Plant Molecular Biology **13**: 573–582, 1989.

Characterization of cDNA clones encoding the extrinsic 23 kDa polypeptide of the oxygen-evolving complex of photosystem II in pea

Richard Wales,[1] Barbara J. Newman,[1,3] Stephanie A. Rose,[1] Darryl Pappin[2] and John C. Gray[1]*
[1] *Botany School, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK (*author for correspondence);* [2] *Department of Biochemistry, University of Leeds, Leeds LS2 9JT, UK;* [3] *Present address: National Institute of Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK*

Received 4 April 1989; accepted in revised form 26 June 1989

Key words: photosynthesis, oxygen evolution, pea chloroplasts, thylakoid lumen, 23 kDa extrinsic polypeptide

Abstract

The 23 kDa polypeptide of the oxygen-evolving complex of photosystem II has been extracted from pea photosystem II particles by washing with 1 M NaCl and purified by anion-exchange chromatography. The *N*-terminal amino acid sequence has been determined and specific antisera have been raised in rabbits and used to screen a pea-leaf cDNA library in λgt11. Determination of the nucleotide sequence of two clones provided the nucleotide sequence for the full 23 kDa polypeptide. The deduced amino acid sequence showed it to code for a mature protein of 186 amino acid residues with an *N*-terminal presequence of 73 amino acid residues showing a high degree of conservation with previously reported 23 kDa sequences from spinach and *Chlamydomonas*. Southern blots of genomic DNA from pea probed with the labelled cDNA gave rise to only one band suggesting that the protein is encoded by a single gene. Northern blots of RNA extracted from various organs indicated a message of approximately 1.1 kb, in good agreement with the size of the cDNA, in all chlorophyll-containing tissues. Western blots of protein extracted from the same organs indicated that the 23 kDa polypeptide was present in all major organs of the plant except the roots.

Abbreviations: bis-Tris, bis (2-hydroxyethyl) imino-tris (hydroxymethyl)-methane; pfu, plaque-forming units

Introduction

The ultimate source of electrons for photosynthesis in green plants is the oxidation of water by photosystem II, with the production of molecular oxygen and the release of protons into the thylakoid lumen. Photosystem II is situated mainly in the appressed regions of the thylakoid membrane and is composed of three main functional sub-complexes, a core complex carrying the reaction centre, a light-harvesting chlorophyll-protein complex and an extrinsic water-

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X15552 *psb*2, *wox*B.

oxidizing complex. Three hydrophilic polypeptides, of 33 kDa, 23 kDa and 16 kDa, have been shown to play a role in the oxidation of water [2, 29], and by immunochemical studies [4] to be present in stoichiometric amounts with either one or two molecules of each present per reaction centre.

Although the precise functions of the 33 kDa, 23 kDa and 16 kDa polypeptides are not yet fully understood, the 33 kDa polypeptide is thought to be involved in the stabilization of the tetranuclear manganese centre of the oxygen-evolving complex [2, 29]. It has been shown that the 23 kDa and 16 kDa polypeptides can be replaced in oxygen-evolving photosystem II preparations (after removal by NaCl washing) by non-physiological levels of Ca^{2+} or Cl^- [1, 3, 13, 27, 28, 30]. For this reason, and the fact that, unlike the 33 kDa polypeptide, these proteins are not present in cyanobacteria [37], they are thought to play regulatory or structural roles rather than having any direct catalytic involvement in oxygen evolution.

The 33 kDa, 23 kDa and 16 kDa polypeptides are all encoded by nuclear genes and synthesized as precursor forms with long *N*-terminal amino acid extensions which function in targeting the polypeptides to the chloroplast and the thylakoid lumen [41, 43]. The presequences of lumenal polypeptides are thought to have two domains, an *N*-terminal chloroplast import domain which is cleaved off in the stroma, and a *C*-terminal thylakoid transfer domain which is cleaved by a thylakoid protease distinct from that in the stroma [15]. In contrast, the polypeptides of the reaction centre complex are all encoded by chloroplast genes [14] and are inserted directly into the thylakoid membrane.

In order to understand the regulatory features necessary to coordinate the expression of the nuclear and chloroplast genes to produce a functional photosystem II we have started to characterize nuclear genes for the extrinsic polypeptides of the oxygen-evolving complex. Study of the genes and their flanking regions may also provide some information on the processes regulating the expression of the oxygen-evolving polypeptides at the level of cell type as has been shown to occur in maize [34]. The presence of a small multi-gene family for the 33 kDa polypeptide was reported recently [41]. In this paper the nucleotide sequence of cDNA clones for the 23 kDa polypeptide and the derived amino acid sequence are presented, and compared to analogous sequences from spinach [16] and *Chlamydomonas* [25]. Evidence is also presented to show that the 23 kDa polypeptide in pea is encoded by a single-copy nuclear gene which is expressed in all chlorophyll-containing organs of the plant.

