Organization and expression of nitrogen-fixation genes in the aerobic nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain RF-1

Tan-Chi Huang,1 Rong-Fong Lin,1 Mu-Kuei Chu1 and Horng-Ming Chen2

Author for correspondence: Tan-Chi Huang. Tel. +886 2 2782 9590. Fax: +886 2 2782 7954.
e-mail: tanchi@gate.sinica.edu.tw

Sixteen nif and ‘nif-associated’ genes (expressed only under conditions of nitrogen fixation) in Synechococcus sp. strain RF-1 have been cloned and sequenced. All of the nif and nif-associated genes identified in Synechococcus RF-1 were arranged in a continuous cluster spanning approximately 18 kb and containing seven operons. The nifH operon (nifH-nifD-nifK) has been reported previously. nifB, fdxN, nifS, nifU and nifP were found to be located upstream of the nifH operon. nifB-fdxN-nifS-nifU were expressed as an operon. A nifP-like gene was found to be located just upstream of nifB. nifE, nifN, nifX, nifW and the nif-associated hesA, hesB and ‘fdx’ were found to be located downstream from nifK. The genes located downstream from nifK are arranged nifE-nifN-nifX-orf-nifW-hesA-hesB-'fdx' and span approximately 7 kb. The function of the ORF situated between nifX and nifW is not known. However, it was identified as a counterpart of ORF-2 in Anabaena sp. strain PCC 7120 based on the deduced amino acid sequence. Northern hybridization and primer extension analysis indicated that the nif and nif-associated genes are organized in nifE-nifN-nifX-orf-nifW-hesA-hesB-'fdx' containing operons, respectively. According to the results of this study and previous reports, the genes are expressed in a rhythmic pattern with peaks during the dark phase when the culture is grown in a 12 h light/12 h dark regimen. The rhythm persisted after the culture was transferred to continuous illumination.

Keywords: cyanobacteria, nitrogen fixation, gene organization, rhythmic expression, Synechococcus RF-1

INTRODUCTION

Cyanobacteria may contribute significant nitrogen into the global nitrogen cycle since they are widely distributed in different environments, especially in paddy rice fields and marine environments. Morphologically, cyanobacteria can be divided into unicellular, non-heterocystous filamentous and heterocyst-forming types. Nitrogen fixation is carried out by all of the heterocyst-forming type, while only a few species of unicellular and non-heterocystous filamentous types can fix nitrogen aerobically. To protect the nitrogenase from oxygen inhibition, about 5–10% of cells of the heterocyst-forming type differentiate into specialized nitrogen-fixing cells. The correlation between heterocyst differentiation and structure of nif genes has been intensively studied in Anabaena sp. strain PCC 7120 (Haselkorn & Buikema, 1992). Species of the non-heterocystous filamentous type usually fix nitrogen only under low ambient oxygenic pressure (Fay, 1992). The unicellular nitrogen-fixing cyanobacteria can fix nitrogen and carbon dioxide within the same cell. These microorganisms have developed a unique strategy to protect nitrogenase from oxygen inactivation by temporal separation of nitrogen fixation and photosynthesis. Structural analysis of the nif genes is important for understanding nitrogen fixation at the biochemical level and for revealing the genetic relationships among different diazotrophs (Postgate & Eady, 1988). About 20 genes identified in Klebsiella pneumoniae were found to be involved in nitrogen-fixing activity. Characterization
of the structural genes (nifH, nifD and nifK) for nitrogenase has been performed in various diazotrophs. However, the structure and organization of nitrogen fixation-related genes has been documented only in some diazotrophs such as Klebsiella pneumoniae (Arnold et al., 1988), Azotobacter vinelandii (Jacobson et al., 1989), Bradyrhizobium japonicum (Ebeling et al., 1987; Fischer et al., 1988; Noti et al., 1986), Rhodobacter capsulatus (Klipp, 1990), Clostridium pasteurianum (Chen et al., 1986; Wang et al., 1989) and Anabaena PCC 7120 (Haselkorn et al., 1987; Fischer et al., 1988; Noti et al., 1986), Bradyrhizobium japonicum (Ebeling et al., 1988), Azotobacter vinelandii (Jacobson et al., 1988), while those in the sheathless form have been investigated in Synechococcus sp. strain RF-1 (Huang & Grobbelaar, 1995), Synechococcus BG43511 (Mitsui et al., 1986) or Cyanobacteria sp. strain ATCC 51142 (Colon-Lopez et al., 1997). In the unicellular cyanobacterium Synechococcus sp. strain RF-1, the nitrogen-fixing activity in a diurnal light/dark-entrained culture was found to exhibit a circadian rhythm (Huang & Grobbelaar, 1995). To understand the control mechanism of rhythmic nitrogen fixation, it is important to characterize the structure of the nif genes. Therefore, the nif genes in Synechococcus sp. strain RF-1 were cloned and sequenced. The nucleotide sequence of the nifHD operon has been reported previously (Chen et al., 1996). In this report, the nucleotide sequence and expression of nifH, nifD, fdxN, nifS, nifU, nifE, nifN, nifX, nifW and some 'nif-associated' genes (expressed only under conditions of nitrogen fixation) are presented. In addition, the organization of all nif and nif-associated genes cloned from Synechococcus sp. strain RF-1 is reviewed and discussed.

