Cloning and transcriptional analysis of the *nifUHDK* genes of *Trichodesmium* sp. IMS101 reveals stable *nifD*, *nifDK* and *nifK* transcripts

Benny Dominic, Yi-Bu Chen and Jonathan P. Zehr

Author for correspondence: Jonathan P. Zehr. Tel: +1 518 276 8386. Fax: +1 518 276 2162.
e-mail: zehrj@rpi.edu

Department of Biology, Rensselaer Polytechnic Institute, 110 8th St, Troy, NY 12180-3590, USA

***Trichodesmium* spp. are marine filamentous, non-heterocystous cyanobacteria capable of aerobic nitrogen fixation. In this study, the nitrogenase structural genes (*nifHDK*) and *nifU* gene of *Trichodesmium* sp. IMS101 were cloned and sequenced. The *Trichodesmium* sp. IMS101 *nifH*, *nifD* and *nifK* amino acid sequences showed only 79%, 66% and 68% identity, respectively, to those of *Anabaena* sp. strain PCC 7120. A potential transcription start site for *nifH* was found 212 bases upstream of the *nifH* start codon. Promoter-like nucleotide sequences upstream of the transcription start site were identified that were very similar to those identified for the nitrogenase genes of *Anabaena* spp. Sequence analysis revealed regions of DNA that may form stem–loop structures in the intercistronic regions downstream of *nifH* and *nifD.* RNA analysis by Northern hybridization revealed the presence of transcripts corresponding to *nifH*, *nifHD* and *nifHDK*. Surprisingly, Northern hybridization also revealed the presence of transcripts that corresponded to *nifD*, *nifDK* and *nifK*, which have not been previously reported as transcripts in contiguous *nifHDK* genes of cyanobacteria. Transcription of the *nifHDK* genes was not significantly repressed in the presence of nitrate at a final concentration of 20 mM or at oxygen concentrations of up to 40%, whereas ammonium and urea inhibited *nifHDK* transcription. The transcription of the *nifHDK* genes was not affected by darkness, which suggests that transcription of these genes in *Trichodesmium* is not directly regulated by light.**

Keywords: cyanobacteria, *Trichodesmium* sp., nitrogen fixation, *nifHDK* gene expression, *nifH* promoters

INTRODUCTION

It is estimated that *Trichodesmium*, the most abundant free-living diazotroph of tropical and subtropical oceans (Carpenter & Romans, 1991), contributes significantly to global nitrogen fixation (Michaels *et al*., 1996; Gallon *et al*., 1996; Capone *et al*., 1997). Interestingly, filamentous non-heterocystous marine cyanobacteria of the genus *Trichodesmium* present a paradox with respect to paradigms of regulation of nitrogen fixation in photosynthetic prokaryotes (Gallon *et al*., 1996). Nitrogenase, the enzyme that catalyses nitrogen fixation, is highly sensitive to inactivation by oxygen (Fay, 1992; Postgate, 1982). The most common mechanisms for protection in cyanobacterial diazotrophs are spatial and/or temporal separation of oxygen-sensitive nitrogen fixation from oxygenic photosynthesis (Bergman *et al*., 1997; Gallon, 1992). By contrast, *Trichodesmium* has the unique ability to fix nitrogen aerobically while photosynthesizing, without any obvious separation of these two incompatible processes (Capone *et al*., 1997). Furthermore, *Trichodesmium* will only fix nitrogen during the day, apparently cued by a circadian rhythm (Chen *et al*., 1996). Recently, slow-growing cultures of *Trichodesmium* sp. IMS101, grown in defined liquid medium, have become available (Chen *et al*., 1996; Prufert-Bebout *et al*., 1993), which facilitates basic molecular biology investigations.

Nitrogenase is composed of the iron protein (dinitrogenase reductase) encoded by the *nifH* gene, and the molybdenum iron protein (dinitrogenase) encoded by the *nifDK* genes. In many nitrogen-fixing microorganisms the *nifHDK* genes are contiguous. Known

The GenBank accession numbers for the nucleotide sequences reported in this paper are AF016484 and AF055034.

exceptions are *Bradyrhizobium japonicum*, *Rhizobium* sp. strain IRc 78 and *Frankia* strain FaC1 (Alvarez-Morales *et al.*, 1986; Ligon & Nakas, 1987). In all the nitrogen-fixing non-heterocystous cyanobacteria examined, including *Trichodesmium*, the *nif* structural genes are contiguous (Bergman *et al.*, 1997; Zehr *et al.*, 1991). In the vegetative cells of the heterocystous cyanobacterium *Anabaena* sp. PCC 7120, the *nifD* gene is interrupted by an 11 kb DNA element that is excised and removed upon differentiation of nitrogen-fixing heterocystous cells, making the rearranged *nifHDK* genes contiguous (Golden *et al.*, 1985). This rearranged *nifHDK* operon is transcribed from a promoter upstream of the *nifH* gene as a single transcriptional unit (Brusca *et al.*, 1989; Golden *et al.*, 1985, 1991; Haselkorn *et al.*, 1986).

The *nif* genes in many non-cyanobacterial diazotrophs, including *Klebsiella pneumoniae*, are transcribed by an RNA polymerase that utilizes a sigma (σ) factor known as RpoN or σ^{54}. The RpoN-dependent promoters usually have a highly conserved -24, -12 sequence, instead of the consensus sequence located around positions -35 and -10 recognized by σ^{70} and most other σ factors (Merrick, 1992). Activation of *nif* promoters in many non-cyanobacterial diazotrophs is NifA dependent (Merrick, 1992). The *nifA* product is required for transcription of all *nif* operons (except *nifLA*) and the *nifL* product turns off transcription of all *nif* operons other than *nifLA*, in response to oxygen and fixed nitrogen (Collins & Brill, 1985; Roberts & Brill, 1980). The *nifA* protein activates transcription by binding to the consensus DNA sequence TGT (N_{10})ACA, which is also known as the upstream activator sequence (UAS) and is typically located between 80 and 150 bp upstream of the transcription start site (Dixon *et al.*, 1987; Merrick, 1992). The UAS is known to be orientation independent and positively influences transcription of genes located 2–4 kb upstream or downstream (Merrick, 1992). The transcription of the *nifLA* operon itself is under the control of a regulatory cascade involving the *ntrB* and *ntrC* gene products and the P_{II} protein (Magasanik, 1988). Though the P_{II} protein has been identified in cyanobacteria, investigations have failed to identify homologues for *ntrB*, *ntrC*, *rpoN* and *nifLA* (Flores & Herrero, 1994; Merrick, 1992; Tandeau de Marsac & Houmard, 1993). However, the search for the global nitrogen regulator led to the identification of the *ntcA* gene, which belongs to the *crp* gene family (Vega-Palas *et al.*, 1990). Hence, the regulation of *nif* genes in cyanobacteria is thought to be under the control of the global nitrogen regulator *ntcA* (Luque *et al.*, 1994; Flores & Herrero, 1994). In general, however, cyanobacterial *nif* operons for which the transcription start site(s) have been characterized do not possess a consensus NtcA-binding site [GTA (N_8)TAC or TGT(N_{9-10})ACA] at the proposed -35 position (Flores & Herrero, 1994; Vega-Palas *et al.*, 1992) of the promoter or upstream of the transcription start site. It has also been documented that under nitrogen-fixing conditions *ntcA* expression (at the transcriptional level) was inversely proportional to *nifHDK* transcript abundance (Bradley & Reddy, 1997).