Materials and methods

Purification of the 23 kDa polypeptide

Chloroplasts were isolated from 50 g of 7–10-day-old pea leaves (*Pisum sativum* cv. Feltham First) by the method of Kalberer *et al.* [17] except that sorbitol was replaced by sucrose. Thylakoid membranes enriched in PSII prepared as described by Ford and Evans [11], were washed in 1 M NaCl, 10 mM MES pH 6.5 for 30 minutes at 4 °C at a chlorophyll concentration of 0.5 mg ml^{-1}. The 23 kDa polypeptide was purified from the supernatant (by FPLC anion-exchange chromatography) after centrifugation at 40 000 × ***g*** for 15 minutes at 4 °C and dialysis against 20 mM piperazine-HCl pH 9.5. The supernatant was applied to a Mono Q HR5/5 column (Pharmacia) equilibrated with 20 mM piperazine-HCl pH 9.5 and proteins eluted with a 20 ml gradient of 0–350 mM NaCl in 20 mM piperazine-HCl pH 9.5. Fractions identified as containing the 23 kDa polypeptide by SDS-polyacrylamide gel electrophoresis were pooled and dialysed against 20 mM bis-Tris pH 6.5 and applied to a Mono Q HR5/5 column (Pharmacia) equilibrated in 20 mM bis-Tris pH 6.5. The protein was eluted with a 20 ml gradient of 0–350 mM NaCl in the same buffer, and fractions identified as containing the 23 kDa polypeptide by SDS-polyacrylamide gel electrophoresis were dialysed against double-distilled water. The *N*-terminal amino acid sequence was determined as described by Wales *et al.* [41].

Antisera

Antibodies were raised against the 23 kDa polypeptide in New Zealand White rabbits by injection of 50 μg protein in 0.5 ml H_2O emulsified with 0.5 ml Freund's complete adjuvant, intramuscularly in the back four times at 14-day intervals. Specificity of antisera was evaluated by the use of western blotting according to Towbin *et al.* [39] except that the transfer buffer contained 0.1% SDS.

Antibody screening of cDNA library

A pea leaf cDNA library in λgt11 was provided by Dr J.S. Gantt [12]. The library was screened at a density of 30000 pfu per 9 cm Petri dish according to Young and Davis [44], except that the filters were incubated with the specific antiserum overnight at 4 °C. Bound antibodies were detected with 0.4 μCi ^{125}I-labelled protein A (35 μCi mg^{-1}, Amersham International) per plate and visualized by autoradiography using Kodak X-Omat S film. Plaques giving a positive signal were picked and rescreened twice at lower density (ca. 200 pfu per dish) after which all plaques gave a positive signal.

Screening of cDNA library with DNA probes

The cDNA library as described above was screened at a density of 10000 pfu per 9 cm Petri dish by hybridization with a cDNA probe. Dry HybondN nylon membranes (Amersham) were placed on the surface of plates for 1 min followed by denaturing, probing and washing as described by the manufacturers. The cDNA fragments used as probes were labelled by the method of Feinberg and Vogelstein [10].

Characterization of cDNA clones

DNA was prepared from recombinant λgt11 clones as described by Maniatis *et al.* [23], and after digestion with *Eco* RI the cDNA inserts were isolated by 1% agarose gel electrophoresis onto DE81 paper (Whatman Ltd.) followed by elution with 1.5 M NaCl [8]. Isolated cDNA fragments were inserted into pUC18 and used to transform *Escherichia coli* strain TG1. DNA was prepared by the method of Birnboim and Doly [5] and used for restriction mapping with enzymes from Amersham International and New England Biolabs. Restriction fragments were inserted into M13mp18 and mp19 and single-stranded M13 DNA was used for sequencing by the dideoxynucleotide chain-termination method of Sanger *et al.* [33].

Preparation of genomic DNA and Southern blotting

Leaves (2 g) of 7–8-day-old peas were ground with a small amount of acid-washed sand in 5 ml of 2.5% (w/v) SDS, 100 mM Tris-HCl pH 7.0, 50 mM EDTA.Na_2, 50 mM NaCl and 100 mM sodium diethyldithiocarbamate. The mixture was extracted with phenol/chloroform (1 : 1 v/v) twice and the nucleic acid precipitated with 2 vol of absolute ethanol. Genomic DNA was spooled out, washed in 70% and absolute ethanol and resuspended in 10 mM Tris-HCl pH 8.0, 1 mM EDTA.Na_2 pH 8.0. RNase A was added to 20 μg ml^{-1} and after incubation at 37 °C for 2 h the DNA was extracted with phenol/chloroform and reprecipitated. Genomic DNA (10 μg) was digested with 200 units of *Bam* HI followed by electrophoresis in a 0.8% agarose gel for 8 h at 150 V. After ethidium bromide staining and visualization on a short-wave UV transilluminator the DNA was depurinated by soaking for 10 min in 0.25 M HCl. The DNA was capillary-blotted [35] to GeneScreen Plus (Du Pont) in 0.4 M NaOH overnight [32]. Following neutralization in 2 × SSC the filter was pre-hybridized in 15 ml of 0.6 M NaCl, 20 mM PIPES pH 6.8, 4 mM EDTA.Na_2, 0.2% gelatin, 0.2% Ficoll-400, 0.2% polyvinyl pyrrolidone-360, 1% SDS, 1 mM $Na_2P_2O_7$ and 0.5 mg ml^{-1} single-stranded fish-milt DNA for 2 h at 65 °C. This was replaced with 5 ml of the same solution containing ca.

