DNA Research **9**, 71–78 (2002)

Biochemical Examination of the Potential Eukaryotic-type Protein Kinase Genes in the Complete Genome of the Unicellular Cyanobacterium *Synechocystis* sp. PCC 6803

Ayako Kamei, Takashi Yuasa, Xiaoxing Geng, and Masahiko Ikeuchi*

Department of Life Sciences (Biology), The University of Tokyo, Komaba 3-8-1, Meguro, Tokyo 153-8902, Japan

(Received 1 May 2002; revised 31 May 2002)

Abstract

The complete genome of the unicellular motile cyanobacterium *Synechocystis* sp. PCC 6803 harbors seven putative genes for a subfamily Pkn2 of the eukaryotic-type (or "Hanks-type") protein kinase. Previously, SpkA and SpkB were shown to have protein kinase activity and to be required for cell motility. Here, the other five genes were examined. These genes, except for *spkG* (*slr0152*), were successfully expressed in *Escherichia coli*. Eukaryotic-type protein kinase activity of the expressed SpkC (Slr0599), SpkD (Sll0776) and SpkF (Slr1225) was demonstrated as autophosphorylation and phosphorylation of the general substrate proteins. SpkE (Slr1443) did not show any activity, a finding consistent with its lack of several key amino acid residues in its kinase motif. Gene-disrupted mutants showed no discernible defect in phenotype except that *spkD* was apparently essential for survival.

Key words: autophosphorylation; cyanobacterium; motility; phototaxis; *pkn2*; protein kinase; signal transduction

1. Introduction

Protein phosphorylation is the principal mechanism with which extracellular signals are translated into cellular responses. In eukaryotes, signal transduction is materialized through cascades of protein phosphorylation mainly on Ser, Thr and Tyr residues catalyzed by a huge superfamily of protein kinases. They are highly differentiated into a number of families and often called eukaryotic-type or Hanks-type.[1,2] On the other hand, in prokaryotes phosphorylation mainly occurs on His and Asp residues via His kinases and response regulators as the two-component signal transduction system. However, it is known that the eukaryotic-type protein kinases are also present in some bacteria and archea,[3–5] while the elements of the two-component signal transduction system have been reported in archaea and eukaryotes.[6,7]

The first putative gene for a eukaryotic-type protein kinase in the bacterial world was identified in *Myxococcus xanthus*, a Gram-negative soil bacterium. This kinase is required for normal development of multicellular fruiting bodies.[3] Since then, similar eukaryotic-type protein kinase genes have been found in many bacteria. Now, a greater number and wider variety of genes are emerging as genome projects proceed.[8] It has been found that some bacterial genomes, including cyanobacteria, carry a few to over 50 eukaryotic-type protein kinase genes, while others do not have any such genes. These genes in bacteria and archaea belong to the eukaryotic-type protein kinase superfamily and are categorized into several subfamilies based on sequence homology, although each of them needs to be confirmed biochemically.[8] However, only a few of them have been demonstrated to encode active kinases and/or to be involved in biological processes.

Cyanobacteria are unique prokaryotes in the sense that they developed chloroplasts during the evolution of algae and plants, and the chloroplasts very likely contributed a considerable part of their nuclear genomes.[9] At present, complete genome sequences have been published for two species: *Synechocystis* sp. PCC 6803 and *Anabaena* sp. PCC 7120.[10,11] Sequence analysis suggested that *Synechocystis* has 7 ORFs for the Pkn2 family and 4 ORFs for the ABC1 family, while *Anabaena* has 35 ORFs for Pkn2, 13 for HstK and 4 for ABC1.[12] Apparently, the Pkn2 family is universal among many bacteria, including cyanobacteria. In an earlier study, we demonstrated that two hypothetical ORFs *sll1574* and *sll1575* in the sequenced strain of *Synechocystis* resulted from frameshift mutation in a single gene (*spkA*) of the Pkn2

Communicated by Satoshi Tabata

* To whom correspondence should be addressed. Tel. +81-3-5454-6641, Fax. +81-3-5454-4337, E-mail: mikeuchi@bio.c.u-tokyo.ac.jp

subfamily.[13] The original *spkA* encodes a functional protein kinase, which is essential for the twitching motility of cyanobacterial cells. We also showed that *spkB* (*slr1697*) encodes a functional protein kinase and is essential for optimal motility.[14] It is now important to study all genes of the *Synechocystis* Pkn2 family to gain a comprehensive understanding of the cyanobacterial genome and its cellular signaling network. Here, we examined protein kinase activity and the physiological roles of the other five genes belonging to the Pkn2 family.

2. Materials and Methods

2.1. Strains and culture conditions

The unicellular motile cyanobacterium *Synechocystis* sp. PCC 6803 was obtained from the Pasteur Culture Collection. A specific clone showing positive phototaxis toward the white light was established as wild type (denoted PCC-P) and subjected to gene disruption or used as a genetic source for cloning and expression.[15] Standard strain and mutants were grown in BG11 medium buffered with N-tris(hydroxymethyl)methyl-2-aminoethanesulfonic acid-KOH, pH 7.8 at 31°C at a light intensity of 20 μmol photon $\cdot$ m^{-2} $\cdot$ s^{-1}. Solid medium was supplemented with 0.8% (w/v) agar and 0.3% (w/v) sodium thiosulfate and used for examination of motility by colony morphology. Kanamycin (20 μg/ml) was added to maintain gene-disrupted mutants; antibiotics were not included for characterization of the mutant phenotype. For analysis of phototactic motility, 1 μl of concentrated cell suspension (ca. 2×10^9 cells/ml) was spotted on the agar plate, dried and incubated for 48 hr under lateral illumination with a white fluorescent lamp at 10 μmol photon$\cdot$m$^{-2}\cdot$s^{-1}. For cloning and subcloning of DNA in plasmids, strains XL10 (Epicurian Coli XL10-Gold Kan, StrataGene) and JM109 of *Escherichia coli* were used, while BL21(DE3) pLysS was used for expression with pET28a (Novagen).

