Bot. Bull. Acad. Sin. (1996) **37:** 99-105

Nucleotide sequence of the *nifHDK* operon in the aerobic nitrogen-fixing unicellular *Synechococcus* RF-1

Horng-Ming Chen[1,2], Tan-Chi Huang[2,3], and Chiu-Yuan Chien[1]

[1]*Department of Biology, National Taiwan Normal University, Taipei, Taiwan 117, Republic of China*

[2]*Institute of Botany, Academia Sinica, Nankang, Taipei, Taiwan 115, Republic of China*

(Received February 5, 1996; Accepted March 19, 1996)

Abstract. A 5 kb nucleotide sequence which included *nifHDK* operon was determined. The amino acid sequences deduced from the *nifH, nifD,* and *nifK* of *Synechococcus* RF-1 and from *Anabaena* 7120 were compared. The start site of transcription was located by primer extension analysis. A region of 405 bp just upstream from the 5' end of *nifH* showed considerable sequence similarity to the *nifU* sequences of *Anabaena* 7120.

Keywords: Cyanobacteria; *nifHDK* operon; Nitrogen fixation; Nucleotide sequence; *Synechococcus* RF-1.

Introduction

The nitrogenase which catalyzes the reduction of atmospheric nitrogen to ammonia is a complex consisting of three different kinds of polypeptide encoded by *nifH, nifD,* and *nifK,* respectively. The molecular structure of *nifHDK* operon has been analyzed in eubacteria, photobacteria, and cyanobacteria. The nucleotide sequence for the coding region of *nif* genes among different microorganisms is highly conserved. However, the arrangement of *nifH, nifD,* and *nifK* is either in a contiguous or noncontiguous array. The *nifH*, *nifD,* and *nifK* in the eubacteria system such as in *Klebsiella pneumoniae* (Roberts and Brill, 1981), or *Rhizobium meliloti* (Corbin et al., 1982), and in the nonoxygenic photosynthetic bacteria, such as in *Rhodopseudomonas capsulata* (currently, *Rhodobacter capsulatus*), are linked in a single *nifHDK* operon (Avtges et al., 1983). In the cyanobacteria, two types of arrangement are found. Those in the nonheterocystous anaerobic nitrogen-fixing cyanobacteria such as in the *Synechococcus* 7335, *Synechococcus* 7425, or *Pseudoanabaena* 7409 are linked in a contiguous array (Kallas et al., 1985). However, those in heterocystous cyanobacteria such as in *Anabaena* 7120, *nifH* is linked to *nifD*, whereas *nifK* is separated from *nifD* by about 11 kb in the vegetative cells (Mazur et al., 1980). A rearrangement of the *nif* genes takes place during the initiation of heterocyst development (Golden et al., 1985).

Among the unicellular cyanobacteria, a few isolates are able to fix nitrogen aerobically. They can be morphologically grouped into sheathed and sheathless forms (Waterbury and Rippka, 1989). These aerobic nitrogen-fixing cyanobacteria fix nitrogen almost exclusively during the dark periods when grown in a diurnal 12 h/12 h light-dark regimen. The properties of the nitrogenase activity in *Synechococcus* RF-1, one of the sheathless aerobic nitrogen-fixing unicellular cyanobacteria have been well characterized. When the culture was exposed to a diurnal 12 h light/12 h dark (L/D) regimen, the nitrogenase activity exhibited a rhythmic pattern with the nitrogenase activity peak occurring in the dark phase. This rhythmic pattern persisted for several days after the culture was transferred to continuous light (L/L) (Grobbelaar et al., 1986). The nitrogenase activity rhythm was also found to be temperature compensated when the temperature was changed within the physiological range. These results indicate that the nitrogenase activity in *Synechococcus* RF-1 is controlled by a "circadian clock" (Huang et al., 1990), a phenomenon which does not occur with other nitrogen-fixing microorganisms.

The nucleotide sequence of the *nifHDK* operon in heterocystous cyanobacteria has been well characterized. However, there have been no reports on the molecular structure of the *nif* genes in unicellular cyanobacteria. In this report, the nucleotide sequence of *nifHDK* operon in *Synechococcus* RF-1 is examined.