METHODS

Organisms and cultivation. Axenic culture of Synechococcus sp. strain RF-1 (PCC 8001) was grown in BG-11 or BG-110 medium under the conditions described previously (Huang & Chow, 1988). Cultures were grown in light (about 35 μmol m−2 s−1) from white fluorescent tubes at 28 ± 0.5 °C. They were incubated either in continuous light (L/L) or in a diurnal light/dark (L/D) regimen, depending on the requirement of the experiment.

DNA isolation and construction of a genomic library. Genomic DNA in Synechococcus sp. strain RF-1 was extracted as described previously (Chen et al., 1996). The genomic DNA was partially digested with Samb3A and DNA fragments with sizes ranging from 10 to 23 kb were isolated from low-melting-point agarose gels (Sambrook et al., 1989). The size-fractionated fragments were ligated into BamHI-digested EMBL4, according to the manufacturer's procedures (Stratagene).

Cloning of nif genes located upstream and downstream from nifHDK operon. The genomic library of Synechococcus sp. strain RF-1 prepared in EMBL4 was used for screening the nif genes by the standard plaque hybridization technique (Sambrook et al., 1989). It is known that nifU in Synechococcus sp. strain RF-1 is located upstream of nifH, based on our previous report (Chen et al., 1996). Therefore, a synthetic oligonucleotide complementary to part of nifU was used as a probe for screening the nif genes located upstream of the nifHDK operon. From our previous results (Chen et al., 1996), one subclone, NF-3, is known to contain nifK and some other genes. Subsequent sequencing revealed that nifE was located immediately downstream from nifK. Therefore, a synthetic oligonucleotide complementary to part of nifE was used as a probe to screen DNA fragments from the genomic library. The probe was end-labelled with digoxigenin (DIG) and plaque hybridization was carried out using the DIG Luminescent Detection Kit according to the manufacturer's instructions (Boehringer Mannheim). Positive plaques were selected and subcloned into pGEM-7Zf (−) for sequencing.

Nucleotide sequencing and comparison. The nucleotide sequence of both strands of the plasmid DNA was determined by the dideoxy chain-termination method on an automated DNA sequencer (Pharmacia), according to the manufacturer's procedures. Sequencing of dsDNA was first carried out with universal sequencing primers using the Auto Read Sequencing Kit (Pharmacia), following the supplier's instructions. Other oligonucleotide primers were synthesized and used to complete sequencing of both strands. The extent of identity of the deduced amino acid sequences of nif genes in Synechococcus sp. strain RF-1 with the products encoded by corresponding genes from other diazotrophs was searched for by using the PCGENE program.

Isolation of total RNA. Thirty millilitres of exponentially growing cells (about 2 × 10⁸ cells ml⁻¹) of Synechococcus sp. strain RF-1 were collected by centrifugation, washed twice with distilled water and then collected in a 1.5 ml screw-top vial. Teflon-coated sea-sand (0.2 g, 0–1 mm diameter; Merck), 0.5 ml solution D (4 M guanidine thiocyanate, 25 mM sodium citrate, 0.5% laurylsarcosine and 0.72% β-mercaptoethanol), 0.5 ml phenol/chloroform (4:1, v/v) and 0.1 ml 2 M sodium chloride (pH 40) were then added. The cells were broken by vibration with a Mini-Beadbeater (Biospec) set at 5000 r.p.m. for 120 s. The broken cells were centrifuged at 13000 g at 4 °C for 5 min. The supernatant was re-extracted twice with phenol/chloroform (1:1). An equal volume of ice-cold absolute alcohol was added to the supernatant. Total RNA was collected after freezing at −70 °C for more than 1 h and then resuspended in an appropriate volume of diethyl pyrocarbonate (DEPC)-pretreated water.