2.2. Construction of spkC*,* spkD*,* spkE*,* spkF*, and* spkG*-disrupted mutants*

To inactivate *spkC* (*slr0599*), an upstream fragment of *spkC*, and a part of *spkC* were separately amplified by PCR using primer-1 (5′-CGGTGCTGACAGTCATGG-3′) and primer-2 (5′-CTCGAGCTTGTCCAGCTTT-CAA-3′), and primer-3 (5′-GAATTCAAACTACTCA-ACAATCGC-3′) and primer-4 (5′-CTGCAGAATA-GGCACTACGGCC-3′), respectively, and cloned into pT7Blue-T. These fragments were ligated with the kanamycin-resistance cassette, resulting in an insertional disruption construct of *spkC* near the translation initiation codon. A part of *spkD* (*sll0776*) was amplified by PCR using primer-5 (5′-ACTGGAATGTGGCGGCTG-3′) and primer-6 (5′-GCAGTGACCACAATAATCC-3′), cloned and interrupted at the *Hpa* I site by insertion of the cassette. A part of *spkE* (*slr1443*) was amplified by PCR using primer-7 (5′-CCTCGATACCTGGCAGAG-3′) and primer-8 (5′-TGGCGATCGCCTACAAGC-3′), cloned and interrupted at the *Msc* I site by the cassette. A part of *spkF* (*slr1225*) was amplified by PCR using primer-9 (5′-CCAGTTGTCGGATGCCAC-3′) and primer-10 (5′-TGCTACTACGGCGATCGC-3′), cloned and interrupted at the *Hpa* I site with the cassette. A part of *spkG* (*slr0152*) was amplified by PCR using primer-11 (5′-CAGACCGGGAAATTAAACC-3′) and primer-12 (5′-ATTGCCCTGCCAGGTAGC-3′), cloned and interrupted at the *Msc* I site by insertion of the cassette. The Tn5-derived kanamycin-resistance cassette without a transcription terminator was inserted in the same direction as the target ORFs to allow expression of downstream genes cotranscribed, if any.[13] Mutants were generated by transformation of strain PCC-P cells with these DNAs and selected on BG11 plates containing kanamycin (20 μg/ml) according to Hihara et al.[16] Complete segregation was confirmed by PCR with the same primers as described above (data not shown), except that the wild-type genome could not be excluded in the *spkD* mutant.

2.3. Cloning of spkC*,* spkD*,* spkE*,* spkF*, and* spkG *for expression*

The coding regions of *spkC*, *spkD*, *spkE*, *spkF*, and *spkG* were amplified from the genomic DNA of the motile PCC-P strain by PCR with primer-13 (5′-CATATGGTTACCCCACTCAA-3′) and primer-14 (5′-TAGAATTCCCTAATTTTGCTCGGG-3′), primer-15 (5′-ATAAACAAACCATATGAATGTCCAAGTACT-3′) and primer-16 (5′-CAGAATTCCCACTAACAGAGCG-3′), primer-17 (5′-CCCATATGACCCCATCCTTGTCC-3′) and primer-18 (5′-GAGAATTCTCCCGATTCG-TCG-3′), primer-19 (5′-AGCATATGGATCTGCTCTG-CACC-3′) and primer-20 (5′-ATGAATTCCCGGCCC-AAACTGGC-3′), and primer-21 (5′-GTCATATGGA-ATCACTCCCCGTC-3′) and primer-22 (5′-AGGA-ATTCAATCTTACAGTCTGTGG-3′), which were designed to contain *Nde* I and *Eco*RI restriction sites, respectively. PCR was performed with *Pfu* DNA polymerase (Stratagene) according to the manufacturer's instructions. Following initial denaturation at 95°C for 1 min, DNA was amplified with 35 cycles consisting of denaturation at 95°C for 30 sec, annealing at 57°C for 1.5 min and elongation at 72°C for 4 min. PCR products were cloned into the *Srf* I-digested pPCR-Script (Stratagene) using XL10 according to the manufacturer's instructions. The DNA sequence of these fragments was determined by the BigDye terminator fluorescence detection method (PE Applied Biosystems) using a capillary sequencer (PRISM 310 Genetic Analyzer; PE Applied Biosystems) and compared with the complete genome sequence of the Kazusa strain.[10] The cloned DNAs

were excised with *Nde* I and *Eco*RI and inserted into pET28a (Novagen) to allow expression by fusion with the N-terminal His-tag.

