Analysis of the genes of the OEE1 and OEE3 proteins of the photosystem II complex from *Chlamydomonas reinhardtii*

Mayfield, Stephen Patrick; Rahire, Michèle; Frank, Gerhard; Zuber, Herbert; Rochaix, Jean-David

How to cite

MAYFIELD, Stephen Patrick et al. Analysis of the genes of the OEE1 and OEE3 proteins of the photosystem II complex from *Chlamydomonas reinhardtii*. In: Plant Molecular Biology, 1989, vol. 12, n° 6, p. 683–693. doi: 10.1007/BF00044159

This publication URL: https://archive-ouverte.unige.ch/unige:148876

Publication DOI: 10.1007/BF00044159

Plant Molecular Biology **12**: 683–693, 1989.

Analysis of the genes of the OEE1 and OEE3 proteins of the photosystem II complex from *Chlamydomonas reinhardtii*

Stephen P. Mayfield,[1,3] Michèle Schirmer-Rahire,[1] Gerhard Frank,[2] Herbert Zuber[2] and Jean-David Rochaix[1] *

[1] *Departments of Molecular Biology and Plant Biology, University of Geneva, Geneva, Switzerland (*author for correspondence); [2] Eidgenössische Technische Hochschule, Hönggerberg, CH-8093 Zürich, Switzerland; [3] Present address: Department of Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Rd, La Jolla, CA 92037, USA*

Received 8 November 1988; accepted in revised form 1 March 1989

Key words: Chlamydomonas reinhardtii, chloroplast, oxygen evolving complex, photosystem II, transit peptides

Abstract

The sequences of the nuclear genes of the 33 kDa (OEE1) and the 16 kDa (OEE3) polypeptides of the oxygen evolving complex of *Chlamydomonas reinhardtii* have been established. Comparison between the OEE1 protein sequences of *C. reinhardtii* and higher plants and cyanobacteria reveals 67 and 47% homology. In contrast, *C. reinhardtii* and higher plants have only 28% overall homology for OEE3 which is mostly limited to the central portion of the protein. The transit peptides of the *C. reinhardtii* proteins consist of 52 (OEE1) and, most likely, 51 (OEE1) amino acids. They have a basic amino terminal region and, at least in the case of OEE1, a hydrophobic segment at their carboxy terminal end typical of thylakoid lumen proteins. Comparison of the genomic and cDNA clones indicates that the OEE1 and OEE3 genes contain five and four introns, respectively, some of which are located within the coding sequences of the transit peptides.

Introduction

The photolysis of water and subsequent evolution of oxygen during photosynthesis take place within the chloroplasts of higher plants and algae on membrane-bound photosystem II (PSII) particles. These particles are composed of a membrane embedded core, which includes the reaction center, an oxygen-evolving enhancer complex and a light-harvesting system. Isolated oxygen-evolving particles are complex assemblies containing chlorophylls a and b, carotenoids, quinones, lipids, iron, a catalytic manganese cluster and an array of at least nine polypeptides [1, 2]. PSII proteins can be grouped into two distinct classes. The first includes the core PSII proteins which are integral membrane proteins encoded by the chloroplast genome. The second includes the three extrinsic proteins of the oxygen-evolving complex called oxygen-evolving enhan-

* The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X13826 (OEE1) and X13832 (OEE3).

cer proteins 1, 2 and 3 (OEE1, 2 and 3) [1, 3]; these proteins are coded for by nuclear genes, synthesized as precursors on cytoplasmic ribosomes [4] and post-translationally transported across the two chloroplast envelope membranes and the thylakoid membrane into the lumen of the thylakoid where they bind to the core PSII particle. It has recently been shown that additional polypeptides are associated with PSII [5, 6, 7].

In vitro reconstitution experiments with isolated oxygen-evolving particles have shown that the PSII extrinsic proteins play important roles in photosynthetic oxygen evolution (reviewed in [1, 8]). These studies have suggested that two of the three extrinsic polypeptides (OEE2 or 24 kDa protein and OEE3 or 16 kDa protein) are involved in the binding of calcium and chloride ions which are required cofactors in oxygen evolution. Removal of the OEE2 and OEE3 proteins from the PSII particle by salt treatment results in a decrease in oxygen-evolving activity, but not in a reduction of manganese ion content. Removal of the OEE1 (or 33 kDa protein) results in both the loss of oxygen-evolving activity and a reduction of manganese ions bound to the thylakoid membrane although a high concentration of chloride can partially substitute for OEE1 [1]. Restoration of oxygen-evolving activity accompanies the rebinding of these extrinsic peptides to the PSII reaction center [1].

Using a molecular genetic approach we have shown that in *Chlamydomonas reinhardtii* the OEE1 and OEE2 proteins are required for high levels of photosynthetic oxygen evolution *in vivo* [3, 10]. The absence of the OEE2 protein does not affect the accumulation of the other PSII core or peripheral proteins [3]. The absence of the OEE1 protein greatly decreases the stability of the PSII core particle and only about 25% of the core PSII proteins accumulate in mutants deficient for the expression of the OEE1 polypeptide [10]. The other peripheral proteins, OEE2 and OEE3, accumulate in OEE1-deficient strains, although they are probably not anchored to the photosynthetic membrane.

As a first step in identifying domains within the OEE proteins which are involved in photosynthetic oxygen evolution and in stabilizing the PSII core complex we undertook the characterization of the OEE proteins and their genes in *C. reinhardtii*. Here we present the sequence of the OEE1 and OEE3 cDNAs, the deduced amino acid sequence of these proteins, and the characterization of the single psb1 (OEE1) and psb3 (OEE3) genes from *C. reinhardtii*. The mature OEE1 and OEE3 proteins from *C. reinhardtii* are compared with the homologous proteins from spinach [10, 12, 13] and cyanobacteria. We also present the transit peptide sequence of the OEE1 and OEE3 proteins and a comparison of these sequences with two other transit peptides from the OEE2 protein [3] and the small subunit of ribulose bisphosphate carboxylase (Rubisco) [14, 15] of *C. reinhardtii*.