In this study, the *nifU* and *nifHDK* genes of *Trichodesmium* sp. IMS101 (Chen *et al.*, 1996; Prufert-Bebout *et al.*, 1993) were cloned and sequenced to compare the nitrogenase proteins and *nif* regulatory region of *Trichodesmium* to other cyanobacterial diazotrophs. The *nifHDK* transcripts were also characterized, and expression of *nifHDK* in response to different nitrogen sources, oxygen concentration and light intensity was examined.

METHODS

Strains and growth conditions. Liquid cultures of *Trichodesmium* sp. IMS101 (Chen *et al.*, 1996; Prufert-Bebout *et al.*, 1993) were grown in YBCII medium (Chen *et al.*, 1996) under a regular 12 h light/12 h dark regime, at 26 °C. The 12 h light regime is the subjective light phase which begins at 1000 and ends at 2200 h. For ammonium, urea and nitrate experiments, cultures were transferred to YBCII medium supplemented with 2 mM NH_4Cl, 2 mM urea or 20 mM $NaNO_3$, respectively, for 2 h. For the light intensity experiment, YBCII medium-grown cultures were exposed to zero (complete darkness), 100 and 300 $\mu E\ m^{-2}\ s^{-1}$ light intensity, respectively, for 3 h. For the oxygen experiment, the head-space of each incubation flask was flushed and replaced with air containing oxygen at the indicated concentration (20%, v/v, in control; 40%, v/v, or 80%, v/v, in experiments) and the flasks were sealed. The duration of the oxygen experiment was 3 h.

Cultures of *Escherichia coli* strain XL-1 Blue MRF′ (Stratagene) were grown and/or maintained in liquid LB medium or on LB agar (1·5% w/v Difco agar) containing 12·5 μg tetracycline ml^{-1} at 37 °C.

DNA extraction and genomic library construction. Genomic DNA from *Trichodesmium* sp. IMS101 was extracted following the protocol of Zehr *et al.* (1991) and purified using caesium chloride density-gradient centrifugation (Sambrook *et al.*, 1989). The genomic DNA was then subjected to *Sau*3AI partial restriction digestion, using standardized conditions to obtain a size distribution between 23 and 14 kb. The partially digested DNA was cloned into Lambda GEM-11 vector (Promega) using the *Xho*I half-site strategy (Promega). The genomic library was amplified and screened for *nif* genes.

Screening of the library for *nif* genes. The λ library plaques were lifted with Hybond-N membranes (Amersham). Membranes were denatured for 2 min in denaturation solution (1·5 M NaCl and 0·5 M NaOH), neutralized for 5 min in neutralization solution (1·5 M NaCl and 0·5 M Tris/HCl pH 8·0), and rinsed for 30 s in a solution of 0·2 M Tris/HCl (pH 7·5) and 2 × SSC. The membranes were vacuum-dried for 90 min at 80 °C. A 359 bp *nifH* DNA fragment cloned from *Trichodesmium* sp. (Zehr & McReynolds, 1989) was labelled with [α^{32}P]dCTP, using the Decaprime kit (Ambion) following the manufacturer's protocols, and used to screen the library using standard protocols (Sambrook *et al.*, 1989). A positive plaque was identified and purified by secondary and tertiary rounds of screening. The size of the insert in the recombinant-phage clone containing the *nifH* gene was estimated to be approximately 13 kb.

Subcloning. The recombinant phage clone was amplified by culturing on a large scale and recombinant DNA was extracted using a phage DNA extraction kit (Qiagen). Recombinant-

Table 1. Plasmids containing *Trichodesmium* sp. IMS101 gene(s) used or reported in this study

The source of the 5·1 kb insert for pBD5N is the *λ-nif* clone obtained by screening *Trichodesmium* IMS101 genomic library. The source of insert for pBD511, pBD52, pBD509 and pBD535 is the 5·1 kb insert for pBD5N (see Methods). The locations of *Hin*dIII and *Sac*I restriction sites are shown in Fig. 5(a).

Plasmid	Properties	Reference
pBD5N	pUC18 containing 5·1 kb insert in *Sac*I site	This work
pBD511	pUC18 containing 1·1 kb insert in *Hin*dIII site	This work
pBD52	pUC18 containing 2 kb insert in *Hin*dIII site	This work
pBD509	pUC18 containing 0·9 kb insert in *Hin*dIII site	This work
pBD535	pUC18 containing 1·1 kb insert between *Hin*dIII and *Sac*I site	This work
pBDHU	pGEM-T vector, 0·7 kb insert obtained by PCR amplification	This work
pTR	M13mp19 vector, 359 bp insert obtained by PCR amplification	Zehr & McReynolds (1989)

phage DNA insert fragments were obtained using the restriction enzymes *Sac*I and *Hin*dIII (Promega), and the fragments purified after gel electrophoresis using a Qiaex II Gel Extraction kit (Qiagen). These DNA fragments were ligated to *Sac*I-digested pUC18 and *Hin*dIII-digested pUC18, respectively, using T4 ligase (Promega) and standard conditions (Sambrook *et al*., 1989). *E. coli* strain XL1B-MRF′ (Stratagene) was transformed with the ligation mixtures. Transformants were selected on LB agar plates with ampicillin (100 μg ampicillin ml^{-1}), X-Gal and IPTG. A 5·1 kb *Sac*I fragment was extracted from a gel and subcloned in the plasmid vector pUC18 (identified as pBD5N). Recombinant-plasmid DNA was extracted from pBD5N, and *Hin*dIII fragments used to create four minor subclones (pBD511, pBD52, pBD509 and pBD535) (Table 1).

PCR amplification and cloning of the *nifU* gene. Since the λ clone isolated from the genomic library did not contain DNA beyond 222 bp upstream of the *nifH* start codon, a PCR-based strategy was adopted to expedite cloning of the region upstream of *nifH*. A 17 base long degenerate oligonucleotide primer, BDNU17, was synthesized based on the amino acid sequence VMGQEA. The primers BDNU17 [5′ GT(GATC)-ATGGG(GATC)CA(AG)GA(AG)GC 3′] and HPRO-17 were used as forward and reverse primers, respectively, in the PCR reaction. The PCR conditions were as described by Zehr & McReynolds (1989) using *Trichodesmium* sp. IMS101 genomic DNA as template. A 0·7 kb DNA fragment was amplified using the degenerate *nifU* primer (BDNU17) and the *Trichodesmium nifH* upstream sequence-specific primer HPRO-17 (5′ CAGCAACATTAGAGTGC 3′) complementary to +42 to +1 and −177 to −193 of the *nifH* coding sequence. This DNA fragment was cloned in pGEM-T vector (Promega) and the plasmid identified as pBDHU.

DNA sequencing. The major subclone pBD5N was partially sequenced to determine orientation of the clone. The pBD5N-derived minor subclones (pBD511, pBD52, pBD509 and pBD535) and pBDHU were then used for subsequent complete sequencing of both DNA strands. The plasmids were sequenced by the primer walking method. The regions at all *Hin*dIII restriction sites were sequenced on plasmid pBD5N to ensure correct orientation and completeness of sequence generated from the subclones. Sequencing was performed by the University of Maine DNA Sequencing Facility (Orono, USA).

