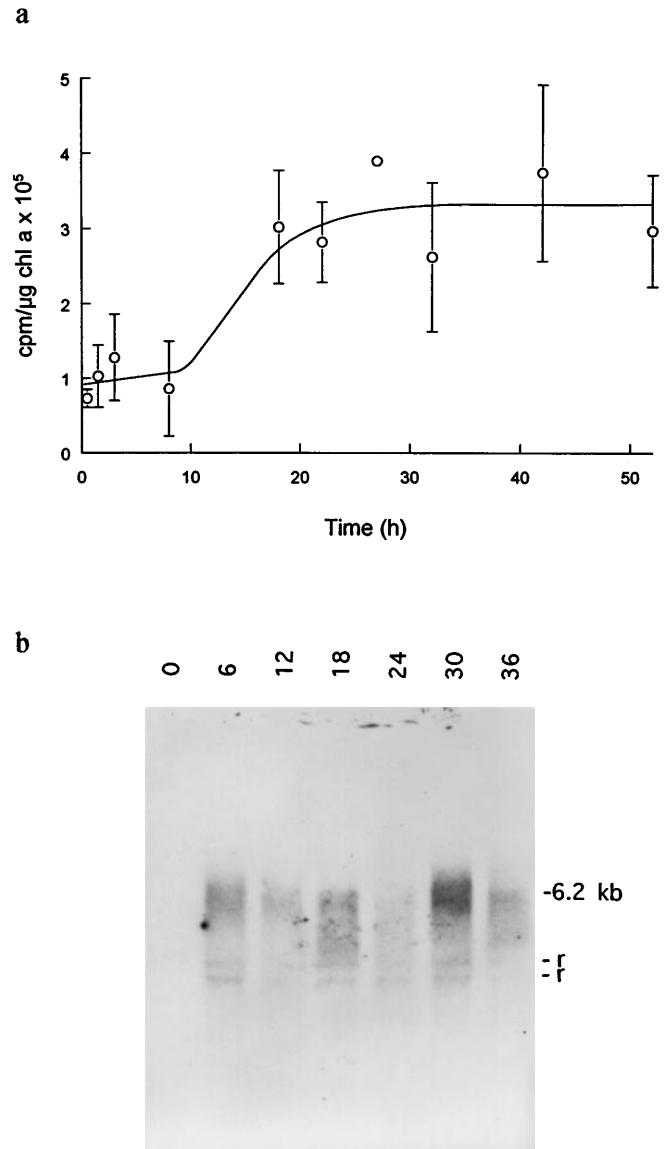


Fig. 4a, b Assays of combined nitrogen-dependent induction of *hglE* transcription. **a** Luciferase activity as a function of time following N-step-down; reported as 10⁵ counts/min normalized to *Nostoc punctiforme* strain UCD 307 biomass (μg chlorophyll *a*). The symbols are means of three replicates; error bars are SE. **b** Hybridization to an RNA blot with a 1.5-kb *Hind*III fragment containing the acyl transferase active site of *hglE* as a probe. RNA was isolated from *N. punctiforme* strain 29133 at the hours indicated across the top following N-step-down, and 10 μg of total RNA was loaded in each lane. The 6.2-kb band is based on the migration of RNA standards that were electrophoresed in the same gel and stained with ethidium bromide (*r* region of rRNA migration)

specific defect is thought to be a lack of structural support for glycolipid deposition. The polysaccharides of the heterocyst envelope are also synthesized in akinetes (Wolk et al. 1994), and since strain UCD 307 differentiates visibly normal akinetes (Cohen et al. 1994), we assume that the mutation in strain UCD 307 does not affect the synthesis of this structural component. A Fox⁻ phenotype occurs in

