ELSEVIER

Biochimica et Biophysica Acta 1237 (1995) 189–192

BBA Biochimica et Biophysica Acta

Short Sequence-Paper

A *Synechococcus* gene encoding a putative pore-forming intrinsic membrane protein ☆

Seiji Kashiwagi, Kengo Kanamaru [1], Takeshi Mizuno *

Laboratory of Molecular Microbiology, School of Agriculture, Nagoya University, Chikusa-ku, Nagoya 464, Japan

Received 1 February 1995; revised 12 April 1995; accepted 20 April 1995

Abstract

A cyanobacterium, *Synechococcus* species PCC7942, has a gene encoding a copper-transporting P-type ATPase, which is located in the thylakoid membrane. At the 5′-upstream of this ATPase gene, we identified another gene, which was supposed to be implicated in a copper-transport process. This novel gene was found to encode a putative pore-forming membrane protein that belongs to a growing family of homologous intrinsic membrane proteins (the MIP family of proteins), which include the major intrinsic protein (MIP) from animal lens fibre junction membranes, the tonoplast intrinsic protein (TIP) from vacuolar membranes of higher plants, and the *Escherichia coli* glycerol facilitator (GlpF) in the cytoplasmic membrane. The deduced product, named SmpX (*Synechococcus* membrane protein), is highly homologous throughout its entire sequence to these intrinsic membrane proteins which were postulated to be pore-forming proteins involved in a variety of transport processes. The primary amino acid sequence of SmpX shares all properties characteristic for members of the MIP family. SmpX is more similar to the eukaryotic members (e.g., nodulin-26 from soybean) than to the prokaryotic ones.

Keywords: ATPase, P-type; Major intrinsic protein (MIP); Protein; DNA sequence; Cyanobacterium; (*Synechococcus*)

Synechococcus is a Gram-negative bacterium, but it harbors a photosynthetic apparatus (thylakoid) similar in structure and function to that located in the chloroplasts of phototrophic higher plants. Thus, this bacterium is an organism of choice to study, at the molecular level, the fundamental processes involved in the oxygen-evolving photosynthesis. As such an approach, we have recently been studying the ion-transporting processes in response to environmental stimuli in this particular microorganism [1–5], and have cloned two distinct genes each encoding an ion-transporting P-type ATPase, named *pacS* and *pacL* [2,5]. The *pacS* gene is particularly intriguing in the sense that it encodes a unique metal-transporting ATPase. Furthermore, PacS is a thylakoid membrane-located copper-transporting ATPase, which is suggested to be involved in intracellular copper-homeostasis in the photosynthetic cyanobacterium [5].

The structure of a *Synechococcus* choromosomal region (a 5.1-kb *Bgl*II–*Eco*RI fragment) is shown in Fig. 1A, in which the *pacS* gene is located. The coding sequence for *pacS* consists of 2403 nucleotides, extending between the *Bam*HI and *Pvu*II sites [2]. We had previously constructed an insertional inactivation mutant of *pacS*, in which the *pacS* gene was replaced by a kanamycin-resistance gene on the chromosome [2] (see Fig. 1). When the resultant *pacS*-deletion strain, named DEL-S-I, allowed to grow on a conventional BG11-plates supplemented with varied concentrations of copper ($CuSO_4$), the colony-forming-efficiency (i.e., cell-viability) was remarkably reduced in proportion to the copper-concentration added, as compared with in the case of its parental strain (wild-type) (Fig. 1B). It was thus assumed that PacS-ATPase is involved in a copper-transporting system in *Synechococcus*, so a defect in this particular copper-transporting system results in copper-hypersensitivity as to growth [5]. Nevertheless, we constructed in this study another insertional inactivation mutant, in which the upstream border was further extended near to the *Vsp*I site, as shown in Fig. 1A (named DEL-S-

☆ The sequence data reported in this paper will appear in the DDBJ, EMBL and NCBI nucleotide sequence databases under the accession number D43774.

* Corresponding author. Fax: +81 52 7894091.

[1] Present address: National Institute of Genetics, Yata 1, 111, Mishima, Shizuoka 411, Japan.