2.4. Expression of spk genes in E. coli

pET28a carrying *spk* genes was introduced into *E. coli* BL21 (DE3) pLysS. Cells were grown at 37°C in 10 ml Luria broth medium containing kanamycin (20 μg/ml) and chloramphenicol (37 μg/ml) to an A_{600} of about 0.5. Then, isopropyl-β-D-thiogalactoside (IPTG) was added to a final concentration of 0.5 mM, and the cultures were incubated for 2 hr at 25°C. The cells were harvested by centrifugation, washed with 50 mM Tris-HCl (pH 7.5) containing 100 mM NaCl and 1 mM phenylmethylsulfonyl fluoride and resuspended in 100 μl of the same medium plus 10% (w/v) glycerol. The cell suspension was sonicated at 4°C for 2 min in a sonicator (Ultrasonic Generator U0300FB). The cell extracts were centrifuged at $13,000 \times g$ for 10 min. Precipitates including insoluble materials and cellular membranes were resuspended in 100 μl of the same medium. Proteins were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by staining with Coomassie brilliant blue R-250. Alternatively, proteins in the SDS-gel were blotted to a polyvinylidene difluoride membrane (Immobilon; Millipore) and the His-tag was detected with His-Probe (Pierce) followed by color development using 3,3′-diamino benzidine (DAB) or chemiluminescence using SuperSignal West Pico^{TM} (Pierce) as instructed by the manufacturer.

2.5. Assay of protein kinase activity

Autophosphorylation of SpkC, SpkD, SpkE, and SpkF and phosphorylation of bovine myelin basic protein ("MBP," Sigma), calf thymus histone (type-IIIS, Sigma) or bovine casein (partially dephosphorylated, Sigma) were assayed *in vitro* with [γ-^{32}P]ATP according to Kamei et al.[13] These are general substrates for typical Ser/Thr protein kinases. About 10 μg protein of each cell extract of *E. coli* was added to 10 μl phosphorylation buffer containing 20 mM 2-morpholinoethanesulfonic acid (pH 6.5), 10 mM $MgCl_2$ and 0.1 mM [γ-^{32}P]ATP (3000 Ci $\cdot$ mol^{-1}), with or without 2.5 μg of each substrate protein, and incubated for 15 min at 30°C. Control phosphorylation experiments were done with a crude extract from *E. coli* before induction. SDS (final 1%) and dithiothreitol (final 60 mM) were added to stop the reaction. After boiling for 5 min, proteins were resolved by SDS-PAGE. The gels were stained with Coomassie brilliant blue R-250, dried and then subjected to autoradiography with X-ray film (X-Omat Blue XB-1, Eastman Kodak).

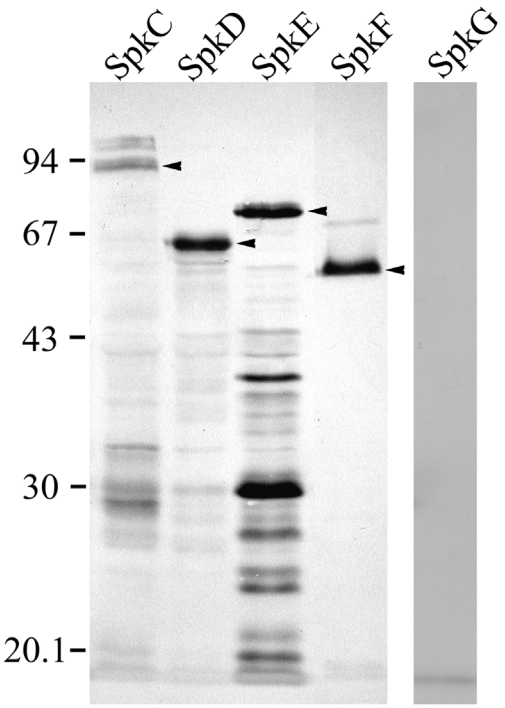

Figure 1. Expression of *Synechocystis* protein kinases in *E. coli*. Proteins of soluble fractions (SpkC, SpkD, SpkE, and SpkG) or insoluble fraction (SpkF) were resolved by SDS-PAGE (12% gel), blotted, probed with His-Probe and detected by DAB (SpkC $\sim$ SpkF) or chemiluminescence (SpkG). Arrowheads show bands for the native protein kinases. Molecular markers; phosphorylase b, bovine serum albumin, ovalbumin, carbonic anhydrase and trypsin inhibitor.

3. Results

3.1. Expression of spk *genes in* E. coli

It is known that the sequenced strain carries several mutations which are not present in a motile wild-type strain PCC-P.[17] Thus, we cloned and determined nucleotide sequences of the five *spk* genes from strain PCC-P. They completely matched with those from the non-motile sequenced strain, in contrast to a frameshift mutation in *spkA*.[13]

We tried to express in *E. coli* functional proteins of SpkC, SpkD, SpkE, SpkF, and SpkG with the N-terminal His-tag fusion under the T7 promoter. Only after induction could we detect the His-tagged proteins in crude extracts of *E. coli* by Western blotting with His-probe and by DAB color development (Fig. 1). The major 93 kDa, 66 kDa, 75 kDa, and 55 kDa bands were detected for SpkC, SpkD, SpkE, and SpkF, respectively. The products of SpkC, SpkD, and SpkE were mostly found in the soluble fraction, while SpkF was recovered from the insoluble fraction. Expression of SpkG could not be confirmed in *E. coli* even with the sensitive chemiluminescence method. The apparent molecular masses were roughly consistent with their predicted values except for

