Plant Gene Register

Sequence of a psaC Gene from the Cyanobacterium Synechococcus sp. PCC 6301

Patricia L. Herman*, Kartika Adiwilaga2, John H. Golbeck, and Donald P. Weeks
Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, Nebraska 68583-0718

The psaC gene encodes Psac, the apoprotein for the terminal iron-sulfur clusters, FA and FB, in the PSI reaction center of cyanobacteria, algae, and higher plants. Psac functions as a membrane-bound oxidoreductase, accepting electrons from the F, iron-sulfur cluster located on the PsaA/PsaB heterodimer and donating them to the soluble electron transfer proteins Fd and flavodoxin (reviewed in Bryant, 1992). The objective of our work is to clarify the role of Psac in linear and cyclic electron transfer. The experimental organism is Synechococcus sp. PCC 6301, a unicellular, freshwater cyanobacterium that is used extensively for physiological, biochemical, and spectroscopic studies of photosynthesis. The PSI complex of Synechococcus sp. PCC 6301 is widely used in structural and functional studies of the FA and FB iron-sulfur clusters (Parrett et al., 1990).

We became interested in the psaC sequence of Synechococcus sp. PCC 6301 because there are subtle differences in the electron paramagnetic resonance spectra of the FA and FB iron-sulfur clusters between cyanobacteria and higher plants. This suggests that there may be a correlation with structure: in all organisms studied thus far, the higher-plant Psac protein differs from the cyanobacterial Psac protein in two regions: at position 37, higher plants contain a Lys whereas cyanobacteria contain a neutral amino acid (Gly or Ala), and at positions 70 and 71, dicots contain a Trp-His pair, monocots contain a Gly-Pro pair, and cyanobacteria contain a Gly-Ala pair. No sequence for the Psac protein of Synechococcus sp. PCC 6301 is available, even though entire sequences are available for PsacD and PsacE and partial sequences are available for Psal and Psaf. This information is relevant because it has been reported that two psaC genes exist in the cyanobacterium Synechocystis sp. PCC 6803, one with a higher-plant sequence (Anderson and McIntosh, 1991) and the other with a cyanobacterial sequence (Steinmüller, 1992). We were interested in determining whether the psaC gene of Synechococcus sp. PCC 6301 was higher-plant-like or cyanobacterial-like and whether the gene was flanked by the mitochondrial ndh genes that encode subunits of the NADH dehydrogenase complex (Schantz and Bogorad, 1988).

* Corresponding author; fax 1-402-472-7842.

1 This research was supported by a grant from the National Science Foundation (MCB-9205756) to J.H.G. This is Journal Series No. 10540 of the Agricultural Research Division, University of Nebraska.2 Present address: Institute of Molecular and Cell Biology, National University of Singapore, Singapore 0511.

We report here the isolation and nucleotide sequence of a psaC gene from Synechococcus sp. PCC 6301 (Table 1). Oligonucleotide primers based on conserved sequences in the 5' and 3' coding regions of previously sequenced cyanobacteria psaC genes were used to generate a fragment from Synechococcus sp. PCC 6301 genomic DNA by PCR amplification. The PCR product was cloned into pBluescript II KS+ (Stratagene), sequenced to confirm its identity, and labeled with digoxigenin by the random-prime method (Genius System, Boehringer). The labeled probe was then used in Southern hybridization experiments (using chemiluminescent detection) to map the position of a number of restriction sites relative to an AccI site located 50 bp from the 3' end of the psaC coding region. Southern analysis showed that the PCR product hybridized to a single 4.2-kb XhoI/HindIII genomic fragment. A size-fractionated genomic library consisting of XhoI/HindIII fragments from 3.5 to 5 kb in length was constructed in pBluescript II KS+ and transformed into DH5α cells. Ninety-seven bacterial colonies were screened with the digoxigenin-labeled PCR product, and two putative positive colonies were identified. Restriction enzyme analysis confirmed that both colonies contained clones with the expected genomic fragment. When several restriction enzyme sites in the 4.2-kb clone were mapped, the putative psaC gene was localized to a 327-bp Eagl/BamHI fragment. This fragment was subcloned into pBluescript II KS+, and both strands were sequenced. To sequence the regions flanking the psaC gene, a set of unidirectional nested deletions was generated for each strand of the 4.2-kb genomic clone using the Erase-a-Base system (Promega).

The nucleotide sequence of the 246-bp ORF that encodes the psaC gene of Synechococcus sp. PCC 6301 is highly homologous to previously reported psaC genes from six other cyanobacteria. A sequence comparison using the FastA program of the Genetics Computer Group (1991) showed the following results: 82.1% identity to Synechococcus sp. PCC 7002 (Rhiel et al., 1992), 81.2% identity to Synechococcus vulcanus (Shimizu et al., 1990), 80.8% identity to a Synechococcus species (GenBank, X63767), 80.5% identity to an Anabaena species (GenBank, X57153), 80.1% identity to Anabaena sp. PCC 7120 (Mulligan and Jackman, 1992), and 80.1% identity to Synechocystis sp. PCC 6803 (Steinmüller, 1992). The identity of the psaC coding sequence from Synechococcus sp. PCC 6301 to previously sequenced psaC genes

Abbreviation: ORF, open reading frame.
from 12 other photosynthetic organisms is approximately 10% lower, ranging from 69.1% for the liverwort Marchantia polymorpha to 73.6% for maize. Interestingly, there is only 72% identity to another psaC sequence from the cyanobacterium Synechocystis sp. PCC 6803 (Anderson and McIntosh, 1991). This result supports the conclusion of Steinmüller (1992), who suggested that the psaC gene isolated by Anderson and McIntosh (1991) was the result of a rare transformation event by chloroplast DNA from another plant species.