Primer extension analysis. The primer extension analysis was performed as described previously (Chen et al., 1996). About 30 μg total RNA prepared from Synechococcus sp. strain RF-1 was used for each primer extension analysis.

Northern hybridization. About 5 μg total RNA was added into 20 μl RNA loading buffer (50% formamide, 65% formamide, 2.5 μg ethidium bromide ml⁻¹, 1× MOPS), denatured at 65 °C for 15 min, cooled in ice and then fractionated by 1.2% agarose gel electrophoresis (containing 3.7% formaldehyde). The RNA gel was immersed in 1× phosphate buffer (50 mM NaH₂PO₄, 50 mM NaOH, 5 mM EDTA, pH 6.5) for 30 min and then blotted onto GeneScreen Plus membrane (NEN) with 20× SSPE overnight (1× SSPE is 0.15 M NaCl, 0.01 M NaH₂PO₄, H₂O, 1 mM EDTA, pH 7.4).
The membrane was prehybridized in 15 ml/10 x 10 cm² hybridization solution [0-2 x SSPE, 0-5 % laurylsarcosine, 1 % SDS] at 65 °C for at least 3 h. Hybridization was carried out at 65 °C overnight in 5 ml hybridization solution by adding about 25 ng denatured DIG-labelled DNA probe. The membrane was washed twice with solution containing 0-1 x SSPE and 0-1 % SDS at 65 °C for 10 min. Detection of the DIG-chemiluminescent reaction was carried out as described by the manufacturer (Boehringer Mannheim).

RESULTS
The genes located upstream of the nifH operon
Cloning and sequencing. About 1 x 10⁴ plaques from the Synechococcus RF-1 genomic library were screened using a synthetic oligonucleotide complementary to part of the DNA fragments from the subcloned recombinants (NF-4, NF-5, NF-6 and NF-7) used for sequence analysis. The location of the polypeptide encoded by nifU. Ten positive clones exhibiting strong hybridization signals were isolated. The phage DNA was restricted with EcoRI, HindIII, or ApaLI and then subcloned into pGEM-7Zf(-). Four independent subclones, NF-4, NF-5, NF-6 and NF-7, were selected (Fig. 1) and both strands of DNA from each subclone were sequenced. Five ORFs in a 6-6 kb fragment were identified after searching for all three possible reading frames with ATG as the initiation codon.

Identification of the nifB operon by the deduced amino acid sequence. Among the five ORFs, the reading direction of the first is the reverse of the other four. Therefore, it is characterized in another section below. The second ORF beginning from position 1961 consists of 1677 bp. It encodes a polypeptide of 558 aa. The predicted amino acid sequence exhibits a high degree of identity (73.89 %) to that of the polypeptide encoded by nifB from Anabaena PCC 7120 (Mulligan & Haselkorn, 1989).

The third ORF (position 3652-4014) is present 15 bp downstream from the termination codon of nifB. It encodes a polypeptide of 120 aa. The predicted amino acid sequence shows 43.1 % identity to the product of fdxN from Anabaena PCC 7120. Although the similarity of the amino acid sequence is not very high, it contains two cysteine motifs, C-X₇-C-X₇-C-X₇-CP at the N terminus and C-X₇-C-X₇-C-X₇-CP at the C terminus, which were found in the nif-associated ferredoxins (Fig. 2) (Mulligan et al., 1988).

The fourth ORF (position 4119-5321) is present 104 bp downstream from the termination codon of fdxN. The deduced amino acid sequence has 74-25 % identity with the polypeptide encoded by nifS from Anabaena PCC

Table 1. Comparison of nif genes in Synechococcus RF-1 with their corresponding counterparts in other diazotrophs

