Cloning and Nucleotide Sequence of a frxC-ORF469 Gene Cluster of Synechocystis PCC6803: Conservation with Liverwort Chloroplast frxC-ORF465 and nif Operon

Yutaka Ogura,† Miho Takemura, Kenji Oda, Katsuyuki Yamato, Eiji Ohta, Hideya Fukuzawa, and Kanji Ohyama‡

Department of Agricultural Chemistry, Faculty of Agriculture, Kyoto University, Kyoto 606-01, Japan
Received December 2, 1991

A gene, frxC, which is unique to the chloroplast genome of the liverwort Marchantia polymorpha, has sequence similarity to nifH, the product of which is an iron protein of a nitrogenase. Although frxC is expressed to produce a protein in liverwort chloroplasts, its function is not known. Using a probe of liverwort chloroplast DNA, a 10.1-kb region containing a gene cluster consisting of open reading frames (ORF278-frxC-ORF469-ORF248) was isolated from the cyanobacterium Synechocystis PCC6803. In this region, frxC and ORF469 showed sequence similarities to liverwort chloroplast frxC (83%) and immediately downstream ORF465 (74%), respectively. Synechocystis frxC showed 31% amino acid sequence identity with nihH from Clostridium pasteurianum. Additionally, Synechocystis ORF469 showed a sequence similarity (19% identity) to C. pasteurianum nihB product, which is the β subunit of a molybdenum-iron protein of a nitrogenase complex. Conservation of the gene arrangement between liverwort and Synechocystis suggests that the liverwort chloroplast frxC-ORF465 cluster may have evolved from an ancestor common to Synechocystis, and that these two genes may have been transferred to the nuclear genome in tobacco and rice during evolution.

The complete nucleotide sequences of the chloroplast genome have been reported in three kinds of land plants, liverwort,† tobacco,‡ and rice.‡ Although more than 130 possible coding regions have been predicted in comparative analysis with nucleotide and amino acid sequence databases, over 30 open reading frames (ORFs) remain to be identified.§ Gene arrangements on chloroplast genomes from three organisms show significant similarity except for some inversions of parts of the large single copy region. Deducing amino acid sequences of gene products also show significant similarity.† In the liverwort chloroplast genome, two ORFs having structural similarity to 4Fe-4S-type ferredoxin have been designated as frxA and frxB. These two ORFs are also found in tobacco and rice chloroplast genomes, and have recently been identified as a gene for the 9-kDa protein in photosystem I, psaC, and a gene for a component of NADH dehydrogenase, "ndhK" or "ndhJ". In addition, an ORF named frxC, the deduced amino acid sequence of which shows significant similarity to an iron protein of nitrogenase (nihH product) in nitrogen-fixing bacteria, is uniquely located on the small single copy region of the liverwort chloroplast genome.‡ ORF465, which could encode 465 amino acid residues, is also a liverwort-specific ORF 42 bp downstream of the coding region of frxC on the liverwort chloroplast genome. Although functions of these ORFs are not yet known, frxC is actively expressed to make a polypeptide that crossreacts with an antibody against the truncated frxC gene product expressed in E. coli.‡ To know more about the evolution and functions of these ORFs, we chose to analyze the genome of the cyanobacterium Synechocystis PCC6803, because cyanobacteria have plant-like oxygen-evolving photosynthesis and are thought to be a kind of ancestor of plant chloroplasts in the symbiotic theory.‡ Here we report the presence of a frxC-ORF469 (ORF465 homologue) gene cluster in the genome of Synechocystis PCC6803 and discuss the phylogenetic relationship among nitrogen-fixing bacteria, cyanobacteria, and the liverwort chloroplasts.

Materials and Methods

Culture condition and DNA isolation from Synechocystis PCC6803. Synechocystis PCC6803 was kindly provided by Dr. N. Murata, National Institute for Basic Biology, Okazaki, Japan. Cells of Synechocystis PCC6803 were grown in the BG-11 medium on a rotary shaker under cool-white fluorescent light at 26 °C. Synechocystis PCC6803 DNA was prepared as described in reference 15.

Southern hybridization. The 0.39-kb Rsal fragment containing a part of the liverwort frxC and the 0.57-kb Rsal–NcoI 52541 fragment containing a part of the liverwort ORF465 were isolated from a plasmid pMP323 and labeled with [α-32P]-dCTP using a labeling kit with random primer (Boehringer Mannheim). Synechocystis DNA digested with HindIII or EcoRI were transferred to a Zeta-probe membrane (Bio-Rad) in 0.4 N NaOH/0.6 M NaCl. Hybridization was done at 37 °C for 16 hr in a hybridization buffer containing 20% formamide, 1 × Denhardt's solution, 5 × SSC, 0.1% SDS, and 0.2 mg/ml heat denatured calf thymus DNA with 32P-labeled probes prepared as above. After hybridization filters were washed in a solution containing 5 × SSC and 0.1% SDS, then 0.1 × SSC and 0.1% SDS at 37 °C.

