Molecular Cloning and Nucleotide Sequence Analysis of the Gene Coding for Chloroplast-type Ferredoxin from the Dinoflagellates *Peridinium bipes* and *Alexandrium tamarense*

Takeshi Yoshikawa,[†] Kiyotaka Takishita,
Yuzaburo Ishida,^{††} and Aritsune Uchida

Laboratory of Marine Microbiology, Department of Applied Bioscience,
Graduate School of Agriculture, Kyoto University, Kyoto 606-01, Japan

(Received December 16, 1996)

We cloned and sequenced the *fedQ* encoding chloroplast-type ferredoxins of dinoflagellates *Peridinium bipes* and *Alexandrium tamarense*. The phylogenetic relationships inferred from their amino acid sequences show that dinoflagellates are closely related to members of chlorophytes. Transit peptides of dinoflagellate ferredoxins show the common feature of land plants, however, they contain signal peptides at amino-termini. The signal peptide may function to translocate the 3 membranes chloroplast envelope of dinoflagellates. The eukaryotic promoter consensus motifs such as TATA box and CAAT box were not found in the upstream regions of ferredoxin genes as far as they are sequenced. Instead GC box which is found in housekeeping genes was recognized. These findings are very interesting to show the peculiarity and complexity of dinoflagellates.

Key words: [2Fe–2S] ferredoxin gene, phylogenetic relationship, promoter, transit peptides, *Peridinium bipes*, *Alexandrium tamarense*

The dinoflagellate is a unicellular, eukaryotic alga that is a very diverse group containing autotrophic, heterotrophic, phagocytic, and symbiotic species. It is noticeable that certain species periodically cause toxic red-tides that give rise to economically serious problems to the aquacultural industries.

The taxonomical and evolutionary relationships between dinoflagellates and other eukaryotes have been controverted because of their peculiar features of nuclei suggesting the prokaryotic states. For instance, the chromosomes always condense throughout the cell cycle^{1,2)} and are segregated by attaching the nuclear envelope at mitosis like bacterial cell divisions;³⁾ the chromatin fibrils show arched whorls similar to bacterial nucleoids;⁴⁾ and in particular, their chromatin is devoid of typical histones⁵⁾ and therefore it lacks the nucleosome structure.⁶⁾ It follows from these ancestral traits that dinoflagellates are the intermediates of the prokaryote-eukaryote lineage, so called mesokaryotes.⁷⁾ However, the nucleotide sequence data accumulated and molecular evolutionary analyses of ribosomal RNA confirm the eukaryotic origin of this Phylum that shares phylogenetically common ancestor with apicomplexans and ciliates.^{8,9)} Moreover, the evidence for the presence of *cdc2*-like histone H1 kinase activity,¹⁰⁾ and eukaryotic nuclear matrix with lamin-like proteins and topoisomerase II¹¹⁾ also make it clear that dinoflagellates possess the same mechanisms as eukaryotes except for the lack of histones, which is explained as the secondary loss after acquisition.¹²⁾

Ferredoxin is a ubiquitous low molecular-weight protein which is involved in diverse electron-transfer systems of photosynthetic organisms. According to the number of iron and sulfur atoms per molecule, ferredoxins are classified as [2Fe–2S], [4Fe–4S], and [3Fe–4S] ferredoxin. Chloroplast-type ferredoxins contain two atoms each of iron and sulfur and four cysteine residues. The [2Fe–2S] ferredoxin is an electron terminal acceptor of electron transport pathways in photosynthesis: it is reduced by the iron-sulfur center of photosystem I complex and functions as an electron donor of NADP photoreduction, cyclic photophosphorylation, glutamate synthesis and nitrite reduction. The [2Fe–2S] ferredoxin is nuclear-encoded and transcribed as a pre-ferredoxin containing the plastid-transfer signal, a transit peptide at its amino-terminus. The characteristics of the transit peptide are well investigated and the importance of the secondary structure for its transport system are pointed out.¹³⁾ The amino acid and nucleotide sequences of the [2Fe–2S] ferredoxin from many photosynthetic organisms have been determined and their phylogenetic relationships are discussed from the sequence comparison in view of the molecular evolution.¹⁴⁾ The expression of the [2Fe–2S] ferredoxin gene is regulated by light under phytochrome control¹⁵⁾ and its promoter sequences containing G-I box and CCAC box are well documented.^{16,17)}

In this report, we amplified the [2Fe–2S] ferredoxin gene from red tide-forming dinoflagellates, *Peridinium bipes* and *Alexandrium tamarense* using the PCR technique and

[†] Present address: Microbiology Laboratory, Faculty of Fisheries, Kagoshima University, Shimoarata, Kagoshima 890, Japan.

^{††} Present address: Research Center on Environments, Research Organization of Science and Engineering, Ritsumeikan University, Nojihigashi, Kusatsu, Shiga 525–77, Japan.

their phylogenetic relationships, the transit peptide characteristics and deduced promoter sequences are discussed.