10 ng of the cDNA fragment labelled by the method of Feinberg and Vogelstein [10] and incubation continued at 65 °C overnight. The filter was washed at 65 °C 3 times in 2 × SSC, 1% SDS for 15 min followed by 0.4 × SSC, 1% SDS for 10 min at 65 °C and then exposed to Kodak X-Omat S film with an intensifying screen at − 70 °C.

Preparation of total RNA and northern analysis

Total RNA was prepared from various tissues of light-grown peas by the method of Covey and Hull [7]. Total RNA (20 μg) was denatured with glyoxal [26] and electrophoresed for 6 h at 150 V on a 1% agarose gel in 10 mM sodium phosphate pH 7.0. The RNA was capillary blotted onto GeneScreen Plus (Du Pont) in 10 × SSC and probed according to the manufacturer's instructions with the cDNA fragment labelled by the method of Feinberg and Vogelstein [10]. The membrane was then exposed to Kodak X-Omat S film with an intensifying screen at − 70 °C.

Preparation of tissue extracts for western blotting

To prepare tissue extracts for western blotting plant tissue (200 mg fresh weight) was ground in 1 ml of 10% glycerol, 10% SDS, 80 mM Tris-HCl pH 6.8, 0.05% bromophenol blue with a small amount of acid-washed sand and immediately boiled for 2 min. The extracts were then centrifuged at 10000 ***g*** for 2 min and the supernatants stored at − 70 °C. SDS-PAGE was carried out by the method of Laemmli [20] and blotting to nitrocellulose by the method of Towbin *et al.* [39] except that the transfer buffer contained 0.1% SDS.

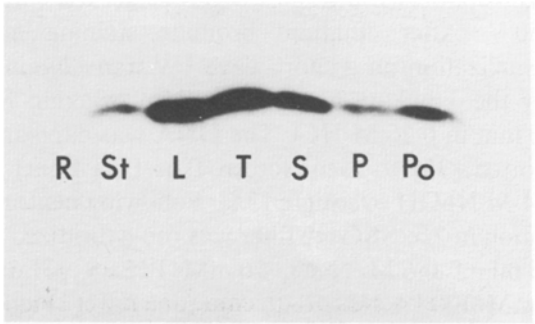


Fig. 1. Western blot of total protein from pea roots (R), stems (St), leaves (L), tendrils. (T), sepals (S), petals (P) and young pods (Po) separated by SDS-PAGE and probed with antisera specific for the 23 kDa polypeptide.

Results

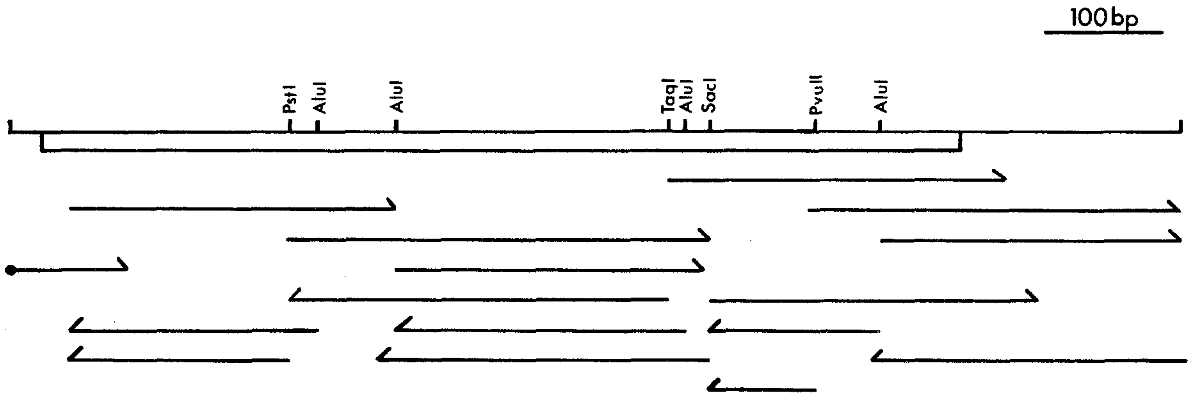
Characterization of the 23 kDa polypeptide

The 23 kDa polypeptide was extracted with 1 M NaCl, 10 mM MES pH 6.5 from photosystem II-enriched thylakoid membranes and purified by FPLC on Mono Q columns. In the first chromatographic step in 20 mM piperazine pH 9.5, the 23 kDa polypeptide was eluted at 170 mM NaCl. Further chromatography after dialysis against 20 mM bis-Tris-HCl pH 6.5 resulted in the elution of a single peak at 130 mM NaCl in 20 mM bis-Tris-HCl pH 6.5, and corresponding fractions were shown by SDS-PAGE to contain a single 23 kDa polypeptide. Determination of the *N*-terminal sequence of the protein using a solid-phase sequencer gave a unique sequence of 30 amino acids (Fig. 3). Preparations of total protein from a number of plant organs separated by SDS-PAGE and transferred to nitrocellulose were probed with antibodies raised against the purified protein and a single band of 23 kDa was obtained with protein extracts from all chlorophyll-containing organs but not roots (Fig. 1).