Materials and Methods

Isolation of DNA from Synechococcus RF-1

Axenic *Synechococcus* RF-1 were grown in $BG\text{-}11_0$ according to the conditions described previously (Chou et al., 1989). Since the surface of *Synechococcus* RF-1 contained an unknown compound which will interact with lysozyme, the cells collected by centrifugation were washed twice in distilled H_2O by vibration with Vortex (Model K-550-G) at the maximum speed for 1 min at room temperature. The cells were transferred to 5 ml 6 M guanidine-HCl and vibrated for 1 min (Votex) at room temperature. The guanidine-HCl treatment was repeated once. The cells were then washed with distilled H_2O three times to remove the guanidine-HCl and then suspended in 10 mM Tris-EDTA buffer (TE), pH 8.0. Lysozyme was added to have a final concentration of 10 mg/ml. The cells were then in-

[3]Corresponding author. Fax: 886-2-7827954.

cubated at 37°C with gentle shaking (100 rpm/min) for 1 h. The cells were collected by centrifugation and then resuspended in solution containing 2% SDS, 0.05 M EDTA, and 20 μg/ml RNase A. The cell suspension was incubated at 37°C for 1 h. Proteinase K was added after the incubation to a final concentration of 100 μg/ml followed by re-incubation at 60°C for 1 h. The final suspension was extracted gently with phenol:chloroform (1:1 v/v). The phenol contaminated in the DNA containing aqueous phase was removed by chloroform: isoamylalcohol (24:1 v/v). DNA in the liquid phase was precipitated with 2 times the volume of ethanol (-20°C) in the presence of 0.2 M NaCl. The precipitated DNA was collected with a glass stirring rod and then solubilized in TE buffer.

Construction of a Genomic Library

A genomic DNA library was prepared using DNA from *Synechococcus* RF-1. Genomic DNA was partially digested with *Sau*3A, and the DNA fragments with sizes from 10 to 23 kb were isolated from low melting-point agarose (Sambrook et al., 1989). The size-fractionated fragments were ligated into *Bam*H1-digested EMBL4 according to the protocol provided by Stratagene Company, La Jolla, California, USA.

Cloning of the nif Genes

The genomic library of *Synechococcus* RF-1 prepared in λ EMBL4 from a partial *Sau*3A digest was used for the *nif* gene cloning. Plasmids pBR322 containing *nifH* or *nifK* cloned from *Anabaena* (Rice et al., 1982) were used as probes for cloning the *nif* genes. The *nifH* or *nifK* probes were labeled with the digoxigenin using the DIG luminescent detection kit according to the instructions of the manufacturer Boehringer Mannheim. About 3,000 recombinant plaques of the genomic library were immobilized on nylon membranes (Micron Separations Inc.) and pretreated according to the procedures described by Sambrook et al. (1989). The membranes were hybridized with the digoxigenin (DIG)-labeled *nifH* or *nifK* probe according to the instructions of the manufacturer Boehringer Mannheim. Positive plaques were purified and amplified for DNA extraction.

Nucleotide Sequencing

Restriction fragments containing the *nif* gene were subcloned into pGEM7(-). The nucleotide sequence of both strands of the plasmid DNA was determined by dideoxy chain termination on an A.L.F. Pharmacia automated DNA sequencer according to the manufacturer's protocols. Sequencing of the double-stranded DNA was first carried out with the universal sequencing primer using the Auto Read Sequencing Kit (Pharmacia-LKB) following the supplier's instructions. Oligonucleotide primers were utilized to complete sequencing of both strands. The extent of identity of the derived amino acid sequence of *Synechococcus nif* genes to those encoded by other *nif* genes was searched for in the PC Gene, Intelli Genetics.

Isolation of Total RNA

Exponentially growing cells (about 2×10^7 cells/ml) of *Synechococcus* RF-1 were collected by centrifugation, washed with 1 ml of distilled water, and then collected in a 1.5 ml screw-top vial (about 6×10^8 cells per vial). Teflon-coated sea-sand (0.1 g; 0.1~0.3 mm diameter; Merck), 0.5 ml solution D (4 M guanidine thiocyanate, 25 mM sodium citrate, 0.5% sodium lauryl sarcosinate, and 0.72% mercaptoethanol), 0.5 ml water saturated hot phenol solution (80°C), 0.1 ml 2 M sodium acetate (pH 4.0), and 0.1 ml Chloroform-isoamyl alcohol (24:1) were added. The cells were broken by vibration with a Mini-Beadbeater (Biospec, USA) set at high speed for 50 sec followed by 300 sec at low speed. The broken cells were cooled in ice water for 5 min and then centrifuged at 13,000 *g* for 5 min. The supernatant was collected and an equal volume of absolute ethanol was added. The total RNA was collected after cooling at -70°C for more than 1 h, and then suspended in 0.3 ml diethyl pyrocarbonate (DEPC)-pretreated water in the presence of 10% (v/v) 3 M sodium acetate (pH 4.8). The RNA solution was repeatedly extracted with an equal volume of phenol-chloroform (1:1). The RNA containing aqueous phase was collected and maintained at -70°C after adding absolute ethanol (-20°C) in an amount 2.5 times its volume.