<table>
<thead>
<tr>
<th>Gene</th>
<th>Anabaena PCC 7120</th>
<th>Klebsiella pneumoniae</th>
<th>Azotobacter vinelandii</th>
</tr>
</thead>
<tbody>
<tr>
<td>nifB</td>
<td>73.9</td>
<td>44.2</td>
<td>51.3</td>
</tr>
<tr>
<td>nifS</td>
<td>74.3</td>
<td>56.0</td>
<td>35.4</td>
</tr>
<tr>
<td>nifU</td>
<td>64.0</td>
<td>45.4</td>
<td>49.2</td>
</tr>
<tr>
<td>nifE</td>
<td>75.6</td>
<td>53.6</td>
<td>55.4</td>
</tr>
<tr>
<td>nifN</td>
<td>59.5</td>
<td>36.7</td>
<td>44.8</td>
</tr>
<tr>
<td>nifX</td>
<td>59.7</td>
<td>20.7</td>
<td>41.5</td>
</tr>
<tr>
<td>nifW</td>
<td>59.0</td>
<td>16.3</td>
<td>34.2</td>
</tr>
</tbody>
</table>

Fig. 1. Schematic representation of a 6-6 kb fragment of the Synechococcus RF-1 chromosome containing the nifB-fdxN-nifS-nifU operon and nifP located upstream of nifB. The location of the DNA fragments from the subcloned recombinants (NF-4, NF-5, NF-6 and NF-7) used for sequence analysis is shown by white boxes. The restriction sites used for cloning are indicated by arrows.

Fig. 2. Comparison of the deduced amino acid sequence of fdxN in Synechococcus RF-1 with that of Anabaena PCC 7120. The cysteine residues probably involved in [4Fe-4S] clusters are indicated by asterisks.
7120 (Mulligan & Haselkorn, 1989). The deduced amino acid sequence of the fifth ORF (position 5486–6370) has 63.95% identity with the polypeptide encoded by \textit{nifU} from \textit{Anabaena} PCC 7120 (Mulligan & Haselkorn, 1989). Therefore, we conclude that the four ORFs [2, 3, 4 and 5 in AF001780 (GenBank)] in \textit{Synechococcus} RF-1 correspond to \textit{nifB}, \textit{fdxN}, \textit{nifS} and \textit{nifU} in \textit{Anabaena} PCC 7120, respectively. A comparison of \textit{nifB}, \textit{nifS} and \textit{nifU} with the corresponding \textit{nif} genes in other diazotrophs is shown in Table 1.

**Transcription start site and the expression of the \textit{nifB} operon.** Primer extension analysis was used to locate the start site of transcription. As shown in Fig. 3(a), a single transcript for \textit{nifB} was detected. The transcription start site is nucleotide ‘A’ which lies 206 bp upstream of the start codon of the \textit{nifB} gene. Expression of the \textit{nifB} operon was examined by Northern hybridization. As shown in Fig. 4, four bands with sizes of 1.8, 2.2, 3.5 and 4.6 kb were detected. Based on molecular size, the four bands correspond to the mRNAs of \textit{nifB}, \textit{nifB-fdxN}, \textit{nifB-fdxN-nifS} and \textit{nifB-fdxN-nifS-nifU}, respectively. The mRNA level of the \textit{nifB} operon changed in a 24 h rhythmic pattern after the 12 h L/12 h D-entrained culture was transferred to L/L, with the active synthesis period at about 14–17 h. Since the time periods from 0 to 12 h and from 12 to 24 h correspond to the L and D phases, respectively, before the transfer to constant illumination, the expression of \textit{nifB} peaked early in the dark phase and was very weak during the light phase under the diurnal L/D regimen. The expression of the \textit{nifHDK} operon in \textit{Synechococcus} RF-1 also exhibits a circadian rhythm, with the mRNA occurring exclusively during the dark phase when the culture grows in a diurnal L/D regimen (Huang & Chow, 1990). Results in Fig. 4 indicated that the expression phase of the \textit{nifB} operon was the same as that of the \textit{nifHDK} operon.

**The gene located upstream of \textit{nifB}**. A complete ORF with a reverse reading direction to the \textit{nifB} operon was found.
upstream of nifB. The inverted ORF (position 832–95) encodes a polypeptide of 245 aa. Based on the similarity to reported protein sequences, the deduced amino acid sequence shows 31-43% identity with the product of cysE from Escherichia coli, which encodes a serine acetyltransferase (Denk & RocK, 1987). The ORF was temporarily assigned as a cysE-like gene ('cysE'). The results of primer extension analysis indicated that the transcription start site of 'cysE' is nucleotide 'A' located 30 bp upstream of the start codon (Fig. 3).