Materials and methods

Isolation of cDNAs and genomic clones encoding OEE1 and OEE3

A *C. reinhardtii* DNA library constructed in λ-gt11 [3] was independently screened with antisera raised against mature *C. reinhardtii* OEE1 and OEE3 proteins as described by Young and Davis [16]. Several plaques containing phage expressing one or the other of the proteins were identified and purified to homogeneity. The largest of the three cDNAs identified for the OEE1 protein was nick-translated and used to identify additional cDNAs from a *C. reinhardtii* λ-gt10 library [15]. One of these cDNAs was of sufficient length (1.6 kb) to be a full-length clone. This cDNA was subcloned into the *Eco* RI site of plasmid pUC19 to form plasmid pPII-12P4.

Three phage were identified which expressed the OEE3 protein. The largest cDNA obtained (pPII-24.3) was 1.3 kb in size and was subcloned into the *Eco* RI site of plasmid pUC8. It was used to isolate a full-length cDNA clone (pPII-24.5) from the *C. reinhardtii* λ-gt10 library.

To isolate genomic clones encoding either the OEE1 or OEE3 genes, plasmids pPII-12P4 and pPII-24.3 were nick-translated and used as probes to identify genomic clones [17] from a

C. reinhardtii genomic library cloned into λ-EMBL3 [18]. Several overlapping clones were identified for each probe.

Sequencing of the OEE1 and OEE3 cDNAs, and determination of the primary amino acid sequence of the OEE1 and OEE3 proteins

The cDNA inserts of plasmids pPII-12P4, pPII-24.3 and pPII-24.5 were isolated, digested with restriction endonucleases (as diagrammed in Fig. 1), labeled at either the 5′ or 3′ end, and sequenced by the method of Maxam and Gilbert [19]. Both strands of the cDNAs were sequenced and all restriction site junctions were sequenced across by using other restriction enzymes. The sequences for the 5′ ends of the psb1 and psb3 genes were determined by sequencing genomic clones containing that portion of the gene. A composite sequence containing the 5′ genomic and the entire cDNA sequence are shown for OEE1 and OEE3 in Figs 2 and 3 respectively.

Identification of mRNA-coding regions of the psb1 and psb3 genes

To determine the mRNA-coding regions of the psb1 and psb3 genes, genomic clones were digested with restriction enzymes which recognized common sites in both the genomic clones and cDNAs and then labeled at the 5′ and 3′ ends. The labeled DNA was heat denatured and allowed to renature in the presence of an excess of *C. reinhardtii* RNA. The resulting RNA/DNA hybrids were digested with S1 nuclease by the method of Berk and Sharp [20]. The 5′ ends of the transcripts were determined by using a sequence ladder as reference.

Isolation of DNA, gel electrophoresis and blotting

For Southern analysis, DNA was isolated from wild type strain 137c by pronase digestion and phenol extraction followed by banding on CsCl gradients as described [21]. The DNA was digested with restriction endonucleases, separated on agarose gels and blotted to nylon membrane filters as described [22]. Prehybridization and hybridization was as described by Johnson *et al.* [23]. All filters were washed in 1 × SSPE, 0.1% SDS at 60 °C. Autoradiography was carried out at − 70 °C with the use of an intensifying screen (Dupont, Cronex).

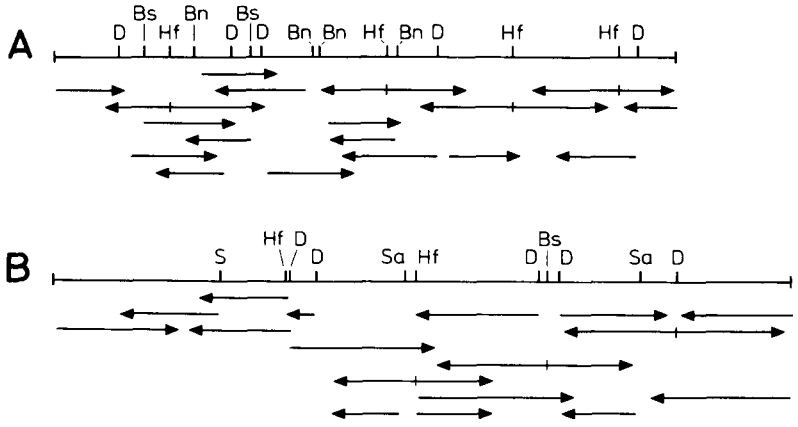

Fig. 1. Sequencing strategy for OEE1 (A) and OEE3 (B) cDNAs. The cDNA inserts were digested with the enzymes as indicated, labeled at either the 5′ or 3′ end and sequenced by the method of Maxam and Gilbert. Restriction endonuclease sites are indicated by D, *Dde* I; Bs, *Bst* NI; Hf, *Hinf* I; Bn, *Ban* I; S, *Sal* I, Sa, Sau 3A.

Protein sequencing

Mature OEE1 polypeptide was isolated from the *C. reinhardtii* mutant F54–14 which lacks both the PSI and the chloroplast ATPase complexes [24] as described [3]. The amino acid sequence was determined by using on Applied Biosystem model 470 A protein sequenator according to standard procedures.

