DNA RESEARCH 5, 221–228 (1998)

The *psbO* Gene for 33-kDa Precursor Polypeptide of the Oxygen-Evolving Complex in *Arabidopsis thaliana* — Nucleotide Sequence and Control of its Expression

Pradeep K. JAIN, Anju KOCHHAR, Jitendra P. KHURANA, and Akhilesh K. TYAGI*

Centre for Plant Molecular Biology and Department of Plant Molecular Biology, University of Delhi South Campus, New Delhi-110021, INDIA

(Received 30 June 1998)

Abstract

The 33-kDa polypeptide of the oxygen-evolving complex of photosystem II is nuclear-encoded. The single *psbO* gene of *Arabidopsis thaliana*, as suggested by Southern hybridization, has been isolated from the genomic library and sequenced. The sequence analysis has revealed that the *psbO* gene harbors two introns and encodes a precursor polypeptide of 332 amino acid residues; the first 85 amino acid residues represent the transit peptide and the following 247 amino acids constitute the mature polypeptide. The hydrophilic nature of the 33-kDa protein is confirmed by the presence of 27% charged residues. Northern analysis of the total RNA from *Arabidopsis* indicates that a 1.2-kb transcript represents the *psbO* gene. It is expressed in a tissue-specific manner — the steady-state transcript levels being highest in the leaves and virtually undetectable in the roots. Also, expression of the *psbO* gene is development-dependent and regulated by light in young *Arabidopsis* seedlings. In a constitutively photomorphogenic mutant of *Arabidopsis*, *pho*2 (*p*lumular *h*ook *o*pen 2), the *psbO* gene is de-repressed in young, dark-grown seedlings, resulting in increased transcript abundance compared to the wild-type. These studies, thus, define the influence of at least one regulatory component for *psbO* expression.

Key words: *Arabidopsis thaliana*; gene expression; oxygen-evolving complex; photomorphogenic mutant; *psbO* gene

1. Introduction

The photosystem II complex of chloroplasts plays a vital role for the sustenance of life on earth since it is responsible for the light-driven transfer of electrons from water to plastoquinone pool, resulting in the evolution of oxygen. It consists of three sub-complexes: a core complex harboring the reaction centre, the light-harvesting chlorophyll-protein complex (LHCII), and water-oxidizing/oxygen-evolving complex. It is well established that, although the chloroplast harbors a genome of its own, more than 50% of the resident polypeptides of thylakoids are encoded by the nucleus.[1] The nuclear-encoded proteins are synthesized in the cytoplasm as precursor proteins with variable transit peptide sequences that target the proteins into the chloroplast at their respective sites of action. The oxygen-evolving/water-oxidizing complex situated on the lumenal side of the thylakoid membrane consists of at least three polypeptide species with molecular weights of 33-, 23- and 16-kDa.[2,3,4] The 23- and 16-kDa polypeptides are present in higher plants and green algae but are absent in cyanobacteria. However, the 33-kDa protein is present in all oxygen-evolving organisms, including cyanobacteria, underlining its importance in photosynthetic oxygen evolution. It has been suggested that the 23- and 16-kDa polypeptides play regulatory roles and are not directly involved in water oxidation.[3] The 33-kDa polypeptide has been implicated in the stabilization of the tetranuclear manganese centre.[4,5] Furthermore, the *Chlamydomonas reinhardtii* mutant which lacks the 33-kDa polypeptide has no oxygen evolution despite having normal levels of 23- and 16-kDa polypeptides.[6,7] Thus, the 33-kDa polypeptide appears to play a central role in the assembly and function of PSII.