An ORF located immediately downstream from nifV in Azotobacter chroococcum has been identified and designated nifP (Evans et al., 1991). It is 41% identical to the product of cysE from E. coli. The nifP gene has been suggested to encode a nif-specific serine acetyltransferase required for optimizing expression of nitrogenase activity. The inverted 'cysE' in Synechococcus RF-1 is 39.5% identical to nifP of Azotobacter chroococcum. According to Northern hybridization (Fig. 5), the 'cysE' gene shows expression only under nitrogen-fixing conditions. Based on its similarity to the Azotobacter gene and its expression limited to nitrogen-fixing conditions, the 'cysE' gene was designated nifP. Fig. 5 also shows that the transcription activity of nifP exhibits the same circadian rhythm as nifB (Fig. 4) in Synechococcus RF-1.

Comparison of nif promoter sequences. The nucleotide sequences of the presumptive nifB and nifH promoters in Synechococcus RF-1 were compared with the promoter regions of the counterparts in Anabaena PCC 7120 reported by Mulligan & Haselkorn (1989). No significant similarity was found for nifH. However, two regions containing a number of identical base pairs were found in nifB (Fig. 6). One is located about 10 bp upstream of the transcription start site and the other is located from position -61 to -73.

The genes located downstream from nifK

Cloning and sequencing. Synthetic oligonucleotides complementary to parts of nifK and nifE were used to screen DNA fragments containing nif genes from the genomic library cloned in EMBL4. The positive clones were isolated and then subcloned into pGEM-7Zf(--). As shown in Fig. 7, four independent subclones, NF-3, NF-8, NF-9 and NF-10, were selected for DNA analysis. The nucleotide sequence covering a region of about 7 kb located downstream from nifK was revealed by subsequent sequencing of the four subclones. Nine ORFs reading in the same direction were found based on searching for all three possible translation frames with ATG as the initiation codon.

Identification of nif genes by the deduced amino acid sequences. The nucleotide sequences for the nine possible ORFs along with their 5' flanking regions were identified. The first ORF (position 328–1737) encodes a polypeptide of 469 aa. The deduced amino acid sequence exhibits a high degree of identity (75.6%) to the polypeptide encoded by nifE from Anabaena PCC 7120. The second ORF (position 1806–3191) encodes a polypeptide of 461 aa. The deduced amino acid sequence shows a 59.5% identity with the product of nifN from Anabaena PCC 7120. The third ORF (position 3314–3721) encodes a polypeptide of 135 aa. The deduced amino acid sequence shows a 59.5% identity with the product of nifX from Anabaena PCC 7120. These results indicate that the three ORFs in Synechococcus RF-1 are nifE-, nifN- and nifX-like, respectively.

Based on data submitted recently to GenBank by Buikema, Scapino & Haselkorn (U47055), two functionally unidentified genes, designated ORF-2 and ORF-1, are located downstream from nifX in Anabaena PCC 7120. Both ORF-2 and ORF-1 in Anabaena PCC 7120 encode polypeptides of 158 and 71 aa, respectively. The amino acid sequences deduced from the fourth (position 3821–4291) and fifth (position 4404–4634) ORFs downstream from nifK in Synechococcus RF-1 are 57.6 and 67.2% identical to the products of ORF-2 and ORF-1 in Anabaena PCC 7120, respectively. Therefore, the fourth
and fifth ORFs in *Synechococcus* RF-1 were identified as the counterparts of ORF-2 and ORF-1, respectively, in *Anabaena* PCC 7120.

The amino acid sequence deduced from the sixth ORF (position 4636-4971) has 59% identity with the product of *nifW* from *Anabaena* PCC 7120. The amino acid sequences deduced from the seventh (position 4980-5546) and eighth (position 5875-6234) ORFs show 75% and 58.4% identity with the products of *hesA* (U47055) and *hesB* (Borthakur et al., 1990) from *Anabaena* PCC 7120, respectively.

The ninth ORF (position 6273-6617) encodes a polypeptide of 114 aa. This protein is identified as a ferredoxin and the gene was designated 'fdx' because of the presence of the characteristic cysteine cluster: CX4CX3C. The deduced amino acid sequence is 47% identical to ferredoxin 1 (PetF) from the vegetative cells of *Anabaena* PCC 7120 (Alam et al., 1986). It is also 47% identical to the nitrogen-fixation-specific ferredoxin (FdxH) from *Anabaena* PCC 7120 (Böhme & Haselkorn, 1988). Further studies are needed to identify a specific role for the product of 'fdx' in *Synechococcus* RF-1.

A comparison of *nif* genes in *Synechococcus* RF-1 with the corresponding counterparts of *Anabaena*, *Klebsiella* and *Azotobacter* is shown in Table 1.

