ELSEVIER

Biochimica et Biophysica Acta 1188 (1994) 439–442

BBA Biochimica et Biophysica Acta

Short Sequence-Paper

Nucleotide sequence of the *psbE*, *psbF* and *trnM* genes from the chloroplast genome of *Chlamydomonas reinhardtii*

Sedigheh Alizadeh, Rachel Nechushtai [1], James Barber, Peter Nixon *

Photosynthesis Research Group, Wolfson Laboratories, Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7 2AY, UK

Received 2 August 1994

Abstract

We have determined the nucleotide sequences of the *psbE*and *psbF* genes, which encode the α and β subunits, respectively, of cytochrome *b*-559, from the chloroplast genome of the green alga *Chlamydomonas reinhardtii*. In contrast to other organisms *psbE* is not co-transcribed with *psbF*. The primary structures of the gene products are very similar to the equivalent proteins in cyanobacteria and plants. Each subunit contains a single histidine residue that is thought to ligate haem. Upstream of the *psbE* gene, a *trnM* gene is located which encodes an elongator $\mathrm{tRNA^{Met}}$ molecule.

Keywords: Photosystem II; Cytochrome; *psbE*; *psbF*; *trnM*; Green alga; Chloroplast genome; (*Chlamydomonas reinhardtii*)

The *psbE* and *psbF* genes encode the α and β subunits, respectively, of cytochrome *b*-559, a transmembrane component of the Photosystem II reaction centre which performs the light-induced oxidation of water to molecular oxygen (reviewed in [1]). The function of this cytochrome is still uncertain but recent evidence suggests that it probably plays a role in protecting the reaction centre from photoinhibitory damage (reviewed in [2]). Each of the subunits is thought to span the thylakoid membrane once [3]. However, the structure of the cytochrome is still a matter of debate with both αβ heterodimers and αα, ββ homodimers possibilities [2,4]. The construction of mutants in the transformable green alga *Chlamydomonas reinhardtii* may help us to understand the function of this cytochrome. Here we report the sequence of the *psbE* and *psbF* genes from the chloroplast genome of this organism.

Southern hybridisation experiments using a *psbE*-specific probe from tobacco allowed us to map the equivalent gene from *C. reinhardtii* onto the chloroplast genome. Fig. 1 shows the location of the *psbE* gene with respect to the *psbF* gene which has previously been mapped by Fong and Surzycki [5] to one end of the 15.1-kb *Pst*I-4 fragment of the *C. reinhardtii* plastome. In all plant and cyanobacterial species so far examined the *psbE* and *psbF* genes are tightly linked and are co-transcribed with the downstream *psbL* and *psbJ* genes [6]. Fig. 1 shows that this gene organisation is not conserved in *C. reinhardtii* as *psbE* lies approximately 7.3 kb away from *psbF* and is transcribed in the opposite direction. The *psbE*/*psbF* cluster has also recently been found to be disrupted in the related alga, *Chlamydomonas moewusii* [7].

Fig. 2A shows the nucleotide sequence of the 0.87 kb *Hin*dIII/*Xho*I DNA fragment present in plasmid pSA2 which was constructed by subcloning a 1.9-kb *Hin*dIII fragment from plasmid p78 [8] into a pBluescript vector (Stratagene, Cambridge, UK). Plasmid p78 which consists of the 15.1-kb *Pst*I-4 fragment from the chloroplast genome of *C. reinhardtii* (strain CC-125) cloned into pBR322 was kindly provided by Professor E.H. Harris of the *Chlamydomonas* Genetics Center, Department of Botany, Duke University, Durham, NC, USA. The *psbE* gene was identified by the high degree

---

The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers X80195 (*psbF*), X80196 (*psbE* and *trnM*).

* Corresponding author. Fax: + 44 71 2250960.

[1] Present address: The Hebrew University of Jerusalem, The Institute of Life Sciences, Department of Botany, Gnat Ram, 91904 Jerusalem, Israel.