SSDI 0005-2736(95)00124-7

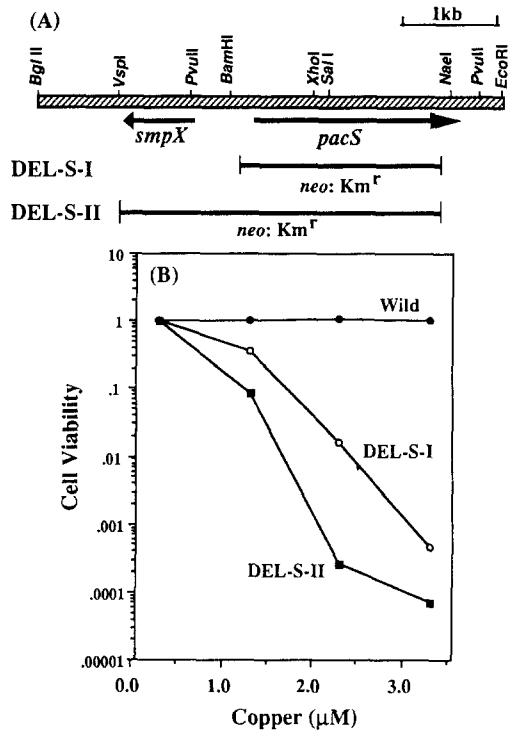

Fig. 1. A schematic representation of insertional inactivation mutants of the *pacS* gene. Two such deletions (named DEL-S-I and DEL-S-II, respectively) were constructed by replacing the *Synechococcus* chromosomal regions, indicated by horizontal bars, with a kanamycin-resistance gene (*neo*, Km^r) (panel A). These deletion mutants as well as their parental strain (wild-type) were grown in a standard liquid medium (BG11) [1], and then portions of them were spread on solid BG11 medium containing the indicated concentrations of $CuSO_4$. Note that BG11 medium contains a basal level of $CuSO_4$ (0.3 μM). The plates were incubated at 32° C for 96 h, and then numbers of colonies were counted (panel B).

II). The reason we constructed this new deletion mutant was that a set of functionally related genes is often clustered in prokaryotic chromosomes. When DEL-S-II was assessed in terms of copper-hypersensitivity as to growth, this deletion mutant was found to be even more (more or less ten times) sensitive to copper added in plates, as compared with in the case of DEL-S-I (Fig. 1B). We thus supposed that the about 1.5-kb region upstream of the *pacS* gene may contain another gene(s), which might be implicated in the presumed copper-homeostasis. Here we determined the nucleotide sequence for the 2.1-kb region extending upstream from the *pacS* gene to the *Bgl*II site (see Fig. 1).