Materials and Methods

Cells

The axenic culture of *Alexandrium tamarense* OFX151 isolated from our laboratory was maintained in SWIM medium at 20°C under the 14 h light and 10 h dark cycle¹⁸ and was harvested 25–30 days after inoculation for nucleic acid extraction. The cells were collected by gentle centrifugation.

Cells of *Peridinium bipes* were collected by net from water blooms in Gokawa Reservoir, Wakayama prefecture, in order to obtain a unialgal sample. Microscopic observation indicated that the samples used to extract DNA contained more than 99% *P. bipes*. The cells were kept at –90°C until used.

Total DNA Isolation

The cell pellet was homogenized in liquid nitrogen. The homogenate was suspended in NET buffer (150 mM NaCl, 100 mM ethylenediamine tetraacetate, 10 mM Tris-HCl, pH 8.0) containing 1.0 mg/ml Proteinase K (Merk, FRG) and 1.0% sodium dodecylsulfate, and incubated at 65°C. Cell debris was separated by centrifugation and the supernatant was extracted once with phenol saturated by TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM ethylenediamine tetraacetate), twice with phenol:chloroform:iso-amyl alcohol (25:24:1), and once with chloroform:iso-amyl alcohol (24:1). Nucleic acids were precipitated with ethanol and the precipitates were dissolved in TE buffer containing 100 µg/ml RNase A (Sigma Chemical Co., USA) and incubated at 37°C. RNase was removed with phenol:chloroform:iso-amyl alcohol extraction and DNA was precipitated with ethanol and dissolved in TE buffer. For small-scale preparation, the cell pellet derived from the centrifugation of the 200 ml culture was homogenized in liquid nitrogen and DNA was extracted with the nucleic acid purification kit SepaGene (Sanko Junyaku, Japan).

Total RNA Isolation

After the cell pellet was homogenized in liquid nitrogen, the homogenate was suspended in GTC solution (4 M guanidium thiocyanate, 25 mM sodium citrate). The suspension was layered onto a cushion of 5.7 M CsCl solution (5.7 M cesium chloride, 0.1 M ethylenediamine tetraacetate) and ultracentrifuged at 100,000 g for 24 h. The RNA pellet was dissolved in TE buffer containing 1.0% sodium dodecylsulfate, and then RNA was extracted with phenol saturated by TE buffer (pH 9.0) and chloroform:iso-amylalcohol (24:1). RNA was concentrated with ethanol. Ten M LiCl was added to RNA solution to bring concentration of LiCl to 2 M and the precipitated RNA was collected by centrifugation. The RNA was resuspended in H₂O.

Polymerase Chain Reaction Amplification of *fedQ*

Oligonucleotide primers FXDTP-I and FXTCV-RI for the polymerase chain reaction (PCR) amplification of the *P. bipes fedQ* were designed (Table 1) according to the amino acid sequence described before¹⁴ and synthesized with the DNA synthesizer Gene Assembler Plus (Pharmacia

LKB Biotech., Sweden). Synthesized oligonucleotides were purified with the Oligonucleotide Purification Cartridges (Applied Biosystems, Inc., USA). PCR amplifications were carried out under the conditions of 2.0 mM MgCl₂, 200 µM dNTPs, 1.0 µM primers, 25 units/ml *Taq* DNA polymerase (Perkin Elmer, USA) and 100 ng/ml total DNA at the final concentration in the buffer recommended by manufacturers. The reaction mixture was subjected to 30 cycles of 1 min at 94°C, 2 min at 50°C, and 3 min at 72°C. Amplified products were checked by 6.0% polyacrylamide gel or 2.5% agarose gel electrophoreses.

Oligonucleotide primers PBFDCSS and PBFDMGMR for the *A. tamarense fedQ* amplification were designed (Table 1) according to the nucleotide sequence of *P. bipes fedQ*. Primer syntheses and PCR amplifications were the same as described above.

5' and 3' Flanking Region Amplification of *fedQ* Using the TAIL-PCR Methods

TAIL-PCR amplifications were performed according to Liu *et al.*¹⁹ For 5' and 3' flanking region amplification of the *P. bipes fedQ*, oligonucleotide specific primers PBFDTAIL1, 2, 3, 4, 5, and 6 were designed (Table 1) according to the nucleotide sequence of PCR-amplified DNA fragment PBFDDY. For 5' and 3' flanking region amplification of the *A. tamarense fedQ*, specific primers ATFDTAIL2R, 3R, and 4R directed to 5' terminus flanking region were designed (Table 1) according to the nucleotide sequence of PCR-amplified DNA fragment ATFDYCM1 with primers PBFDCSS and PBFDMGMR, and specific primers ATFDTAIL5, 6, and 7 directed to 3' terminus flanking region were designed (Table 1) according to the nucleotide sequence of the TAIL-PCR products of 5' terminus flanking region. Arbitrary degenerate primers were designed according to Liu *et al.*¹⁹ Amplified products were checked by 1.5% agarose gel electrophoresis.