RNA isolation and Northern blot analysis. Cells were harvested and lysed in a modified STET buffer (only 25 mM EDTA and with 100 mM DTT). STET buffer (Ausubel *et al*., 1990) contains 8% (w/v) sucrose, 5% (v/v) Triton X-100, 50 mM EDTA and 50 mM Tris pH 7. The total lysate was extracted once with equal-volume low-pH (4·5) equilibrated phenol/chloroform (1:1, v/v) followed by extraction with TE pH 8 saturated phenol/chloroform. The supernatant was then extracted with an equal volume of chloroform. RNA was ethanol precipitated and resuspended in nuclease-free water. The RNA extracted from equivalent amounts of biomass (1 μg chlorophyll *a*) was loaded on 1% (w/v) formaldehyde gel, transferred to a Nytran membrane (Schleicher & Schuell), and Northern hybridization carried out essentially following the protocols in Ausubel *et al*. (1990). Probes specific to *nifH* and *nifK* genes were obtained by *Hin*dIII restriction and gel extraction of appropriate DNA fragments (as described previously) from pBD5N. A gel-purified, PCR-amplified 1050 bp *nifD* DNA fragment was used as the *nifD*-specific probe. PCR conditions were as described by Zehr & McReynolds (1989) using pBD5N as template and the synthetic oligonucleotides 5′ GGTGGAGACAAAAAGCT 3′ and 5′ AGTTGGGCTATTGAGAG 3′ as the forward and reverse primers, respectively. The probes were labelled with [$\alpha^{32}P$] dCTP using the Decaprime kit (Ambion), following the manufacturer's suggestions. The same membrane was sequentially hybridized to the *nifK*, *nifH* and *nifD* probes to characterize the transcripts, or only to the *nifH* probe in the experiments to study *nifHDK* expression. Hybridizations were incubated overnight in hybridization solution containing 50% (v/v) formamide, 5× SSC, 0·1% SDS and sonicated denatured salmon sperm DNA (100 μg ml^{-1}). Membranes were washed at room temperature twice in 2× SSC/0·1% SDS solution and then twice in 1× SSC/0·1% SDS solution followed by two high-stringency washes at 65 °C in 0·1× SSC/0·1% SDS

solution. Each wash was performed for 10 min. The probes were stripped between hybridizations by rinsing the membrane in boiling 0·1× SSC and 0·1% SDS solution, at 42 °C.

Primer extension analysis. To determine the *nifH* transcription start site, 25 µg total RNA (free of contaminating DNA) was used for the template of a reverse transcription reaction using a synthetic oligonucleotide (PXT-1) primer (5′ TAGCCATTGCAGCTAGAGTATTCTGAGAAGTAGTGGACTTAC 3′) complementary to nucleotide positions +79 to +37 of the *nifH* coding sequence. After hybridization, dNTPs minus dCTP (2·5 µl of a 0·5 mM stock solution), 25 µCi ($9{\cdot}25 \times 10^5$ Bq) [α^{32}P]dCTP, reaction buffer and AMV reverse transcriptase (Promega) were added at 42 °C (the total reaction volume was adjusted to 25 µl). The reaction was stopped after 45 min by adding 0·4 vol. stop solution (USB Sequenase kit). The reaction products were heat-denatured and separated on a sequencing gel, followed by autoradiography. The primer extension experiments were repeated with primers PXT-2 (5′ CTTACCGATACCACCTTTTCCGTAAAATGCAATCTGACGCAT 3′) and HPRO-17 (5′ CAGCAACATTAGAGTGC 3′) complementary to +42 to +1 and −177 to −193 of the *nifH* coding sequence, using the same reaction conditions. To determine the size of the extension product, dideoxynucleotide sequencing reactions were run on the same gel using the same primers and pBD5N as template. The sequencing reactions were performed using Sequenase, following the manufacturer's protocols (USB).

RESULTS

DNA sequence analysis

The nucleotide sequence of the *nifH* coding region reported by Sroga *et al*. (1996) for natural populations of *Trichodesmium thiebautii* was 95·2% identical to the sequence reported in this study. The *Trichodesmium* sp. IMS101 *nifH*-deduced amino acid sequence of 296 amino acid residues has striking differences compared to the carboxy-terminus of the deduced 294 amino acid residues of the *T. thiebautii nifH* sequence (Sroga *et al*., 1996). The deduced amino acid sequences for the *nifH*, *nifD* and *nifK* genes were compared to those of other known diazotrophs (Table 2). The *Trichodesmium* sp. IMS101 *nifH*, *nifD* and *nifK* amino acid sequences were 79%, 66% and 68% identical, respectively, to those of *Anabaena* sp. PCC 7120 *nifH*, *nifD* and *nifK*. The nucleotide sequence of the *Trichodesmium* sp. IMS101 *nifH* upstream and *nifHDK* intergenic regions is given in Fig. 1. A putative ribosome-binding site with reasonable similarity to a cyanobacterial ribosome-binding site (Stricker *et al*., 1997; Thiel, 1993) was identified in front of the coding sequence of the *nifH*, *nifD* and *nifK* genes (Fig. 1). The *nifUHDK* genes are contiguous, indicating unidirectional transcription. The sequence, TGT (N_{10})ACA, centred about 157 bases upstream of the *nifH* translational start codon (Fig. 1a), but downstream of the transcript start site, is identical to the consensus-binding sequence [TGT(N_{10})ACA] for the product of *nifA* (Dixon *et al*., 1987; Merrick, 1992). There are regions of DNA sequence that may form stem–loop structures in the mRNA. One (Fig. 1b) sequence starts 22 bases beyond the stop codon (TAA) of *nifH*. This sequence is followed by eight 'A' residues. The other (Fig. 1c) possible stem–loop structure starts 95 bases beyond the stop codon of *nifD*.

Transcription start site mapping

The results of the three primer extension experiments indicated that a transcription start site for *nifH* is located 212 bases upstream of the *nifH* start codon (Fig.

Table 2. Comparison of *Trichodesmium* sp. IMS101 *nifHDK* deduced amino acid sequences with those of *T. thiebautii* (only *nifH* is available), *Anabaena* sp. strain PCC 7120, *Azotobacter vinelandii*, *Clostridium pasteurianum*, *K. pneumoniae* and *Synechococcus* sp. strain PCC 8801

nif **gene**	**GenBank accession no.**	**Percentage identity**
nifH (*T. thiebautii*)	U23507	90
nifH (*Anabaena* sp. PCC 7120)	P00457	79
nifH1 (*A. vinelandii*)	P00459	71
nifH1 (*C. pasteurianum*)	P00456	65
nifH (*K. pneumoniae*)	P00458	71
nifH (*Synechococcus* sp. strain PCC 8801)	U22146	79
nifD (*Anabaena* sp. PCC 7120)	P00464	66
nifD (*A. vinelandii*)	P07328	66
nifD (*C. pasteurianum*)	P00647	42
nifD (*K. pneumoniae*)	P00466	67
nifD (*Synechococcus* sp. strain PCC 8801)	U22146	75
nifK (*Anabaena* sp. PCC 7120)	P00468	68
nifK (*A. vinelandii*)	P07329	55
nifK (*C. pasteurianum*)	P11347	41
nifK (*K. pneumoniae*)	P09772	50
nifK (*Synechococcus* sp. strain PCC 8801)	U22146	70