Primer Extension Analysis

The RNA prepared according to the procedures described under "Isolation of Total RNA" above was used as template. An oligonucleotide with the sequence 5'GACTTACCGATACCGCCTTTTCCGTAA3', which is complementary to the coding sequence from positions 44 to 18 of the *nifH* gene (Figure 1), was used as primer. About 5 to 20 μg RNA dissolved in DEPC-pretreated water was mixed with 2 μl primer (10 μg/ml) in a 1.5 ml Eppendorf tube. The mixture was kept in a 65°C water bath for 10 min and then reset at 42°C. After the temperature had decreased from 65°C to 42°C naturally (about 30 min), 1.5 μl of 10xRT buffer (reverse transcription buffer, Stratagene), 2 μl of diluted dNTP mix (Sequenase Version 2.0 DNA sequencing kit, United States Biochemical), 2

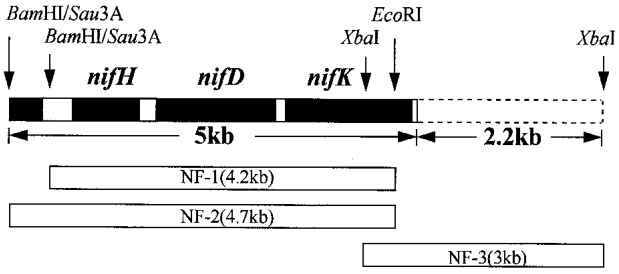

Figure 1. Schematic representation of a 7.2 kb locus of the *Synechococcus* RF-1 chromosome containing the *nifHDK* operon. The locations of the DNA fragment from the subcloned recombinants (NF-1, NF-2, and NF-3) used for the sequence analysis are shown by open bars. The 5 kb region was sequenced (Figure 2). The unsequenced downstream region (2.2 kb) is indicated by the dashed line.

```
1     GATCTGACCACAGCCGAACAAGTTACCAACTATGTTAAAGCCGGTGGAGGATGTAGTTCT
61    TGTCTGTCTGATATCGATGATATTTTAGCCGATATTACTCAAGAAAAAGCCACCGCCGTC
121   ACAGCAGCCACCGAACTTGTTCAAAGTAAATTAACTCCCCAAAAACCCCTAAATAACTTA
181   CAAAAAATCACCCTCATTCAACAAATTCTCGACGAAGAAATTAAACCAGCCTTGGCAAAA
241   GATGGAGGAGATGTAGAGTTATTTGATGTCGAAGGAGATTTGGTCAAAGTGATATTACAA
301   GGAGCCTGTGGTTCCTGTGCCAGCAGTACCCAAACCTTAAAAATGGGAATCGAAGCCAGA
361   TTACGAGAGCGTGTTTCTCCTGAGTTAACGGTTATTTCTGTGTAAATTCTTGCCTTATCT
421   TGCTCCGCAAGTCCACCCGAATTATGACCCAATTAAACACAAAAGTCGAGAGCGGTCGCC
481   ACCATAAGCATCGCTTTCACCCCAGACATTTCAAGAACAAAAATCACCTAAAAACCCCGT
541   ACTTTTTGTTGCTTCATCTATCCCTAACCATGACGTAAAGAGCGAGGCTTCGACGGGACG
601   TAATTTCCCACTGAATTAAGCAAAGAACGAATCATCTTTACGTCGAAATTGGAGAGAAAT
661   CTCCGACCTTGTATTCACACTAAACTCTCAACAACAAGCGAGAAATCAACCATGCGTCAA
721   ATTGCATTTTACGGAAAAGGCGGTATCGGTAAGTCTACCACTTCTCAAAACACCTTAGCC
781   GGAATGGCTCAAGCTGGCAACCGCATCATGATTGTTGGTTGTGACCCCAAAGCTGACTCT
841   ACCCGTTTGATCCTCAACTGTAAAGCTCAGGTAACGGTATTACACTTAGCTGCCGAACGT
901   GGTGCTGTTGAAGATCTCGAACTCTCAGATGTATTACTCACTGGGTTTGAAAACATCAAG