Results and discussion

Cloning and sequence of the cDNAs encoding the OEE1 and OEE3 polypeptides of photosystem II

Both the OEE1 and OEE3 polypeptides associated with oxygen-evolving photosystem II particles (called the 33 kDa and 16 kDa proteins in spinach) had been shown to be nuclear encoded in spinach [4] as was indicated in *C. reinhardtii* [25]. We therefore constructed a library in the λgt11 expression vector containing *C. reinhardtii* cDNA inserts [3]. The library was screened with rabbit polyclonal antisera specific for either the OEE1 or OEE3 polypeptides. Several plaques containing phage expressing one or the other of these polypeptides were identified and purified to homogeneity. In order to obtain full-lenght cDNA clones the inserts obtained were used as probes to identify additional clones from a *C. reinhardtii* λgt10 cDNA library [15]. Two full-length clones corresponding to OEE1 and OEE3 were identified and their inserts subcloned into pUC19 giving rise to plasmids pPII-12P4 and pPII-24.5, respectively. The nucleotide sequence of 12P4 and the deduced amino acid sequence of the OEE1 protein are presented in Figure 2, and the nucleotide sequence of cDNA pPII-24.5 and the deduced amino acid sequence of the OEE3 protein are presented in Figure 3.

As a way of positively identifying the cDNAs as those encoding the authentic OEE1 and OEE3 polypeptides, and as a means of identifying the amino terminus of the mature protein, we isolated mature OEE1 and OEE3 protein for sequence analysis [3]. The amino terminal end of the mature OEE1 protein corresponds to amino acid 52 of the precursor OEE1 protein and is marked by an * in Fig. 2. We were unable to obtain an amino acid sequence for the mature OEE3 protein, but have estimated the end of the transit peptide (amino terminus of the mature protein) from a comparison with the ends of transit sequences of nuclear-encoded chloroplast proteins that are localized on the lumenal side of the thylakoids: OEE1 and OEE2 proteins of *C. reinhardtii* (see Fig. 7) and plastocyanin of higher plants [26, 27]. The carboxy terminal part of the transit peptides of these proteins is hydrophobic and ends with AXA.

Examination of the cDNAs reveals that both share the unusual codon bias found in other *C. reinhardtii* nuclear genes [3, 15, 28]. Note that A is not used in the third position of any amino acid codon. This codon bias, with a few deviations, has been observed in all of the *C. reinhardtii* nuclear-encoded genes sequenced to date that give rise to abundant transcripts.

Both OEE1 and OEE3 are encoded by single-copy genes

To determine the number of genes encoding OEE1 and OEE3 in *C. reinhardtii*, genomic DNA was digested with restriction endonucleases, fractionated on agarose gels, blotted to nitrocellulose and hybridized with either the labeled 750 bp *Hin*d III fragment (at the 5′ end of psb1), Fig. 5) or pPII-24.3 plasmid. As shown in Figure 4, both of these probes hybridize to only one or two fragments for each of the restriction enzymes used, suggesting that both OEE1 and OEE3 are encoded by single genes. To further characterize the psb1 (encoding the OEE1 protein) and psb3 (encoding the OEE3 protein) genes we isolated genomic clones for both of these genes. An EMBL3 *C. reinhardtii* genomic library [18] was screened using nick-translated pPII-12P4 or pPII-24.3 as probes. Several overlapping clones were isolated for each of the cDNAs and a map of the chromosomal section containing the psb1 and psb3 genes was compiled using com-