Transcriptional Region Determination of *A. tamarense fedQ* by RACE

The RACE (Rapid amplification of cDNA ends) method was used to determine the 5' and 3' ends of cDNA as described by Frohman *et al.*²⁰ Specific amplification of cDNA ends was carried out by nested PCR under the same reaction mixture condition as described above with the primers shown in Table 1. The hybrid primer RACEHB contains 17 nts oligo (dT) to anneal poly(A) strands, and the adaptor sequence recognized by the adaptor primer RACEAP. For 5' end amplification, poly(A)-tailed cDNA was subjected to PCR with primers RACEHB, RACEAP, and ATRC2 under the condition, which was 5 cycles of 1 min at 94°C, 2 min at 40°C, and 3 min at 72°C following 30 cycles of 1 min at 94°C, 2 min at 50°C, and 3 min at 72°C. The 3' end amplification was carried out with primers RACEAP and ATRC3 for the first amplification, and with primers RACEAP and ATRC4 for the second amplification under the condition of 30 cycles of 1 min at 94°C, 2 min at 50°C, and 3 min at 72°C. The amplified products were checked by 6.0% polyacrylamide gel.

DNA Sequence Analysis

PCR-amplified DNA fragments were cloned with the

Table 1. Primer sequences used for PCR, TAIL-PCR, and RACE

Primers	Primer sequences (5'→3')
<i>Peridinium bipes</i>	
for PCR	
FXDTP-I	GA(C/T)ACICCGA(C/T)GGIAA(A/G)AA*
FXTCV-RI	GTIGG(A/G)TAIGTIAC(A/G)CAIGT*
for TAIL-PCR	
PBFDTAIL1	CCCTGGGCACTCAAA(A/G/T)(A/G)(A/G)CTTCTTC
PBFDTAIL2	TGAGGACCTTCTCGCA(A/G)CTCGA
PBFDTAIL3	GATC(A/G)(A/C/T)ACCCGTGAGGACCTTTC
PBFDTAIL4	CTCGAG(C/T)TGCGCAGGAAAGGTCCTCA
PBFDTAIL5	GTCCCTCACGGGT(A/G/T)(C/T)GATCGACCA(A/G)
PBFDTAIL6	AACGGCTA(C/T)TGCCTCACCTGCGTCA
<i>Alexandrium tamarense</i>	
for PCR	
PBFDCSS	TGCTC(A/G)TCCGATCAGGCATTCTTG
PBFDMGMR	GGTGAGGCA(A/G)TAGCCGTTGCCCAT
for TAIL-PCR	
ATFDTAIL2R	CTGGTC(A/G)ATGGA(G/T)CCGGAGAGGACC
ATFDTAIL3R	CCTGGTCAGACTGGTC(A/G)ATGGA(G/T)CC
ATFDTAIL4R	GGTGAGGCAGTAGCCGTTGCCCATC
ATFDTAIL5	CTC(A/C/G/T)TGCCG(C/T)GC(A/G/T)GGCTC(C/T)TG(C/T)TCC
ATFDTAIL6	(A/G)(C/T)TGCC(A/G)TACTC(A/C/G/T)TGCCG(C/T)GC(A/G/T)GGC
ATFDTAIL7	CT(C/G/T)GA(C/T)CAGGC(C/G/T)GA(A/G)GAGGA(G/T)GGC
for RACE	
RACEHB	GACTCGAGTCGACATCGATTTTTTTTTTTTTTTT
RACEAP	GACTCGAGTCGACATCGA
ATRC1	T(A/G)TGGTGAAGGC(G/T)GCATC
ATRC2	(C/T)GTCACCTTGAAGTG
ATRC3	AT(C/T)GACCAGTCTGACCAG
ATRC4	CACGTGAGGATGAGCTG

* I indicates inosine.

plasmid vector, pCRTMII (Invitrogen, USA) and sequenced with the DNA sequencer 373A (Applied Biosystems, Inc., USA) using the *Taq* DyeDeoxyTM Terminator Cycle Sequencing Kit (Applied Biosystems, Inc., USA). The obtained sequences were analyzed by the DNA sis software (Hitachi Software Engineering, Japan).

Phylogenetic Analysis of *fedQ*

Amino acid sequences of the ferredoxin were aligned by the multiple sequence alignment program Clustal W version 1.5.²¹⁾ Genetic distance matrices calculations and phylogenetic tree constructions by Nei and Saitou's neighbor-joining method²²⁾ were performed by the PROT-DIST, NEIGHBOR, and DRAWTREE programs, all of which were conducted by the PHYLIP package version 3.5c.²³⁾

Signal Peptide Analysis of *fedQ* Transit Peptides

Amino acid sequences of dinoflagellate transit peptides were sent as electronic mails to the PSORT e-mail server in order to predict the protein sorting signals. Predictions of signal peptide cleavage sites were performed by the McGeoch's²⁴⁾ and von Heijne's²⁵⁾ methods.