961   TGTGTTGAATCTGGTGGTCCCGAACCTGGGGTTGGTTGTGCTGGACGTGGTATTATCACC
1021  TCCATCAACTTCCTTGAAGAAGAAGGTGCTTATGAAGATCTAGACTTCGTATCCTATGAC
1081  GTATTAGGAGACGTTGTTTGTGGTGGTTTCGCTATGCCTATCCGTGAAGGAAAAGCACAA
1141  GAAATCTACATCGTTACCTCTGGGGAAATGATGGCGATGTATGCAGCTAACAACATCGCT
1201  CGTGGGATCTTAAAATATGCCCACACTGGTGGTGTTCGTTTAGGTGGTTTAATTTGTAAC
1261  AGCCGTAACGTTAACAAAGAGATCGAATTGATCGAAGAGTTAGCTGAACGCTTAAACACC
1321  CAAATGATCCACTTCGTACCCCGTTCCAAACAGGTACAAGAAGCTGAATTACGTCGTCAG
1381  ACCGTTATTCAATACTCCCCTGAGCACCCCCAAGCTCAAGAATACCGTGATTTAGGTGAC
1441  AAGATCGTTAACAACACCAAACTCACCATCCCCACTCCTATCGACAACGACGAACTCGAA
1501  GAACTGTTGATCAACTACGGTTTACTTGGCTCAGAAGAAGAGTACAAGAAAGTTATGGAA
1561  GCTGACATGGCTGCCCAAGCTTTAACCAGAGGCGCGAAGTAATGTAAGGGCTTAATTACC
1621  TTAACCCCCTACTATCGGGAAATGGGGGAGTTTAAGAAGCCAAAAGGCAAAAGGCAGAAG
1681  GCAACACGGACAATCATTATCCCTAAGTCCTCACTCCTAACCCCTAACCCCTAACTCCCT
1741  TTACCCCCATCCTCATCAAAATTTCCTACTCGTCATTATTGACGAAGGGGGTGCGTCCGA
1801  TCCTAATCGCCAATTCATCACTGAGGAACACTATGTCAACAGTAGAAGACAGAAAGCAGC
1861  TTATCCAAGACGTTCTTGATACCTATCCTGAGAAGTTAGCCAAGAAACGGTCTAAACACC
1921  TCAATGTTTACGAAGAAGGCAAAGACGATTGTGGAGTAAAATCTAACATTAAGTCTGCAC
1981  CTGGTGTAATGACCGCTCGTGGTTGTGCTTATGCAGGATCTAAAGGGGTGGTTTGGGGTC
2041  CTATCAAAGATATGATCCATATCTCCCACGGACCTGTTGGTTGCGGTTACTACTCTTGGT
2101  CTGGTCGTCGTAACTATTACATCGGAACCACTGGGGTTGATACCTTTGGTACGATGAACT
2161  TTACCTCTGACTTCCAAGAAAAAGACATCGTTTTTGGTGGAGACAAAAAACTCCTCAAAA
2221  TCACCGAAGAAATCGAAGAATTATTCCCCCTCCACAATGGGATTTCCATTCAGTCTGAAT
2281  GTCCTGTTGGATTAATTGGGGATGACATCGAAGGTGTTGCCAAAAAAGCGCAAAAAATTA
2341  CTGGCAAACCCGTATTCCCGTTCCGTTGTGAAGGATTCCGTGGCGTTTCCCAATCCTTAG
2401  GACACCACATCGCTAACGACGCAGTGCGTGACTGGGTATTTAGCCGTGATGATGCTCAAG
2461  AAATCGAAACCACTCCCTATGATGTTGCCATCATTGGAGACTACAACATCGGTGGAGATG
2521  CTTGGTCTAGCCGTATTCTTCTCGAAGAAATGGGTCTGCGCGTCGTTGCTCAATGGTCTG
2581  GAGACGGAACCATCAACGAAATGATGCAAACCCCCAAAGTGAAACTCAACCTGATTCACT
2641  GTTACCGTTCCATGAACTACATCAGTCGTCACATGGAAGAAAAATACGGTATTCCCTGGT
2701  TTGAGTACAACTTCTTTGGTCCTACCAAGATTGCTGAATCCTTACGCGCGATCGCTGCTC
2761  TGTTTGATGACACCATCAAAGAAAATGCAGAGAAAGTAATTGCTAAGTACGAACAACAAA
2821  CCGCAGAAGTCTTAGCCAAATACCGTCCTCGTTTGGAAAACAAAACCGTCATGATGATGG
2881  TGGGTGGACTACGTCCTCGTCACGTTGTTCCTGCTTTCACAGACTTAGGCATGAAAATGA
2941  TCGGAACCGGATATGAGTTCGCTCACGGTGACGACTATAAACGTACCACTGAGTATGTTG