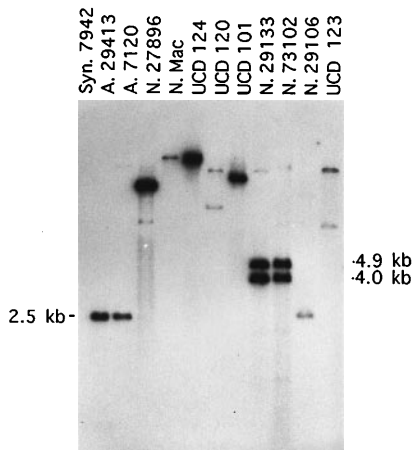


Fig. 5 Southern hybridization with the acyl transferase domain of *hglE* from *Nostoc punctiforme* strain 29133 to a blot of *SpeI*-digested genomic DNA isolated from heterocyst-forming cyanobacteria and one nonheterocyst-forming unicellular cyanobacterium. The probe is depicted in Fig. 3. Sizes of the migration bands were determined relative to *HindIII*- and *BstEII*-digested phage DNA electrophoresed in the same gel and stained with ethidium bromide [*Syn. 7942* unicellular *Synechococcus* sp. strain PCC 7942, *A. 29413* *Anabaena variabilis* strain ATCC 29413, *A. 7120* *Anabaena* sp. strain PCC 7120, *N. 27896* *Nostoc* sp. strain ATCC 27896, *N. Mac* *Nostoc* sp. strain MAC (sibling to PCC 8009), *N. 29133* *Nostoc punctiforme* strain 29133, *N. 73102* *N. punctiforme* strain PCC 73102 (parental to *N. 29133*); *N. 29106* *Nostoc* sp. strain ATCC 29106; *UCD* strains are various *Nostoc* spp. isolates from a soil habitat (*123*) or symbiotic associations with plants (*101, 120, 124*)

a mutant of *Anabaena* strain 7120 that synthesizes heterocyst glycolipids having a typical TLC migration property, but they are not assembled in the heterocyst envelope (Black et al. 1995); in this case, the affected gene product (*HglK*) is thought to function in localization of the glycolipids to the envelope region. These results also indicate that heterocyst glycolipids are completely synthesized in the cytoplasm and are then translocated through the cytoplasmic membrane and assembled interior to the polysaccharide layer of the heterocyst envelope. Strain UCD 307 clearly failed to complete the synthesis of any detectable glycolipid characteristic of heterocysts at any time following N-step-down and, thus, has an *Hgl* phenotype different from *hglK*. Since strain UCD 307 has no growth defect except when it is cultured under N_2 -fixing conditions, the *hglE* product is not essential for membrane lipid or LPS synthesis and must be specific for synthesis of the heterocyst glycolipids. *hglE* also appears to be found in other heterocyst-forming cyanobacteria examined but is lacking in a unicellular nonheterocyst-forming strain. Both mRNA and reporter gene analyses indicated that *hglE* is transcribed only after an approximately 6- to 10-h lag period following N-step-down. These kinetics of induction and taxonomic distribution are consistent with the involvement of *hglE* in the synthesis of a product (e.g., glycolipids) required for the formation of a mature heterocyst that functions in N_2 fixation in an oxic environment. However, since we do not know where *hglE* is located in

the 6.2-kb mRNA, a polar effect of the transposition in *hglE* on a possible downstream gene cannot be ruled out unequivocally.

Heterocyst glycolipids are very long-chained (typically C_{26} or C_{28}) di- or tri-hydroxy (or keto and mono- or di-hydroxy) saturated fatty acid chains, ether- (rarely, if at all, ester-) linked to a glucose (rarely galactose) residue (Wolk 1975; Soriente et al. 1993; Wolk et al. 1994). While the heterocyst glycolipids have been analyzed from only a few strains, at least six (or as many as nine) different structures have been described; the variability depends on the chain length and on uncertainty in the positioning and orientation of the keto and hydroxy groups (Bryce et al. 1972; Lambein and Wolk 1973; Davey and Lambein 1992; Soriente et al. 1993). Based on TLC R_f values in comparison to extracts from *Anabaena cylindrica* (Lambein and Wolk 1973), *N. punctiforme* strain 29133 appears to synthesize primarily glycolipid III (1-[0- α -D-glucopyranosyl]-3,25-hexacosanediol) (E. L. Campbell, unpublished work). If *N. punctiforme* strain 29133 or strain UCD 307 synthesized other glycolipids in concentrations too low to have been detected by TLC, it would appear either that the insertion in *hglE* also interfered with their synthesis or that insufficient amounts were synthesized to replace the structural contribution from glycolipid III.