Results

Amplification of the *P. bipes* and *A. tamarense* *fedQ* by PCR Methods

In order to clone the *P. bipes* ferredoxin gene, the oligonucleotide primers FXDTP-I and FXTCV-RI were

designed (Table 1) according to its amino acid sequence reported by Uchida *et al.*¹⁴⁾ and PCR was performed. The DNA fragment PBFDDT of 220 bp was amplified and confirmed from its nucleotide and deduced amino acid sequences that PBFDDT was derived from the *fedQ*. To obtain the DNA sequence information of the flanking region of PBFDDT, TAIL-PCR technique was applied (for details, see Materials and Methods) and the amplified products of 660 bp for 5' terminus flanking region and 640 bp for 3' terminus flanking region were obtained. The fact that their nucleotide sequences contain the same sequences as PBFDDT confirmed that they also derived from *fedQ*. As a result, 921 bp nucleotide sequence containing the mature *fedQ* and its flanking regions was obtained (Fig. 1).

For the sequence determination of the *A. tamarense* *fedQ*, we adopted the same approach as *P. bipes*. We designed the primers PBFDCSS and PBFDMGMR for PCR according to the nucleotide sequence of PBFDDT, and the specific primers ATFDTAIL2R, 3R, 4R, 5, 6, and 7 for TAIL-PCR (Table 1). In consequence, 821 bp nucleotide sequence of *fedQ* containing its flanking regions was obtained (Fig. 2).

Nucleotide and Amino Acid Sequence Analyses of the *P. bipes* and *A. tamarense* *fedQ*

The nucleotide and deduced amino acid sequences of the ferredoxin genes for *P. bipes* and *A. tamarense* are shown in Figs. 1 and 2, respectively. The mature protein coding regions were decided from the comparison to the

5'	AGG ACC TTT CCT GCG CAG CTC GAG AAG CCA CTA GCA CGG CAT AAG CCG ACA GAG	-220
	TTT ATC ATC AGT TGT CGA TCA TTT CAA GTC GCA TTA GTC AAC AGG CCG GTC ATG	-166
	CCC GAT CCC CAC ATT TCA CTT CCC AGT CTG CGA CGT AGC TTA AGC TTC ACA CAT	-112
	TTT AGC AGG AAG <u>TCG GCC GGG</u> TGT GTG GGC AGG CAA CTC AGG GGA TGC GTT GCG	-58
	ATG TGG TAC AAG CCC CCT TTT TGA AAG GCT GCC ACA ATC TTG ATT TGC AAG CCT	-4
	GCC ATG GCG TCC AAG AGC ATT CTG GCC CCA ACT CTG CTC GTG GCG GCG CTC GGC	50
	M A S K S I L A P T L L V A A L G	
	ACT GTT CTC CTC CGC CGC TCC ATG CTG GCG CCG ACC TTC ATT GCA CCT CAC TCA	104
	T V L L R R S M L A P T F I A P H S	
	AGC AGC ACT TTG AGT CTA CAC CAG GGC TGC GTT CAT GTG GCT GGC GCG GGC GAC	158
	S S T L S L H Q G C V H V A G G G D	
	ACC GCC ACC CTT GCC GGT GCT GTG CCG GCG ATC GTG AGC AGC GTG CCT ACG CGA	212
	T A T L A G A V P A I V S S V P T R	
	CAA GGA GTC GCA GCA CAC TTC AAA GTG ACT TTG GAC ACC CCG GAC GGC AAG AAG	266
	Q G V A A H F K V T L D T P D G K K	
	TCC TTC GAG TGC CCA GGG GAC TCC TAC ATC CTC GAC GAG GCG GAG GAG GAG GGC	320
	S F E C P G D S Y I L D E A E E E G	
	CTG GAG CTG CCG TAT TCG TGC CGC GCA GGC TCG TGC TCG AGC TGC GCA GGA AAG	374
	L E L P Y S C R A G S C S S C A G K	
	GTC CTC ACG GGT GTG ATC GAC CAG TCC GAT CAA GCA TTC TTG GAC GAC GAC CAG	428
	V L T G V I D Q S D Q A F L D D D Q	
	ATG GGC AAC GGC TAT TGC CTC ACC TGC GTC ACC TAC CCC ACA TCC GAC GTG ACG	482
	M G N G Y C L T C V T Y P T S D V T	
	ATC ATG ACA CAC TGC GAG AGC GAA CTT TGA GCG GGG GTG GCG CGT CCG GAG GGA	536
	I M T R C E S E L *	
	GAA ATA AAT TCT ATG ATT GCG GCC CAA AAA ATA TAT ATA AAA TAC AGA ACA TTT	590
	TGG GGG GCT TGT TCA ACA TTT TGA ATT ATT GAC CAA GAT CCT TCC ATA TTC ACC	644
	GTA T 3'	

Fig. 1. Nucleotide and deduced amino acid sequences of *Peridinium bipes* fedQ.