(a)

```
GTCTTCAATCAAATTTAAAAAATTAACTGTTGAGTGAAAGCCACAAAAAA

CCAATAAAGCTAAAACATTAATCACAATTTTATATATGCAGGTAATACCT
         ↓
GATCAAGCTTTAGGGCGAGCGCCTCCGGGCACTCTAATGTTGCTGCTAGA

TACACGCTGTTCTTAAGTGAACAAAATTACGAATGTTGTGCTGTAAAGTC

AAAAGCCCTAGAAACAGCAGCCACCAAACATAAAATCTAAGGAAACCCTT

ACACAGCAAAAAAAAATACATTCGTTCATAATTTTAAGTCAAATACGCTA

ACAAAACGGAGAGAATTTATCatg-------------------------nifH ORF//
       ******
```

(b)

```
GCTTAAAATCTTTGAGCACAAAAGTGGGGATGTACAAAATACACCCCATC

TAAAAAAAACAATTAGTGATGAGTTTTTCTTAATTAATTTGTAGGGGTAG

ATTTTGTATATACTCAAAAGGTAACTGCCCCTATTAATACAAAACTCAAA

TACCAAAATATTCAAATATCCGAAAGGAGATATGTAACatg---------nifD ORF//
                         *****
```

(c)

```
AACCAGCCGCTACGCGGAAGTTATAGATTTAAAAGTTAAAAGTTAAAAGT

ATTAATTTAACCTTTTGGCTTTGAAGATCAATTTAGGAATGGAAATGGGG

TAATGCTATTATTGCTACGGCAACCCCAATACACAAACAAAAGATTTGAC

TGGGAATTGGGAGTTAGAAGTTGTATTTCTTTCTTCCCCCTTCCTTCTGA

AGAATACTATAGAGGATAAACACCatg-----------------------nifK ORF//
           *****
```

Fig. 1. Nucleotide sequences (a) upstream of the *nifH* translational start, (b) between *nifH* and *nifD* and (c) between *nifD* and *nifK* of *Trichodesmium* sp. IMS101. The translational start is indicated by lower case letters. The arrow (↓) indicates the transcription start site determined in this study. The presumed ribosome-binding sites are indicated by asterisks. The putative *nifA*-binding site-like sequence, TGT(N_{10})ACA, is indicated by a bold underline. The nucleotide sequences that may form stable stem–loop structures in the mRNA are shown in bold type and overlined. The putative −10 and −35 sequences are underlined.

2). There were additional bands in the primer extension reaction that corresponded to base positions closer to the *nifH* start codon. Although a few of the prominent bands could potentially be alternative transcription start sites, these are more likely due to RNA secondary structures that caused premature termination of the primer extension reactions. A putative −10 sequence (TAATAC) was located 11 bases upstream of the transcription start site and a putative −35 sequence (CACAAT) was located 31 bases upstream of the transcription start site (Fig. 3).

Transcription of *nifHDK* operon/genes

Transcription of the *nifHDK* genes in a transcriptionally active nitrogen-fixing culture of *Trichodesmium* sp. IMS101 was studied by Northern analysis. Total RNA from a mid-day *Trichodesmium* sp. IMS101 culture was hybridized with probes for *nifH* (BD511), *nifD* (BD1000), and *nifK* (BD509) (Fig. 4). The relative location of these probes is shown in Fig. 5(b). Hybridization with the *nifH* probe showed three distinct bands that correspond to approximately 1·1, 2·8 and 4·5 kb. The 1·1 kb signal had the highest signal intensity and the *nifH* probe did not hybridize to any bands less than 1·1 kb. The discrete bands indicate that sample RNA was of high quality with minimal degradation. Hybridization with the *nifD* probe showed four distinct bands that corresponded to approximately 2·8, 3·4 and 4·5 kb with varying signal intensity. The *nifD* probe also hybridized and produced a slightly diffuse band centred at 1·7 kb. Hybridization with the *nifK* probe showed two distinct bands that corresponded to 3·4 and 4·5 kb and a slightly diffuse band centred at 1·7 kb, which is very similar to that observed with the *nifD* probe.

Transcription of *nifHDK* in the presence of different nitrogen sources, and varying oxygen concentrations and light intensities

Northern hybridization experiments revealed that the *nifHDK* transcript signal intensities on the autoradiogram decreased significantly after 2 h when the *Trichodesmium* culture medium contained 2 mM NH_4Cl or 2 mM urea. It was also observed that the NH_4Cl treatment led to the largest reduction in transcript signal intensity (Fig. 6b). Nitrogenase activity as measured by acetylene reduction also showed significantly lower values after 3 h treatment with the different nitrogen sources mentioned above and the lowest value for acetylene reduction was obtained from cultures treated with NH_4Cl (data not shown). The decrease in *nifHDK* transcripts was minimal when the culture medium contained 20 mM $NaNO_3$. When the *Trichodesmium* culture was exposed to air with 40% oxygen for 3 h, the *nifHDK* transcript abundance (Fig. 6c) was approximately the same as the control. By contrast, at 80% oxygen concentration, *nifHDK* transcripts were absent (Fig. 6c) and acetylene reduction was low (data not shown). Exposure to complete darkness or 300 $\mu E\ m^{-2}\ s^{-1}$ light intensity did not result in any significant change in *nifHDK* transcript abundance (Fig. 6a) when compared with that of the control culture maintained at a light intensity of 100 $\mu E\ m^{-2}\ s^{-1}$.

DISCUSSION

The deduced amino acid sequences for the products of the *nifH*, *nifD* and *nifK* genes of *Trichodesmium* sp. IMS101 were found to have considerable identity to those of other known diazotrophs (Table 2), although the percentage identity varied significantly. Interestingly, it appears that the *Trichodesmium* sp. IMS101 NifD and NifK sequences have higher percentage identity to the NifD and NifK sequences of unicellular cyanobacteria than to the NifD and NifK sequences of filamentous heterocystous cyanobacteria.

The amino acid sequence VMGQEA is conserved in the *nifU* deduced amino acid sequences of *K. pneumoniae*, *Anabaena* sp., *Anabaena azollae*, *Nostoc commune*, *Cyanothece* sp. PCC 8801 and *Plectonema boryanum* (GenBank accession numbers: 128319, J05111, L34879,

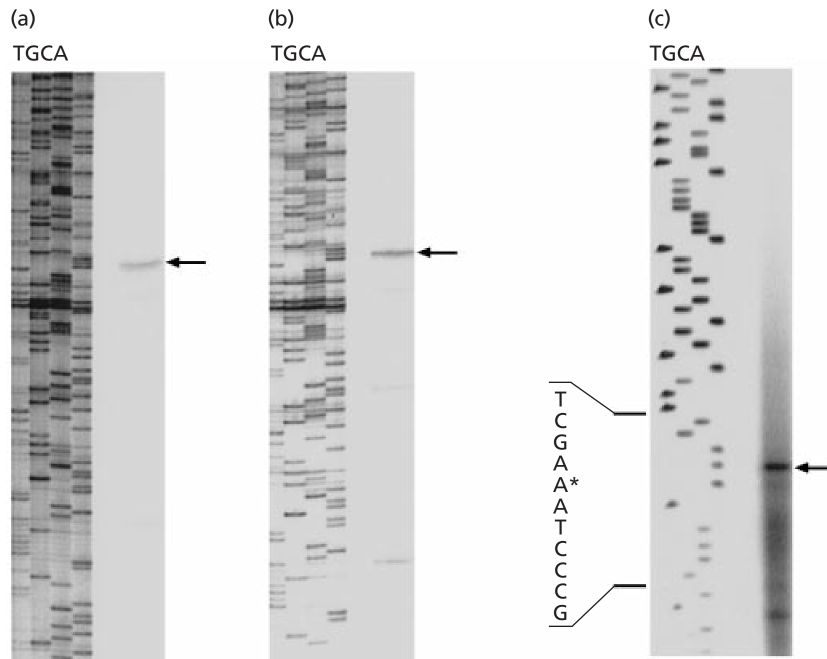

Fig. 2. Primer extension analysis of the 5′ end of the *Trichodesmium* sp. IMS101 *nifH* transcript. Panels (a), (b) and (c) show autoradiograms of the reverse transcriptase products using primers PXT-1, PXT-2 and HPRO-17, respectively. Lanes A, C, G and T show the dideoxynucleotide sequencing ladder generated using the same primers and pBD5N DNA as template. The uppermost (longest) primer extension product (indicated by an arrow) corresponded to the same nucleotide (A), indicated by an asterisk, for all three primers.