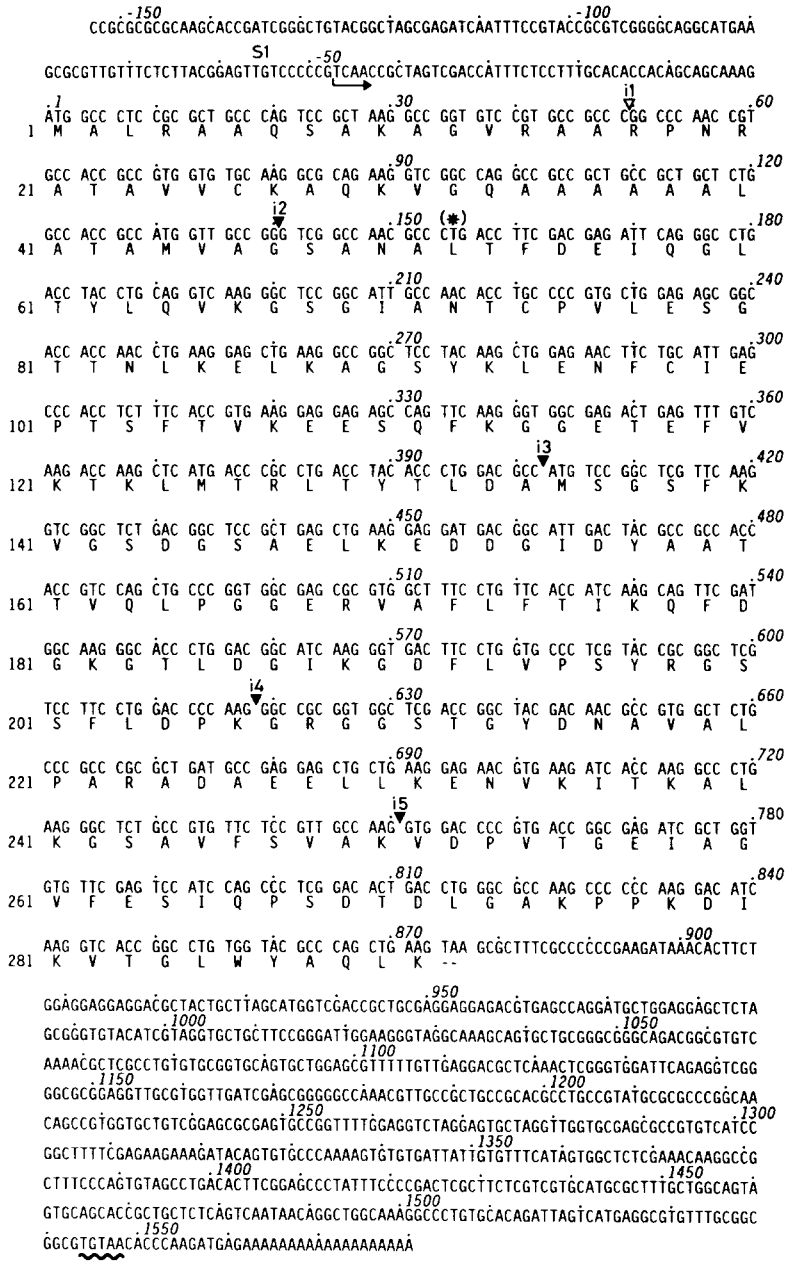

Fig. 2. Complete nucleotide and amino acid sequence of the OEE1 cDNA and protein. The nucleotide sequence was derived from cDNA pPII-12P4 and is shown from the arrow at − 50 bp to the poly(A) tail, while the 5′ end of the psb1 gene was derived from sequencing genomic clones and is shown to − 150 bp from the initiation codon. The location of the transcription start is indicated by the S1 mark located 55 bp from the initiation codon. The start of the mature OEE1 protein is designated by an * at amino acid 52. The locations of intervening sequences within the single psb1 gene are shown by arrowheads numbered i1 through i5. The likely poly(A) addition sequence TGTAA is underlined by a wavy line.

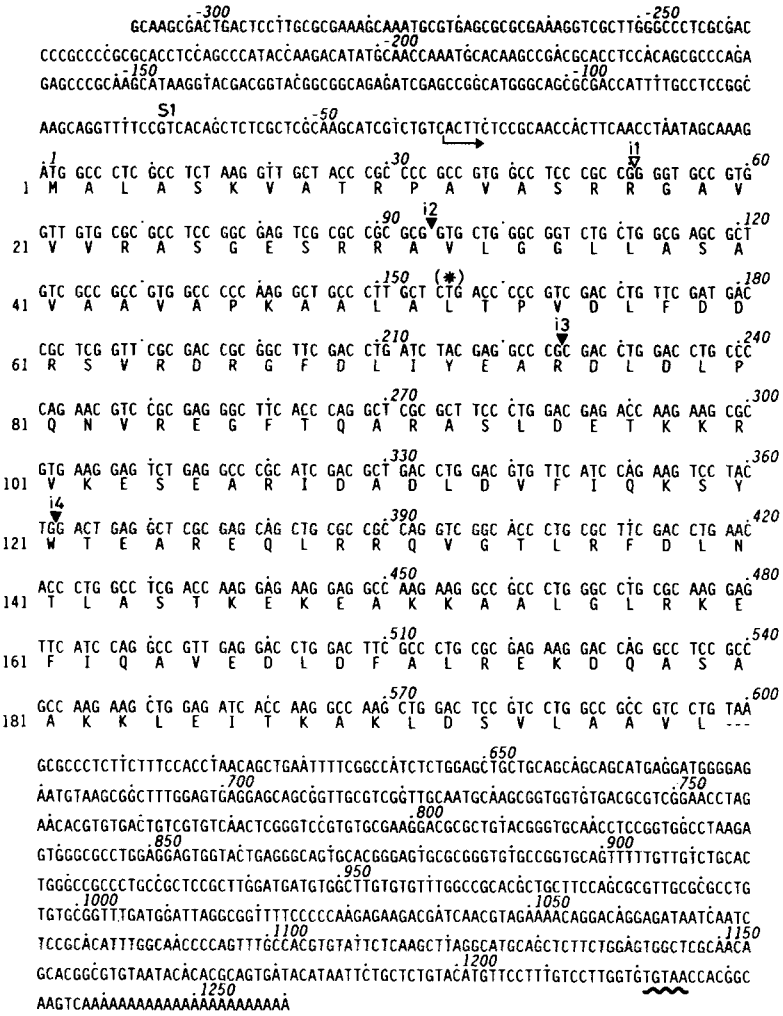

Fig. 3. Complete nucleotide and amino acid sequence of the OEE3 cDNA and protein. The nucleotide sequence was derived from cDNAs pPII-24.3 and pPII-24.5 and is shown from the arrow at − 35 bp to the poly(A) tail, while the 5′ end of the psb3 gene was derived from sequencing genomic clones and is shown to − 300 bp from the initiation codon. The location of the transcription start is indicated by the S1 mark located 65 bp from the initiation codon. The likely start of the mature OEE3 protein is designated by an * at amino acid 52. The locations of intervening sequences within the single psb3 gene are shown by arrowheads numbered i1 through i4. The likely poly(A) addition sequence TGTAA is underlined by a wavy line.

mon restriction enzymes (Fig. 5). Comparison of the restriction map with the pattern of the DNA fragments on the Southern blots confirmed that both psb1 and psb3 are single-copy genes, as is the psb2 (OEE2 protein) gene in *C. reinhardtii* [3]. These results agree with the finding that a mutation in the OEE1 gene segregates as a single nuclear mutation [10].

To define the coding regions and to determine the 5′ and 3′ boundaries of the two genes, genomic fragments were digested with restriction enzymes that recognized sites in both the cDNA and the genomic clones, labeled at either the 5′ or 3′ end, hybridized with an excess of *C. reinhardtii* mRNA, and digested with S1 nuclease. The transcription initiation site, as defined by these S1 protection experiments, lies about 55 base pairs upstream of the ATG start codon of the psb1 gene (shown in Fig. 2), while the transcription start site of the psb3 gene lies 65 base pairs upstream of the