Isolation and characterization of cDNA clones

A pea leaf cDNA library in λgt11 was screened with the antibodies to the 23 kDa polypeptide, and 85 plaques out of a total of 300000 gave a positive signal. Five of these were further purified by two more rounds of screening, after which two clones, λg23.3 and λg23.4 still gave a positive antibody reaction. Both clones contained inserts of 950 bp and gave rise to similar restriction en-

A



B

```
                                                                                        -20      -10
                                                                                   ATTAAACTATAGCATACTCTTAAGAAAA

MetAlaSerThrGlnCysPheLeuHisHisGlnTyrAlaIleThrThrProThrArgThrLeuSerGlnArgGlnValValThrThrLysProAsnHisIleValCysLysAlaGlnLys
ATGGCATCTACACAATGTTTCTTGCACCACCAATATGCTATTACAACTCCAACTAGAACCTTATCTCAACGCCAAGTAGTTACCACCAAACCAAACCATATTGTTTGCAAGGCACAGAAA
       10        20        30        40        50        60        70        80        90       100       110       120

GlnAspAspValValAspAlaValValSerArgArgLeuAlaLeuSerValLeuIleGlyAlaAlaAlaValGlySerLysValSerProAlaAspAlaAlaTyrGlyGluAlaAlaAsn
CAAGATGATGTTGTTGATGCTGTTGTGTCTCGTAGGTTAGCACTTTCTGTTCTCATTGGTGCTGCTGCTGTTGGCTCCAAGGTTTCACCTGCAGATGCTGCTTATGGAGAAGCTGCTAAT
      130       140       150       160       170       180       190       200       210       220       230       240

ValPheGlyLysAlaLysThrAsnThrAspTyrLeuProTyrAsnGlyAspGlyPheLysLeuLeuValProAlaLysTrpAsnProSerLysGluArgGluPheProGlyGlnValLeu
GTGTTTGGAAAGGCAAAGACAAACACAGATTACTTGCCATACAATGGAGATGGATTCAAGCTGCTGGTGCCAGCAAAATGGAATCCAAGCAAAGAGAGAGAGTTCCCTGGTCAGGTTCTT
      250       260       270       280       290       300       310       320       330       340       350       360

ArgTyrGluAspAsnPheAspAlaThrSerAsnValSerValLeuValGlnThrThrAspLysLysSerIleThrAspTyrGlySerProGluGluPheLeuSerLysValAspTyrLeu
AGATATGAGGATAATTTTGATGCAACAAGCAATGTTTCGGTTTTAGTCCAAACAACTGATAAGAAGTCCATCACTGACTATGGTTCACCTGAAGAGTTCCTATCTAAGGTGGACTATTTG
      370       380       390       400       410       420       430       440       450       460       470       480

LeuGlyLysGlnAlaPhePheGlyGlnThrAspSerGluGlyGlyPheAspThrAsnAlaValAlaValAlaAsnIleLeuGluSerSerAlaProValIleGlyGlyLysGlnTyrTyr
CTTGGAAAACAAGCCTTCTTTGGACAAACTGATTCAGAGGGAGGTTTCGATACAAATGCTGTAGCTGTTGCAAACATCTTAGAGAGCTCAGCACCAGTGATTGGTGGAAAACAGTATTAC
      490       500       510       520       530       540       550       560       570       580       590       600

AsnIleSerValLeuThrArgThrAlaAspGlyAspGluGlyGlyLysHisGlnLeuIleThrAlaThrValLysAspGlyLysLeuTyrIleCysLysAlaGlnAlaGlyAspLysArg
AACATTTCGGTATTGACAAGGACTGCTGATGGAGATGAAGGTGGAAAGCATCAGCTGATTACAGCAACTGTAAAGGATGGAAAATTATACATCTGCAAGGCTCAAGCTGGTGACAAGAGG
      610       620       630       640       650       660       670       680       690       700       710       720

TrpPheLysGlyAlaArgLysPheValGluAspThrAlaSerSerPheSerValAla
TGGTTTAAGGGAGCAAGAAAGTTTGTGGAGGACACTGCAAGTTCTTTCAGTGTTGCCTAATCAAGAACTTTGGTGAAAGGCTGTTATTTTGCATGTAAATCTTATTCTTAAGTAAGCAAG
      730       740       750       760       770       780       790       800       810       820       830       840

CACTTGGAGTTTTAATTTATAATGGCTTTTTAATCTGCTTTTTAATTAAAAATTTCAAGTTGTTACAAGCCGTTTAAGAAGCACTATAATATGAAAAATCATTGTTGGAAAAAAAAAn
      850       860       870       880       890       900       910       920       930       940       950       960
```

Fig. 2. A. Strategy used in the sequencing of the cDNA inserts of λg23.3 and λg23.106 isolated from the λgt11 pea leaf cDNA library with antisera specific for the 23 kDa polypeptide and a labelled cDNA probe respectively. Arrow heads indicate the direction and extent of sequencing, arrows marked ● indicating sequence derived from λg23.106. B. Nucleotide sequence and deduced amino acid sequence of the cDNA for the 23 kDa polypeptide. Underlining indicates sequence obtained from λg23.106.