Sequence	
AAAACATTAATCACAATTTTAT-AT-ATGCAGG-TAATACCTGATCAAGCT**T**	T.sp.
TTTCACAACTA..T..CGAACCC..C...A.CAC....T.TACTGGTTTTT**C**	A.sp.1
TTTCACAACTA..T..CGAACCC..C...A.CAC....T.TACTG**G**	A.sp.2
TCACTTAATTA..T..CGAACCC..C...A.CAA....T.TACCG**G**	A.az
-35 -10	

Fig. 3. Alignment of *nif* gene sequences upstream of the transcription start site. Letters in bold indicate the transcription start site. T.sp., *nifH* of *Trichodesmium* sp. IMS101 (this work); A.sp.1, *nifH* of *Anabaena* sp. strain PCC 7120 (Mulligan & Haselkorn, 1989); A.sp.2, *nifH* of *Anabaena* sp. strain PCC 7120 (Haselkorn *et al.*, 1983); A.az, *nifH* of *Anabaena azollae* (Jackman & Mulligan, 1995). The consensus −10 and −35 promoter elements are indicated by bold lines. The dashes (-) indicate gaps introduced in the sequence for the purpose of alignment, and periods (.) denote identity to the *Trichodesmium* sp. IMS101 sequence.

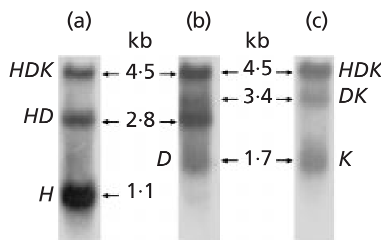

Fig. 4. Transcription of the *nifHDK* operon in *Trichodesmium* sp. IMS101. Northern blot of RNA obtained from an actively nitrogen-fixing *Trichodesmium* culture, probed with DNA probes for (a) *nifH*, (b) *nifD* and (c) *nifK*. The *nifH* probe hybridized to three bands corresponding to 1·1 kb (*H*), 2·8 kb (*HD*) and 4·5 kb (*HDK*) transcripts. The *nifD* probe produced a diffuse band centred at 1·7 kb (*D*), and discrete bands corresponding to 2·8 kb (*HD*), 3·4 kb (*DK*) and 4·5 kb (*HDK*) transcripts. The *nifK* probe also produced a diffuse band centred at 1·7 kb (*K*), and discrete bands corresponding to 3·4 kb (*DK*) and 4·5 kb (*HDK*) transcripts.

L23514, AF001780, D00666). Moreover, the *nifU* gene is located directly upstream of *nifH* in all cyanobacterial species for which the gene arrangement is known. Using the degenerate *nifU* primer, BDNU17, and the *nifH* upstream-sequence-specific primer, HPRO-17, part of the *nifU* gene was amplified. The sequence of the 0·7 kb insert of pBDHU overlapped with that of the 5·1 kb insert of pBD5N by 45 bases. Therefore, the *nifU* gene is located directly upstream of *nifH*.

The transcription start site identified was located 212 bases upstream of the *nifH* translational start (Fig. 2). Therefore, *Trichodesmium nifH* mRNA has a long untranslated *nifH* leader sequence compared to other cyanobacterial *nifH* leaders (Haselkorn *et al.*, 1983; Jackman & Mulligan, 1995). The leader sequence and the *nifH* coding region span 1·1 kb, which is in agreement with the 1·1 kb band detected with the *nifH* probe (Fig. 4). The transcription start sites and promoter elements of cyanobacterial (*Anabaena* spp.) *nifH* genes are still not defined. *Anabaena* sp. PCC 7120 and *A. azollae* have almost identical sequences upstream of *nifH* as shown in Fig. 3 (Jackman & Mulligan, 1995; Mulligan & Haselkorn, 1989) but the proposed transcription start sites differ by at least six bases (Tumer *et al.*, 1983; Haselkorn *et al.*, 1983; Mulligan & Haselkorn, 1989; Jackman & Mulligan, 1995). Significant identities were observed between the sequence $_{-17}$TAATAC$_{-12}$ for *Trichodesmium* sp. IMS101 and the $_{-17}$TAATTC$_{-12}$ for *Anabaena* sp. PCC 7120, since the '**C**' (of A.sp.1, Fig. 3) is considered to be the transcription start site as reported by Mulligan & Haselkorn (1989). The sequence TAATTC is also present in *A. azollae* (Fig. 3). The sequence TAATCT was identified to be the −10 promoter element of the *vnfD* gene of *Anabaena variabilis* (Thiel,

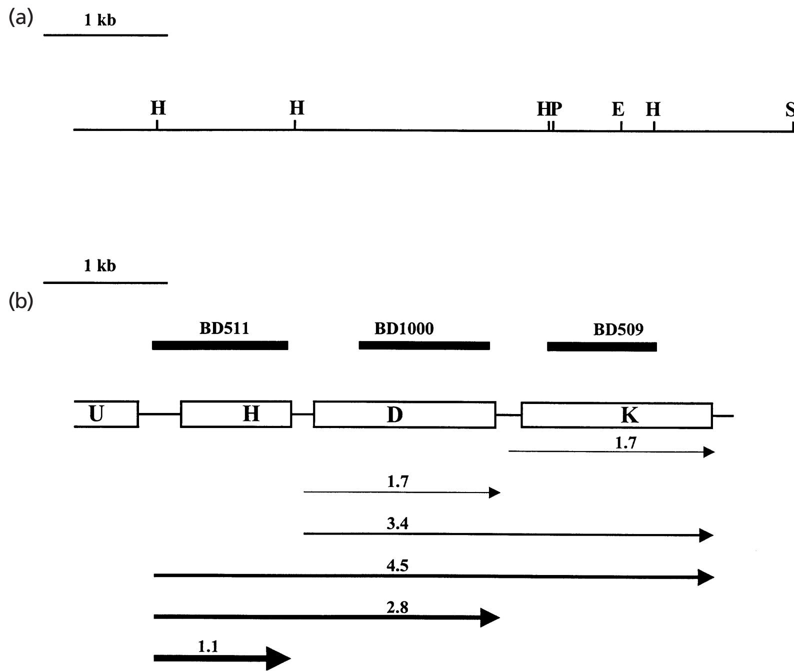

Fig. 5. The cloned *nifUHDK* region of *Trichodesmium* sp. IMS101. (a) Restriction site map of the cloned 5·8 kb *nifUHDK* region. E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; S, *Sac*I. (b) Physical map of the *nifUHDK* genes. The open reading frames of *nifU*, *H*, *D* and *K* genes are shown as open rectangles. The DNA fragments (BD511, BD1000 and BD509) used as *nifH*, *nifD* and *nifK* probes, respectively, are shown as labelled bars. Stable transcripts (as deduced from Northern blot experiments) are drawn as horizontal arrows, with the thickness of the arrow reflecting the relative abundance of the transcripts, the number above the arrow indicating the estimated size of the transcript in kb and arrow head indicating the 3′ end of transcript.