Cloning and sequencing of genomic DNA from Synechocystis PCC6803. Restriction enzyme-digested fragments of Synechocystis genomic DNA

† Present address: Research Institute for Bioreources, Okayama University, Kurashiki 710, Japan.
‡ To whom correspondence should be addressed.

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank Nucleotide Sequence Database [accession no. D10474].
were separated in 0.6% agarose gel (Seakem GTG, FMC) and isolated by electrophoresis. DNA fragments were ligated to a plasmid BluescriptII KS+ and used to transform *E. coli* DH5α MCR (BRL). Colony hybridization was done under the same conditions as in the case of Southern hybridization. Inserts of positive clones were subcloned into BluescriptII KS+, and nested deletion clones were made by a deletion kit (Takara, Kyoto) following the manufacturer's protocol. Single stranded DNA was obtained using helper phage M13 KO7. [16]DNA was sequenced by the dideoxy method using the Klenow fragment of DNA polymerase I.[18]

### Results

**Presence and cloning of frxC and ORF465 homologues of *Synechocystis PCC6803***

In the *Synechocystis* genomic DNA we detected DNA fragments which show sequence similarity to liverwort chloroplast frxC and ORF465 by Southern hybridization using liverwort DNA probes. A 9.1-kb *HindIII* fragment was detected with both probes for frxC and ORF465 (Fig. 1, lanes 1 and 2) when the blotted filter was washed under stringent conditions using washing solution containing 0.1 x SSC and 0.1% SDS at 37°C. This indicates that homologues of the liverwort frxC and ORF465 are present in a closely related manner on the *Synechocystis* genome, possibly as single copy genes. Therefore, we cloned the 9.1-kb *HindIII* fragment from a genomic DNA library of *Synechocystis PCC6803*. A size-fractionated genomic DNA library of 8.5—9.4 kb *HindIII* inserts was constructed in a plasmid vector, BluescriptII KS+, and screened by colony hybridization under the same conditions as described above using two kinds of probes for frxC and ORF465. Since all the positively hybridized clones with the two probes showed an identical restriction pattern (data not shown), it was concluded that isolated plasmids should contain frxC and ORF465 homologues of *Synechocystis PCC6803*.

This 9.1-kb *HindIII* fragment contained only 118 bp downstream from ORF469 (see below). To clone the downstream region from the coding region corresponding to the ORF469 (see below), Southern hybridization was done using ORF465 probe with EcoRI-digested *Synechocystis* genomic DNA. A 3.9-kb EcoRI fragment, which overlaps with the 9.1-kb *HindIII* fragment, was detected in Southern hybridization and cloned in BluescriptII KS+ by colony hybridization using the liverwort ORF465 probe by the same procedures.

**Nucleotide sequence analysis of a region containing frxC and ORF469 (ORF465 homologue)**

The restriction map of the cloned *Synechocystis* DNA and positions of frxC and ORF465-homologues are shown in Fig. 2. From Southern hybridization data (Fig. 1) the frxC homologue and ORF465 homologue were in the 6.5-kb BamHI—EcoRI region that is covered by the cloned 9.1-kb *HindIII* and 3.9-kb EcoRI fragments. Therefore, the 6.5-kb BamHI—EcoRI region was sequenced on both strands as shown in Fig. 3. In the sequenced region, four significant ORFs were predicted from the nucleotide sequence and named ORF(n) with the number (n) of amino acids encoded. An ORF31 having sequence similarity with the liverwort frxC was found at position 1539—2474. However, the TTG codon for 24 th leucine at 1608—1610 was assigned as the translation initiation codon for this reading frame, since the TTG codon has been reported to be used as initiation codon in *E. coli* and a ribosome-binding sequence (AGGAG) was located 8 bp upstream from the TTG codon. Therefore, a reading frame consisting of 288 amino acids was designated as *Synechocystis frxC*. An ORF of 469 amino acids that shows sequence similarity with liverwort ORF465 was also found 518 bp downstream from frxC and named as ORF469. This ORF469 was also proceeded by a typical ribosome binding sequence (AGGAG). Additionally, ORF278 was found upstream from frxC and ORF248 was downstream from ORF469. All these four ORFs were in the same orientation in the following order; ORF278—frxC—ORF469—ORF248.