Given the importance of the 33-kDa polypeptide, it is not surprising that the cDNA clones have been isolated and characterized from a number of higher plants and cyanobacteria. The first cDNA clone for the 33-kDa polypeptide was isolated from spinach.[8] Subsequently,

Communicated by Masahiro Sugiura

* To whom correspondence should be addressed. Tel. +91-11-4673216, 4671208, Fax. +91-11-6886427, E-mail: pmb@dusc.ernet.in

† The nucleotide sequence data reported in this paper will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number Y15433.

cDNA clones have been isolated and characterized from a number of plants including *Arabidopsis*,[9] potato,[10] wheat,[11] pea,[12] tomato[13] and tobacco.[14] Surprisingly, to the best of our knowledge, the genomic clone for the 33-kDa polypeptide has not been isolated so far from any of the higher plants. However, the genomic clone has been isolated from *Synechocystis* 6803.[15] Expression of the *psbO* gene has been shown to be regulated by light at the transcriptional as well as translational levels in maize[16,17] and tomato.[9] A dark-adaptive decrease in transcript is also clearly observed in *Arabidopsis* and tomato, 3 days after the transfer of light-grown seedlings to dark.[9] In addition, the *psbO* transcript and the protein product exhibit tissue-specific differences in their steady-state levels in tomato.[9]

In the present investigation, the complete *psbO* gene of *Arabidopsis thaliana* was isolated and sequenced. The effect of certain intrinsic and extrinsic factors on the expression of this gene has been analysed. The steady-state transcript levels of *psbO* have also been examined in a constitutively photomorphogenic mutant, *pho*2 (*p*lumular *h*ook *o*pen 2), of *A. thaliana*. The *pho*2 mutant, like *pho*1, exhibits a few characteristics of the light-grown seedlings even when grown in dark and also flowers precociously under normal light conditions.[18,19] Allelism tests have shown that the mutant *pho*2 represents a novel locus not described earlier,[18,20] although it resembles phenotypically to some of the less pleiotropic *cop* mutants.[21]

2. Materials and Methods

A heterologous cDNA probe from spinach[8] was used to screen a genomic library of *A. thaliana*, ecotype Columbia, plated at a density of 50,000 pfu per 22.5×22.5 cm^2 culture dish (Nunc, Denmark). Replica filters were prepared from the plate and subjected to prehybridization and hybridization exactly as described for analysis of eukaryotic genomic DNA in Sambrook et al.[22] The probe was labelled using the Megaprime DNA Labelling Kit (Amersham International Inc., UK). The clones identified during primary screening were subjected to secondary, tertiary and quaternary screening using identical conditions, except that the clones were plated at a low density in 9.0-cm Petri dishes. After final screening, stock lysates were prepared from the purified clones and phage DNA isolated.[23] Subcloning of gene-specific restriction fragments was done in pBluescript SK$^+$ vector using standard protocols. Sequencing was done using T3 and T7 primers with an automated DNA sequencer (ABI, Prism).

The growth conditions of *A. thaliana*, ecotype Estland, and its photomorphogenic mutant, *pho*2, have been described earlier.[19] The procedures for DNA isolation,[24] Southern hybridization, RNA extraction[25] and northern analysis have also been given previously.[19] Each experiment was repeated at least twice with similar results but the data of only a representative experiment are given.