**Gene expression examined by Northern hybridization.** DIG-labelled DNA probes specific to *nifE, nifN, nifX, orf2, orf1, nifW, hesA, hesB* and 'fdx', prepared by PCR, were used to investigate gene expression. Samples for total RNA extraction were taken at 4 h intervals for a period of 46 h after the *Synechococcus* RF-1 culture grown in a 12 h L/12 h D diurnal regimen was transferred to continuous light. The results of Northern hybridization indicated that *nifE, nifN, nifX, orf2, orf1, nifW, hesA, hesB* and 'fdx' were all expressed in a circadian rhythmic pattern (Fig. 8). The transcripts of these genes were detected mainly during the dark periods when grown in a diurnal L/D regimen and the rhythms persisted after the culture was transferred to continuous illumination.

As shown in Fig. 8(a) and (b), the signals detected by the DNA probe specific to *nifE* or *nifN* were smeared and the sizes were smaller than predicted. This may have been caused by the degradation of mRNA during isolation or poor specificity of the DNA probe. However, defined mRNA bands were detected when the same RNA preparation was characterized with other DNA probes (Fig. 8), so the smeared signals were not necessarily due to the breakdown of mRNA during isolation. They are also unlikely to be caused by the poor specificity of the DNA probe, since no RNA–DNA hybridization band was detected during the light periods (Fig. 8). Therefore, we speculate that the transcripts of *nifE* and *nifN* are relatively unstable and can be degraded immediately after translation. However, two major bands with sizes of 0·5 and 1·1 kb were detected by the *nifX* probe (Fig. 8c). The lengths of these two bands correlate well with those predicted from the transcript of *nifX* and that extending from *nifX* to *orf2*, respectively. As shown in Fig. 8(d), the band of 1·1 kb was present as a major band when a DNA probe specific to *orf2* was used. The results of Fig. 8(a–d) indicate that *nifE, nifN, nifX* and *orf2* are unlikely to be transcribed as a single operon. The observation of a major band of 0·5 kb correlating to *nifX* suggests that *nifX* is expressed as a unit distinct from *nifE*. The possibility is further supported by the results of primer extension analysis described below.

When the DNA probes specific to *orf1* and *nifW* were used, the hybridization patterns were the same. As shown in Fig. 8(e) and (f), two major bands with sizes of 0·6 and 2·1 kb were detected. The presence of a major transcript of 0·6 kb suggests that *orf1* and *nifW* are expressed as a distinct unit from *nifX* and *orf2*. The length of 0·6 kb correlates closely with the predicted length of the transcript extending from *orf1* to *nifW*. It appeared that *orf1*, instead of encoding a polypeptide, was just an untranslated leader region of *nifW* after using oligonucleotides specific to *orf1* and *nifW* for primer extension analysis (described below). When total RNA was hybridized by *hesA* and *hesB* probes, respectively, a major band of 2·1 kb was detected (Fig. 8g, h). Since no transcription initiation site was detected for *hesA* or *hesB* (see below), *nifW, hesA* and *hesB* are probably organized as a single operon. The size of 2·1 kb correlates closely to the predicted length of the transcript extending from *nifW* to *hesB* or from *nifW* to 'fdx'. However, the latter possibility was excluded because 'fdx' seemed to be transcribed in a unit different from *nifW*, according to Northern analysis. As shown in Fig. 10(d), two bands with sizes of 0·4 and 1·4 kb were detected by an 'fdx' probe. The size of the major band (0·4 kb) correlated well with the predicted length of the mRNA transcribed from 'fdx'. It is suggested that 'fdx' is organized in a separate operon from *nifW*.

A clear band of 4·4 kb was detected when using *nifX* or *orf2* as probe (Fig. 8c, d). It is possible that the band is a transcript extending past the 'fdx' gene. If this is the case, then we would expect to detect the 4·4 kb band when *orf1, nifW, hesA, hesB* or 'fdx' are used for Northern hybridization, as revealed by the results of the study of the *nifHK* operon in *Anabaena* PCC 7120 (Haselkorn et al., 1986) and *Synechococcus* RF-1 (Huang & Chow, 1990). However, the 4·4 kb band was not detected when *orf1, nifW, hesA, hesB* or 'fdx' were used.