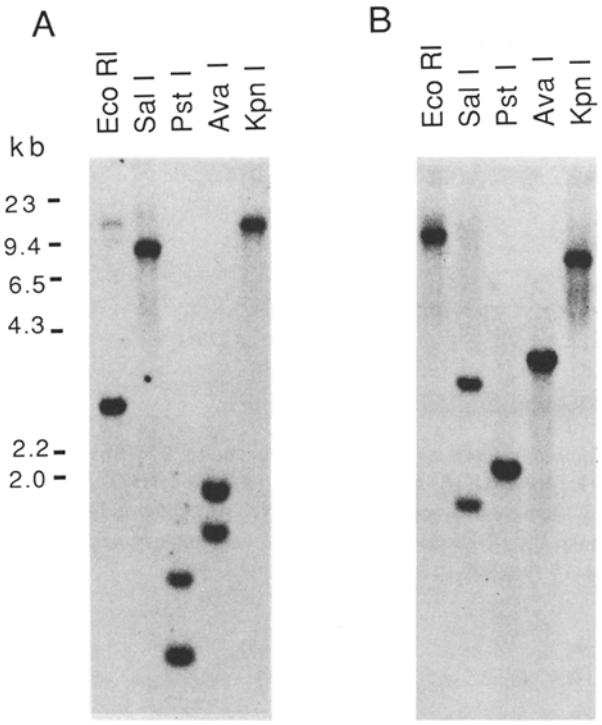

Fig. 4. Southern blot analysis of wild type *C. reinhardtii* genomic DNA. The DNA was digested with the restriction enzymes as indicated at the top of the figure. Fragments were separated by agarose gel electrophoresis, blotted to nylon membrane and probed with labeled nick-translated cDNA inserts (A) from plasmid pPII-24.3 (whose 5′ end is near the *Sal* I site shown in Fig. 5) and from the 5′ end of the psbI gene (750 bp genomic *Hind* III fragment in Fig. 5).

initiation codon. The psb1 gene has two 3′ boundaries which were visible as discrete S1 nuclease protected bands 485 and 685 base pairs downstream from the stop codon (data not shown). Two distinct mRNAs are also observed when denatured RNA is blotted to nitrocellulose and probed with nick-translated pPII-12P4 cDNA [10]. A single 3′ boundary was observed for the psb3 gene approximately 640 base pairs downstream from the stop codon. Short 5′ leader sequences and long 3′ untranslated sequences have also been observed in other *C. reinhardtii* nuclear genes coding for OEE2 [3] and the small subunit of carboxylase [15]. However, this feature has not been observed for the nuclear α and β tubulin genes [28]. The 3′ boundaries of both OEE1 and OEE3 are adjacent to the sequence TGTAA which has been proposed as a poly(A) addition sequence in *C. reinhardtii* [28]. Comparison of the cDNA and corresponding genomic sequences reveals that the psb1 and psb3 genes of *C. reinhardtii* contain 5 and 4 exons, respectively. Since there are no published reports on the characterization of genomic clones of the OEE1 and OEE3 genes of higher plants, it remains to be seen whether the sites of introns within these genes are the same in *C. reinhardtii* and higher plants.

OEE1 proteins from C. reinhardtii *and spinach share large regions of homology, while OEE3 proteins share limited homology*

Recently the amino acid sequence of the OEE1 protein from spinach has been deduced by protein sequencing of isolated mature OEE1 protein [11] and by the sequence of a spinach cDNA clone of OEE3 [12]. The 241 amino acid OEE1 protein of *C. reinhardtii* is 6 residues shorter that its spinach homologue. These missing amino acids are mostly located at the N-terminus of the spinach sequence. Comparison of the primary amino acid sequence of the OEE1 protein of *C. reinhardtii* with those of spinach [11, 12], *Synechocystis* 6803 [29] and *Anacystis nidulans* [30] reveals an overall homology of 67, 48 and 46%, respectively. The two cyanobacterial sequences are 43 and 48% homologous to their higher-plant counterpart [29, 30]. As shown in Figure 6A several regions can be recognized where the homology is much greater, in particular a stretch of 33 identical residues (139 to 171 in Fig. 6A) between *C. reinhardtii* and spinach which is 70% conserved in cyanobacteria. Two cysteine residues at positions 23 and 47 are also conserved in these four organisms. It has been noticed previously that some sequence homology exists between a region of the OEE1 protein of spinach and a segment of bacterial superoxide dismutases surrounding the Mn ligand [11]. The corresponding region in *C. reinhardtii* which includes residues 13 to 28 is strikingly less well preserved raising questions as to the

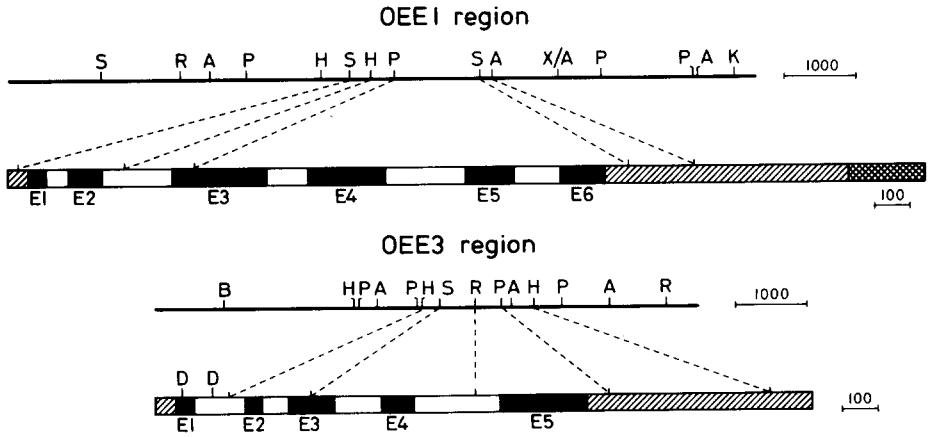

Fig. 5. Restriction map of the chromosomal region of wild-type *C. reinhardtii* containing the OEE1 (psb1) and OEE3 (psb3) genes. The restriction enzymes shown are (R) *Eco* RI, (A) *Ava* I, (P) *Pst* I, (H) *Hind* III, (S) *Sal* I, (X) *Xho* I, (K) *Kpn* I, (D) *Dde* I, and (B) *Bam* HI. The coding regions are shown in dark bars, the 5′ and 3′ non-coding regions of the mRNA are shown in lined and dotted bars (corresponding to the two 3′ ends of psbI mRNA; see text), and the intervening sequences are shown in white bars. Scales are indicated at the right of each diagram in base pairs.