The biochemical pathway for heterocyst glycolipid synthesis is unknown. Fatty acid and polyketide synthesis have common biosynthetic pathways involving acyl carrier protein, acetyl and malonyl transferase, ketoacyl synthase, ketoreductase, dehydrase, enoyl reductase and thioesterase activities [see Hopwood and Sherman (1990) and Hutchinson and Fujii (1995)]. Polyketide synthesis differs from fatty acid synthesis in that (1) the carbon chains of the final product are typically much longer (up to C_{50}), (2) propionate or butyrate can serve as donors for chain elongation, and (3) the keto and hydroxy groups of the growing chains are often not eliminated. Therefore, the carbon chain may have a variety of reactive groups that tend to yield complex cyclic structures with saturated and unsaturated bonds.

In terms both of end product and gene structure, the hydrocarbon chains of the heterocyst glycolipids and *hglE* appear to be intermediate between polyketides and fatty acids, and Type I (large multicatalytic domain proteins) and Type II (single catalytic domains in single peptides) synthases. The heterocyst glycolipid hydrocarbon chains are longer than those characteristic of the fatty acids of membrane lipids, lipoproteins, gram-negative bacterial lipopolysaccharides (Magnuson et al. 1993), or rhizobial Nod factors (Carlson et al. 1994); similar to polyketides all contain hydroxy groups, and some retain keto groups. Although they do not appear to form a cyclic structure and are not as complex as most polyketides, we have chosen to refer to their properties and synthesis as being more similar to those of polyketides than to those of fatty acids. *HglE* has the conserved amino acid sequences characteristic of the active sites of acyl transferases (ATs), which function in transfer of acetyl- or malonyl-CoA units to initiate or continue the elongation process, and sequences

characteristic of chain length factor (CLF) activity. With two active site domains, HglE is not typical of the Type II fatty acid synthases (FASs) that characterize fatty acid synthesis in most bacteria where the activities are in separately encoded polypeptides (Magnuson et al. 1993). Type II proteins containing both AT and β -ketoacyl synthase (KS) active site domains occur in some polyketide synthases (PKSs) (McDaniel et al. 1993), although a function for the AT active site has been questioned (Meurer and Hutchinson 1995). Thus, *hglE* also has an organizational similarity to some genes of polyketide synthesis.

All Type II CLFs that lack the active site cysteine (e.g., HglE) have been suggested to function as heterodimers (Siggaard-Andersen 1993). Since we can detect a mRNA species of 6.2 kb, i.e., approximately twice the size of *hglE*, it is possible that a second KS protein is encoded as part of an operon with *hglE*; tandem KS genes are relatively common in Type II PKSs (Hutchinson and Fujii 1995). Unfortunately, we were unable to subclone fragments that would allow us to conveniently characterize the DNA flanking *hglE*. We have no facile explanation for why the region was recalcitrant to subcloning, even in low-copy pACYC vectors, or why the cosmid clones were stable. An inability to clone acyl carrier protein genes in *Escherichia coli* has been observed previously (Hill et al. 1995). The observation that the only stable clones we recovered had portions of the gene inserted antiparallel to the transcriptional frame of *lacZ* in the screenable plasmids implies that products toxic to *E. coli* were made and that these were suppressed in the cosmids. A cluster of FAS or PKS genes has been identified in *Anabaena* strain 7120 (Bauer 1994); the analyzed region contains a partially sequenced ORF showing similarity to KS domains with a cysteine in the putative active site (*hglD*; Fig. 2b), followed 3' by a 3.3-kb ORF that has putative AT and CLF active site domains (*hglC*; Fig. 2a,b). This region in *Anabaena* strain 7120 also contains a gene with active site domains for putative acyl carrier protein and keto reductase functions [*hglB* (Bauer 1994), also called *hetM* (Black and Wolk 1994)]. The active site domains in *hglC* and *hglE* are very similar, but we suggest the two genes are likely to be analogues rather than homologues. The *hglE* probe hybridizes strongly to a band in genomic DNA from *Anabaena* strain 7120 (Fig. 5), and *hglC* hybridizes to *N. punctiforme* strain 29133 genomic DNA (data not shown); probes of the two genes, however, hybridize to different restriction fragments in *N. punctiforme* strain 29133 genomic DNA; under stringent conditions, *hglC* does not hybridize to the cosmids carrying *hglE*; the predicted gene products share only 48% amino acid sequence similarity (25% identity) and are of different molecular mass. We interpret these results to suggest that multiple KS proteins, and perhaps associated modifying enzymes of chain elongation, occur as multiprotein complexes in synthesis of the polyketides of heterocyst glycolipids. Whether different complexes are specific for a single polyketide or more than one synthase complex is required for synthesis of one polyketide remains to be resolved.