```

```
3001 ATGATGCAACCCTCATCTATGATGACGTAACTGCCTACGAGTTCGAGAAATTCGTTCAAG
3061 AACTGAAACCCGACTTAGTTGCTTCTGGCGTTAAAGAGAAGTATGTCTTCCAGAAAATGG
3121 GACTACCTTTCCGTCAAATGCACTCTTGGGATTACTCTGGTCCTTACCACGGTTATGATG
3181 GGTTCGCTATCTTTGCACGGGATATGGACTTAGCTCTCAATAACCCGACCTGGGGATTAA
3241 TCAAATCTCCTTGGAATAAGTAAGAGGGAATTAGGGAATAGGGAACAGGGAACAGGGAAT

3301 AGTGAATTGGGAAGTCCTTCACTAACTCTCTAGTTACGAAGCCACCCCCTCACCCCCTCC
3361 CCCCCCTCATCTCCTCACCTCCTCCCCCTTTCTCCACAACTGCATCATTAATCCTTGACC
                          nifK
3421 GAACAACGGAGTATCACGAATGTCTCAGAAAATTGATAAAATCCAAGACCACGTTGAGTT
3481 ATTCCACCAACCAGAGTACCAAGAGCTATTTGAAAACAAGAAAGCTCTCCAAGGAATGGC
3541 TTCTGATGAGAAAGTCGCTGAAATAGCCGAATGGACCAAAACCTGGGAATATCGGGAAAA
3601 GAACTTCGCTCGTGAAGCTCTGACCATCAACCCCGCTAAAGCTTGTCAACCTTTGGGTGC
3661 TATCTTAGCTGCGGTTGGTTTTGAAGGAACCCTCCCCTTTGTGCATGGATCACAAGGTTG
3721 TGTGGCTTACTTCCGTACCCACTTTACCCGTCACTTCAAAGAGCCTTTCAGTGGTGTTTC
3781 TTCTTCCATGACTGAAGATGCAGCCGTCTTCGGTGGACTGAAAAACATGATCGAAGGGTT
3841 ACAGAATGCTTATAGTCTCTATCAACCCAAAATGATTGCTGTCTGTACAACTTGTATGGC
3901 AGAAGTTATTGGGGATGACTTAGGTTCCTTCATTGGCAATGCTAAGGCTGACGGTTCTGT
3961 TCCTAAAGATTTCCCCGTTCCCTTTGCTCACACTCCTTCTTTCGTGGGTTCTCATATCAC
4021 GGGATATGACAACATGATGAAAGCCATCCTGTTGAACTTAACCGACGGCAAGAAACCTAC
4081 CACCAGCAACGGTAAAGTTAACTTTATTCCTGGGTTTGAAACCTATGTTGGTAACCTACG
4141 CGAACTGAAGCATTTAACCAGTGCTATGGGGGTTGATGCTACCATTTTAGGAGACAACGA
4201 ACTCTATTTAGATTCTCCTAACGATGGCGAGTTCAAAATGTACCAAGGTGGTACTACTCT
4261 AGAAGAAGGTGCTGATGCTATCAATGCAACCAAAACCATCGCACTGCAAACCTATCCCAC
4321 CGTTAAAACCCTCGAATACATCGAGAAAGAATGGGCAGCAACCCACCGCTACCTATCGTC
4381 CTTGGGGGGTATTAAAGGAACGGATGAGTTCGTCATGGCTTTATCTGAACTCACTGGGAA
4441 TCCTGTTCCTCCCGAATTGGAACTAGAACGGGGACGCGCAGTGGATGCTATGACCGATAG
4501 TCATGCTTGGTTACATGGTAAAAAAGCGGCTATCTATGGCGACCCTGACTTAGTCATGGG
4561 AATGCTGCAATTCATGTTAGAGATGGGTGTTGAACCTGTTCACGTTTTGGTTCACAACTC
4621 TACCACTGAATTTGAAGAAGAAGCCAAAGCTCTCTTAGCTTCTAGTCCTTATGGTCAAAA
4681 AGCCACCGTTTGGGGCGGTAAATACCTCTGGCACCTCCGTTCCTTACTGTTTACTGAACC
4741 TGTTGACTTCTTAATCGGGAATTCCTACGGTAAATACCTCTGGCGTGATACCAAGATTCC
4801 TTTAATCCGCATCGGGTATCCTATCTTTGATCGCCACCACTTACATCGCTATTCTACCAT
4861 TGGTTACAATGGCGCGATTAACCTGCTCAATTGGATTGTTAATGGTCTGTTTGAAGAAAT
4921 CGACCGCAACACCAATATCCCCTCGAAGACCGACATTTCCTTCGATTTAGTTCGTTAATC
4981 AAAGGTAGACAAGGAAGTAA
```