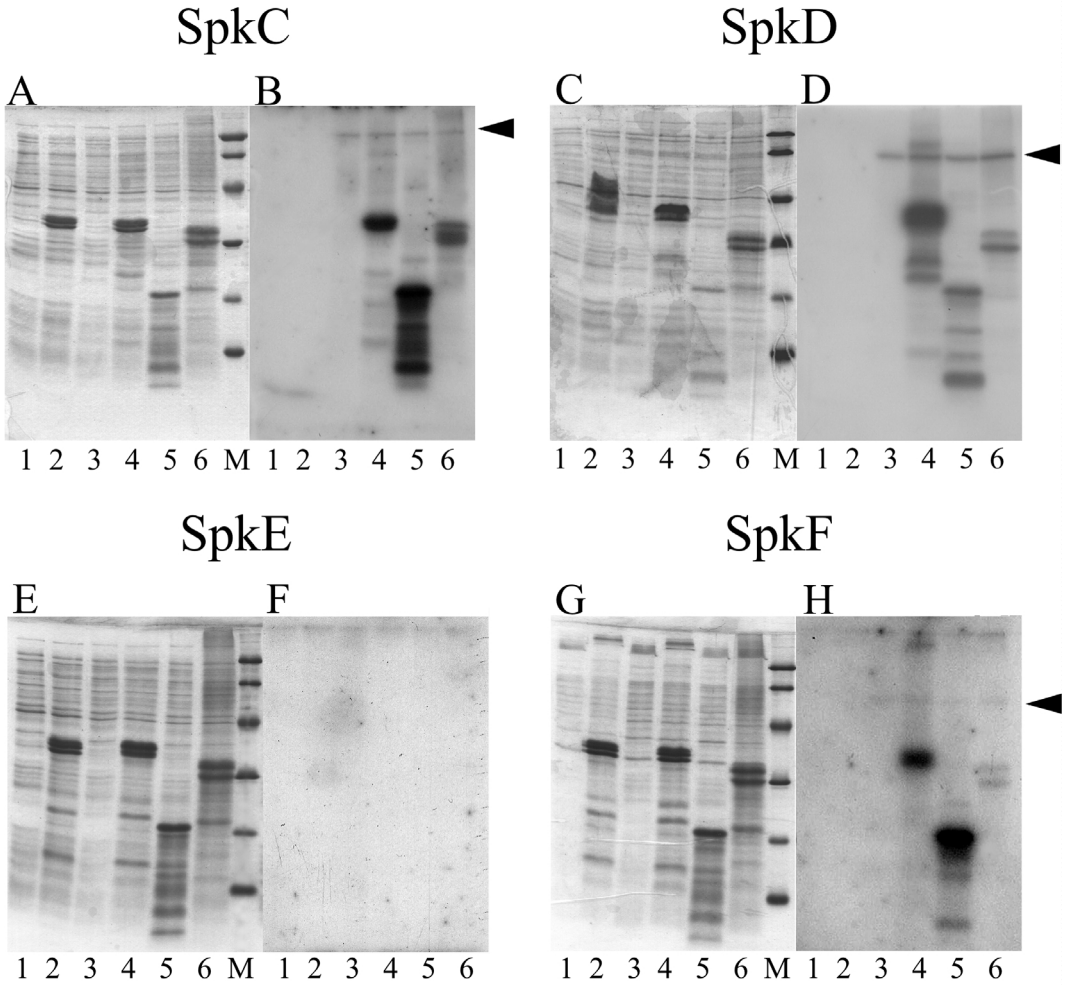

Figure 2. Detection of protein kinase activity. Phosphorylated proteins were resolved on 15% polyacrylamide gel and visualized by Coomassie brilliant blue (panels A, C, E, and G) and autoradiography (panels B, D, F, and H). Lanes 1 and 2, cell extract of non-induced *E. coli* without (lane 1) or with histone (lane 2). Lanes 3 to 6, cell extracts of IPTG-induced *E. coli* without substrate (lane 3), with histone (lane 4), MBP (lane 5), and casein (lane 6). Lane M, molecular markers. The arrowhead shows autophosphorylation of each expressed protein kinase. Note that the soluble fraction of the extracts was used for SpkC, SpkD, and SpkE, while the insoluble fraction was used for SpkF.

SpkC (93 kDa in Fig. 1 vs. 60 kDa by prediction). This discrepancy may be accounted for by the high content of Pro residues (15.7%) in SpkC. There are faint bands in the low-molecular-weight region in SpkC and SpkE, which were assumed to be degradation products retaining the tag.

3.2. Protein kinase activity of SpkC, SpkD, SpkE, and SpkF

Previously, we showed that cell extract of *E. coli* carrying cyanobacterial *spkA* before induction did not phosphorylate any bacterial proteins *per se* or the substrate proteins.[13] Thus, we examined protein kinase activity of the expressed proteins in crude extracts from *E. coli* with or without substrate proteins (Fig. 2). Expectedly, cell extracts carrying *spk* genes before induction did not show any phosphorylation (Fig. 2, lanes 1 and 2). After induction, significant autophosphorylation was detected in the presumptive undegraded proteins of SpkC, SpkD, and SpkF. Nevertheless, none of the degraded products of Spk proteins or any proteins derived from *E. coli* were labeled at all (Fig. 2, lane 3). Strong phosphorylation was observed in histone, MBP or casein, when they were added to SpkC, SpkD, or SpkF (Fig. 2, lanes 4, 5, and 6). Efficiency of labeling appeared to depend on both kinase and substrate species. Notably, kinase activity was detected in the insoluble fraction of SpkF, while the soluble fraction containing SpkC or SpkD was active in autophosphorylation and substrate phosphorylation. On the other hand, the SpkE protein or its degraded products, which were mostly recovered in the soluble fraction, did not show any phosphorylation (Fig. 2). Thus, we concluded that SpkC, SpkD, and SpkF are active eukaryotic-type