zyme fragments. DNA sequence analysis revealed that both λg23.3 and λg23.4 encoded the 23 kDa polypeptide by comparison to the determined *N*-terminal sequence but both lacked an initiating ATG codon. To obtain clones containing the nucleotide sequence for the complete *N*-terminus of the polypeptide the cDNA insert of λg23.3 was radioactively labelled and used to reprobe the pea leaf cDNA library. A positive signal was obtained with approximately 100 plaques out of a total of 50 000 screened. Ten of these were further purified by two more rounds of screening at a lower plaque density. Agarose gel electrophoresis after digestion with *Eco* RI to cut out the inserts revealed that two clones contained longer inserts. DNA sequence analysis showed one of these clones,

```
                          10          20          30          40
Pea              MASTQCFLHHQYAITTP--TRTLS-QR-QVVTTKPNHIVCKAQKQDD
Spinach          ::::A:::::HA::SS:AAG:GSAA::Y:A:SI:::Q::::::::::
Chlamydomonas    ::TAL:NKAFAA:---:VA-:--PAS:RSA:----------------

                              50          60          70  ▼      80
Pea Protein                                                XYGEAANVFG
Pea              ----VVDA---VVSRRLALSVLIGAAAVGSKVSPADAAYGEAANVFG
Spinach          NEAN:LN-SG--:::::::T:::::::::::::::::::::::::::
Chlamydomonas    ----:VR:SGSD::::A::AGFA::::LV:S-:::N::::DS:::::

                       90         100         110         120         130
Pea Protein      KAKTNTDYLPYNGDGFKLLV
Pea              KAKTNTDYLPYNGDGFKLLVPAKWNPSKEREFPGQVLRYEDNFDATS
Spinach          :P:K::EFM:::::::::::::S:::::::K::::::::::::::::::
Chlamydomonas    :VTNKSGFV::A::::A::L::::::::::ND:::VI::::::::RVN

                         140         150         160         170
Pea              NVSVLVQTTDKKSITDYGSPEEFLSKVDYLLGKQAFFGQTDSEGGFD
Spinach          :L:::::P::::::::F::::D:::Q::::::::::Y::K::::::::
Chlamydomonas    :LV:IA:D::::A:A:F::QDK::ES:S::::::::YS:E:Q:::::A

                 180         190         200         210         220
Pea              TNAVAVANILESSAPVIG-GKQYYNISVLTRTADGDEGGKHQLITAT
Spinach          SGV::S::V::::T::VD-:::::S:T:::::::::::::::::V:A::
Chlamydomonas    P:R:SA:SL:DV:TTTDKK::T::KYEL:V:S:::::::R::::G::

                     230         240         250
Pea              V-KDGKLYICKAQAGDKRWFKGARKFVEDTASSFSVA
Spinach          :-::::::::::::::::::::::K::::SAT::::::
Chlamydomonas    :GS:N::::I:I:I::::::::::K:EAMGAFD::T:V
```

Fig. 3. Comparison of the determined *N*-terminal amino acid sequence with the sequence deduced from the nucleotide sequence of the pea cDNA clone. Also included in the figure are the published sequences of the spinach [16] and *Chlamydomonas* [25] proteins. The putative Box III type sequence and thylakoid transfer domain are underlined and the final site of processing to the mature polypeptide is indicated (▼). The amino acids are numbered with regard to the pea protein.

λg23.106 to contain an essentially full-length cDNA for the precursor of the 23 kDa polypeptide. The nucleotide sequence obtained contained an open reading frame of 777 bp encoding a polypeptide of 259 amino acid residues (Fig. 2b). Comparison to the determined *N*-terminal sequence of the 23 kDa polypeptide showed the *N*-terminus of the mature protein to be Ala74 indicating that the polypeptide is synthesized with an *N*-terminal extension of 73 amino acid residues. The mature polypeptide is predicted to consist of 186 amino acid residues, giving a protein of molecular weight 20265 (Fig. 3).

Messenger RNA

A northern blot of total RNA isolated from various organs of light-grown peas was probed with radioactively-labelled cDNA insert. Bands of around 1.1 kb were obtained with RNA from young pods, petals, sepals and leaves whereas roots did not appear to contain any RNA capable of hybridizing to the 23 kDa cDNA probe (Fig. 4).

Genomic Southern blot

In a Southern blot of total genomic DNA digested with *Bam* HI hybridization was obtained to one fragment of around 6.6 kbp (Fig. 5) suggesting that the 23 kDa polypeptide is encoded by a single-copy gene in pea. In other genomic Southern blots hybridization was obtained to a 13.8 kb *Hin*d III fragment (results not shown).

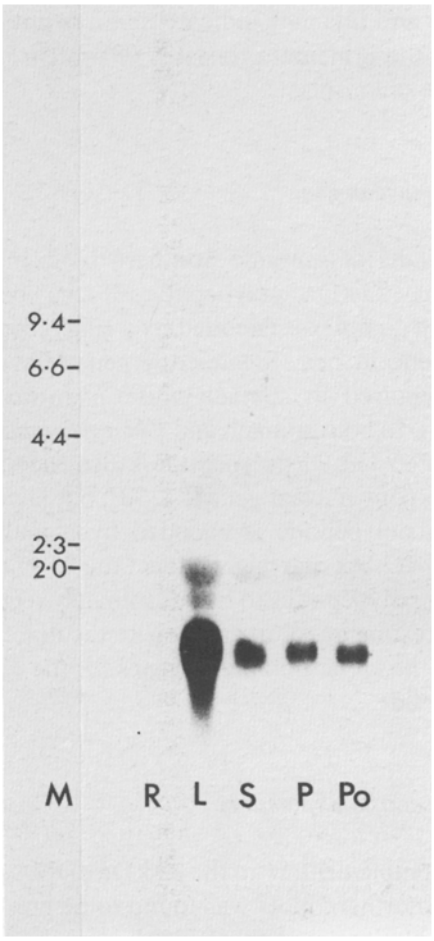


Fig. 4. Northern blot of total RNA (20 μg) from pea roots (R), leaves (L), sepals (S), petals (P) and young pods (Po) probed with the labelled cDNA insert for the 23 kDa polypeptide from λg23.3. Marker sizes (M) are in kb.