0005-2728/94/$07.00
SSDI 0005-2728(94)00153-7

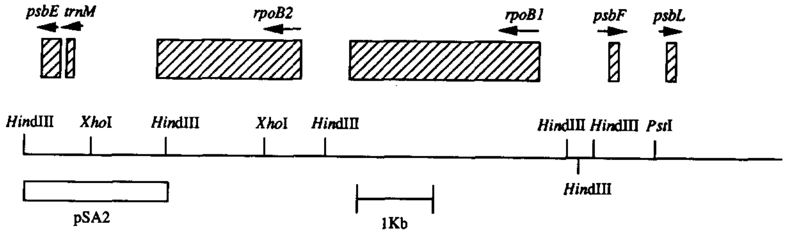

Fig. 1. Portion of the chloroplast genome of *Chlamydomonas reinhardtii* showing the location of the *psbE* and *trnM* genes. The location of the *psbF*, *psbL*, *rpoB1* and *rpoB2* genes within this region is taken from Fong and Surzycki [5,8]. Arrows indicate direction of transcription. The insert of plasmid pSA2 which was used to obtain the sequence in Fig. 2 is shown as the open bar.

A

```
 XhoI
CTCGAGGGTTTAACTCATTCTACTGTTATTATGGCGGGTTATATTTTAGTTCTTGGATATCTGGGAGTAAGAGGCTAGTATTAAAACAGT   90

AGTTGTAGTCAAAATTCTGGTAAATTGGTCTTTTTTATGCTCAAAAACTTAAAACTTGATAAAAAAAAAAATTTTATGTTATAATATTATA   180
    >trnM  ___________________________________________________________________________
TAATCTTTATAGCCTGCTTAGCTCAGTTGGTTAGAGCGTCCGTTTCATAAGCTGATTGTCACTAGTTCAAATCTAGTAGCAGGCATTAGT   270

ATTTTCAATAACATTAATTTATGGTATAATATTTAATATAATACATTGATTATAAAGCAGACAAATTGTTGAAAAAGCTAAGTAAGCACA   360
                   >psbE                                                             *
     SD          MetAlaGlyLysProValGluArgProPheSerAspIleLeuThrSerIleArgTyrTrpValIleHisSerIle
CTTAGGAGGTATTTTTATGGCTGGTAAACCAGTAGAGCGTCCTTTCTCAGATATTTTAACTAGTATTCGTTACTGGGTTATTCACAGTA   450

  ThrValProAlaLeuPheIleAlaGlyTrpLeuPheValSerThrGlyLeuAlaTyrAspValPheGlyThrProArgProAsnGluTyr
TTACAGTACCTGCTTTATTTATTGCAGGTTGGTTATTTGTAAGTACAGGTCTTGCTTATGACGTATTTGGTACTCCACGTCCAAATGAAT   540

  PheThrGluAspArgGlnGluAlaProLeuIleThrAspArgPheAsnAlaLeuGluGlnValLysLysLeuSerGlyAsnEnd
ATTTCACAGAAGATCGTCAAGAAGCTCCATTAATTACTGACCGTTTCAATGCTTTAGAACAAGTTAAAAAATTATCTGGTAACTAATTAA   630

TCTAAAAACTTTTTTTACTTAGAAAATTTTTTCTAAGTAAAAAAAGTACCTTTTCAAACCATCTATATAAATCTTGGCAAAGGCATTTAA   720
       -------------------> <------------------
TTTGGGTCTACCACACATAATGATTAAAAAAAAACGTAAATCCTAAAGAAAAAGTCCATTGATGTCAGTTAGCATAACTCACGTACAAGAA   810

ATAATAATATGATAGTAATATAAGATATATATAGGAGTTTCCATAAAGATAAAGAAAGCTT  871
                                                      HindIII
```

B

```
                >psbF                                                               *
                MetThrThrLysLysSerAlaGluValLeuValTyrProIlePheThrValArgTrpLeuAlaIleHisGlyIle
TTTATTTTTATTAATTATGACAACAAAAAAATCAGCTGAAGTACTTGTTTATCCTATTTTCACAGTTCGTTGGTTAGCTATCCACGGTAT   90

 AlaValProThrIlePhePheLeuGlyAlaIleThrAlaMetGlnPheIleGlnArgEnd
TGCAGTACCAACAATTTTCTTCTTAGGAGCTATTACTGCTATGCAATTCATTCAACGTTAAGTATTCGTCCCTAAAAGAACGTTAGTACG   180
```