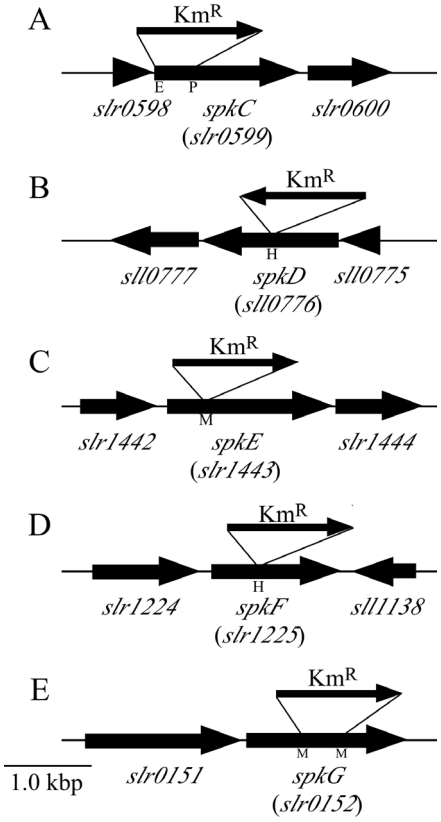

Figure 3. Gene map showing the position of the *spkC*, *spkD*, *spkE*, *spkF*, and *spkG* genes. A 0.54-kbp fragment was excised from *spkC* with *Eco*RI (E) and *Pst* I (P) and replaced with the kanamycin-resistance cassette (panel A). The kanamycin-resistance cassette was inserted into an *Hpa* I (H) site of *spkD* (panel B) and *spkE* (panel C), and an *Msc* I (M) site of *spkF* (panel D). A 0.43-kbp fragment was excised from *spkG* with *Msc* I and replaced with the kanamycin-resistance cassette (panel E).

protein kinases like SpkA.[13] To identify intrinsic substrates for them, we performed *in vitro* phosphorylation of *Synechocystis* proteins by these Spk proteins as already described for SpkA.[13] However, no specific labeling was detected in either the soluble or membrane fractions, even though a large amount of crude Spk proteins was added (data not shown). We could observe only some intrinsic phosphorylation of several bands without the Spk proteins, as described previously. This suggests that *Synechocystis* has as yet unidentified protein kinase(s) other than the Spk proteins.

3.3. Phenotype of gene-disrupted mutants

To know the functional roles of these Spk proteins, we constructed gene-disrupted mutants from the motile wild-type strain PCC-P that shows positive phototaxis (Fig. 3).[15] Complete segregation of the mutant genome was achieved for disruption of *spkC*, *spkE*, *spkF*, and *spkG*. On the other hand, *spkD*-disrupted genome was always accompanied with the wild-type DNA, even though the heteroplasmic cells grew at a rate comparable to that of the wild type. All mutants showed positive phototaxis similar to the wild-type strain (Fig. 4), indicating that *spkC*, *spkD*, *spkE*, and *spkF* are not involved in motility. We examined photoautotrophic growth in liquid or on solid media under low (20 μmol photon$\cdot$m$^{-2}\cdot$s^{-1}) or high light (300 μmol photon$\cdot$m$^{-2}\cdot$s^{-1}), content of chlorophyll *a*, carotenoids and phycocyanin during growth under low or high light, and photosystem stoichiometry (ratio of photosystem II to I). However, we could not detect any discernible difference between wild type and the mutants (data not shown).

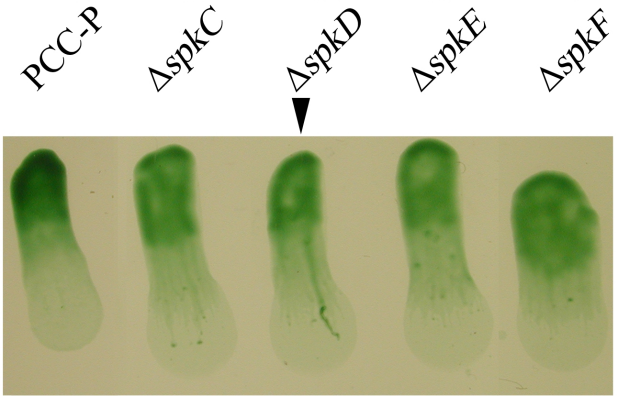

Figure 4. Phototactic movement of gene-disrupted mutants generated in the motile strain PCC-P of *Synechocystis* sp. PCC 6803. Cell suspension was spotted on the agar plate and incubated for 48 hr under lateral illumination (arrowhead).

4. Discussion

In this communication, we demonstrated that at least SpkC, SpkD, and SpkF are functionally active protein kinases of the eukaryotic type, even though disruption of their genes did not show any recognizable phenotype in *Synechocystis*. On the other hand, SpkE is likely a nonfunctional mutant of a protein kinase. Fortunately, *E. coli* did not show any stable phosphorylations on intrinsic proteins or the general phosphorylatable proteins under our experimental conditions, even though it intrinsically has a number of His kinases and other ATPases. Thus, putative protein kinase genes can be efficiently evaluated in crude extracts of *E. coli*. The present results are summarized in Table 1 together with our previous work on SpkA[13] and SpkB.[14]

SpkC consists of the N-terminal protein kinase domain and a C-terminal Pro-rich domain. The latter domain is somehow conserved in several potential protein kinases of Pkn2 in *Anabaena* (*alr2412*, *alr4368*, and *all4668*).[12] Although we could not recognize any phenotype of the *spkC*-disrupted mutants, overexpression of *spkC* was apparently lethal since we could not obtain any transfor-

Table 1. Summary of "eukaryotic-type" protein kinases in *Synechocystis* sp. PCC 6803.

	molecular size (kDa)		kinase	physiological	unique	homolog in	
	theoretical	SDS-PAGE	activity	role	feature	Anabaena 7120	ref
SpkA	59	63	+	motility	-	All4518	10
SpkB	63	77	+	motility	Cys motif pentapeptide repeats	Alr3268	11
SpkC	60	93	+	?	Pro rich	-	this work
SpkD	55	66	+	?	Glu rich	-	this work
SpkE	69	75	-	?	-	Alr2412	this work
SpkF	54	55	+	?	Cys motif TM motif	-	this work
SpkG	71	-	n.d.	?		-	this work

TM, Transmembrane.