Acknowledgements We thank Kari Hagen for help in the preparation of the figures and, together with Tom Hanson and Francis Wong, for critical review of the manuscript. We thank W. J. Buikema for a gift of DNA containing *Anabaena* strain 7120 *hglC* and *hglD* sequences. This work was supported by US National Science Foundation grants IBN 92-06139 and 95-14787.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DL (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Bauer C (1994) PhD Thesis, University of Chicago, Chicago, Illinois, USA
- Berg CM, Berg DE (1987) Uses of transposable elements and maps of known insertions. In: Neidhardt FC, Ingraham JL, Low KB, Magasanik B, Schaechter M, Umberger HE (eds) *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, DC, pp 1071–1109
- Black TA, Wolk CP (1994) Analysis of a *Het⁻* mutation in *Anabaena* sp. strain PCC 7120 implicates a secondary metabolite in the regulation of heterocyst spacing. *J Bacteriol* 176:2282–2292
- Black K, Buikema WJ, Haselkorn R (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
- Bryce TA, Welti D, Walsby AE, Nichols BW (1972) Monohexoside derivatives of long-chained polyhydroxy alcohols; a novel class of glycolipid specific to heterocystous algae. *Phytochemistry* 11:295–302
- Cai Y, Wolk CP (1990) Use of a conditionally lethal gene in *Anabaena* sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. *J Bacteriol* 172:3138–3145
- Campbell EL, Meeks JC (1992) Evidence for plant-mediated regulation of nitrogenase expression in the *Anthoceros-Nostoc* symbiotic association. *J Gen Microbiol* 138:473–480
- Campbell EL, Hagen KD, Cohen MF, Summers ML, Meeks JC (1996) The *devR* gene product is characteristic of receivers of two-component regulatory systems and is essential for heterocyst development in the filamentous cyanobacterium *Nostoc* sp. strain ATCC 29133. *J Bacteriol* 178:2037–2043
- Capone DG, O'Neil JM, Zehr J, Carpenter EJ (1990) Basis for diel variation in nitrogenase activity in the marine planktonic cyanobacterium *Trichodesmium thiebautii*. *Appl Environ Microbiol* 56:3532–3536
- Carlson RW, Price NPJ, Stacey G (1994) The biosynthesis of rhizobial lipo-polysaccharide nodulation signal molecules. *Mol Plant Microbe Interact* 7:684–695
- Cohen MF, Wallis JG, Campbell EL, Meeks JC (1994) Transposon mutagenesis of *Nostoc* sp. strain ATCC 29133, a filamentous cyanobacterium with multiple cellular differentiation alternatives. *Microbiology* 140:3233–3240
- Davey MW, Lambein F (1992) Semipreparative isolation of individual cyanobacterial heterocyst-type glycolipids by reverse-phase high-performance liquid chromatography. *Anal Biochem* 206:226–230
- Donadio S, Staver MJ, McAlpine JB, Swanson SJ, Katz L (1991) Modular organization of genes required for complex polyketide biosynthesis. *Science* 252:675–679
- Elhai J, Wolk CP (1990) Developmental regulation and spatial pattern of expression of the structural genes for nitrogenase in the cyanobacterium *Anabaena*. *EMBO J* 9:3379–3388
- Enderlin CS, Meeks JC (1983) Pure culture and reconstitution of the *Anthoceros-Nostoc* symbiotic association. *Planta* 158:157–165
- Ernst A, Black T, Cai Y, Panoff J-M, Tiwari DN, Wolk CP (1992) Synthesis of nitrogenase in mutants of the cyanobacterium *Anabaena* sp. strain PCC 7120 affected in heterocyst development or metabolism. *J Bacteriol* 174:6025–6032