The *Bgl*II–*Eco*RI *Synechococcus* chromosomal segment was previously cloned onto an *Escherichia coli* plasmid (pUC119), and the nucleotide sequence of the 3.0-kb *Bam*HI–*Eco*RI region encompassing the entire *pacS* gene was reported (see Fig. 1A) [2]. We subcloned the *Bgl*II–*Bam*HI segment which contains the upstream sequence of *pacS*. A set of segments, deleted successively from both the ends, was constructed from the plasmid. These plasmid-carried segments were subjected to sequencing with the conventional dideoxy chain termination method. The nucleotide sequence for the *Bgl*II–*Bam*HI region (2100 bp) was determined, and the 1560-bp sequence extending upstream from the coding sequence of *pacS* was shown in Fig. 2. Analyses of this sequence revealed an open reading frame (ORF) consisting of 269 amino acids, which could encode a protein with a calculated molecular mass of 30098. This putative ORF has a head-to-head orientation relative to *pacS*-ORF (Fig. 2). An about 600-bp sequence, residing between these ORFs, is most likely a non-coding regulatory sequence, in which

```
GTTTGCTGATTGACCATGACTGGACTCCAGACTCGAACGTAGCTCTAATCGTCCCCAGTCTTAAGTCTCTAGTTGACTGCAGAGTCAAGACGAGTGACAAAACTTCTCTTAAAAGCGCGA 120
 T  Q  Q  N  V  M <- PacS

TCGCCCTGCCCCAGCCCAATTGGGCCTCTTTCGACCTGCTCTACTGCTTGCTCCCAGACGTAGCCAAGCTGAGGAGAGAACGATCGGAGTCAAGCACATGTTAAGGCTCAAGTGCTTTGC 240
CCGCTTGAGCGACGGCCCCTCGTCAATGCCTCTTTCGTTTTGTCGAACCGCATTGGATCCACTAGCAACTATGAGCTTGAGCAGGACGTTTTAAGCATGTTGAGCGCTTCTATTGCCCCA 360
GATTCAGCGCTCAGGCCTTCGTTACGCTAGTGGCATTGAAGTCAGTTAGTTGAATGAGCTGTAATCAAGGTGCGATCGCCTTTCCTTTGGTTATTGGCTCACAACCTGAAGGGTAAACAT 480
TAGGCTGACAGTCCGTGCTGGCTTGAGTCAATCAGCCTGATCCGGTCAGTTGATCCCAGCCATCGCCTATCTCTGACTCGACTGGGATCATCATTCGTTCCTGCCTCTGCGCAACGACAA 600
CCCACTCCATTCCTCTGCGATCCAGTCTCTATGAAGATGCTTAGAGCACTTAAACACCACTGGCCAGAATACTTAATTGAAGCTTGGGGGCTTGGGCTGTTTATGGTGGCAGCTGGGGTT 720
                  SmpX -> M  K  M  L  R  A  L  K  H  H  W  P  E  Y  L  I  E  A  W  G  L  G  L  F  M  V  A  A  G  V

GTCGGAACGCTTGTCTTCTATCCCCAATCTCCTGCCTATCAAGCGATCGCAGACCCTTTCCTACAGCGAGTCGTGATGGGCTTGGGGATGGGTTTGACGGCAATGATCATCATGTATTCA 840
V  G  T  L  V  F  Y  P  Q  S  P  A  Y  Q  A  I  A  D  P  F  L  Q  R  V  V  M  G  L  G  M  G  L  T  A  M  I  I  M  Y  S
                                                                                                50
CCTTGGGGAAAGCGATCGGGGGCTCATATCAACCCAGCCGTAACGCTCACGTTCTACCGGCTTAAAAAAATTGCTGCCTGGGATGCTTTTTTCTACGTTGTTTTTCAGTTTATTGGTGGA 960
P  W  G  K  R  S  G  A  H  I  N  P  A  V  T  L  T  F  Y  R  L  K  K  I  A  A  W  D  A  F  F  Y  V  V  F  Q  F  I  G  G
                                                                                                100
CTTCTAGGAGTCGTGCTGGTTGCCTTTTTACTACAAACTCCCTTTACCCAAGCACCGGTCAACTATGTAGTCACGGTCCCTGGAAAACAGGGGGCGATCGTTGCATGTATCGCAGAGTAT 1080
L  L  G  V  V  L  V  A  F  L  L  Q  T  P  F  T  Q  A  P- V  N  Y  V  V  T  V  P  G  K  Q  G  A  I  V  A  C  I  A  E  Y
                                                                                                                        150
TTCATTGCTGTACTCATGATGAGTATGGTTCTCTTCACCAGCAATCAGCCAAAGCTAGAGCGATTCACTCCATTTTTTGCGGGCTGTCTCATTGTCAGCTACGTCATTTTTGAATCGCCG 1200
F  I  A  V  L  M  M  S  M  V  L  F  T  S  N  Q  P  K  L  E  R  F  T  P  F  F  A  G  C  L  I  V  S  Y  V  I  F  E  S  P

CTCTCAGGGTTTGGCATGAACCCTGCTCGAACCGTTGCCTCAGCTTTACCCTCGGGTATTTGGACAGCTATCTGGCTTTATTTCCTTGCACCGATCGCAGGCATGTTAACTGCAGCAGAA 1320
L  S  G  F  G  M  N  P  A  R  T  V  A  S  A  L  P  S  G  I  W  T  A  I  W  L  Y  F  L  A  P  I  A  G  M  L  T  A  A  E
                                200
CTCTATCTGCGAATGATCGGGCCTCGCAAAATCTTCTGTGCCAAGCTCTACCATGATCCTTTGTATCGCTGTATACACTGCGGCCATCTCATTCATTGGCATCGTCCTCATCTGAGATAG 1440
L  Y  L  R  M  I  G  P  R  K  I  F  C  A  K  L  Y  H  D  P  L  Y  R  C  I  H  C  G  H  L  I  H  W  H  R  P  H  L  R  *
                                                                                                250
GTCAAAAATATGGCTGCCAGACAGATTGGATGCGATCGCAGCCTAACAAAAGGCCTAGTTCACCTTATTAATAGCAATAATTGTGAATTTTTCTTATCCTCTAGGAGATAAGAAACTTTA 1560
```

Fig. 2. The nucleotide sequence of the *smpX* gene and the deduced amino acid sequence of SmpX from *Synechococcus* sp. PCC7942.