Figure 2. The 5 kb nucleatide sequence containing the *nifHDK* operon of the *Synechococcus* RF-1 Chromosome (Gen Bank U 22146). The open reading frame is represented by bold characters, and the non-coding region is shown in italics. The start point (underlinded G) of transcription determined by primer extension analysis is marked by an arrow. The start codon for each open reading frame is underlined. The segment from position 729 to 755 selected for primer extension analysis is also underlined.

μl [α-^{35}S]-dATP (10 μCi/μl), 1 μl RNase inhibitor (RNase Block, 25,000–40,000 U/ml, Stratagene), and 1 μl RNase H$^-$ reverse transcriptase (MMLV-RT, 50,000 U/ml, Stratagene) were added. Autoclaved water was added up to a final volume of 14 μl. After 10 min of incubation at 42°C, 1 μl of 25 mM dNTP was added and incubated at 42°C for 1 to 2 h. The reaction was terminated by adding an equal volume of sequencing loading dye (United States Biochemical). For the control experiment, the same synthetic oligonucleotide was used as a primer in dideoxy chain termination DNA sequencing reactions with the *nifH* ORF as a template. The final reaction mixtures were analysed by 6% DNA sequencing gel.

Results and Discussion

Among the *nif* gene positive recombinants subcloned from the λ EMBL4 genomic library, recombinant NF-1, NF-2, and NF-3 were selected for DNA analysis (Figure 1). The NF-1 and NF-2, both hybridized with probe *nifH* and *nifK,* contained a 4.2 kb and 4.7 kb *Eco*RI fragment respectively. The NF-3 contained a 3 kb *Xba*1 fragment, which hybridized only with probe *nifK*. Results of DNA sequencing revealed that the 4.7 kb fragment overlaps with the 4.2 kb fragment, and the 4.7 kb fragment has 555 base pairs longer than the 4.2 kb at the 5'-end (Figure 1). The 5'-end of the 3 kb *Xba*1 fragment has 507 base pairs that

overlap with the 3'-end of the 4.7 kb *Eco*RI fragment. Based on the nucleotide sequence for fragment 4.2 kb, 4.7 kb, and 3.0 kb, and the sequence overlapping between them, a 5 kb segment containing *nif* gene was constructed (Figure 1).

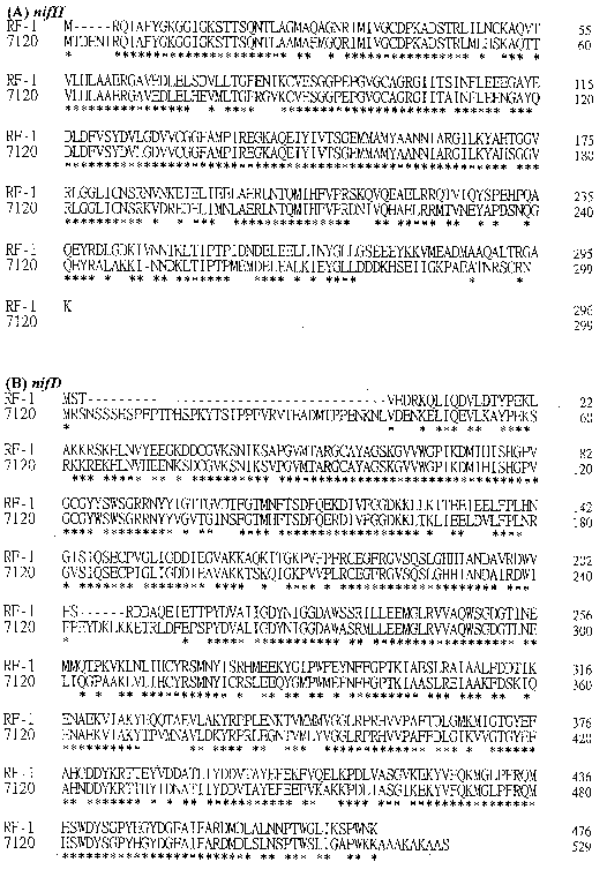

Figure 3. Comparison of the deduced amino acid sequences of the *nifH* genes (A), the *nifD* genes (B), and the *nifK* genes (C), from *Synechococcus* RF-1 and *Anabaena* 7120. The character to show that two aligned residues are identical is "*". Gaps (shown as dashes) have been introduced in the sequences where necessary to give better alignment.

The nucleotide sequence of the *nif* genes containing fragment along with its 5'-flanking sequence is given in Figure 2. Putative open reading frame (ORF) in the sequenced region was searched in all three possible translation frames with ATG as the initial codon. Three complete ORFs are present in the constructed *nif* gene locus. As shown in Figure 2, the first ORF beginning from the position 712 consists of 891 base pairs (bp). It encodes a polypeptide consisting of 296 amino acids (Figure 3A). The predicted amino acid sequence exhibits a high degree of similarity (76.7%) to that of the polypeptide encoded by the *nifH* from *Anabaena* 7120 (Mevarech et al., 1980). Therefore, we concluded that the ORF is the *nifH* from *Synechococcus* RF-1.