**Identification of the start site by primer extension analysis.** To define the transcriptional unit more precisely, primer extension analysis using synthetic oligonucleotides specific to the coding regions of *nifE, nifN, nifX, orf2, orf1, nifW, hesA* and *hesB*, respectively, was performed. A significant extended product was detected when the oligonucleotide specific to *nifE, nifX, orf1* or *nifW* was assayed. However, no reasonable result was found when *nifN, orf2, hesA* or *hesB* was examined. When a synthetic oligonucleotide complementary to the coding
sequence from position 56 to 33 of nifE was used, the 'G' at position 222 upstream of the start codon of nifE was found to be the transcription start site (Fig. 9a). When an oligonucleotide complementary to the coding sequence from position 58 to 29 of nifX was used, the 'G' at position 45 upstream of the start codon of nifX was shown to be the start site (Fig. 9b). These results indicate that nifE-nifN and nifX-orf2 are organized in separate operons.

When an oligonucleotide complementary to the coding region from position 198 to 178 of orf1 was used, the extension site was found to be the 'G' situated at position 7 within the coding region of orf1 (Fig. 9c).

When a synthetic oligonucleotide complementary to the coding sequence from position 61 to 38 of nifW was used, as shown in Fig. 9(d), the same 'G' identified in orf1 was also found to be the start site of nifW. Thus, it appears that the region of 'orf1' is possibly an untranslated leader region of nifW rather than a gene. The primer extension data support the suggestion that nifE-nifN, nifX-orf2 and nifW-hesA-hesB, respectively, are organized in a transcriptional unit.

hesA, hesB and 'fdx' genes are expressed only under conditions of nitrogen fixation. Fig. 8(g) and (h) show that the hesA-like and hesB-like genes in Synechococcus RF-1 are expressed in a rhythmic pattern similar to the nif genes; however, to reduce the size of the figure, the data for 'fdx' are not included.

Fig. 8. Rhythmic expression of nifE (a), nifN (b), nifX (c), orf2 (d), orf1 (e), nifW (f), hesA (g) and hesB (h) in Synechococcus RF-1 examined by Northern hybridization. Samples for RNA isolation were taken at 4 h intervals (as indicated above each lane) after the 12 h L/12 h D-entrained culture was transferred to L/L. An equal amount of RNA (about 5 µg) was loaded on each lane. DIG-labelled DNA probe specific to each gene was prepared by PCR. The estimated sizes of major hybridizing bands are shown by arrowheads. (The expression pattern of 'fdx' was similar to the nif genes; however, to reduce the size of the figure, the data for 'fdx' are not included.)
Fig. 9. Primer extension analysis using the nifE- (a), nifX- (b), orf1- (c) and nifW-specific (d) primers. Lanes 1, product of primer extension (arrow) when a synthetic oligonucleotide was annealed with total RNA and extended as described in Methods. The synthetic oligonucleotides specific to nifE, nifX, orf1 and nifW are described in Methods.

Fig. 10. Comparison of the expression patterns of hesA (a), hesB (b) and ‘fdx’ (d) between cultures grown in nitrate-containing (BG-11) and nitrate-free (BG-11o) media. The transcript levels of hesA, hesB and ‘fdx’ were assayed by Northern hybridization. Samples for RNA isolation were taken at 18 and 22 h after the 12 h L/12 h D-entrained cultures were transferred to LL. Northern hybridization of nifX (c) under the same conditions was performed as control.

DISCUSSION
Repeating sequences in the non-coding regions

The occurrence of repeating sequences in the non-coding regions of nif genes in heterocyst-forming cyanobacteria has been documented in Anabaena PCC 7120 (Mulligan & Haselkorn, 1989), Anabaena azollae (Jackman & Mulligan, 1995) and Calothrix sp. PCC 7611 (Mazel et al., 1990). The function of the repeats is unknown; however, it has been suggested that they may be important for the regulation of transcription (Mulligan & Haselkorn, 1989), chromosome maintenance (Mazel et al., 1990) or responsible for mRNA stability or processing (Jackman & Mulligan, 1995).

The repeating sequences that occur in nif operons of heterocyst-forming cyanobacteria were not found in Synechococcus RF-1. However, an unusual DNA sequence, (CCCCCACTCTCCCCACT)₅ (the underlined ‘T’ is replaced by ‘C’ in the fifth copy), was found to be situated in the non-coding region between nifB and nifP (transcribed in opposite directions). The function of the repeating sequence is not known. However, in view of the high G + C ratio and the location at the middle of the non-coding region between nifB and nifP, it might
Synechococcus RF-1 has been studied in this paper and containing nip, fdxN, nifS and nip, is located upstream of Anabaena PCC 7120 and other cyanobacteria tested, but approximately that the degradation may occur immediately following the translation and the introduction of such an unusual feature in the untranslated leader may have a role in mRNA processing. On the other hand, since the transcripts of nifE and nifN are relatively unstable, it is also possible that the unusual feature may play a role in increasing the translation efficiency.