A

```
                 .10        .20        .30        .40        .50        .60
C.r.    LTFDEIQGLTYLQVKGSGIANTCPVLESGTTNLKELKAGSYKLENFCIEPTSFTVKEESQ
        || ||||  ||| ||| | || ||  | |       | | |    || ||| | || |
Sp.  EGGKRLTYDEIQSKTYLEVKGTGTANQCPTVEGGVDSF-AFKPGKYTAKKFCLEPTKFAVKAEGI

                 .70        .80        .90        .100       .110       .120
C.r.    FKGGETEFVKTKLMTRLTYTLDAMSGSFKVGSDGSAELKEDDGIDYAATTVQLPGGERVA
         |     | ||||||||||||   | | | |||     | |||||| ||||||||||
Sp.     SKNSGPDFQNTKLMTRLTYTLDEIEGPFEVSSDGTVKFEEKDGIDYAAVTVQLPGGERVP

                 .130       .140       .150       .160       .170       .180
C.r.    FLFTIKQFDGKGTLDGIKGDFLVPSYRGSSFLDPKGRGGSTGYDNAVALPA--RADAEEL
        |||||||    |       |||||||||||||||||||||||||||||||||||  | | |||
Sp.     FLFTIKQLVASGKPESFSGDFLVPSYRGSSFLDPKGRGGSTGYDNAVALPAGGRGDEEEL

                 .190       .200       .210       .220       .230       .240
C.r.    LKENVKITKALKGSAVFSVAKVDPVTGEIAGVFESIQPSDTDLGAKPPKDIKVTGLWYAQLK
         ||| |     ||    ||    | |||  ||| | ||||||||||| ||| |  | |||||
Sp.     QKENNKNVASSKGTITLSVTSSKPETGEVIGVFQSLQPSDTDLGAKVPKDVKIEGVWYAQLEQQ
```

B

```
                .10        .20        .30        .40        .50        .60
C.r.    LTP--VDLFDDRSVRDRGFDLIYEARDLDLPQNVREGFTQARRSLDETKKRVKESEARID
         | |      |   |       |||  ||       |         |  | | | |
Sp.  EARPIVVGPPPPLSGGLPGTENSDQARDGTLP-YTKDRFYLQPLPPTEAAQRAKVS-ASEI

                .70        .80        .90        .100       .110       .120
C.r.    ADLDVFIQKSYWTEAREQLRRQVGTLRFDLNTLASTKEKEAKKAALGLRKEFIQAVEDLD
           ||    |      ||     || || |  | | |  ||    |           ||
Sp.     LNVKQFIDRKAWPSLQNDLRLRASYLRYDLKTVISAKPKDEKKSLQELTSKLFSSIDNLD

                .130       .140       .150
C.r.    FALREKDQASAAKKLEITKAKLDSVLAAVL
         |   |    | |    |      || | |
Sp.     HAAKIKSPTEAEKYYGQTVSNINEVL-AKLG
```

Fig. 6. A (top). Comparison of the primary amino acid sequences of the OEE1 proteins from *C. reinhardtii* (C.r) and spinach (Sp.) [11, 12, 13]. Residues that are also conserved in the OEE1 proteins from *Synechocystis* 6803 [30] and *Anacystis nidulans* [31] are marked with dots. B (bottom). Comparison of the primary amino acid sequences of the OEE3 proteins from *C. reinhardtii* (C.r) and spinach (Sp.). Protein comparison was made by the method of Needleman and Wünsch [37].

significance of the homology found in higher plants. Similar conclusions have been reached from the sequence comparison of the OEE1 proteins of *Synechocystis* 6803 [29] and *Anacystis nidulans* [30] with the higher-plant homologue.

Comparison of the amino acid sequence of the OEE3 proteins of spinach [13] and *C. reinhardtii* show them to contain limited, though significant, sequence homology (Fig. 6B). Both the spinach and *C. reinhardtii* OEE3 proteins contain the same number of residues (149) with a high proportion of charged amino acids (36/149 for *C. reinhardtii*). Overall, the OEE3 proteins share only 28% homology, and in only one section of the proteins (from residues 86–104) does the homology exceed 50%.

These comparisons of algal and higher plant OEE proteins reveal that although these proteins belong to the same complex they have been conserved to a variable extent: 67% homology for OEE1, 61% for OEE2 [3, 13] and 28% for OEE3. In contrast the primary amino acid sequences of the PSII core polypeptides of *C. reinhardtii* and higher plants are significantly more conserved with a homology ranging between 86 and 93% [31, 32]. The OEE1 protein is closely associated with the core proteins. Loss of the OEE1 protein destabilizes the PSII core complex resulting in a more rapid turnover of the reaction center proteins [10]. This is not the case for the OEE2 protein. Loss of OEE2 does not affect the accumulation of the core PSII proteins [3].

Amongst the OEE proteins OEE3 is the least conserved. Its loose association with the PSII complex may have allowed for a greater sequence divergence. It thus appears that the amount of homology between the OEE proteins of *C. reinhardtii* and higher plants could reflect the constraints imposed by the specific interactions of these proteins with the conserved PSII core.