The second ORF is present 230 bp downstream from the termination codon of the *nifH* (from position 1833 to 3263). The ORF encodes a polypeptide of 476 amino acids (Figure 3B). As shown in Figure 3B, the predicted

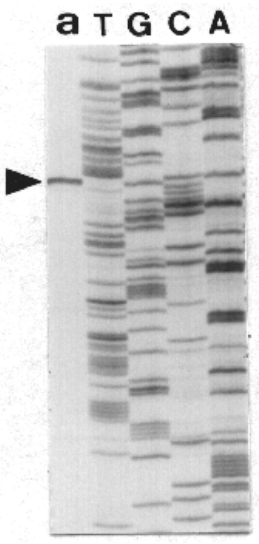

Figure 4. Primer extension analysis of the 5' end of *nifH* gene. Lane a: A 27 oligonucleotide was annealed with total RNA and extended as described under "Materials and Methods". Lanes T,G,C,A: Electrophoretic separation of the products of dideoxynucleotide sequencing.

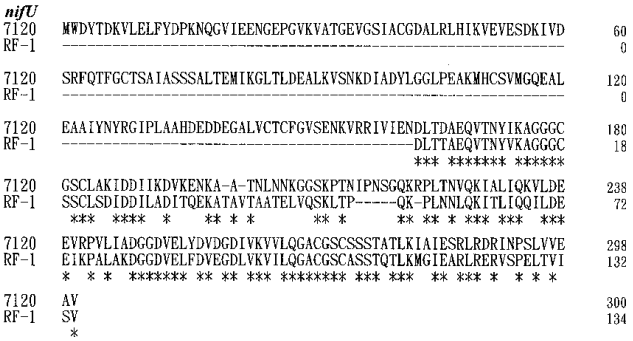

Figure 5. Alignment of amino acid sequences deduced from the *nifU* gene of *Anabaena* 7120 and the partial sequence of the *nifU* of *Synechococcus* RF-1.

amino acid sequence exhibits a high degree of similarity (77.3%) to that of the polypeptide encoded by the *nifD* from *Anabaerna* 7120 (Lammers and Haselkorn, 1983; Golden et al., 1985). The third ORF is present 176 bp downstream from the termination codon of the *nifD* (from position 3440 to 4978). As shown in Figure 3C, the amino acid sequence deduced from the third ORF has a 69% similarity with the polypeptide encoded by the *nifK* from *Anabaena* 7120 (Mazur and Chui, 1982). Therefore, we concluded the second and the third ORF correspond to the *nifD* and *nifK* of *Synechococcus* RF-1, respectively.

Transcription of the *nifHDK* operon of *Synechococcus* RF-1 was examined previously by Northern analysis (Huang and Chow, 1990). When mRNA from *Synechococcus* RF-1 was hybridized with *nifK* probe, a transcript of about 4.5 kb was detected. The size of the transcript corresponds to the size that includes the *nifH*, *nifD,* and *nifK* sequences. When the mRNA was probed with *nifH*, three bands of the sizes 1.3, 2.8, and 4.5 kb were detected, with the 1.3 kb as the most abundant band. These results indicate that *nifH* is the first gene of the operon. Primer extension analysis was used to locate the start site of transcription. As shown in Figure 4, a single transcript was detected. The transcription start site lies 247 bp upstream from the start codon of *nifH* gene (Figure 2).

A *nifU* gene was shown to be located upstream from the 5' end of the *nifH* gene from *Nostoc muscorum* (DeFrancesco and Potts, 1988), *Anabaena* 7120 (Mulligan and Haselkorn, 1989), and *Plectonema boryanum* (Fujita et al., 1991). The *nifU* gene has been suggested to be required for maturation of the MoFe protein of nitrogenase (Orme-Johnson, 1985). A partial ORF (405 bp) located 5' to the *nifH* gene of *Synechococcus* RF-1 (Figure 2) was sequenced. As shown in Figure 5, the deduced amino acid sequence (294 residues) of the partial ORF has considerable homology (higher than 64%) to the C-terminal region of the *nifU* product from *Anabaena* 7120 (Mulligan and Haselkorn, 1989). The results indicate that the same *nifU-nifH* order probably occurs in *Synechococcus* RF-1.

Acknowledgements. We thank Professor R. Haselkorn, University of Chicago, for providing the pAn207.8 and pAn154.3 plasmids. This work was funded by the Academia Sinica and the National Science Council of the Republic of China.

Literature Cited

Avtges, P., P. A. Scolnik, and R. Haselkorn. 1983. Genetic and physical map of the structural genes (*nifH, D, K*) coding for the nitrogenase complex of *Rhodopseudomonas capsulata.* J. Bacteriol. **156:** 251–256.

Chou, H. M., T. J. Chow, T. Jenn, H. R. Wang, H. C. Chou, and T.C. Huang. 1989. Rhythmic nitrogenase activity of *Synechococcus* RF-1 established under various light-dark cycles. Bot. Bull. Acad. Sin. **30:** 291–296.