Organization of nif and nif-associated genes

Comparison of the physical organization of nif and nif-associated genes between Klebsiella pneumoniae and Anabaena PCC 7120 (Dean & Jacobson, 1992) or among Klebsiella pneumoniae, Azotobacter vinelandii, Bradyrhizobium japonicum, Rhodobacter capsulatus and Clostridium pasteurianum (Haselkorn & Buikema, 1992; Dean & Jacobson, 1992) has been reviewed. Characterization of 16 nif and nif-associated genes in Synechococcus RF-1 has been studied in this paper and our previous report (Chen et al., 1996). These genes are clustered in a continuous arrangement spanning approximately 18 kb with seven transcriptional units (Fig. 11). They all have the same reading direction except nifP. The structural genes for nitrogenase (nifH, nifD and nifK) are arranged as a nifHDK operon, an arrangement common among diazotrophs. The nifB operon, containing nifB, fdxN, nifS and nifU, is located upstream of nifH. The organization of the nifB operon (nifB-fdxN-nifS-nifU) in Synechococcus RF-1 is the same as in Anabaena PCC 7120 and other cyanobacteria tested, but different from the other diazotrophs. The nif and nif-associated genes located downstream from nifK in Synechococcus RF-1 were found to be arranged nifE-nifN-nifX-orf-nifW-hesA-hesB-'fdx'. The organization is nearly the same as in Anabaena PCC 7120 except that ORF-3 located between nifK and nifE in Anabaena PCC 7120 (U47055) was not found in Synechococcus RF-1. The operon containing nifP, located upstream of nifB, and that containing ‘fdx’, located downstream from hesB, in Synechococcus RF-1 have not been fully characterized. Recently, a nif gene encoding an oxidoreductase in Anabaena has been identified (Bauer et al., 1993; Schmitz et al., 1993). Therefore, it is likely that more nif and nif-associated genes in Synechococcus RF-1 will be revealed after further investigation.

Rhythmic expression of nif and nif-associated genes

Based on the results of this study and our previous report (Huang & Chow, 1990), all nif and nif-associated genes examined in Synechococcus RF-1 were expressed in a rhythmic pattern with transcription activity occurring mainly within the dark phase when the culture was grown in a diurnal L/D regimen. It has been suggested that the rhythmic nitrogenase activity is controlled by a ‘circadian oscillator’ (Huang et al., 1990). Several possibilities for an ‘oscillator’ to regulate the circadian rhythm of nif gene expression have been put forward. They may be directly regulated by the product of the ‘oscillator’, or they may be controlled stepwise. The non-coding regions of nifB, nifH and nifE of Synechococcus RF-1 were examined and analysed; however, no significant structural similarity among them was found. The results suggest that the rhythmic expression of nif genes is unlikely to be regulated directly by the product of an ‘oscillator’. It is known that the expression of nitrogen-fixation genes is affected by oxygen concentration. There is a close association between the rhythmic nitrogenase activity and the dark respiration rate in Synechococcus RF-1 (Grobblelaar & Huang, 1991). The dark respiration rate increased considerably a few hours before the increase in nitrogenase activity when Synechococcus RF-1 was grown in a diurnal L/D regimen. Therefore, the rhythmic expression of nif genes can be controlled by the rhythmic change of oxygen concentration which in turn is

---

**Fig. 11.** Physical organization of the nif and nif-associated genes from Synechococcus RF-1. Arrows indicate the position and direction for the transcription initiation sites revealed by Northern hybridization and primer extension analysis. nifP, hesA, hesB and ‘fdx’ were assigned as nif-associated genes because they were expressed only under nitrogen-fixing conditions. The ORF downstream of nifX was identified as a counterpart of ORF-2 in Anabaena PCC 7120. Information concerning the organization is presented in this paper and in our previous publication (Chen et al., 1996).
ACKNOWLEDGEMENTS

This work was funded by the Academia Sinica and the National Science Council of the Republic of China.

REFERENCES


Received 27 May 1998; revised 22 September 1998; accepted 11 November 1998.