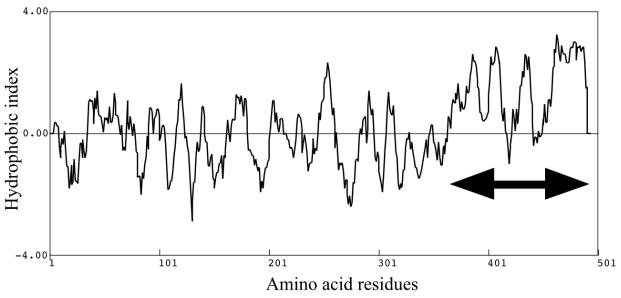

Figure 5. Hydropathy plot of SpkF. Hydrophobic indices were calculated with the Kyte-Doolittle scale.[18] An arrow indicates a C-terminal hydrophobic region.

mants with a construct carrying the strong promoter of *psbAII* instead of the native promoter. This implies that SpkC either is not expressed or is inactive under normal conditions. By contrast, we could not achieve complete segregation of the *spkD*-disrupted genome, whereas the disrupted mutant carrying wild-type genome or overexpression mutant with the *psbAII* promoter did not show any recognizable phenotype. SpkD may play an essential role in *Synechocystis*, although no clear homolog was found in the *Anabaena* genome.

Sequence motif analysis showed that SpkF has the whole protein kinase in the N-terminal part, while it has four transmembrane signatures in the C-terminal part (Fig. 5). Consistent with this, SpkF exclusively accumulated in the insoluble fraction of *E. coli* as an active protein kinase. Judging from the membrane topology, the N-terminal protein kinase domain is located in the cytoplasm, while anchored on the cytoplasmic membrane or the thylakoid membrane. The SpkF homolog is absent in *Anabaena* but present in *Nostoc punctiforme* ATCC 29133 (ORF466, http://www.jgi.doe.gov/JGI_microbial/html/nostoc/nostoc_mainpage.html) and a thermophilic cyanobacterium *Thermosynechococcus elongatus* strain BP-1 (ORF503, Kaneko, T. and Tabata, S., personal communication). They show very similar hydropathy profiles to *Synechocystis* SpkF, indicative of the same membrane topology (data not shown). Moreover, these SpkF proteins shared a unique motif including four Cys residues on the N-terminus, which is different from the N-terminal four-Cys motif of SpkB and other protein kinases. On the other hand, active proteins of SpkC and SpkD were recovered from the soluble fractions of *E. coli*, even though not a single transmembrane region is predicted in any of them.