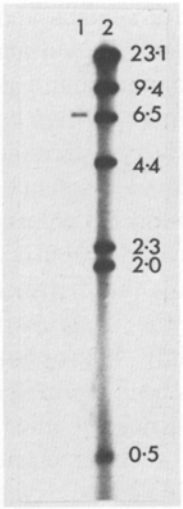


Fig. 5. Southern blot of pea genomic DNA restricted with *Bam* HI (lane 1) and run alongside λ DNA digested with *Hin*d III (lane 2) probed with the labelled cDNA insert for the 23 kDa polypeptide from λg23.3.

Discussion

Mature polypeptide

The derived amino acid sequence of the mature 23 kDa polypeptide from pea shows it to consist of 186 amino acid residues and predicts a molecular weight of 20265. This is slightly less than the molecular weight of 23000 predicted by gel electrophoresis and gel filtration [19]. The polypeptide is rich in charged amino acids (25.2%) with a total polar amino acid content of 62.8% and the primary structure is very similar to that determined for spinach [16] with 83% of the residues in the mature polypeptide being identical and most other changes being conservative (Fig. 3). The proportion of residues identical to the 23 kDa polypeptide from *Chlamydomonas* [25] is less at 58% but once again many of the changes are conservative.

N-terminal presequence

It has been shown by Westhoff *et al.* [43] that the 23 kDa polypeptide is synthesized from poly(A) RNA on cytosolic ribosomes as a precursor with a cleavable *N*-terminal extension. The long *N*-terminal extension presumably contains the information necessary for transferring the protein across both the chloroplast envelope and the thylakoid membrane. Comparison of the presequences of the three 23 kDa polypeptides so far elucidated [16, 25] (Fig. 3) reveals less conservation of primary structure than is seen in the mature polypeptides. The *N*-terminal extension of the pea 23 kDa polypeptide is 8 amino acids

shorter than that of spinach and 16 amino acids longer than that of *Chlamydomonas* with only 75% and 44% of the amino acids respectively being identical. There are two conserved regions, the first being the very *N*-terminus. This region is similar to the Box I sequence reported to be present in most imported chloroplast proteins by Karlin-Neumann and Tobin [18]. The second conserved region is that between residues Val49 and Ala73 and the latter part of this region consists of mainly hydrophobic residues. A similar region has been reported at the *C*-terminal end of the presequences of other nuclear-encoded lumenal proteins such as plastocyanin and the 33 kDa polypeptide of the oxygen-evolving complex and is thought to constitute a thylakoid transfer domain [15, 36, 41]. This region also bears strong similarities to the proposed uncleaved thylakoid transfer domain located at the *C*-terminus of the lumenal 10 kDa polypeptide of PSII [42]. The site specificity of the stromal peptidase, carrying out intermediate processing in the case of proteins bound for the thylakoid lumen is unknown but seems likely to recognize secondary structural features. Tyagi *et al.* [40] have suggested that a region of β-sheet, predicted for the presequences of all the spinach oxygen-evolving polypeptides may play a role in the intermediate processing step in the stroma. Such a structure is also predicted in pea at residues 24–30 (Arg-Gln-Val-Val-Thr-Thr-Lys) and this structure is immediately followed by a β-turn, which is also predicted in a similar position in both spinach [16] and *Chlamydomonas* [25] around residues 31 and 30 respectively. The conserved Box III of Karlin-Neumann and Tobin (Gly-Gly-Arg-Val) [18] likely to form a β-turn structure is also thought to be involved in the stromal processing step. In the presequences of the 23 kDa polypeptides of pea and spinach this could possibly be represented by the sequence Val-Cys-Lys-Ala (residues 35–38 in pea and residues 39–42 in spinach) while in *Chlamydomonas* residues 26–29, (Val-Val-Arg-Ala) may be equivalent. The decreased length of the *Chlamydomonas* presequence is due mostly to the absence of a region immediately prior to the proposed Box III sequence and this may indicate that it is not essential for the principal processes of protein import into the chloroplast.

Single nuclear gene

The results of genomic Southern blots indicate that the 23 kDa polypeptide of the oxygen-evolving complex is encoded by a single-copy nuclear gene in pea. Single-copy genes have also been reported in spinach and *Chlamydomonas* [16, 25]. In both spinach and *Chlamydomonas* the associated 33 kDa polypeptide is also encoded by a single-copy nuclear gene [24, 38], but in pea the 33 kDa polypeptide is encoded by a multigene family [41]. As one might expect the 23 kDa and 33 kDa polypeptides to be coordinately regulated an interesting question arises as to the function, if any, of the larger number of genes for the 33 kDa polypeptide.