3. Results and Discussion

3.1. Isolation and characterization of the psbO *gene*

The genomic library of *A. thaliana* in EMBL3 (bacteriophage lambda) replacement vector was screened with a 1.2-kb *Eco*RI fragment from a *psbO* cDNA clone of spinach[8] to identify the corresponding gene from *Arabidopsis*. In the first round of screening, 12 distinct signals were identified. The number was reduced to eight and three after secondary and tertiary screening, respectively. The purity of these three clones was confirmed in quaternary screening and they were designated as *At 33a*, *At 33b* and *At 33c* (for *A. thaliana*, 33-kDa polypeptide). The DNA of all the three clones was isolated and restricted with *Eco*RI, *Bam*HI, *Kpn* I, *Pst* I, *Sal* I and *Sac* I. Upon Southern analysis with the heterologous cDNA probe, various gene-specific fragments of varying sizes were identified in the digested samples. However, *Sal* I-digested DNA resulted in only one gene-specific fragment of about 4.3-kb in all the three clones. This fragment from all the three phage clones was subcloned and the resultant plasmids were designated as *pAt 33a*, *pAt 33b* and *pAt 33c*. The recombinant plasmids were digested with *Bam*HI, *Kpn* I, *Pst* I, *Sac* I and *Sal* I to determine the distribution of restriction sites within the fragment hybridizing to the heterologous probe. Surprisingly, the digestion pattern of the recombinant plasmids containing genomic DNA did not reveal any new restriction sites for these enzymes besides those already known to be present in the earlier reported *psbO* cDNA clone from *Arabidopsis*.[9] Similar to the cDNA clone of *Arabidopsis*, each of these enzymes had only one site in the genomic clones. The Southern blot revealed essentially a similar pattern for all the three clones based on which restriction maps of these clones were prepared. One gene-specific fragment each of size 4.3-kb, 4.0-kb and ~1-kb was obtained on digestion with *Sal* I, *Kpn* I and *Sac* I, respectively. Two positive fragments each were observed on restriction with *Bam*HI (~6.8-kb and 0.4-kb) and *Pst* I (~6.6-kb and 0.6-kb). It was also evident from the hybridization pattern that the coding region of the gene is located on one end of the cloned fragments and also that the three clones may represent the same gene with some additional region at the 3′ ends.

Genomic Southern blot probed with a 1.0-kb fragment, representing the 3′ end of the gene, revealed the presence of two gene-specific bands (~5.7- and 1.8-kb) in *Bam*HI digested sample whereas a single band each of ~9.5- and 6.3-kb appeared in *Hin*dIII and *Sal* I digests, respectively (Fig. 1). Although this pattern also conforms to the possible existence of a single *psbO* gene in *A. thaliana* as suggested earlier,[9] an additional faint fragment of ~4.6-kb of

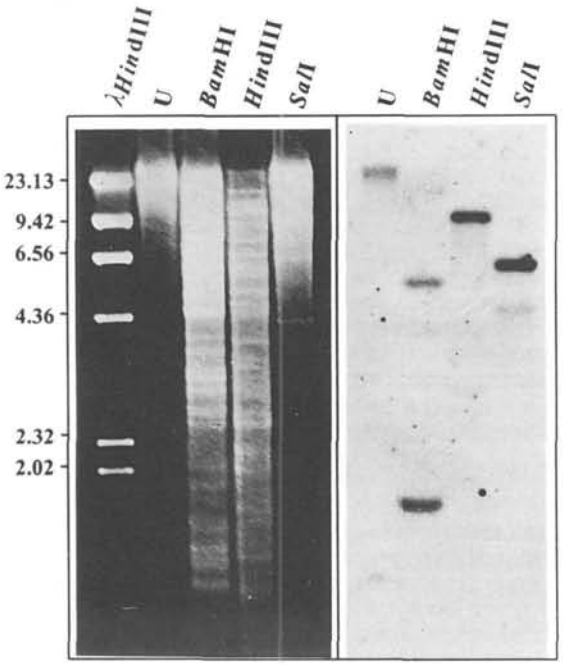


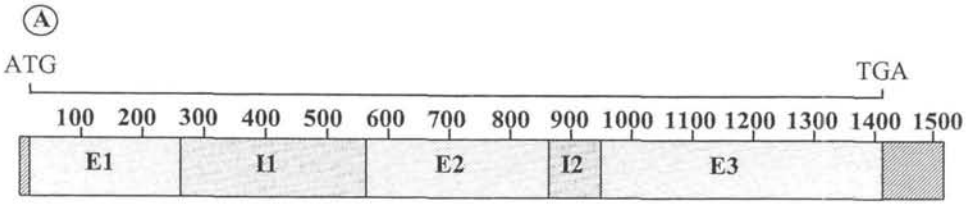
Figure 1. Southern analysis of *Arabidopsis* genomic DNA. Undigested and restricted DNA samples have been resolved on 1% agarose gel and processed for Southern analysis. The right panel shows an autoradiogram obtained after hybridization of the blot from the gel shown in the left panel with the gene-specific 1.0-kb fragment. The fragments obtained on digestion of λDNA with *Hin*dIII serve as DNA fragment length markers whose sizes are given on the left.