Expressed SpkE did not show any protein kinase activity under the experimental conditions employed. $MnCl_2$, which often enhances the activity of protein kinase,[19,20] was not effective for SpkE. In agreement with this, the deduced product of *spkE* lacked several key residues for the catalytic reaction, such as (1) Gly in subdomain I, (2) Lys in Subdomain II, (3) Asp and Asn in Subdomain VI and (4) Gly in subdomain VII, although its overall sequence is highly homologous to the typical protein kinase motif (Fig. 6). Gly in subdomain I is a member of the Gly-X-Gly-X-X-Gly motif for ATP binding, and Lys in subdomain II is known to be essential for protein kinase activity.[8,21] By contrast, the *Anabaena* homolog Alr2412 retained all the key amino acid residues, which were replaced in *Synechocystis* SpkE (Fig. 6). Taking into account the fact that those key residues are also conserved in the active protein kinases of *Synechocystis* (SpkA $\sim$ SpkD, SpkF), we must conclude that *Synechocystis* SpkE is a pseudogene. In the complete genome of the sequenced strain (Kazusa strain) of *Synechocystis* sp. PCC 6803, several pseudogenes have been documented, which resulted from frameshift mutation or transposition of the insertion sequence (ISY203).[13,17,22] Instead of split pseudogenes,

```
S.6803 SpkE                                        MTPSLSLGQRLDDRYLIQLHLGQNSLGQQYLAQDTH
A.7120 Alr2412  MTKLYCSKGHENYPGTRFCLTCGEKLVDNFITQSIQPGLTLSDRYLIVRQVGQGGFGRTYLAEDIN
                                                  * *   *  * *****    **   *  *** *

S.6803 SpkE     RFDEPCTIT-----VLTAPPGDKAGNLFKQQAEILYALDHPQIANFREFFAL---GSE-LYLVQDF
A.7120 Alr2412  RFRELCILKEFSPQVQTAYVLQKAEELFQRBAKVLHQLQHPQIPRFREIFRVNLAGKEYLFLVQDY
                ** * *        * **    **  **   *  *  * ****  *** *     * * * ****

S.6803 SpkE     IEGQSYFDLLKDRRLQGKAFTELEVRQWCRQLLPVLHYLHIQKICHGQISPSSIMRRTVDGLPVLV
A.7120 Alr2412  VEGETYSGLLNHRIQQGLRFTEAEIRQLLQQILPVLDYIHSLGVIHRDISPDNLILRSVDKLPVLI
                 **  *  **  *  **  *** * **   * **** * *     *  ***     * ** ****

S.6803 SpkE     DF---------THSQYYSG-----------------TPGEDRR--------DLHGLGLGAIALLTG
A.7120 Alr2412  DFGGVKQVVAVVASQYYQPGVVASPPAATLLGKVGFAPPEQMQTGNVSPHSDLYALAVTAVVLLTG
                **            ****                  * *          **  *   *  ****

S.6803 SpkE     EV------NPNPPWENLLSSVALSQEFRQLLLKLLAWPPLVDWPT---LTDNLNNLPQHFLPTVAE
A.7120 Alr2412  KQPQELLDTYNLSW-NWQREISLSPILGQVLDKMLAARPGDRYQSAQQVLHALNPAPTNYPPTQAP
                          *  * *      **    * * * **  *             **  *    ** *

S.6803 SpkE     FNLTSTIPEQTDNGVGKSSTGEPPFPTVHQSPESSSSAKVKNMVRVDGLKGLVKKSFILLGLMAIA
A.7120 Alr2412  TPTSVPIPPPVPTPDTIAIAPSLPTPTISPAPQ--------------------RKNWL---IPTAA
                      **               * **    *                      *         *

S.6803 SpkE     MALGWGAGQLWL--DNQQRAALEKLAQEEPGDDVP-------EKTDLEIKNEIRARRLNLGISPQR
A.7120 Alr2412  LLLVGTAGLVWLGMSNSDRHSISVEPTPTPTETQPTNPLDKYSPTERRRKQRLNDRRQQLGID---
                  *   **  **   *  *          *    *         *    *     **  ***

S.6803 SpkE     FQTLMDDGLAFQLGIDPQQRLEKNPNTNGAPLSLGSPEQQM-----AMTITVLDALESLSREATRD
A.7120 Alr2412  --------FNFYINVVNQIFWDRNPSERGRTLSDGAEDENLRAEWDAVAAELLDKLKPLSNKARRQ
                         *      *     **   *  ** *           *     ** *  **  * *

S.6803 SpkE     FAQNNSGDRRRWIPRVNQLRLSSRSFYDLVNARFHHYLPMVSPALLGEPEFEQRPLAQVWNAMAFD
A.7120 Alr2412  LGTYTAAERDRWKVEVNQINVGSRSLYDLGDAAFFKAFPEQK-----GKNFIEQPIGQVWQAFVSD
                       * **   ***    *** ***  * *    *           *   *  *** *   *

S.6803 SpkE     SLASLEDESHYQRLSFD-DTNQLNLNGTLEPGEGYAYAVTIPPTEEFSLQLTAPPTARISFYPPTG
A.7120 Alr2412  KLSAIVAGSAFQKIVFDPGAINKTVSGTLQPAGGKVFIADLAQDQSLELQLKANPKVLLSVYSPSG
                 *      *  *   **         *** *  *              *** * *    * * * *

S.6803 SpkE     PEVILQNSAMHRWSGPTDQAGYYELVITSVAEEPIAFELELSIIPASAQPF*
A.7120 Alr2412  KTIFLEDSEKRSLSTELPESGFYEFVVVSTASTSADYRLTLTAENPTPPPEPTPTPSETPTPTPTP
                    *  *     *      * ** *  * *       * *

A.7120 Alr2412  TPSETPTETPTPTPTTTP*
```

Figure 6. Amino acid sequence alignment of *Synechocystis* SpkE (S. 6803 SpkE) with Alr2412 from *Anabaena* sp. PCC 7120 (A. 7120 Alr2412). Residues conserved with SpkE are shown with asterisks. The highly conserved residues in eukaryotic-type protein kinase are indicated with reverse type.

functional genes (*spkA*, *sll1473–75*, *pilC*) were found in a substrain close to the original wild-type isolate (strain PCC-P), indicative of recent inactivation during propagation of the culture strain. On the other hand, *spkE* in strain PCC-P shared the same missense mutations with the sequenced strain. This suggests a rather ancient origin for the inactivation of *spkE*. However, it is surprising that the whole ORF of *spkE* is conserved despite multiple mutations in the limited number of key residues, when compared with the *Anabaena* homolog *alr2412*. In general, nonsense or frameshift mutations are found in many pseudogenes, as exemplified in various fragmented ORFs derived from transposase genes in the *Synechocystis* genome.[10] Obviously, this is not the case with *spkE*. Conservation of the whole ORF in *spkE* suggests that the C-terminal domain of SpkE still retains some physiological activity. To confirm this, we created both disruption and overexpression mutants for *spkE* but could not detect specific phenotypes like other mutants. Potential domain inactivation by multiple missense mutation is also suggested in a *lexA*-like gene (*sll1626*) in *Synechocystis*.[23] The authentic LexA in *E. coli* encodes a transcriptional factor that represses *recA* and other SOS-related genes. The repression is released by RecA-induced autoproteolysis of LexA. A similar gene encoding the DNA-binding domain and the autoproteolytic domain is also found in many cyanobacterial genomes (*sll1626* in *Synechocystis*, *alr4908* in *Anabaena*).[10,11] DNA microarray analysis showed that Sll1626 represses several motility-related genes instead

of SOS-related genes in *Synechocystis*.[23] Interestingly, at least two of the key amino acid residues in the autoproteolytic domain are replaced by missense mutations only in *Synechocystis*, leading to presumptive inactivation of this domain while preserving the repressor activity of the DNA-binding domain. These observations may imply that *spkE* and *sll1626* are evolving towards the acquisition of a novel function instead of their original regulatory mechanisms, which proceeds only in *Synechocystis*. Further biochemical examination of those proteins would provide insights into molecular mechanisms by which genes lose and acquire functions in evolution.