Organ-specific expression

RNA complementary to the 23 kDa cDNA probe used in northern blots was found to be present in total RNA isolated from petals, sepals, young pods and leaves but not roots; the size of the RNA detected at around 1.1 kb corresponding well with the size of the cDNA and the mRNA detected previously in spinach [38]. The 23 kDa polypeptide RNA made up a much greater portion of the RNA in leaves than in other organs as might be expected. Probing western blots of total protein extracts of various organs with antibodies against the 23 kDa polypeptide showed that the polypeptide was present in all tissues containing the RNA and absent from the roots. The 23 kDa polypeptide was found to be present in the petals which are predominantly white although the polypeptide was probably located in the green veins and markings. The tissue specific-expression of the 23 kDa polypeptide which has so far shown a similar pattern of location to the 33 kDa polypeptide (results not shown) is similar to that for other nuclear-encoded chloroplast pro-

teins such as the chlorophyll *a*/*b*-binding proteins [21, 31] the small subunit of ribulose bisphosphate carboxylase [6, 31] and the 10 kDa extrinsic polypeptide of photosystem II [9, 22, 42].

The role of the 23 kDa polypeptide is still not clear. Its absence is known to increase the levels of Ca^{2+} and Cl^- required for the optimal activity of oxygen-evolving photosystem II preparations and this has lead to the speculation that it has a Ca^{2+} chelating role. However following the suggestion that the 33 kDa polypeptide is a calcium-binding protein it seems more likely that the 23 kDa polypeptide has its effect on the optimal Ca^{2+} levels by binding to the 33 kDa polypeptide. The absence of the 23 kDa polypeptide from cyanobacterial thylakoids has led to the suggestion that it may be involved in charge shielding and stacking mechanisms in higher plants. The primary structure of the mature 23 kDa polypeptide gives few clues to this role but the comparison of such data from more plant species may provide vital clues to the important regions of the 23 kDa polypeptide and therefore its function.

Acknowledgements

We would like to acknowledge Dr J.S. Gantt for the pea-leaf cDNA library. This work was supported by a Sainsbury Studentship from the Gatsby Charitable Foundation (R.W.) and by a Research Grant from the Science and Engineering Research Council.

References

1. Akabori K, Imaoka A, Toyoshima Y: The role of lipids and 17-kDa protein in enhancing the recovery of O_2 evolution in cholate-treated thylakoid membranes. FEBS Lett 173: 36–40 (1984).
2. Andersson B: Proteins participating in photosynthetic water oxidation. In: Staehlin LA, Arntzen CJ (eds) Photosynthesis III. Encyclopedia of Plant Physiology 19, pp. 447–456. Springer-Verlag, Berlin (1986).
3. Andersson B, Critchley C, Ryrie IJ, Jansson C, Larsson C, Anderson JM: Modification of the chloride requirement for photosynthetic O_2 evolution: The role of the 23-kDa polypeptide. FEBS Lett 168: 113–117 (1984).
4. Andersson B, Larsson C, Jansson C, Ljungberg U, Åkerlund H-E: Immunological studies on the organization of proteins in photosynthetic oxygen evolution. Biochim Biophys Acta 766: 21–28 (1984).
5. Birnboim HC, Doly J: A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7: 1512–1523 (1979).
6. Coruzzi G, Broglie R, Edwards C, Chua N-H: Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase. EMBO 3: 1671–1679 (1984).
7. Covey SN, Hull R: Transcription of cauliflower mosaic virus DNA: Detection of transcripts, properties and location of gene encoding the virus inclusion body protein. Virology 111: 463–474 (1981).
8. Dretzen G, Bellard M, Sassone-Corsi P, Chambon P: A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. Anal Biochem 112: 295–298 (1981).
9. Eckes P, Rosahl S, Schell J, Willmitzer L: Isolation and characterization of a light-inducible, organ-specific gene from potato and analysis of its expression after tagging and transfer into tobacco and potato shoots. Mol Gen Genet 205: 14–22 (1986).
10. Feinberg AP, Vogelstein B: A technique for radiolabelling restriction endonuclease fragments to high specific activity. Anal Biochem 132: 6–13 (1983).
11. Ford RC, Evans MCW: Isolation of a photosystem 2 preparation from higher plants with highly enriched oxygen evolution activity. FEBS Lett 160: 159–164 (1983).
12. Gantt JS, Key JL: Isolation of nuclear encoded plastid ribosomal protein cDNAs. Mol Gen Genet 202: 186–193 (1986).
13. Ghanotakis DF, Babcock GT, Yocum CF: Calcium reconstitutes high rates of oxygen evolution in polypeptide depleted photosystem II preparations. FEBS Lett 167: 127–130 (1984).
14. Gray JC, Dunn PPJ, Eccles CJ, Hird SM, Last DI, Newman BJ, Willey DL: Synthesis of electron transfer components of the photosynthetic apparatus. In: von Wettstein D, Chua N-H (eds) Plant Molecular Biology, pp. 105–113. Plenum, New York (1987).
15. 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).
16. 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).
17. Kalberer PP, Buchanan BB, Arnon DI: Rates of photosynthesis by isolated chloroplasts. Proc Natl Acad Sci USA 57: 1542–1549 (1967).