unknown origin was also visible in *Sal*I-digested sample on long exposures. Further, the possibility of an additional homologous sequence in *A. thaliana* has also been pointed out.[9]

Since all the three phage clones were giving nearly the same restriction pattern as well as the gene-specific fragments were also of nearly same molecular range, it was decided to subclone the fragments for only two clones, i.e. *pAt 33b* and *pAt 33c*. The restriction with *Bam*HI, *Pst*I and *Sac*I allowed generation of serially deleted clones from the 3′ end of the 4.3-kb *Sal*I fragment. These subcloned fragments, besides the original 4.3-kb *Sal*I fragment, were used for determining the nucleotide sequence of *psbO* gene. Since only a few nucleotides from the 3′ untranslated region of the gene were present in *pAt 33c*, it was thought worthwhile to sequence the *pAt 33b* clone which shows a marginally larger gene-specific fragment in Southern blots. The sequence obtained provided information about 104 bp of the 3′ untranslated region. The sequence comparison of *pAt 33b* and *pAt 33c* revealed 100% sequence identity over a span of 900 nucleotides (only these many base pairs have been sequenced for the clone *pAt 33b*) which again suggests that these two clones harbor the same gene.

Like many eukaryotic genes, the *psbO* gene has a split structure consisting of 3 exons and 2 introns (Fig. 2A) as determined by comparison with the cDNA sequence reported by Ko et al.[9] The nucleotide sequence of the complete *psbO* gene (1520 bp) includes 33 bp of 5′ and 104 bp of 3′ untranslated regions, respectively (Fig. 2B). The GT-AG consensus, found at the 5′ and 3′ splice sites, respectively, of most introns[26] is conserved in both the introns (297 and 87 bp) of *psbO* gene of *Arabidopsis*. However, the introns of *psbO* gene do not possess a pronounced polypyrimidine tract which is essential in vertebrate intron splicing,[27] but there are reports which do suggest that plant intron sequences do not have a very pronounced polypyrimidine tract.[28] The AT complements of introns I and II are 67.3% and 57.4%, respectively, a feature which has been suggested as a potential signal for intron recognition and accessibility of intron sequences to the splicing machinery[29,30,31] without involving a branchpoint motif. However, a recent study provides a convincing and direct evidence for the presence of branchpoint sequences for the efficient splicing of plant introns.[32] Similarly, both the introns of *psbO* gene of *Arabidopsis* contain optimal branchpoint sequence UUGAU, based on animal branchpoint consensus YUNAN [positioned between 20 and 60 nucleotides (nt) of the 3′ splice site], found in 80% of the plant introns examined. Moreover, the putative branchpoint adenosine at 34 and 20 nt upstream of the 3′ splice site for intron I and II, respectively, is also in accord with an earlier observation which indicates that branchpoint adenosine occurs at a mean position of 27 nt upstream of the 3′ splice site.[32] The coding sequence of the gene represented by three exons shows six nucleotide changes in comparison with the cDNA clone of *Arabidopsis*.[9] Four of these do not change the encoded amino acids, but codons 113 and 114 are changed to code for the amino acids glutamine and cysteine, which are in agreement with the sequences reported for other species.[8,12]

The *psbO* gene has an optimal context (GCCATGG) of initiation codon required for translation initiation, as defined by Kozak,[33] according to which a purine (A/G) and a G nucleotide occupy the positions −3 and +4, respectively, with respect to A (+1) of the initiation codon (ATG). The gene has a coding potential for a precursor polypeptide of 332 amino acids (deduced M.W., 35.2-kDa) of which the first 85 amino acids (8.6-kDa) represent transit peptide followed by 247 amino acid (26.6-kDa) long mature polypeptide. The codon usage is characteristic of eukaryotic genes as evident from the more frequent (32.73%) occurrence of nucleotide C at the third position of the codon compared to that of T (22.82%) at analogous position. The amino terminus transit peptide contains the information necessary for translocating the protein across both the chloroplast envelope and the thylakoid membrane. A protein like 33-kDa which resides in the thylakoid lumen is first imported into the stromal compartment and then