Nucleotide position 1 agrees with the predicted translation initiation site. A promoter motif-like sequence is underlined. The predicted transit peptide region is indicated as italics.

amino acid sequence of *P. bipes* ferredoxin reported by Uchida *et al.*¹⁴⁾ Moreover, chloroplastic proteins coded at the nucleus genome like ferredoxin are transcribed and translated as a precursor which contains transit peptides at their amino-termini for translocation to chloroplasts. It is generally said that transit peptides of ferredoxin always start with Met-Ala-. In the case of dinoflagellate ferredoxin, Met-Ala- was found at nucleotide positions 1 for *P. bipes* and *A. tamarensis*. From these results, we interpreted these methionine residue as start codons and the transit peptide regions were determined (Figs. 1 and 2). The translation initiation sites also confirmed that consensus nucleotides A or G at position -3 and G at position +4 were conserved.²⁶⁾ The transcriptional start and stop sites of *A. tamarensis* ferredoxin gene were determined to be 63 bp upstream of translational start site, and 60 and 72 bp downstream of translational stop site (Fig. 2).

Phylogenetic Tree Construction Inferred from the Amino Acid Sequences of fedQ

The amino acid sequences deduced from *A. tamarensis* and *P. bipes* fedQ were aligned by the method of Thompson *et al.*²¹⁾ (Fig. 3) with ones of land plants, chlorophytes, rhodophytes, chromophytes, and cyanobacteria retrieved from the protein sequence database PIR. The highly conserved feature of the ferredoxin amino acid sequences, especially around the cysteine residues which function as ligands of the [2Fe-2S] cluster, was also applied to the dinoflagellates (Fig. 3). In addition, the dinoflagellate-specific cysteine residues which are boxed in Fig. 3 were recognized. The genetic distances were calculated from the multiple-aligned sequences based on the formula of Kimura.²⁷⁾ The phylogenetic tree was constructed based on the distant matrix by the neighbor-joining method.²²⁾ The phylogenetic tree shows the clustering of dinoflagellates and the close relationship between

Fig. 2. Nucleotide and deduced amino acid sequences of *Alexandrium tamarense fedQ*.

Nucleotide position, promoter motif-like sequences, and the predicted transit peptide region are shown as Fig. 1. The transcription initiation and termination sites are also indicated as arrows.

dinoflagellates and chlorophytes, especially Chlorophyceae (Fig. 4). This relationship is proven from the genetic distances: genetic distances between dinoflagellates and chlorophytes are 0.3–0.45, while distances between dinoflagellates, chromophytes, rhodophytes, and cyanobacteria are 0.48–0.55 (Table 2).

Transit Peptides of Dinoflagellate fedQ Contain the Signal Peptides

We compared the amino acid sequences of dinoflagellate ferredoxin transit peptides to those of land plants (Fig. 5). Although there are no homology between the transit peptide amino acid sequences of the dinoflagellates and of land plants, it was shown that dinoflagellates and land plants have common characteristics that transit peptides

Fig. 3. Amino acid sequence alignment of [2Fe-2S] ferredoxins of chlorophytes, euglenophytes, rhodophytes, chromophytes, cyanobacteria, and dinoflagellates.

The cysteine residues which are conserved among [2Fe-2S] ferredoxins and function as ligands of the [2Fe-2S] cluster are dotted above the alignment, and those specific for dinoflagellates are boxed.

Table 2. Genetic distance matrix deduced from the amino acid sequence alignment of chloroplast-type ferredoxins

	At	Pb	Cr	Sq	Ps	Os	At	Bf	Cc	Sp	An
dinoflagellates											
<i>Alexandrium tamarense</i>	0.0000										
<i>Peridinium bipes</i>	0.1570	0.0000									
chlorophytes											
<i>Chlamydomonas reinhardtii</i>	0.3899	0.3498	0.0000								
<i>Scenedesmus quadricauda</i>	0.3899	0.4024	0.1150	0.0000							
<i>Pisum sativum</i>	0.3721	0.4207	0.3620	0.3202	0.0000						
<i>Oriza sativa</i>	0.3899	0.4585	0.4908	0.4213	0.2740	0.0000					
<i>Arabidopsis thaliana</i>	0.4458	0.4981	0.3620	0.3202	0.2592	0.3045	0.0000				
chromophytes											
<i>Bumilleriopsis filiformis</i>	0.4852	0.5186	0.3968	0.4213	0.5572	0.5161	0.5785	0.000			
rhodophytes											
<i>Cyanidium caldarium</i>	0.5481	0.5397	0.3125	0.3362	0.4213	0.5364	0.5161	0.4275	0.0000		
cyanobacteria											
<i>Spirulina platensis</i>	0.4852	0.5186	0.2966	0.3045	0.4962	0.4962	0.4394	0.3761	0.3121	0.0000	
<i>Anacystis nidulans</i>	0.5481	0.5397	0.3125	0.2740	0.3362	0.3525	0.3362	0.3435	0.3121	0.2112	0.0000

are rich with hydroxy amino acids, the helix-breaking residue proline is conserved, and the processing site motif is found in its carboxy-terminus.²⁸⁾ However, the length of dinoflagellate transit peptides were very different from that of land plants: the transit peptides of dinoflagellates were

constructed by 78–80 residues, on the other hand, that of land plants were about 50 residues. As an additional information for protein sorting is expected, the transit peptides of dinoflagellates were subjected to the PSORT e-mail server analysis. As a result, they contain the signal peptide-like

Fig. 4. Phylogenetic relationships inferred from [2Fe-2S] ferredoxin amino acid sequence similarities.