Transit peptides of the proteins from the oxygen-evolving complex of C. reinhardtii

All nuclear-encoded chloroplast-localized proteins characterized to date have been shown to contain N-terminal transit peptides. These presequences appear to be required for the proper sorting of proteins to the chloroplast and for guiding the proteins to their correct location within the chloroplast. The routing of thylakoid lumen proteins is of special interest since they have to traverse the two chloroplast envelope membranes and the thylakoid membrane.

The 51 amino acid transit peptides of OEE1 of *C. reinhardtii* (Fig. 7) is significantly shorter than its spinach homologue which contains 84 [12]. Although there is some uncertainty as to the precise location of the transit sequence cleavage site of OEE3, its transit peptide also appears to be considerably shorther than its higher-plant homologue [13]. Comparison of the transit sequences of the three lumen-localized OEE proteins and of the stromal small subunit of Rubisco of *C. reinhardtii* reveals some conserved features. The VARPA motif or some derivative of it appears in approximately the same position within each the transit peptides of *C. reinhardtii* except for OEE1 where it is only partially present. This

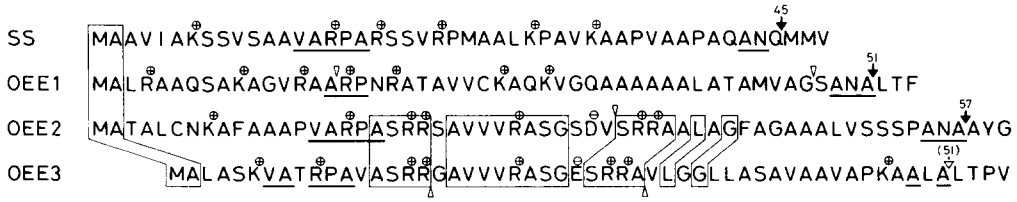

Fig. 7. Comparison of the transit peptides from the small subunit of carboxylase (SS), and the OEE1, OEE2 and OEE3 peptides of *C. reinhardtii*. Positively charged amino acids are indicated as are the sites of cleavage of the transit peptides (arrows with amino acid number), the location of intervening sequences within the gene (white arrowheads), and the position of conserved sequences (underlined or boxed).

short sequence is not found in the transit peptides from other organisms. Other conserved features include the C-end of the transit peptides which is related to the cleavage site of signal sequences [32]. It is remarkable that four of the last five residues of the transit peptide of OEE1 from *C. reinhardtii* are identical to the C-terminal end GSAFA of the signal sequence of OEE1 from the cyanobacterium *Synechocystis* 6803 [30]. Figure 7 reveals that the transit peptides of OEE2 and OEE3 are highly related. However the conserved region is not found within any of the other transit peptides and its significance is not clear. A striking feature of the transit peptides of OEE1 and OEE2 from *C. reinhardtii* proteins is the presence of a basic amino terminal region and a hydrophobic region. This structure also appears to be present in the transit peptide of OEE3 although its precise end could not be determined. A similar two-domain organization was first reported for the transit peptide of plastocyanin, another thylakoid lumen protein [26]. The first domain which includes the amino terminal region is rich in basic and hydroxylated amino acids and appears to play a role in protein import into the chloroplast. The second domain contains mostly hydrophobic residues and has been postulated to play a key role in the routing of protein towards the thylakoid lumen. There is growing evidence that the precursors of these lumen proteins are processed in two steps [33, 34]. Domains I and II of the transit peptide are removed sequentially by two peptidases localized in the stroma and on the thylakoids, respectively [35].

Surprisingly, the coding sequences of the transit peptides of the three OEE proteins of *C. reinhardtii* are interrupted by one or two intervening sequences. It is of interest that some of the introns in psb2 and psb3 are located near the border between domains I and II which appear to correspond to functionally distinct domains of the transit peptide. While such a correlation between exons and protein functional domain is well documented [36], it remains to be seen whether introns will be found a similar positions in other organisms.

Acknowledgements

We thank O. Jenni for drawings and photography, N.H. Chua and L. McIntosh for antibodies, and L.G. Franzen, M. Goldschmidt-Clermont and M. Kuchka for helpful comments. This work was supported by grant 3.328–086 from the Swiss National Foundation.

References

1. Miyao M, Murata N: Photoinactivation of the oxygen-evolving complex of photosystem II. In: Kyle DJ, Osmond CB, Arntzen CJ (eds). Photoinhibition, pp. 289–307. Elsevier Science Publishers, Amsterdam (1987).
2. Barber J: Photosynthetic reaction centers: a common link. Trends Biochem Sci 12: 321–326 (1987).
3. Mayfield SP, Rahire M, Frank G, Zuber H, Rochaix JD: Expression of the nuclear gene encoding oxygen-evolving enhancer protein 2 is required for high levels of photosynthetic oxygen evolution in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 84: 749–753 (1987).
4. Westhoff P, Jansson C, Klein-Hitpass L, Berzborn R, Larsson C, Bartlett SG: Intracellular coding sites of polypeptides associated with photosynthetic oxygen evolution of photosystem II. Plant Mol Biol 4: 137–146 (1985).
5. Ljungberg U, Akerlund HE, Andersson B: Isolation and characterization of the 10 kDa and 22 kDa polypeptides of higher plant photosystem II. Eur J Biochem 158: 477–482 (1986).
6. Ljungberg U, Henrysson T, Rochester CP, Akerlund HE, Andersson B: The presence of low-molecular-weight polypeptides in spinach Photosystem II core preparations. Isolation of a 5 kDa hydrophilic polypeptide. Biochim Biophys Acta 849: 112–120 (1986).
7. Packham NK: Is the 9 kDa thylakoid membrane phosphoprotein functionally and structurally analogous to the H subunit of bacterial reaction centers? FEBS Lett 231: 284–290 (1988).
8. Critchley C: The role of chloride in Photosystem II. Biochim Biophys Acta 811: 33–46 (1985).
9. Andersson B: Protein participating in photosynthetic water oxidation. Encycl Plant Physiol 19: 447–456 (1986).
10. Mayfield SP, Bennoun P, Rochaix JD: Expression of the nuclear encoded OEE1 protein is required for oxygen evolution and stability of photosystem II particles in *Chlamydomonas reinhardtii*. EMBO J 6: 313–318 (1987).
11. Oh-oka H, Tanaka S, Wada K, Kuwabara T, Murata N: Complete amino acid sequence of 33 kDa protein isolated from spinach photosystem II particles. FEBS Lett 197: 63–66 (1986).