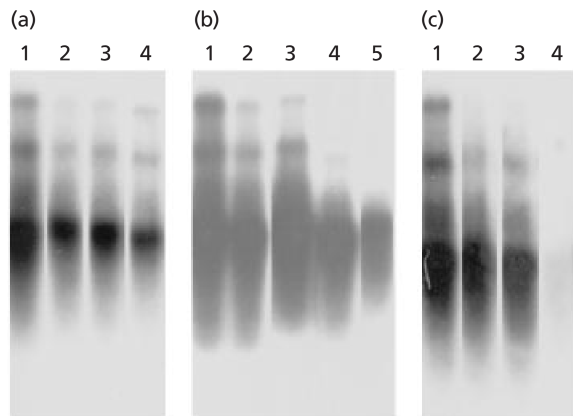

Fig. 6. Transcription of *nifHDK* genes as a function of light, nitrogen source and oxygen level. (a) Light experiment. Lane 1, zero time point 1200; lane 2, control (100 μE m^{-2} s^{-1}) 1500; lane 3, high light (300 μE m^{-2} s^{-1}) 1500; lane 4, darkness 1500. (b) Nitrogen treatment experiment. Lane 1, zero time point 1200; lane 2, control 1400; lane 3, nitrate (20 mM) 1400; lane 4, urea (2 mM) 1400; lane 5, ammonium (2 mM) 1400. (c) Oxygen experiment. Lane 1, zero time point 1200; lane 2, control 1500; lane 3, 40 % (v/v) oxygen 1500; lane 4, 80 % (v/v) oxygen 1500. (Note: 1200, 1400 and 1500 denote time points on a 24 h clock with 12 h light and 12 h dark phases; the subjective light phase was from 1000 to 2200 h.)

1993). It should be noted that the conserved TAAT part of the −10 promoter element reported here is part of the proposed −10 promoter element of *Anabaena* sp. PCC 7120 (ACTAAT; Haselkorn *et al.*, 1983). A −35 promoter-element-like sequence CACAAT was identified in *Trichodesmium* sp. IMS101 (Fig. 3) based on sequence similarity to the −35 sequence (CATAAC) reported for *Anabaena* sp. PCC 7120 (Tumer *et al.*, 1983).

Multiple transcripts could be explained as processed products of one primary transcript or as multiple primary transcripts. The Northern hybridization analysis of *Trichodesmium* mRNA using *nifH*, *nifD* and *nifK* probes (Fig. 4) revealed that each probe hybridized to at least three transcripts. The size of the bands hybridizing to the *nifH* probe (1·1, 2·8 and 4·5 kb) suggests that transcription is initiated upstream of *nifH*, producing transcripts corresponding to *nifH*, *nifHD* and *nifHDK*. The 4·5 kb band is the only hybridization signal that is common to the *nifH*, the *nifD* and *nifK* probes. The 4·5 kb *nifHDK* band is consistent with the suggestion that the *nif* operon is transcribed as a single unit (Golden *et al.*, 1985, 1991; Haselkorn *et al.*, 1986). The slightly diffuse 1·7 kb band observed with the *nifD* and *nifK* probes could possibly be nascent *nifD* transcripts extending into *nifK*, or could be *nifD* transcripts or *nifK* transcripts. The hybridization signal that corresponded to the 1·7 kb transcript was not observed when the *nifH* probe was used and a 1·7 kb transcript is not large enough to be a *nifHD* (2·8 kb) transcript identified with both the *nifH* and the *nifD* probes. Hence, the 1·7 kb band must correspond to a transcript spanning the *nifD* region. Similarly, the *nifK* probe also hybridized to a 1·7 kb band that must correspond to a transcript spanning the *nifK* region. Since the *nifK* probe did not produce any hybridization signal at the 2·8 kb *nifHD* region, the possibility that the *nifK* probe cross-hybridized to the 1·7 kb *nifD* transcripts can be eliminated. The *nifD* and *nifK* genes are very similar in size, making distinctions based on size of transcripts (as evidenced by

the location of the hybridization signal) very difficult. It is likely that some of the *nifD* transcripts terminate between the *nifD* and the *nifK* coding regions. The *nifK* probe hybridizes to three transcripts, *nifK* (1·7 kb), *nifDK* (3·4 kb) and *nifHDK* (4·5 kb). The hybridization signals were not due to non-specific binding to 16S or 23S rRNA or other non-*nif* mRNA, since it was observed that RNA from non-nitrogen-fixing (sampled during the dark phase) cultures failed to give any signal for *nifH*, *nifD* or *nifK* probes (data not shown).

The other possible explanation for multiple transcripts is that these different transcripts are due to processing of the single 4·5 kb long *nifHDK* transcript. Processing of long mRNA by RNA-specific endonuclease(s) and exonuclease(s) has been described or suggested in *E. coli* and *Rhodobacter capsulatus* (Belasco *et al.*, 1985; Heck *et al.*, 1996; Peterson, 1992; Willison *et al.*, 1993). In this study it was shown (Fig. 4) that there are significant amounts of stable *nifHD*, *nifDK*, *nifD* and *nifK* transcripts. The large amount of the 1·1 kb *nifH* mRNA indicates that there is no significant degradation of the mRNA by non-specific $5' \rightarrow 3'$ exonucleases. If the *nifHDK* transcripts were processed by cleavage at specific stem–loop sites and part of the *nifHDK* transcripts remained unprocessed, the observed pattern of transcripts (*nifHDK*, *nifHD*, *nifDK*, *nifH*, *nifD* and *nifK*) could result. However, such complex processing of transcripts is not usual and to our knowledge such processing machinery has not yet been reported in cyanobacteria. The different transcripts and the relative abundance of the transcripts of the *Trichodesmium nifHDK* genes are summarized in Fig. 5(b). To our knowledge, there has not been a transcription start site reported between *nifH* and *nifD* or *nifD* and *nifK* in any diazotroph that has contiguously arranged *nifHDK* genes, although transcriptional regulation of internal promoters in operons has been documented in cyanobacteria. In the filamentous cyanobacteria *Calothrix* sp. strain PCC 7601, multiple transcripts were observed from different polycistronic units such as *apc1*, *cpc1* and *cpc2* (Houmard *et al.*, 1990; Mazel *et al.*, 1988; Tandeau de Marsac *et al.*, 1988).

The multiple *nif* transcripts are consistent with the presence of potential stem–loop structures that could terminate transcription downstream of *nifH* and *nifD*. Similar secondary structures with comparable energy values in the intercistronic regions of polycistronic transcription units have been shown to terminate transcription, even if not preventing read-through completely (Houmard, 1994). Hybridization to the *nifH* probe revealed significantly higher amounts of stable transcripts of *nifH* when compared with stable transcripts of *nifHD* and *nifHDK* (Fig. 4). The *nifHD* transcripts appeared to be more abundant than *nifHDK* transcripts. It is known that the *Trichodesmium nifHDK* transcription starts during early morning hours and stabilizes gradually towards mid-day (Chen *et al.*, 1998; Wyman *et al.*, 1996). The results indicate that transcripts that are initiated from the *nifH* transcription start site do not all contain the complete *nifHDK* mRNA, and that the incomplete transcripts are not randomly terminated, but tend to contain only complete *nifH* or *nifHD* messages. This model is very similar to transcription of the gas vesicle genes (*gvpA1A2C*) of *Calothrix* sp. strain PCC 7601 (Csiszar *et al.*, 1987).