Chlorophytes, euglenophytes, rhodophytes, chromophytes, cyanobacteria, and dinoflagellates are abbreviated as Chl, Eug, Rho, Chr, Cya, and Dino, respectively.

Land plants

<i>A. thaliana</i>	MAS---T-ALSSAIVGTSFIRRSAPISLRSLSANTQSLFGLKSGTARGGRVTAM
<i>S. oleracea</i>	MAA---T-TTTMMGMATTFVFKPQAPPMAALPSNTGRSLFGLKTGS-RGGRMT-M
<i>S. pratensis</i>	MAS---T--LSTLSVSASLLPK-QQPMVASSLPNTMGQALFGLKAGSR--GRVTAM
<i>P. sativum</i>	MAT---TPALYGTAVSTSFRLRTQPMPSVTT-TKAFSNGFLGLKTSLSKRGDLAVAM
<i>Z. mays</i>	MATVLGSPRAPAFFFSSSSLRAAPAPTAV-ALPAKV-GIMGRSASSRR--RLRAQ
	** *

Dinoflagellates

<i>A. tamarense</i>	MASSTRVLPVILVVTLGAMFLR-SFLMPGHINDAAFTISKLATHSVHHGCRPVQSSR-AL
<i>P. bipes</i>	MASKSILAPTLLVAALGTVLLRRSMLAPTTFIAPHSSST-----LSLHQGCVHVAGGGDTA
	*** . * . * . * . * . * . * . * . * . * . *
<i>A. tamarense</i>	LAQGGLSAVVNGVPARQGVAAH
<i>P. bipes</i>	TLGAVPAIVSSVPTROGVAAH
	* . * . * . * . *

Fig. 5. Amino acid sequence alignment of land plant and dinoflagellate [2Fe-2S] ferredoxin transit peptides.

Bold cases are hydroxy amino acids, and underlined cases with single and double lines indicate the conserved helix-breaking proline and processing site-motif Ile/Val-X-Ala/Cys, respectively. Arrows indicate the predicted signal peptide cleavage sites of dinoflagellate transit peptides.

sequences (Fig. 5) which function the translocation through the endoplasmic reticulum membrane at their amino-termini. Their characteristics such as amino-terminal charged residues and central hydrophobic regions were well agreed with the signal peptides.²⁹⁾

Dinoflagellate fedQ Do Not Have Universal Promoters of Eukaryotes

The flanking regions of the ferredoxin genes were sequenced using the TAIL-PCR methods, so the existence of promoter sequence motifs was investigated. However, we could not find the consensus sequences of the eukaryotic promoter such as TATA box and CAAT box in the upstream regions which were sequenced. Also, they lack G box and CCAC box, which are well known as photoregulation factors of *fedQ*. On the other hand, a motif like GC box that functions at housekeeping genes was found at the position -97 for *P. bipes* and *A. tamarense* (Figs. 1 and 2), and also AC box that is conserved among the ferredoxin gene promoter regions of land plants³⁰⁾ were found at the position -136 for *A. tamarense* (Fig. 2).

Discussion

In this report, we have determined *fedQ* sequences of dinoflagellates, *A. tamarense* and *P. bipes* applying the PCR techniques, and discuss about the phylogenetic relationship of dinoflagellates and the peculiarity of transit peptides and promoter sequences.

The peculiar and especially ancient features of dinoflagellate nuclei have suggested that dinoflagellates are intermediates of the prokaryote-eukaryote lineage.⁷⁾ Recently ribosomal RNA and DNA sequences of dinoflagellates were reported and suggested that dinoflagellates are

not an ancestral group and share a common ancestor for apicomplexans and ciliates.^{8,9} On the other hand, when [2Fe-2S] ferredoxin sequences are compared, dinoflagellates are rather close to chlorophytes. The close relationship between dinoflagellates and chlorophytes is contradictory to the chlorophyll contents: that is to say, chlorophytes have chlorophyll *a* and *b*, and on the other hand, dinoflagellates and chromophytes have chlorophyll *a* and *c*, which point out the close relationship between dinoflagellates and chromophytes.³¹