**Sequence comparison of Synechocystis frxC and ORF469 with counterparts in liverwort and Clostridium pasteurianum**

The deduced amino acid sequence of *Synechocystis*
frxC showed 83% sequence identity with that of liverwort frxC and 31% identity with that of nifH product, an iron protein, which has also been called nitrogenase reductase or component II of the nitrogenase complex in nitrogen-fixing organisms. Comparison of amino acid sequences among Synechocystis frxC, liverwort frxC and nifH from Clostridium pasteurianum21,24 was shown in Fig. 4A. In the NH3-terminal region, a putative-binding site whose sequence is Gly-(Xaa)-(Xaa)-Gly-Lys-Ser, was conserved in these three gene products. In addition, 4 cysteine residues, two of which are thought to serve as ligands to Fe-S center in nitrogenase,25 were also conserved in the frxC products in Synechocystis and liverwort.

The amino acid sequence deduced from the Synechocystis ORF469 was 74% identical with that of liverwort ORF465, and 19% identical with nifK product which is the β subunit
of molybdenum-iron protein of nitrogenase complex from *Clostridium pasteurianum*. All 3 of the three cysteine residues which are conserved among nifK products from nitrogen-fixing organisms (see below) were conserved both in *Synechocystis* ORF469 and liverwort ORF465. ORF278, which is upstream from *frxC*, and ORF248, downstream from ORF469, did not show any significant sequence similarities with proteins compiled in the NBRF-PIR database (release 26.0) and swiss-prot (release 15.0).

**Discussion**

To analyze the gene organization of *frxC* and ORF465 homologues of *Synechocystis* PC6803, we determined the nucleotide sequence of the 6.5-kb BamHI-EcoRI genomic region of *Synechocystis* PC6803. The deduced gene organization in the *Synechocystis* genome is shown in Fig. 5 compared with those of liverwort chloroplast and the nitrogen-fixing bacterium *Clostridium pasteurianum*. In the liverwort chloroplast genome, *frxC*-ORF465 gene cluster is located in the small single copy region, but the S-terminal 7 nucleotides of the coding region for the *frxC* are in the inverted repeat region (IR). The promoter elements for *frxC* are postulated to be in the IR region.

In the sequence analysis of *frxC* in *Synechocystis* PC6803, we assumed that TTG codon is used as a translation initiation codon, based on the fact that a possible ribosome binding sequence (AGGAG) is just upstream from the TTG codon. In *Escherichia coli*, triplet codons TTG, GTG and ATG are used as initiation codons in the relative frequency 1:10:89. Therefore, in cyanobacteria GTG and TTG may be used at higher frequencies than in *E. coli* and as efficiently as ATG. Of course a possibility that the 66-bp upstream ATG (indicated by asterisks in Fig. 3) is used as an initiation codon cannot be excluded, although no preceding ribosome-binding sequence has been detected. Although the deduced *frxC* product of *Synechocystis* PC6803 was 83% identical to that in liverwort chloroplasts, it showed about 30% identity with nifH products from various nitrogen-fixing organisms. In the nitrogen-fixing bacterium, *Clostridium pasteurianum*, genes for nitrogen complex (nifH1, nifD, and nifK) are clustered in an operon as shown in Fig. 5.
Fig. 4. Comparison of the Deduced Amino Acid Sequences of Synechocystis frxC, Marchantia frxC and Clostridium nifH1 (A), Synechocystis ORF469, Marchantia ORF465 and Clostridium nifK (B).

Gaps introduced into the sequences for alignment are indicated by hyphens. Amino acid residues in Marchantia and Clostridium ORFs identical to corresponding Synechocystis ORFs are denoted by colons. Sequence identities to Synechocystis ORFs are expressed by percentages with length of gene products. Putative ATP-binding sites are boxed. Conserved cysteine residues are indicated by filled circles.

Fig. 5. Comparison of Gene Organizations among frxC-ORF469 Cluster of Synechocystis PCC6803, frxC-ORF465 Cluster of Marchantia polymorpha (Liverwort) Chloroplast and nifH1, nifD, nifK Cluster of Clostridium pasteurianum.

Coding regions of Synechocystis frxC and its homologues are filled boxes. Those of ORF469 and its counterparts are shown by hatched boxes. IR indicates inverted repeat region in liverwort chloroplast DNA.