12. Tyagi A, Hermans J, Steppuhn J, Jansson Ch, Vater F, Herrmann RG: Nucleotide sequence of cDNA clones encoding the complete 33 kDa precursor protein associated with the photosynthetic oxygen-evolving complex from spinach. Mol Gen Genet 207: 288–293 (1987).
13. Jansen T, Rother C, Steppuhn J, Reinke H, Beyreuther K, Jansson C, Andersson B, Herrmann RG: Nucleotide sequence of cDNA clones encoding the complete 23 kDa and 16 kDa precursor proteins associated with the photosynthetic oxygen-evolving complex from spinach. FEBS Lett 216: 234–240 (1987).
14. Schmidt GW, Devilliers-Thiery A, Desruisseaux H, Blobel G, Chua NH: NH terminal sequences of precursor and mature forms of the ribulose-1,5-bisphosphate carboxylase small subunit from *Chlamydomonas reinhardtii*. J Cell Biol 83: 615–622 (1979).
15. Goldschmidt-Clermont M, Rahire M: Sequence, evolution and differential expression of the two genes encoding variant small subunits of ribulose bisphosphate carboxylase/oxygenase in *Chlamydomonas reinhardtii*. J Mol Biol 191: 421–432 (1986).
16. Young RA, Davis RW: Efficient isolation of genes by using antibody probes. Proc Natl Acad Sci USA 80: 1194–1198 (1983).
17. Benton WD, Davis RW: Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196: 180 (1977).
18. Goldschmidt-Clermont M: The two genes for the small subunit of RuBP carboxylase/oxygenase are closely linked in *Chlamydomonas reinhardtii*. Plant Mol Biol 6: 13–21 (1986).
19. Maxam AM, Gilbert W: Sequencing end-labelled DNA with base-specific chemical cleavages. Meth Enzymol 64: 499–560 (1980).
20. Berk AJ, Sharp PA: Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. Cell 12: 721–732. (1977).
21. Rochaix JD: Restriction fragments from *Chlamydomonas* chloroplast DNA. Meth Enzymol 65: 785–795 (1980).
22. Khandjian EW: UV crosslinking of RNA to nylon membrane enhances hybridization signals. Mol Biol Rep 11: 107–115 (1986).
23. Johnson DA, Gantsch JW, Sportman JR, Elder JH: Gene. Anal Tech 1: 3–8 (1984).
24. Chua NH, Matlin K, Bennoun P: A chlorophyll-protein complex lacking in photosystem II mutants of *Chlamydomonas reinhardtii*. J Cell Biol 67: 361–377 (1975).
25. Delepelaire P: Partial characterization of the biosynthesis and integration of the photosystem II reaction centers in the thylakoid membrane of *Chlamydomonas reinhardtii*. EMBO J 3: 701–706 (1984).
26. Smeekens S, de Groot M, van Binsbergen J, Weisbeek P: Sequence of the precursor of the chloroplast thylakoid lumen protein plastocyanin. Nature 317: 456–458 (1985).
27. Rother C, Jansen T, Tyagi A, Tittgen J, Herrmann RG: Plastocyanin is encoded by an uninterrupted nuclear gene in spinach. Curr Genet 11: 171–176.
28. Silflow CD, Chisholm RL, Conner TW, Ranum PW: The two alpha-tubulin genes of *Chlamydomonas reinhardtii* code for slightly different proteins. Mol Cell Biol 5: 2389–2398.
29. Philbrick J, Zilinskas BA: Cloning, nucleotide sequence and mutational analysis of the gene encoding the Photosystem II manganese-stabilizing polypeptide of *Synechocystis* 6803. Mol Gen Genet 212: 418–425 (1988).
30. Kuwabara T, Teddy KJ, Sherman LA: Nucleotide sequence of the gene from the cyanobacterium *Anacystis nidulans* R2 encoding the Mn-stabilizing protein involved in photosystem II water oxidation. Proc Natl Acad Sci USA 84: 8230–8234 (1987).
31. Erickson JM, Rochaix JD, Delepelaire P: Analysis of genes encoding two photosystem II proteins of the 30–34 kD size class. In: Steinback B, Bonitz S, Arntzen CJ, Bogorad L (eds). Molecular Biology of the Photosynthetic Apparatus, pp. 53–65. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1985).
32. Rochaix JD, Kuchka M, Mayfield SP, Schirmer-Rahire M, Girard-Bascou J, Bennoun P: Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast psbC gene product in *Chlamydomonas reinhardtii*. EMBO J, in press.
33. Robinson C, Ellis RJ: Transport of proteins into chloroplasts. Eur J Biochem 142: 237–342 (1984).
34. Chia CP, Arntzen CJ: Evidence for two-step processing of nuclear-encoded chloroplast proteins during membrane assembly. J Cell Biol 103: 725–731 (1986).
35. Hageman J, Robinson C, Smeekens S, Weisbeek P: A thylakoid processing protease is required for complete maturation of the lumen protein plastocyanin. Nature 324: 567–569 (1986).
36. Brändén CI, Eklund H, Cambillou C, Pryor AJ: Correlation of exons with structural domains in alcohol dehydrogene. EMBO J 3: 1307–1310 (1984).
37. Needleman SB, Wunsch CD: A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 48: 443–453 (1970).