Fig. 2. (A) Nucleotide and deduced amino acid sequences of the *psbE* gene from the chloroplast genome of *Chlamydomonas reinhardtii*. The 12 amino-acid N-terminal sequence determined by deVitry et al. [9] for the purified α subunit is underlined. The histidine residue that is thought to ligate the haem is indicated by an asterisk. A candidate Shine-Dalgarno sequence is denoted SD and underlined. Inverted repeats that may act as transcription terminators are arrowed. The sequence of the upstream *trnM* gene is overlined. The DNA sequence was determined for both DNA strands by the dideoxy method using plasmid pSA2 as the template and custom-made oligonucleotides. (B) Nucleotide and deduced amino acid sequence of the *psbF* gene from *Chlamydomonas reinhardtii*. The histidine residue, at position 23, which ligates the haem molecule is indicated by the asterisk. The DNA sequence of both strands was determined from plasmid p78 [8] using custom-made oligonucleotides and then confirmed by sequencing DNA amplified from the chloroplast genome by PCR.

psbE gene product

```
                                    10        20        30        40        50        60        70        80      %IDENTITY
C. reinhardtii              MAGKPVERPFSDILTSIRYWVIHSITVPALFIAGWLFVSTGLAYDVFGTPRPNEYFTEDRQEAPLITDRFNALEQVKKLSGN     100
Pea                          S STG  S A  I      I     I S                   S         T  GI    G  DS   LDEF RSF     72
Liverwort                    S NTG    A  I            I S                   S         N   V    G   S   IDEFTKSF     76
Euglena gracilis                STG       I          V I S  VG  I     IV  I      S     T  Q    S        MDQFTK       72
Synechocystis sp. PCC 6803   S TTG       V            I M                A      D    QT   L ILQE YDINQEIQEFNQ       65
```

psbF gene product

```
                                    10        20        30        40      %IDENTITY
C. reinhardtii              MTTKKSAEVLVYPIFTVRWLAIHGIAVPTIFFLGAITAMQFIQR      100
Pea                           -----IDRT          V  L    VS   S S              74
Liverwort                     -----IDRT          V  L    V      S              79
Euglena gracilis               N ---DTR     F    V AL I  V    S S              71
Synechocystis sp. PCC 6803   A QNPNQPVT          V TL   SV  V   A              64
```

Fig. 3. Comparison of the deduced amino acid sequences of the *psbE* and *psbF* gene products from *C. reinhardtii* with those from pea [13], liverwort [14], *Euglena gracilis* [15] and *Synechocystis* PCC 6803 [16]. For the latter four sequences only differences to the *C. reinhardtii* sequence are shown. Gaps are indicated by hyphens. Putative membrane-spanning regions are underlined.

of similarity to other published sequences (Fig. 3) and also by the perfect match between the predicted N-terminal sequence and the determined sequence of the purified α subunit from *C. reinhardtii* ([9]; Alizadeh et al., unpublished). The predicted product consists of 82 amino acids, one residue less than higher plant sequences, and has a molecular mass of 9297 Da before removal of the initiating methionine residue. The size of the mature protein (9166 Da) is slightly smaller than the α subunit from pea (9285 Da). In common with other organisms, the α subunit from *C. reinhardtii* contains a single histidine residue (at position 23) which is thought to ligate a haem molecule [10]. A potential ribosome-binding site was identified just upstream of the start codon (sequence AGGAGG, underlined in Fig. 2A) and a potential stem-loop structure that may serve as a transcription terminator is located downstream.

A *trnM* gene which probably encodes an elongator tRNA$^{\mathrm{Met}}$ was also found 186 bp upstream of the initiation codon for *psbE*. It is 74 bp long and shows an identity of 74% with *trnM* from liverwort chloroplast DNA [11] and 69% with the gene from *Euglena gracilis* [12]. As with other chloroplast sequences, the gene does not code for the 3′-CCA terminus of the tRNA [11].