each promoter region for the ORFs should be located. In fact, we previously reported that a major transcript of *pacS* is considered to start from nucleotide [G], which is located at 33-bp upstream from the ATG translation initiation codon of *pacS*. The putative gene specifying the newly identified ORF was tentatively named *smpX*. A Northern-blot analysis was carried out by using appropriate DNA probes, which are specific for *smpX* and *pacS*, respectively. *pacS*-mRNA was revealed to remarkably accumulate upon addition of copper to the growth medium, as reported previously [5]. However, a very low level of *smpX*-mRNA was detected regardless of the growth conditions tested. In any case, *smpX* should be the gene which appears to be absent in DEL-S-II on its chromosome, but DEL-S-I does contain, as shown in Fig. 1A. Therefore, the *smpX* gene was supposed to be implicated in the copper hypersensitivity observed in Fig. 1B.

A computer-aided search for sequences homologous to the amino acid sequence of SmpX was conducted by using the FASTA program (provided by the DDBJ e-mail server, version 1.1). The search revealed a striking feature as to the SmpX sequence, namely, this protein is highly homologous to a large and growing family of homologous intrinsic membrane proteins (often called the MIP family of proteins). Members of this family were found in quite diverse organisms including animals, plants, yeast and bacteria (for reviews, see [6–9]), and they include the major intrinsic protein (MIP) from animal lens fibre junction membrane [10], the tonoplast intrinsic protein (TIP) from plant vacuolar membranes [11], the plant-encoded nodulin-26 (NOD) from peribacteroid membrane of root nodules [12], the *E. coli* glycerol facilitator (GlpF) in the cytoplasmic membrane [13]. Members of the MIP family are similar in size (250–280 amino acids), and show similar hydropathy profiles that reveal six putative membrane-spanning domains (for a review, see [7]). These proteins all share sequence identity throughout their entire regions containing a number of highly conserved amino acids. It was also proposed that the first three of the six transmembrane segments are significantly similar to the second three of the six transmenbrane segments [6]. It was revealed that SmpX shares all these properties with members of the MIP family. Such an example of sequence-alignments was shown for between SmpX and NOD (from soybean) in Fig. 3. SmpX (269 amino acids) and NOD (271 amino acids) exhibit a 24.5% identity in their overlapping 212 amino acid regions. Thus, we concluded that SmpX from *Synechococcus*, found in this study, is a member of the MIP family, although it should be noted that only a few instances of the MIP family were reported for bacteria (e.g., *E. coli* GlpF) [7,13]. A phylogenetic tree depicting the relatedness of the various MIP family proteins has been proposed [7]. According to the tree, the eight plant proteins and the five animal proteins form a cluster, whereas the yeast and bacterial ones appear to be less related. In this context, it is worth mentioning that our brief inspection, using the FASTA program, revealed that SmpX is significantly more similar to eukaryotic homologues, particularly to NOD (see Fig. 3), than to prokaryotic ones.

The function of many MIP family proteins are still unknown, although most of them have been postulated to be involved in transport processes. For examples, plant TIPs and animal ChIP have been shown to be aquaporins [14–16], animal MIPs were suggested to transport ions (possibly Na^+) [17], bacterial GlpFs function as a diffusion facilitator for glycerol [13], and NOD may allow an exchange of metabolites between the bacteroids and cytoplasm of root nodules [18]. Therefore, it can be assumed that SmpX is also a pore-forming protein, although verification of this must await further experimentation. It should be recalled here, however, that SmpX was originally impli-

```
Synechococcus;SmpX:  19 WGLGLFMVAAGVVGTLVFYPQSPAYQAIADPFLQRV  55
                                .**..* . ..... ...  .... . ...
Soybean;Nodulin-26:  34 VPFLQKLVAEAVGTYFLIFAGCASL-VVNENYYNMI  68

VM-GLGM--GLTAMIIMYSPWGKRSGAHINPAVTLTFYRLKKIAAWDAFFYVVFQFIGGL 111
.. *...  **.  ...*.  *. **.:.*****..* . ....  ..  *** *..*..
TFPGIAIVWGLVLTVLVYTV-GHISGGHFNPAVTIAFASTRRFPLIQVPAYVVAQLLGSI 127

LGVVLVAFLLQTPFTQAPVNYVVTVPGKQGAIVACIAEYFIAVLMMSMV-LFTSNQPKLE 170
*.   . .*    *  .. ..  *** . ... * . *.... ..* ..   ..... ..
LASGTLRLL----FMGNHDQFSGTVP-NGTNLQAFVFEFIMTFFLMFVICGVATDNRAVG 182

RFTPFFAGCLIVSYVIFESPLSGFGMNPARTVASALPSGIVTAIWLYFLAPIAGMLTAAE 230
 *. .  *. ..  **...*..* .***.*....*.. *  ..**.*.***..* ...*
EFAGIAIGSTLLLNVIIGGPVTGASMNPVRSLGPAFVHGEYEGIWIYLLAPVVGAIAGAW 242

LYLRMIGPRKIFCAKLYHDPLYRCIHCGHLIHWHRPHLR 269
.*   .   *
VYNIVRYTDKPLSETTKSASFLKGRAASK 271
```