Corbin, D., D. Ditta, and D. R. Helinski. 1982. Clustering of nitrogen fixation (*nif*) genes in *Rhizobium meliloti.* J. Bacteriol. **149:** 221–228.

DeFrancesco, N. and M. Potts. 1988. Cloning of *nifHD* from *Nostoc muscorum* UTEX584 and a flanking region homologous to part of the *Azotobacter vinelandii nifU* gene. J. Bacteriol. **170:** 3297–3300.

Fujita, T., Y. Takahashi, F. Shonai, Y. Ogura, and H. Matsubara. 1991. Cloning, nucleotide sequences and differential expression of the *nifH* and *nifH*-like (*frxC*) genes from the filamentous nitrogen-fixing cyanobacterium *Plectonema boryanum*. Plant Cell Physiol. **32:** 1093–1106.

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

Grobbelaar, N., T. C. Huang, H. Y. Lin, and T. J. Chow. 1986. Dinitrogen-fixing endogenous rhythm in *Synechococcus* RF-1. FEMS Microbiol. Letts. **37:** 173–177.

Huang, T. C. and T. J. Chow. 1990. Characterization of the rhythmic nitrogne-fixing activity of *Synechococcus* sp. RF-1 at the transcription level. Curr. Microbiol. **20:** 23–26.

Huang, T. C., J. Tu, T. J. Chow, and T. H. Chen. 1990. Circadian rhythm of the prokaryote *Synechococcus* sp. RF-1. Plant Physiol. **92:** 531–533.

Kallas, T., T. Coursin, and R. Rippka. 1985. Different organization of *nif* genes in nonheterocystous and heterocystous cyanobacteria. Plant Mol. Biol. **5:** 321–329.

Lammers, P. J. and R. Haselkorn. 1983. Sequence of the *nifD* gene coding for the α subunit of dinitrogenase from the cyanobacterium *Anabaena.* Proc. Natl. Acad. Sci. USA **80:** 4723–4727.

Mazur, B. J. and C. F. Chui. 1982. Sequence of the gene coding for the β-subunit of dinitrogenase from the blue-green alga *Anabaena.* Proc. Natl. Acad. Sci. USA **79:** 6782–6786.

Mazur, B. J., D. Rice, and R. Haselkorn. 1980. Identification of blue-green algal nitrogen fixation genes by using heterologous DNA hybridization probe. Proc. Natl. Acad. Sci. USA **77:** 186–190.

Mevarech, M., D. Rice, and R. Haselkorn. 1980. Nucleotide sequence of a cyanobacterial *nifH* gene conding for nitrogenase reductase. Proc. Natl. Acad. Sci. USA **77:** 6476–6480.

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

Orme-Johnson, W.H. 1985. Molecular basis of biological nitrogen fixation. Annu. Rev. Biophys. Chem. **14:** 419–459.

Rice, D., B. J. Mazur, and R. Haselkorn. 1982. Isolation and physical mapping of nitrogen fixation genes from the cyanobacterium *Anabaena* 7120. J. Biol. Chem. **257:** 13157–13163.

Roberts, G. P. and W. J. Brill. 1981. Genetics and regulation of nitrogen fixation. Ann. Rev. Microbiol. **35:** 207–235.

Sambrook, J., E. F. Fritseh, and T. Maniatis. 1989. Molecular Clonong: A Laboratory Manual. Cold Spring Harbor Laboratory, New York.

Waterbury, J.R. and R. Rippka. 1989. Order *Chroococcales* Wettstein 1924, emend, Rippka et al., 1979, *In* J. T. Staley, M. P. Bryant, N. Pfenning, and J. G. Holt (eds.), Bergey's Manual of Systematic Bacteriology, vol 3. Williams & Wilkins, Baltimore, pp. 1728–1746.

好氣單胞固氮聚球藻之固氮基因座的核甘酸序列

陳弘民[1,2] 黃檀溪[1] 簡秋源[2]

[1] 中央研究院植物研究所
[2] 國立台灣師範大學生物研究所

聚球藻 RF-1 品系 (*Synechococcus* RF-1) 為一具有固氮週期韻律的藍綠藻，本文對其 *nifHDK* 基因座的構造加以研究。一段長約 5,000 個氮鹼基的基因序列已被解讀，與魚腥藻 (*Anabaena* 7120) 之固氮基因操縱組 *nifHDK* 做胺基酸序列比對時，約有 75% 之相似度。 *nifHDK* 之轉錄起始位於轉譯起始之前 247 個氮鹼基位置，在 *nifHDK* 之前 (5' 端) 證明為 *nifU* 基因。

關鍵詞：藍綠藻；固氮基因；固氮作用；核甘酸序列；聚球藻 RF-1 。