We are currently undertaking analysis using a yeast two-hybrid screening system in order to identify the physiological function and the pair of kinase, phosphatase, substrate and other interacting components.

Acknowledgements: This work was supported by a Research Fellowship for Young Scientists from the Japan Society of the Promotion of Science (to A. K.), by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan (to M. I.), by the Program for Promotion of Basic Research Activities for Innovative Biosciences of Japan (to M. I.) and by a Grant for Scientific Research from the Human Frontier Science program (to M. I.).

References

1. Hanks, S. K., Quinn, A. M., and Hunter, T. 1988, The protein kinase family: conserved features and deduced phylogeny of the catalytic domains, *Science*, **241**, 42–52.
2. Hanks, S. K. and Hunter, T. 1995, The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification, *FASEB J.*, **9**, 576–596.
3. Munoz-Dorado, J., Inouye, S., and Inouye, M. 1991, A gene encoding a protein serine/threonine kinase is required for normal development of *M. xanthus*, a gram-negative bacterium, *Cell*, **67**, 995–1006.
4. Zhang, C. C. 1993, A gene encoding a protein related to eukaryotic protein kinases from the filamentous heterocystous cyanobacterium *Anabaena* PCC 7120, *Proc. Natl. Acad. Sci. USA*, **90**, 11840–11844.
5. Kennelly, P. J. and Potts, M. 1996, Fancy meeting you here! A fresh look at "prokaryotic" protein phosphorylation, *J. Bacteriol.*, **178**, 4759–4764.
6. Loomis, W. F., Shaulsky, G., and Wang, N. 1997, Histidine kinases in signal transduction pathways of eukaryotes, *J. Cell Sci.*, **110**, 1141–1145.
7. Kim, D. and Forst, S. 2001, Genomic analysis of the histidine kinase family in bacteria and archaea, *Microbiology*, **147**, 1197–1212.
8. Leonard, C. J., Aravind, L., and Koonin, E. V. 1998, Novel families of putative protein kinases in bacteria and archaea: Evolution of the "eukaryotic" protein kinase superfamily, *Genome Res.*, **8**, 1038–1047.
9. Rujan, T. and Martin, W. 2001, How many genes in *Arabidopsis* come from cyanobacteria? An estimate from 386 protein phylogenies, *Trends Genet.*, **17**, 113–120.
10. Kaneko, T., Sato, S., Kotani, H. et al. 1996, Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions, *DNA Res.*, **3**, 109–136.
11. Kaneko, T., Nakamura, Y., Wolk, C. P. et al. 2001, Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120, *DNA Res.*, **8**, 205–213.
12. Ohmori, M., Ikeuchi, M., Sato, N. et al. 2001, Characterization of genes encoding multi-domain proteins in the genome of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120, *DNA Res.*, **8**, 271–284.
13. Kamei, A., Yuasa, T., Orikawa, K., Geng, X. X., and Ikeuchi, M. 2001, A eukaryotic-type protein kinase, SpkA, is required for normal motility of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803, *J. Bacteriol.*, **183**, 1505–1510.
14. Kamei, A., Yuasa, T., Geng, X., Yoshihara, S., and Ikeuchi, M. 2001, Analysis of eukaryotic-type protein kinase, SpkB in the *Synechocystis* sp. PCC 6803, *J. Plant Res. suppl.*, **114**, 72.
15. Yoshihara, S., Suzuki, F., Fujita, H., Geng, X. X., and Ikeuchi, M. 2000, Novel putative photoreceptor and regulatory genes required for the positive phototactic movement of the unicellular motile cyanobacterium *Synechocystis* sp. PCC 6803, *Plant Cell Physiol.*, **41**, 1299–1304.
16. Hihara, Y. and Ikeuchi, M. 1997, Mutation in a novel gene required for photomixotrophic growth leads to enhanced photoautotrophic growth of *Synechocystis* sp. PCC 6803, *Photosynth. Res.*, **53**, 129–139.
17. Ikeuchi, M. and Tabata, S. 2001, *Synechocystis* sp. PCC 6803—a useful tool in the study of the genetics of cyanobacteria, *Photosynth. Res.*, **70**, 73–83.
18. Kyte, J. and Doolittle, R. F. 1982, A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.*, **157**, 105–132.
19. Udo, H., Inouye, M., and Inouye, S. 1997, Biochemical characterization of Pkn2, a protein Ser/Thr kinase from *Myxococcus xanthus*, a Gram-negative developmental bacterium, *FEBS Lett.*, **400**, 188–192.
20. Peirs, P., De Wit, L., Braibant, M., Huygen, K., and Content, J. 1997, A serine/threonine protein kinase from *Mycobacterium tuberculosis*, *Eur. J. Biochem.*, **244**, 604–612.
21. Zhang, C. C., Gonzalez, L., and Phalip, V. 1998, Survey, analysis and genetic organization of genes encoding eukaryotic-like signaling proteins on a cyanobacterial genome, *Nucl. Acids Res.*, **26**, 3619–3625.
22. Okamoto, S., Ikeuchi, M., and Ohmori, M. 1999, Experimental analysis of recently transposed insertion sequences in the cyanobacterium *Synechocystis* sp. PCC 6803, *DNA Res.*, **6**, 265–273.
23. Kamei, A., Hihara, Y., Yoshihara, S., Geng, X. X., Kanehisa, M., and Ikeuchi, M. 2001, In: PS2001 Proceedings: 12th International Congress on Photosynthesis, CSIRO Publishing, Melbourne, pp. S41–013.