B

```
   1  GAAAAAAAAAAACTCAAGAGAGACTTTGTGGCCATGGCAGCCTCTCTCCAATCCACCGCT      60
                                        M  A  A  S  L  Q  S  T  A
  61  ACATTCCTCCAGTCGGCGAAGATCGCCACCGCTCCTTCTCGCGGAAGTTCTCACCTCCGA     120
       T  F  L  Q  S  A  K  I  A  T  A  P  S  R  G  S  S  H  L  R
 121  TCGACTCAAGCCGTCGGCAAATCTTTTGGGCTCGAAACTTCCTCGGCTCGCCTCACTTGC     180
       S  T  Q  A  V  G  K  S  F  G  L  E  T  S  S  A  R  L  T  C
 181  TCCTTCCAGTCTGACTTTAAGGACTTCACCGGTAAATGCTCCGACGCTGTCAAAATCGCC     240
       S  F  Q  S  D  F  K  D  F  T  G  K  C  S  D  A  V  K  I  A
 241  GGATTCGCTCTTGCCACCTCTGCTCTCGTCGTCTCGGTAACTTTCTTTAATCTCCGACTC     300
       G  F  A  L  A  T  S  A  L  V  V  S
 301  TCCTTTAAAACTATAGATTAATCAAGAAAAAAAAGGAGAATTGAGTAACAGAGAAGAAGA     360
 361  GTATTTTGTTGTGTTTGTGATAACATGTTTATGATTTGCATTACTATGTATCGAAGTTGG     420
 421  CACTTGAGAACACCTAAACCAATAACAGACTCGGATAGTATAATAGAAGGAGAAACCGCT     480
 481  ATTCCAATTCTGTAATCTTCGATTCATGTAGATACAATCTTCTTGGATCAGTGAGTTTGA     540
 541  TTTAAAGTTGCTTGTTTCTCGTGTAATGAAAAGGGAGCAAGTGCGGAGGGAGCTCCAAAG     600
                                        G  A  S  A  E  G  A  P  K
 601  AGATTGACTTATGATGAGATTCAGAGCAAGACATACATGGAAGTGAAAGGAACTGGAACG     660
       R  L  T  Y  D  E  I  Q  S  K  T  Y  M  E  V  K  G  T  G  T
 661  GCTAACCAGTGCCCTACTATTGACGGTGGCTCTGAGACTTTCTCGTTCAAACCCGGAAAG     720
       A  N  Q  C  P  T  I  D  G  G  S  E  T  F  S  F  K  P  G  K
 721  TATGCAGGAAAGAAGTTCTGCTTCGAGCCTACTTCCTTCACCGTCAAGGCAGACAGTGTA     780
       Y  A  G  K  K  F  C  F  E  P  T  S  F  T  V  K  A  D  S  V
 781  AGCAAGAACGCTCCTCCAGAGTTCCAGAACACAAAGCTCATGACCCGTCTTACCTACACC     840
       S  K  N  A  P  P  E  F  Q  N  T  K  L  M  T  R  L  T  Y  T
 841  CTTGACGAGATCGAAGGACCCTTCGAGGTTCGATTTCGCGTCTTTGCTCTAGACTGCAAG     900
       L  D  E  I  E  G  P  F  E
 901  AAAAAACTCTGCTTAGGGTTAGGGTTTGTGATTGATCGTCTCTTTAATGAGCAGGTTGCT     960
                                                          V  A
 961  TCAGACGGAAGCGTCAATTTCAAGGAAGAAGATGGAATCGACTATGCTGCAGTCACAGTC    1020
       S  D  G  S  V  N  F  K  E  E  D  G  I  D  Y  A  A  V  T  V
1021  CAACTTCCAGGAGGTGAACGTGTTCCATTCCTTTTCACAGTCAAACAGCTTGACGCCTCA    1080
       Q  L  P  G  G  E  R  V  P  F  L  F  T  V  K  Q  L  D  A  S
1081  GGCAAACCAGACAGCTTCACCGGAAAATTCTTGGTTCCATCGTACCGTGGCTCTTCCTTC    1140
       G  K  P  D  S  F  T  G  K  F  L  V  P  S  Y  R  G  S  S  F
1141  TTGGACCCAAAGGGCCGTGGTGGATCCACAGGATATGACAACGCCGTGGCATTGCCAGCT    1200
       L  D  P  K  G  R  G  G  S  T  G  Y  D  N  A  V  A  L  P  A
1201  GGAGGCAGAGGAGACGAGGAGGAGCTTGTAAAGGAGAACGTGAAGAACACTGCCGCTTCA    1260
       G  G  R  G  D  E  E  E  L  V  K  E  N  V  K  N  T  A  A  S
1261  GTGGGAGAGATCACTCTGAAAGTGACTAAGAGCAAGCCGGAGACAGGAGAGGTGATCGGA    1320
       V  G  E  I  T  L  K  V  T  K  S  K  P  E  T  G  E  V  I  G
1321  GTGTTCGAGAGTCTTCAGCCGTCGGACACTGACTTGGGTGCTAAGGTACCAAAGGATGTG    1380
       V  F  E  S  L  Q  P  S  D  T  D  L  G  A  K  V  P  K  D  V
1381  AAGATCCAAGGGGTGTGGTATGGTCAACTTGAGTGATCATGTTATTATATTTTCCGTTGA    1440
       K  I  Q  G  V  W  Y  G  Q  L  E  *
1441  TTGTGTGTGATGATAATGATAACATCTTTTGATGCTTTCTTCGTTTATCTCTCTTATATA    1500
1501  AACTGCACATGCCTAAGATC
```