Fig. 2B shows the nucleotide sequence and deduced amino acid sequence of the *psbF* gene which was obtained by sequencing plasmid p78 using the dideoxynucleotide chain-termination method. There are three differences to the previously published nucleotide sequence for this section of DNA [5]. Our sequence data contain an extra T between bases 6999 and 7000 (using the numbering scheme of Fong and Surzycki [5,8]), and lacks the two T's at positions 6925 and 7024. The sequence shown in Fig. 2B was further confirmed by sequencing DNA amplified from the chloroplast genome by the polymerase chain reaction. In contrast to the original sequencing studies of Fong and Surzycki [5] the predicted gene product now shows a greater degree of sequence identity to sequences from other sources (e.g., 74% identity with the pea amino-acid sequence, Fig. 3) and importantly now contains the histidine residue at position 23 that is thought to ligate the haem [4,10]. The predicted product consists of 44 amino acids and in comparison to higher plants contains 5 extra residues at the N-terminus (Fig. 3). The calculated molecular mass of the gene product is 4977 Da compared to 4424 Da for the pea subunit [13].

In a parallel study to ours, T. Mor, I. Ohad and H. Pakrasi have also sequenced the *psbE* and *psbF* genes from *C. reinhardtii* and have similarly found that His23 is conserved in the β subunit (I. Ohad, personal communication).

This work was financially supported by the BBSRC and the Royal Society. S.A. is a recipient of an SERC studentship, P.J.N. is a Royal Society University Research Fellow and R.N. was a Royal Society Visiting Research Professor at Imperial College.

References

[1] Cramer, W.A., Tae, G.-S., Furbacher, P.N. and Böttger, M. (1993) Physiol. Plant. 88, 705–711.

[2] Whitmarsh, J., Samson, G. and Poulson, M. (1994) in Photoinhibition of Photosynthesis (Baker, N.R. and Bowyer, J.R., eds.), pp. 75–93, βIOS Scientific Publishers Limited, Oxford, UK.

[3] Tae, G.-S., Black, M.T., Cramer, W.A., Vallon, O. and Bogorad, L. (1988) Biochemistry 27, 9075–9080.

[4] Herrmann, R.G., Alt, J., Schiller, B., Widger, W.R. and Cramer, W.A. (1984) FEBS Lett. 176, 239–244.

[5] Fong, S.E. and Surzycki, S.J. (1992) Curr. Genet. 21, 527–530.

[6] Erickson, J.M. and Rochaix, J.-D. (1992) in The Photosystems: Structure, Function and Molecular Biology, (Barber, J., ed.), pp. 101–177, Elsevier, Amsterdam.

[7] Boudreau, E., Otis, C. and Turmel, M. (1994) Plant Mol. Biol. 24, 585–602.

[8] Fong, S.E. and Surzycki, S.J. (1992) Curr. Genet. 21, 485–497.

[9] De Vitry, C., Diner, B.A. and Popot, J.-L. (1991) J. Biol. Chem. 266, 16614–16621.

[10] Babcock, G.T., Widger, W.R., Cramer, W.A., Oertling, W.A. and Metz, J.G. (1985) Biochemistry 24, 3638–3645.

[11] Ozeki, H., Ohyama, K., Inokuchi, H., Fukuzawa, H., Kohchi, T., Sano, T., Nakahigashi, K. and Umesono, K. (1987) Cold Spring Harb. Symp. Quant. Biol. 52, 791–804.

[12] Hollingsworth, M.J. and Hallick, R.B. (1982) J. Biol. Chem. 257, 12795–12799.

[13] Willey, D.L. and Gray, J.C. (1989) Curr. Genet. 15, 213–220.

[14] Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.I., Inokuchi, H. and Ozeki, H. (1986) Nature 322, 572–574.

[15] Cushman, J.C., Christopher, D.A., Little, M.C., Hallick, R.B. and Price, C.A. (1988) Curr. Genet. 13, 173–180.

[16] Pakrasi, H.B., Williams, J.G.K. and Arntzen, C.J. (1988) EMBO J. 7, 325–332.