NH_4Cl was the most effective of the three nitrogen sources in turning off transcription. Significant repression of *nifHDK* transcription was not observed even at high concentrations (20 mM) of $NaNO_3$. This observation is consistent with that of the filamentous non-heterocystous cyanobacterium, *P. boryanum*, where nitrogenase activity (as measured by acetylene reduction) is repressed or inhibited more rapidly by ammonium than by nitrate compounds (Rai *et al.*, 1992). One of the reasons suggested for the lack of repression was the slow uptake of nitrate (Rai *et al.*, 1992). However, transcription was not assayed in that study and the inhibition could have been due to post-translational modification of nitrogenase protein.

Even after 3 h of exposure to 40 % oxygen concentration there was no significant reduction in transcription of *nifHDK* genes of *Trichodesmium* sp. IMS101. Hence the decrease in nitrogenase activity above pO_2 of 0·3 observed in *Trichodesmium* sp., could be due to post-translational modification of nitrogenase protein (Ohki & Fujita, 1988; Zehr *et al.*, 1993). The results indicate that inhibition of *nifHDK* transcription occurs at oxygen concentration greater than 40 % but less than or equal to 80 % (after 3 h exposure).

Abrupt changes in light intensity (whether it be from 100 $\mu E\ m^{-2}\ s^{-1}$ to complete darkness or from 100 $\mu E\ m^{-2}\ s^{-1}$ to 300 $\mu E\ m^{-2}\ s^{-1}$) did not result in any significant change in *nifHDK* transcription, suggesting that transcription of *nifHDK* is not directly regulated by light. Conversely, it is known that nitrogenase activity, as measured by acetylene reduction, declines rapidly in darkness and that the rates of acetylene reduction were proportional to the level of illumination in *Trichodesmium* (Ohki & Fujita, 1988).

Sequence analysis revealed regions of DNA that may form stem–loop structures in the intercistronic regions downstream of *nifH* and *nifD* and possible regulatory protein binding sites. Analysis of the transcription of *nifHDK* in *Trichodesmium* indicates several features in common with *nifHDK* transcription in other organisms. The sequences of the *nifHDK* genes are similar to other *nifHDK* sequences and there are striking similarities in the region upstream of the transcription start site of *Trichodesmium* sp. IMS101 and *Anabaena* spp. *nifH* genes. However, stable *nifD*, *nifDK* and *nifK* transcripts could be unique to *Trichodesmium* among cyanobacterial diazotrophs. The absence of a consensus NtcA-binding site upstream of the *nifH* transcription start site indicates that *Trichodesmium* sp. IMS101 *nifHDK* might not be directly regulated by NtcA. Further analysis of the regulation of *nifHDK* transcription in *Trichodesmium* might require the development of suitable mutants, which is not yet feasible.

ACKNOWLEDGEMENTS

This research was supported by NSF (OCE 9503593, OCE 9202106) to J. P. Z. We thank M. Mellon for technical help and J. Collier for encouragement, for helpful suggestions and for reviewing the manuscript.

REFERENCES

Alvarez-Morales, A., Betancourt-Alvarez, M., Klauza, K. & Hennecke, H. (1986). Activation of the *Bradyrhizobium japonicum nifH* and *nifDK* operons is dependent on promoter-upstream DNA sequence. *Nucleic Acid Res* **14**, 4207–4227.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (editors) (1990). *Current Protocols in Molecular Biology*. New York: Greene Publishing and Wiley Interscience.

Belasco, J. G., Beatty, J. T., Adams, C. W., von Gabain, A. & Cohen, S. N. (1985). Differential expression of photosynthesis genes in *R. capsulata* results from segmental differences in stability within polycistronic *rxcA* transcript. *Cell* **40**, 171–181.

Bergman, B., Gallon, J. R., Rai, A. N. & Stal, L. J. (1997). Nitrogen fixation by non-heterocystous cyanobacteria. *FEMS Microbiol Rev* **19**, 139–185.

Bradley, R. L. & Reddy, K. J. (1997). Cloning, sequencing, and regulation of the global nitrogen regulator gene *ntcA* in the unicellular diazotrophic cyanobacterium *Cyanothece* sp. strain BH68K. *J. Bacteriol* **179**, 4407–4410.

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

Capone, G. C., Zehr, J. P., Pearl, H. W., Bergman, B. & Carpenter, E. J. (1997). *Trichodesmium*, a globally significant marine cyanobacterium. *Science* **276**, 1221–1229.

Carpenter, E. J. & Romans, K. (1991). Major role of the cyanobacterium *Trichodesmium* in nutrient cycling in the North Atlantic Ocean. *Science* **254**, 1356–1358.

Chen, Y., Zehr, J. P. & Mellon, M. (1996). Growth and nitrogen fixation of the diazotrophic filamentous nonheterocystous cyanobacterium *Trichodesmium* sp. IMS101 in defined media: evidence for circadian rhythm. *J Phycol* **32**, 916–923.

Chen, Y., Dominic, B., Mellon, M. & Zehr, J. P. (1998). Circadian rhythm of nitrogenase gene expression in the diazotrophic filamentous nonheterocystous cyanobacterium *Trichodesmium* sp. IMS 101. *J Bacteriol* **180**, 3598–3605.

Collins, J. J. & Brill, W. J. (1985). Control of *Klebsiella pneumoniae nif* mRNA synthesis. *J Bacteriol* **162**, 1186–1190.

Csiszàr, K., Houmard, J., Damerval, T. & Tandeau de Marsac, N. (1987). Transcriptional analysis of the cyanobacterial *gvpABC* operon in differentiated cells: occurrence of an antisense RNA complementary to three overlapping transcripts. *Gene* **60**, 29–37.

Dixon, R. A., Austin, S., Buck, M., Drummond, M., Hill, S., Holtel, A., MacFarlane, S., Merrick, M. & Minchin, S. (1987). Genetics and regulation of *nif* and related genes in *Klebsiella pneumoniae*. *Phil Trans R Soc Lond B* **317**, 147–158.

Fay, P. (1992). Oxygen relations of nitrogen fixation in cyanobacteria. *Microbiol Rev* **56**, 340–373.

Flores, E. & Herrero, A. (1994). Assimilatory nitrogen metabolism and its regulation. In *The Molecular Biology of Cyanobacteria*, vol. 1, *Advances in Photosynthesis*, pp. 487–517. Edited by D. A. Bryant. Dordrecht: Kluwer.

Gallon, J. R. (1992). Tansley Review No. 44. Reconciling the incompatible: N_2 fixation and O_2. *New Phytol* **122**, 571–609.

Gallon, J. R., Jones, D. A. & Page, T. S. (1996). *Trichodesmium*, the paradoxical diazotroph. *Algol Stud* (*Arch Hydrobiol*) **83**, 215–243.

Golden, J. W., Robinson, S. J. & Haselkorn, R. (1985). Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium *Anabaena*. *Nature* **314**, 419–423.