Most of the chloroplast proteins including ferredoxin are encoded in the nuclear genome. These proteins are synthesized as precursor molecules which include transit peptides at their amino-termini for posttranslational import into plastids. After translocation, these proteins are processed to their mature sizes by the stromal peptidase.³² The feature of transit peptides was well documented for land plants. In case of the [2Fe-2S] ferredoxin, they are constructed with about 50 amino acids and are characterized by the hydrophobic amino-terminus domains and potentially amphiphilic β -strand forming carboxy-terminus domains,¹³ even though they have little homologies with each other for amino acid sequences. Although the function of the transit peptides is not well understood, their interaction to the chloroplast envelopes or envelope proteins as a receptor of transit peptides are suggested,³³ and it may act as a main recognition and sorting system of chloroplast proteins to chloroplasts. For the dinoflagellate FedQ the transit peptide-like sequences are also recognized at their amino-termini and their characteristics were similar to others, however, they were very different from those of land plants in respect of their peptide lengths and amino acid sequences, which is derived from the signal peptide-like sequences added at amino-termini. These diversities are associated with the membrane structure of dinoflagellate chloroplasts: the chloroplast of dinoflagellates is surrounded by 1 membrane chloroplast endoplasmic reticulum (ER) additional to 2 membranes chloroplast envelope, and signal peptide-like sequences are likely to function to translocate the ferredoxin through the chloroplast ER.

In dinoflagellates, the information for the promoter regions are only reported for the luciferase-binding protein (LBP) gene of *Gonyaulax polyedra*³⁴ and the histone-like protein HCcl gene of *Cryptocodinium cohnii*.³⁵ Similar to the both genes, we could not find consensus sequences of eukaryotic promoters such as TATA box and CAAT box in the 5' terminus flanking regions of [2Fe-2S] ferredoxin genes of the dinoflagellates as far as sequenced. Instead, although it is not known whether it functions as a promoter or not, the GC box, that is frequently found as a GGGCGG motif in housekeeping genes instead of the TATA box, is recognized at -97 for *P. bipes* and *A. tamarensis*. This motif was also found in the LBP gene of *G. polyedra*.³⁴ These facts suggest the peculiarity of the transcription system of dinoflagellates, which will support the non-eukaryotic feature of dinoflagellate nuclei.

Though we need more molecular-biological information to discuss, the results described in this paper illustrate the complexity of dinoflagellates such as phylogenetically close relationship of dinoflagellates to chlorophytes and diverse transit peptides and promoters of [2Fe-2S] ferredoxin.

Acknowledgments We thank Dr. H. Maeda for his supplying cells of *P. bipes*. This research was supported by Grants in Aid for Scientific Research (No. 04660225 and 04301020) from the Ministry of Education, Science, and Culture, Japan and Grant Pioneering Research Project in Biotechnology from the Ministry of Agriculture, Forestry and Fisheries, Japan.

Note. The accession numbers of *fedQ* regions from *A. tamarensis* and from *P. bipes* in the DDBJ, EMBL, and GenBank nucleotide sequence databases are D89337 and D89338, respectively.

References

- 1) J. D. Dodge: The dinophyceae, in "The Chromosomes of the algae" (ed. by M. Godward), St. Martin's Press, New York, 1971, pp. 96-115.
- 2) M.-O. Soyer and O. K. Haapara: Structural changes of dinoflagellate chromosomes by pronase and ribonuclease. *Chromosoma*, **47**, 179-192 (1974).
- 3) D. F. Kubai and H. Ris: Division in the dinoflagellate *Gyrodinium cohnii* (Schiller). A new type of nuclear reproduction. *J. Cell Biol.*, **40**, 508-528 (1969).
- 4) M. Herzog, S. Von Boletzky, and M.-O. Soyer: Ultrastructural and biochemical nuclear aspects of eukaryote classification: independent evolution of the dinoflagellates as a sister group of the actual eukaryotes? *Orig. Life*, **13**, 205-215 (1984).
- 5) P. J. Rizzo: Comparative aspects of basic chromatin protein in dinoflagellates. *ByoSystems*, **14**, 433-443 (1981).
- 6) S. Bodansky, L. B. Mintz, and D. S. Holmes: The mesokaryote *Gyrodinium cohnii* lacks nucleosomes. *Biochem. Biophys. Res. Commun.*, **88**, 1329-1336 (1979).
- 7) J. D. Dodge: Chromosome structure in the dinoflagellates and the problem of the mesokaryotic cell. *Excerpta Med. Int. Congr. Ser.*, **91**, 264-265 (1965).
- 8) A. A. Gajadhar, W. C. Marquardt, R. Hall, J. Gunderson, E. V. Ariztia-Carmona, and M. L. Sogin: Ribosomal RNA sequences of *Sarcocystis muris*, *Theileria annulata* and *Cryptocodinium cohnii* reveal evolutionary relationships among apicomplexans, dinoflagellates, and ciliates. *Mol. Biochem. Parasitol.*, **45**, 147-154 (1991).
- 9) K. L. McNally, N. S. Govind, P. E. Thomé, and R. K. Trench: Small-subunit ribosomal DNA sequence analyses and a reconstruction of the inferred phylogeny among symbiotic dinoflagellates (Pyrrophyta). *J. Phycol.*, **30**, 316-329 (1994).
- 10) Y. Bhaud, M. Barbier, and M.-O. Soyer-Gobillard: A detailed study of the complex cell cycle of the dinoflagellate *Cryptocodinium cohnii* Biecheler and evidence for variation in histone H1 kinase activity. *J. Euk. Microbiol.*, **41**, 519-526 (1994).
- 11) A. Minguez, S. Franca, and S. M. D. de la Espina: Dinoflagellates have a eukaryotic nuclear matrix with lamin-like proteins and topoisomerase II. *J. Cell Sci.*, **107**, 2861-2873 (1994).
- 12) A. G. Hinnebusch, L. C. Klotz, R. L. Blanken, and A. R. Loeblich III: An evaluation of the phylogenetic position of the dinoflagellate *Cryptocodinium cohnii* based on 5S rRNA characterization. *J. Mol. Evol.*, **17**, 334-347 (1981).
- 13) A. Douwe de Boer and P. J. Weisbeek: Chloroplast protein topogenesis: import, sorting and assembly. *Biochim. Biophys. Acta*, **1071**, 221-253 (1991).
- 14) A. Uchida, S. Ebata, K. Wada, H. Matsubara, and Y. Ishida: Complete amino acid sequence of ferredoxin from *Peridinium bipes* (Dinophyceae). *J. Biochem.*, **104**, 700-705 (1988).
- 15) L. S. Kaufman, L. L. Roberts, W. R. Briggs, and W. F. Thompson: Phytochrome control of specific mRNA levels in developing pea buds. Kinetics of accumulation, reciprocity, and escape kinetics of the low fluence response. *Plant Physiol.*, **81**, 1033-1038 (1986).
- 16) G. Giuliano, E. Pichersky, V. S. Malik, M. P. Timko, P. A. Scolnik, and A. R. Cashmore: An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. *Proc. Natl. Acad. Sci. USA*, **85**, 7089-7093 (1988).
- 17) T. Caspar, P. H. Quail: Promoter and leader regions involved in the expression of the *Arabidopsis* ferredoxin A gene. *Plant J.*, **3**, 161-174 (1993).