The key enzymatic reactions in nitrogen fixation are catalyzed by nitrogenase complex, which has two components: the iron protein (also referred to as component II or nitrogenase reductase) and the molybdenum-iron protein, which is a tetramer of two α- and two β-subunits that binds 30–32 Fe, 2 Mo, and inorganic sulfur. Those components are encoded by nifH, nifD, and nifK, respectively. The reduction site of dinitrogen is on the molybdenum-iron proteins, which are encoded by nifD and nifK. On the other hand, the reduced iron protein serves as a specific electron donor to molybdenum-iron proteins. Therefore, it is possible that the frxC product may also serve as an electron donor. Indeed the frxC product has been detected as a dimer in liverwort chloroplasts and can bind ATP, like the iron protein of nitrogenase. 

Nucleotide sequences of frxC and nifH from the nitrogen-fixing filamentous cyanobacterium Plectonema boryanum has been analyzed. Since Plectonema frxC also showed 91% amino acid sequence identity with Synechocystis frxC and showed only 34% identity with the nifH of P. boryanum, frxC could be essential to some fundamental biological function in photosynthetic organisms other than in nitrogen fixation.

In addition to the frxC, we detected a significant amino acid sequence similarity between Synechocystis ORF469 and Clostridium nifK, which is just downstream from nifH1. From structural analysis, three cysteine residues conserved among nifK products are located on the surface of the protein to serve as ligands to redox prosthetic groups. These cysteine residues are also conserved in
Gene Cluster frxC-ORF469 in Synechocystis PCC6803

Synechocystis ORF469 and liverwort chloroplast ORF465 (Fig. 4B), suggesting they may chelate some metal ions.

In *Clostridium*, nifH1, nifD, and nifK form a gene cluster. In *Synechocystis* PCC6803, however, an nifD-like ORF was not found between frxC and ORF469 in the sequenced 6.5-kb BamHI-EcoRI region (Fig. 5). Therefore, it is possible that an nifD-like ORF may be somewhere else on the *Synechocystis* genome. When probed with liverwort ORF465 at lower stringency conditions, some minor bands were detected in Southern hybridization (data not shown). Therefore, other functionally related ORFs may be located far from this region.

Although the deduced product of frxC has significant sequence similarity with nitrogenase iron protein, they do not seem to encode a component of nitrogenase, because *Synechocystis* PCC6803 does not fix nitrogen, and the homology between ORF469 and the β subunit of molybdenum-iron protein seems too low. Moreover no nitrogenase activity is detected in liverwort chloroplasts. Gene tagging analysis using a *Synechocystis* PCC6803 transformation system will be expected to elucidate functions of frxC and ORF469 products.

Recently, Suzuki and Bauer demonstrated that the genes bchL and bchB, which encode subunits of the dark-driven protochlorophyllide reductase in the photosynthetic bacterium *Phodobacter capsulatus*, have high levels of sequence similarities with liverwort frxC and ORF465, respectively. In addition, they showed that gene disruptants for chloroplastic frxC in the unicellular green alga *Chlamydomonas reinhardtii* were yellow in the dark but green in the light.

From gene organization and amino acid sequence similarities, the frxC-ORF469 gene cluster and the nifH, nifD, nifK cluster have a common ancestor, and have acquired different functions in the course of evolution of bacteria. Similar structural characteristics have been reported on between the nifD, nifK cluster and the nifE, nifN cluster from *Azotobacter vinelandii*. The nifE, nifN genes, which are just downstream from the nifD, nifK gene cluster, encode polypeptides involved in the biosynthesis of molybdenum-iron cofactor. Therefore, the gene clusters from common ancestor with an nifH, nifD, nifK cluster may have been fairly spread on bacterial genome, or prone to be rearranged. Since *Synechocystis* PCC6803 does not fix nitrogen, even if they exist, would not be functional. The presence of an frxC-ORF469 gene cluster in the genome of the unicellular cyanobacterium *Synechocystis* PCC6803 would imply that a similar gene cluster had been in the genome of the ancestral organism of chloroplasts. The gene cluster may have become more compact with shorter spacer regions between ORFs as in case of the liverwort chloroplast genome (Fig. 5). Moreover, these genes may have been transferred into the nuclear genomes in tobacco or rice during evolution.

Acknowledgments. We are grateful to Professor Dr. Yasuyuki Yamada, Faculty of Agriculture, Kyoto University, for his encouragement during this investigation. This research was supported partly by Grant-in-Aid for Scientific Research from the Ministry of Education, Science, and Culture of Japan, and partly by NEDO/RITE, Japan, under contract of research commissioned.

References