The predicted sequence of the PsacC polypeptide encoded by the psaC gene from Synechococcus sp. PCC 6301 consists of 81 codons and has a high degree of sequence similarity to the previously reported PsacC proteins from six other cyanobacteria. The Gly residue at position 37 and the Gly-Ala pair at positions 70 and 71 are indicative of a cyanobacterial PsacC protein. A sequence comparison using the TFastA program of the Genetics Computer Group (1991) showed the following results: 98.8% identity to Synechococcus sp. PCC 7002 (Rhiel et al., 1992), 97.5% identity to Synechocystis sp. PCC 6803 (Steinmüller, 1992), 95.1% identity to S. vulcanus (Shimizu et al., 1990) and another Synechococcus species (GenBank, X63767), and 93.8% identity to Anabaena sp. PCC 7120 (Mulligan and Jackman, 1992) and to another Anabaena species (GenBank, X57153). As expected, the identity of the deduced PsacC sequence from Synechococcus sp. PCC 6301 to the PsacC proteins from 13 other photosynthetic organisms is lower, ranging from 85.2% for the liverwort M. polymorpha to 88.9% for tobacco. The only difference between the predicted PsacC sequences from Synechococcus sp. PCC 7002 and Synechococcus sp. PCC 6301 is at position 41, where a Ser residue in the Synechococcus sp. PCC 7002 sequence is replaced by an Ala residue in the Synechococcus sp. PCC 6301 sequence. The Ala residue at position 41 is unique because all of the PsacC proteins that have been sequenced have a Ser residue at this position.

In the chloroplast genome of a number of plants, the psaC gene is located between sequences with homology to two mitochondrial ndh genes that encode subunits of the NADH dehydrogenase complex (Schantz and Bogorad, 1988). The first homolog, ndhE, is upstream from the psaC gene. The second homolog, ndhD, is downstream and actually co-transcribed with the psaC gene from maize (Schantz and Bogorad, 1988). Anderson and McIntosh (1991) reported that this gene arrangement is conserved in the cyanobacterium Synechocystis sp. PCC 6803. The Compare and DotPlot programs of the Genetics Computer Group (1991) were used to compare the nucleotide sequence of the region surrounding the psaC gene of Synechococcus sp. PCC 6301 to the published nucleotide sequence of the region flanking 19 psaC genes from both cyanobacteria and other photosynthetic organisms. This
analysis revealed no significant similarity between the sequence flanking the *Synechococcus* sp. PCC 6301 psaC gene and the sequence surrounding the previously reported psaC genes. In particular, the psaC gene of *Synechococcus* sp. PCC 6301 is not flanked by sequences homologous to the ndhE or ndhD genes. This result agrees with that of Rhiel et al. (1992), who found no evidence of ndhE or ndhD sequences immediately adjacent to the coding region of the psaC gene from the cyanobacterium *Synechocystis* sp. PCC 7002. Ellersiek and Steinmüller (1992) reported a sequence with homology to ndhD located 368 bp upstream from another psaC gene isolated from the cyanobacterium *Synechocystis* sp. PCC 6803. However, this ndhD gene is transcribed on the opposite DNA strand. It can be concluded that most of the available evidence suggests that cyanobacteria do not have the same gene arrangement as plant chloroplasts in the region of the genome that contains the psaC gene.

A sequence analysis using the FastA program of the Genetics Computer Group (1991) detected five ORFs not previously reported for cyanobacteria in the 4.2-kb clone containing the psaC gene of *Synechococcus* sp. PCC 6301 (P.L. Herman and D.P. Weeks, unpublished data). There is a cluster of four consecutive ORFs starting 459 bp upstream of the psaC gene encoded on the opposite strand that is highly homologous to previously sequenced genes that function in polyketide biosynthesis. In addition, there is an ORF that begins 111 bp downstream from the psaC coding region on the same strand. This ORF has a high degree of sequence similarity to *nmdM*, a gene found in the nitrogen-fixing *Rhizobium* bacterium that is involved in the production of a host-specific nodulation signal.

**ACKNOWLEDGMENTS**

The authors gratefully acknowledge the DNA Sequencing Facility supported by the Center for Biotechnology at the University of Nebraska-Lincoln for their help in sequencing the clones reported in this study. The authors also thank Drs. Sylvia Darr and Robert Speirizter (Department of Biochemistry, University of Nebraska-Lincoln) for their critical review of the manuscript and Dr. Don Bryant (Pennsylvania State University) for advice on designing the PCR primers.

Received November 9, 1993; accepted November 22, 1993.

Copyright Clearance Center: 0032-0889/94/104/1459/03.

The GenBank accession number for the sequence reported in this article is U01536.

**LITERATURE CITED**