- 18) Y. Sako, M. Adachi, and Y. Ishida: in "Toxic Phytoplankton Blooms in the Sea" (ed. by T. J. Smayda and Y. Shimizu), Elsevier, Amsterdam, 1993, pp. 87-93.
- 19) Y.-G. Liu, N. Mitsukawa, T. Oosumi, and R. F. Whittier: Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J.*, **8**, 457-463 (1995).
- 20) M. A. Frohman, M. K. Dush, and G. R. Martin: Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA*, **85**, 8998-9002 (1988).
- 21) J. D. Thompson, D. G. Higgins, and T. J. Gibson: CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucl. Acids Res.*, **22**, 4673-4680 (1994).
- 22) N. Saitou and M. Nei: The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406-425 (1987).
- 23) J. Felsenstein: PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, 1993.
- 24) D. J. McGeoch: On the predictive recognition of signal peptide sequences. *Virus Res.*, **3**, 271-286 (1985).
- 25) G. von Heijne: Patterns of amino acids near signal-sequence cleavage sites. *Eur. J. Biochem.*, **133**, 17-21 (1983).
- 26) M. Kozak: An analysis of vertebrate mRNA sequences: Intimations of translational control. *J. Cell Biol.*, **115**, 887-903 (1991).
- 27) M. Kimura: A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, **16**, 111-120 (1980).
- 28) M. Pilon, H. Wienk, W. Sips, M. de Swaaf, I. Talboom, R. van't Hof, G. de Korte-kool, R. Demel, P. Weisbeek, and B. de Kruijff: Functional domains of the ferredoxin transit sequence involved in chloroplast import. *J. Biol. Chem.*, **270**, 3882-3893 (1995).
- 29) L. M. Gierasch: Signal sequences. *Biochemistry*, **28**, 923-930 (1989).
- 30) T. Caspar and P. H. Quail: Promoter and leader regions involved in the expression of the *Arabidopsis* ferredoxin A gene. *Plant J.*, **3**, 161-174 (1993).
- 31) R. Iglesias-Prieto, N. S. Govind, and R. K. Trench: Isolation and characterization of three membrane-bound chlorophyll-protein complexes from four dinoflagellate species. *Phil. Trans. R. Soc. Lond. B*, **340**, 381-392 (1993).
- 32) Q. Su and A. Boschetti: Partial purification and properties of enzymes involved in the processing of a chloroplast import protein from *Chlamydomonas reinhardtii*. *Eur. J. Biochem.*, **217**, 1039-1047 (1993).
- 33) D. J. Schnell, F. Kessler, and G. Blobel: Isolation of components of the chloroplast protein import machinery. *Science*, **266**, 1007-1012 (1994).
- 34) D.-H. Lee, M. Mittag, S. Sczekan, D. Morse, and J. W. Hastings: Molecular cloning and genomic organization of a gene for luciferin-binding protein of the dinoflagellate *Gonyaulax polyedra*. *J. Biol. Chem.*, **268**, 8842-8850 (1993).
- 35) T. Yoshikawa, A. Uchida, and Y. Ishida: There are 4 introns in the gene coding the DNA-binding protein HCC of *Cryptocodinium cohnii* (Dinophyceae). *Fisheries Sci.*, **62**, 204-209 (1996).