18. Karlin-Neumann GA, Tobin EM: Transit peptides of nuclear-encoded chloroplast proteins share a common amino acid framework. EMBO J 5: 9–13 (1986).
19. Kuwabara T, Murata N: Chemical and physiochemical characterization of the proteins involved in the oxygen evolution system. In: Sybesma C (ed) Advances in Photosynthetic Research I, pp. 371–374. Martinus Nijhoff/Dr W Junk, Dordrecht (1984).
20. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685 (1971).
21. Lamppa GK, Morelli G, Chua N-H: Structure and developmental regulation of a wheat gene encoding the major chlorophyll a/b-binding polypeptide. Mol Cell Biol 5: 1370–1378 (1985).
22. Lautner A, Klein R, Ljungberg U, Reilander H, Bartling D, Andersson B, Reinke H, Beyreuther K, Herrmann RG: Nucleotide sequence of cDNA clones encoding the complete precursor for the 10-kDa polypeptide of photosystem II from spinach. J Biol Chem 263: 10077–10081 (1988).
23. Maniatis T, Fritsch EF, Sambrook J: Rapid, small scale isolation of bacteriophage λ DNA. In: Molecular Cloning: A Laboratory Manual p. 371. Cold Spring Harbor Press, Spring Harbor, NY (1982).
24. Mayfield SP, Bennoun P, Rochaix J-D: 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).
25. Mayfield SP, Rahire M, Frank G, Zuber H, Rochaix J-D: 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).
26. McMaster GK, Carmichael GC: Analysis of single- and double-stranded nucleic acids on polyacrylamide and agarose gels by using glyoxal and acridine orange. Proc Natl Acad Sci USA 74: 4835–4838 (1977).
27. Miyao M, Murata N: Calcium ions can be substituted for the 24 kDa polypeptide in photosynthetic oxygen evolution. FEBS Lett 168: 118–120 (1984).
28. Miyao M, Murata N: Role of the 33-kDa polypeptide in preserving Mn in the photosynthetic oxygen-evolution system and its replacement by chloride ions. FEBS Lett 170: 350–354 (1984).
29. Murata N, Miyao M: Extrinsic membrane proteins in the photosynthetic oxygen-evolving complex. Trends Biochem Sci 10: 122–124 (1985).
30. Nakatani HY: Photosynthetic oxygen evolution does not require the participation of polypeptides of 16 and 24 kilodaltons. Biochem Biophys Res Comm 120: 299–304 (1984).
31. Piechulla B, Pichersky E, Cashmore AR, Gruissem W: Expression of nuclear and plastid genes for photosynthesis-specific proteins during tomato fruit development and ripening. Plant Mol Biol 7: 367–376 (1986).
32. Reed KC, Mann DA: Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res 13: 7207–7221 (1985).
33. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5469 (1977).
34. Sheen J-Y, Sayre RT, Bogorad L: Differential expression of oxygen-evolving polypeptide genes in maize leaf cell types. Plant Mol Biol 9: 217–226 (1987).
35. Southern EM: Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503–517 (1975).
36. Smeekens S, Bauerle C, Hageman J, Keegstra K, Weisbeek P: The role of the transit peptide in the routing of precursors toward different chloroplast compartments. Cell 46: 365–375 (1986).
37. Stewart AC, Ljungberg U, Åkerlund H-E, Andersson B: Studies on the polypeptide composition of the cyanobacterial oxygen-evolving complex. Biochim Biophys Acta 808: 353–362 (1985).
38. Tittgen J, Hermans J, Steppuhn J, Jansen T, Jansson C, Andersson B, Nechushtai R, Nelson N, Herrmann RG: Isolation of cDNA clones for fourteen nuclear-encoded thylakoid membrane proteins. Mol Gen Genet 204: 258–265 (1986).
39. Towbin H, Staehelin T, Gordon J: Electrophoretic transfer of proteins from acrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc Natl Acad Sci USA 76: 4350–4354 (1979).
40. Tyagi A, Hermans J, Steppuhn J, Jansson C, Vater F, Herrmann RG: Nucleotide sequence of cDNA clones encoding the complete '33kDa' precursor protein associated with the photosynthetic oxygen-evolving complex from spinach. Mol Gen Genet 207: 288–293 (1987).
41. Wales R, Newman BJ, Pappin D, Gray JC: The extrinsic 33-kDa polypeptide of the oxygen-evolving complex of photosystem II is a putative calcium-binding protein and is encoded by a multi-gene family in pea. Plant Mol Biol 12: 439–451 (1989).
42. Webber AN, Packman LC, Gray JC: A 10-kDa polypeptide associated with the oxygen-evolving complex of photosystem II has a putative *C*-terminal noncleavable thylakoid transfer domain. FEBS Lett 242: 435–438 (1989).
43. Westhoff P, Jansson C, Klein-Hitpaß L, Berzborn R, Larsson C, Bartlett SG: Intracellular coding sites of polypeptides associated with photosystem II. Plant Mol Biol 4: 137–146 (1985).
44. Young RA, Davis RW: Yeast RNA polymerase II genes: Isolation with antibody probes. Science 222: 778–782 (1983).