Fig. 3. SmpX appears to be a member of the MIP family. The amino acid sequence of SmpX of *Synechococcus* (upper) was aligned with a representative of the MIP family, nodulin-26 (NOD) of soybean (lower). Identical and similar amino acids were highlighted by asterisks and dots, respectively. Analyses were done with the FASTA program [19].

cated in copper-hypersensitivity as to growth (see Fig. 1). It is thus tempting to speculate that SmpX might be involved in ion-transport, particularly in metal-transport. We are currently examining this interesting possibility.

In short, the gene-product, SmpX, appears to be a unique example of prokaryotic members of the large MIP family. This cyanobacterial gene may encode a putative pore-forming membrane protein, which is the first example of the MIP family proteins in cyanobateria. Since *Synechococcus* PCC7942 can be easily manipulated genetically, extensive genetic analyses of the *smpX* gene should shed light on not only the physiological function of its gene-product but also the general issue as to the structure and function of the MIP family proteins.

This study was supported by a Grant-in-Aid for Scientific Research on a Priority Area (No. 04273013) from the Ministry of Education, Science and Culture of Japan.

References

[1] Aiba, H., Nagaya, H. and Mizuno, T. (1993) Mol. Microbiol. 8, 81–91.
[2] Kanamaru, K., Kashiwagi, S. and Mizuno, T. (1993) FEBS Lett. 330, 99–104.
[3] Nagaya, M., Aiba, H. and Mizuno, T. (1994) J. Bacteriol. 176, 2210–2215.
[4] Aiba, H. and Mizuno, T. (1994) Mol. Microbiol. 13, 25–34.
[5] Kanamaru, K., Kashiwagi, S. and Mizuno, T. (1994) Mol. Microbiol. 13, 369–377.
[6] Pao, G.M., Wu, L.-F., Johnson, K.D., Hofte, H., Chrispeels, M.J., Sweet, G., Sandal, N.N. and Saier, M.H., Jr. (1991) Mol. Microbiol. 5, 33–37.
[7] Reizer, J., Reizer, A. and Saier, M.H., Jr. (1993) Crit. Rev. Biochem. Mol. Biol. 28, 235–257.
[8] Chrispeels, M.J. and Maurel, C. (1994) Plant Physiol. 105, 9–13.
[9] Chrispeels, M.J. and Agre, P. (1994) Trends Biochem. Sci. 19, 421–425.
[10] Gorin, M.B., Yancey, S.B., Cline, J., Revel, J.-P. and Horwitz, J. (1984) Cell 39, 49–59.
[11] Johnson, K.D., Hofte, H. and Chrispeels, M.J. (1990) Plant Cell 2, 525–532.
[12] Sandal, N.N. and Marcker, K.A. (1988) Nucl. Acids Res. 16, 9347.
[13] Muramatsu, S. and Mizuno, T. (1989) Nucl. Acids. Res. 17, 4378.
[14] Preston, G.M., Carroll, T.P., Guggino, W.B. and Agre, P. (1992) Science 256, 385–387.
[15] Fushimi, K., Uchida, S., Hara, Y., Hirata, Y., Marumo, F. and Sasaki, S. (1993) Nature 361, 549–552.
[16] Maurel, C. Reizer, J., Schroeder, J.I. and Chrispeels, M.J. (1993) EMBO J. 12, 2241–2247.
[17] Nikaido, H. and Rosenberg, E.Y. (1985) J. Membr. Biol. 85, 87–92.
[18] Ouyang, L., Whelan, J., Weaver, C.D., Roberts, D.M. and Day, D.A. (1991) FEBS Lett. 293, 188–190.
[19] Lipman, D.J. and Pearson, W.R. (1985) Science 227, 1435–1441.