Golden, J. W., Whorff, L. L. & Wiest, D. R. (1991). Independent regulation of *nifHDK* transcription and DNA rearrangement during heterocyst differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. *J Bacteriol* **173**, 7098–7105.

Haselkorn, R., Rice, D., Curtis, S. E. & Robbinson, S. J. (1983). Organization and transcription of genes important in *Anabaena* heterocyst differentiation. *Ann Microbiol* **134B**, 181–193.

Haselkorn, R., Golden, J. W., Lammers, P. J. & Mulligan, M. E. (1986). Developmental rearrangement of cyanobacterial nitrogen-fixation genes. *Trends Genet* **2**, 255–259.

Heck, C., Rothfuchs, R., Jäger, A., Rauhut, R. & Klug, G. (1996). Effect of *pufQ–pufB* intercistronic region on *puf* mRNA stability in *Rhodobacter capsulatus*. *Mol Microbiol* **20**, 1165–1178.

Houmard, J. (1994). Gene transcription in filamentous cyanobacteria. *Microbiol* **140**, 433–441.

Houmard, J., Capuano, V., Colombano, M. V., Coursin, T. & Tandeau de Marsac, N. (1990). Molecular characterization of the terminal energy acceptor of cyanobacterial phycobilisomes. *Proc Natl Acad Sci USA* **87**, 2152–2156.

Jackman, D. M. & Mulligan, M. E. (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. *Microbiol* **141**, 2235–2244.

Ligon, J. M. & Nakas, J. P. (1987). Isolation and characterization of *Frankia* sp. strain FaC1 genes involved in nitrogen fixation. *Appl Environ Microbiol* **53**, 2321–2327.

Luque, I., Flores, E. & Herrero, A. (1994). Molecular mechanism for the operation of nitrogen control in cyanobacteria. *EMBO J* **13**, 2862–2869.

Magasanik, B. (1988). Reversible phosphorylation of an enhancer of an enhancer binding protein regulates the transcription of bacterial utilization genes. *Trends Biochem Sci* **12**, 475–479.

Mazel, D., Houmard, J. & Tandeau de Marsac, N. (1988). A multigene family in *Calothrix* sp. PCC 7601 encodes phycocyanin the major component of the cyanobacterial light harvesting antenna. *Mol Gen Genet* **211**, 296–304.

Merrick, M. J. (1992). Regulation of nitrogen fixation genes in free-living and symbiotic bacteria. In *Biological Nitrogen Fixation*, pp. 835–876. Edited by G. Stacey, R. H. Burris & H. J. Evans. New York: Chapman & Hall.

Michaels, A. F., Olson, D., Sarmiento, J. L., Ammerman, J. W., Fanning, K., Jahnke, R., Knap, A. H., Lipshultz, F. & Prospero, J. M. (1996). Inputs, losses and transformations of nitrogen and phosphorus in the pelagic North Atlantic Ocean. *Biogeochemistry* **35**, 181–226.

Mulligan, M. E. & Haselkorn, R. (1989). Nitrogen fixation (*nif*) genes of the cyanobacterium *Anabaena* species strain PCC 7120. *J Biol Chem* **264**, 19200–19207.

Ohki, K. & Fujita, Y. (1988). Aerobic nitrogenase activity measured as acetylene reduction in the marine cyanobacterium *Trichodesmium* spp. grown under artificial conditions. *Marine Biol* **98**, 111–114.

Peterson, C. (1992). Control of functional mRNA stability in

bacteria: multiple mechanisms of nucleolytic and non-nucleolytic inactivation. *Mol Microbiol* **6**, 277–282.

Postgate, J. R. (1982). *The Fundamentals of Nitrogen Fixation*. Cambridge: Cambridge University Press.

Prufert-Bebout, L., Pearl, H. W. & Lassen, C. (1993). Growth, nitrogen fixation, and spectral attenuation in cultivated *Trichodesmium* species. *Appl Environ Microbiol* **59**, 1367–1375.

Rai, A. N., Borthakur, M. & Bergman, B. (1992). Nitrogenase derepression, its regulation and metabolic changes associated with diazotrophy in the non-heterocystous cyanobacterium *Plectonema boryanum* PCC 73110. *J Gen Microbiol* **138**, 481–491.

Roberts, G. P. & Brill, W. J. (1980). Gene–product relationships of the *nif* regulon of *Klebsiella pneumoniae*. *J Bacteriol* **144**, 210–216.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sroga, G. E., Landegren, U., Bergman, B. & Lagerström-Fermér, M. (1996). Isolation of *nifH* and part of *nifD* by modified capture polymerase chain reaction from a natural population of the marine cyanobacterium *Trichodesmium* sp. *FEMS Microbiol Lett* **136**, 137–145.

Stricker, O., Masepohl, B., Klipp, W. & Böhme, H. (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.

Tandeau de Marsac, N. & Houmard, J. (1993). Adaptation of cyanobacteria to new environmental stimuli: new steps towards molecular mechanisms. *FEMS Microbiol Rev* **104**, 119–190.

Tandeau de Marsac, N., Mazel, D., Damerval, T., Guglielmi, G., Capuano, V. & Houmard, J. (1988). Photoregulation of gene expression in the filamentous cyanobacterium *Calothrix* sp. PCC 7601: light harvesting complexes and cell differentiation. *Photosynth Res* **18**, 99–132.

Thiel, T. (1993). Characterization of genes for an alternative nitrogenase in the cyanobacterium *Anabaena variabilis*. *J Bacteriol* **175**, 6276–6286.

Tumer, N. E., Robinson, S. J. & Haselkorn, R. (1983). Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen. *Nature* **306**, 337–342.

Vega-Palas, M. A., Madeuno, F., Herrero, A. & Flores, E. (1990). Identification and cloning of a regulatory gene for nitrogen assimilation in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *J Bacteriol* **172**, 643–647.

Vega-Palas, M. A., Flores, E. & Herrero, A. (1992). NtcA, a global nitrogen regulator from the cyanobacterium *Synechococcus* that belongs to the Crp family of bacterial regulators. *Mol Microbiol* **6**, 1853–1859.

Willison, J. C., Pierard, J. & Hübner, P. (1993). Sequence and transcript analysis of the nitrogenase structural gene operon (*nifHDK*) of *Rhodobacter capsulatus*: evidence for intramolecular processing of *nifHDK* mRNA. *Gene* **133**, 39–46.

Wyman, M., Zehr, J. P. & Capone, D. G. (1996). Temporal variability in nitrogenase gene expression in natural population of the marine cyanobacterium *Trichodesmium thiebautii*. *Appl Environ Microbiol* **62**, 1073–1075.

Zehr, J. P. & McReynolds, L. A. (1989). Use of degenerate oligonucleotides for amplification of the *nifH* gene from the marine cyanobacterium *Trichodesmium thiebautii*. *Appl Environ Microbiol* **55**, 2522–2526.

Zehr, J. P., Ohki, K. & Fujita, Y. (1991). Arrangement of nitrogenase structural genes in an aerobic filamentous nonheterocystous cyanobacterium. *J Bacteriol* **173**, 7055–7058.

Zehr, J. P., Wyman, M., Miller, V., Duguay, L. & Capone, D. G. (1993). Modification of the Fe protein of nitrogenase in natural populations of *Trichodesmium thiebautii*. *Appl Environ Microbiol* **59**, 669–676.

Received 6 April 1998; revised 10 July 1998; accepted 18 August 1998.