- Fay P (1992) Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev* 56:340–373
- Gallon JR (1992) Reconciling the incompatible: N₂ fixation and O₂. *New Phytol* 122:571–609
- Hill RB, MacKenzie KR, Flanagan JM, Cronan JE Jr, Prestegard JH (1995) Overexpression, purification, and characterization of *Escherichia coli* acyl carrier protein and two mutant proteins. *Protein Expr Purif* 6:394–400
- Holland D, Wolk CP (1990) Identification and characterization of *hetA*, a gene that acts early in the process of morphological differentiation of heterocysts. *J Bacteriol* 172:3131–3137
- Hopwood DA, Sherman DH (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu Rev Genet* 24:37–66
- Hutchinson CR, Fujii I (1995) Polyketide synthase gene manipulations: a structure-function approach in engineering novel antibiotics. *Annu Rev Microbiol* 49:201–238
- Lambein F, Wolk CP (1973) Structural studies on the glycolipids from the envelope of the heterocyst of *Anabaena cylindrica*. *Biochemistry* 12:791–798
- Magnuson K, Jackowski S, Rock CO, Cronan JE Jr (1993) Regulation of fatty acid biosynthesis in *Escherichia coli*. *Microbiol Rev* 57:522–542
- McDaniel R, Ebert-Khosla S, Hopwood DA, Khosla C (1993) Engineered biosynthesis of novel polyketides. *Science* 262:1546–1550
- Meurer G, Hutchinson CR (1995) Functional analysis of putative β -ketoacyl:acyl carrier protein synthase and acyltransferase active site motifs in a Type II polyketide synthase of *Streptomyces glaucescens*. *J Bacteriol* 177:477–481
- Murry MA, Wolk CP (1989) Evidence that the barrier to the penetration of oxygen into heterocysts depends upon two layers of the cell envelope. *Arch Microbiol* 151:469–474
- Parkinson JS, Kofoed EC (1992) Communication modules in bacterial signaling proteins. *Annu Rev Genet* 26:71–112
- Rippka R, Herdman M (1992) Pasteur culture collection of cyanobacteria in axenic culture. Institut Pasteur, Paris
- Rippka R, Stanier RY (1978) The effects of anaerobiosis on nitrogenase synthesis and heterocyst development by Nostocacean cyanobacteria. *J Gen Microbiol* 105:83–94
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Siggaard-Andersen M (1993) Conserved residues in condensing enzyme domains of fatty acid synthases and related sequences. *Protein Seq Data Anal* 5:325–335
- Soriente A, Gambacorta A, Trincone A, Sili C, Vincenzini M, Sodano G (1993) Heterocyst glycolipids of the cyanobacterium *Cyanospira rippkae*. *Phytochemistry* 33:393–396
- Summers ML, Wallis JG, Campbell EL, Meeks JC (1995) Genetic evidence of a major role for glucose-6-phosphate dehydrogenase in nitrogen fixation and dark growth of the cyanobacterium *Nostoc* sp. strain ATCC 29133. *J Bacteriol* 177:6184–6194
- Theil T, Lyons EM, Erker JC, Ernst A (1995) A second nitrogenase in vegetative cells of a heterocyst-forming cyanobacterium. *Proc Natl Acad Sci USA* 92:9358–9362
- Walsby AE (1985) The permeability of heterocysts to the gases nitrogen and oxygen. *Proc R Soc Lon B* 226:345–366
- Wilcox M, Mitchison GJ, Smith RJ (1973) Pattern formation in the blue-green alga, *Anabaena*. I. Basic mechanisms. *J Cell Sci* 12:707–725
- Wolk CP (1975) Differentiation and pattern formation in filamentous blue-green algae. In: Gerhardt P, Costilow RN, Sadoff HL (eds) *Spores*, vol 6. American Society for Microbiology, Washington DC, pp 85–96
- Wolk CP, Cai Y, Panoff J-M (1991) Use of a transposon with luciferase as a reporter to identify environmentally responsive genes in a cyanobacterium. *Proc Natl Acad Sci USA* 88:5355–5359
- Wolk CP, Ernst A, Elhai J (1994) Heterocyst metabolism and development. In: Bryant DA (ed) *The molecular biology of cyanobacteria*. Kluwer, Dordrecht, pp 769–823