Figure 2. A. Diagrammatic representation of the genomic fragment (1.52-kb) harboring the *psbO* gene in *Arabidopsis*. The hatched boxes represent the untranslated regions of exons, and I1 and I2 represent the two introns interspersed between the three exons, E1, E2 and E3. The numbers represent the size of the fragments in base pairs. B. Nucleotide sequence and deduced amino acid sequence (single letter code) of the *psbO* gene of *Arabidopsis* encoding the 33-kDa polypeptide of the oxygen-evolving complex of photosystem II. Amino acid residues involved in terminal processing of the transit peptide are shown in bold. The amino acid residues most likely involved in Mn^{2+}-binding are underlined and shown in bold. The Ca^{2+}-binding domain is indicated by the underlined amino acid residues (normal). The nucleotides representing introns are shown in italic.

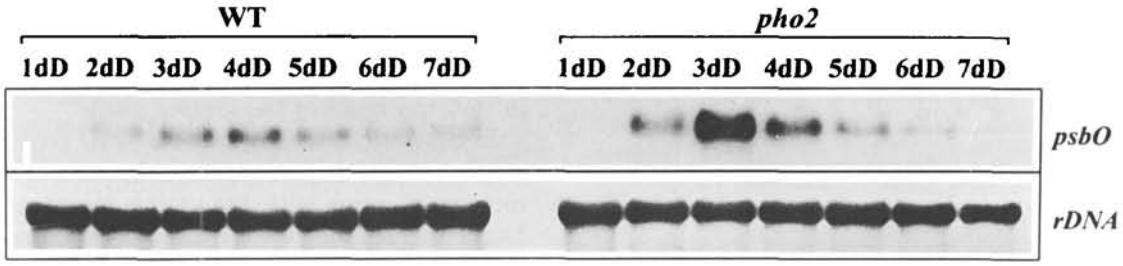


Figure 3. Development-dependent transcript accumulation of the *psbO* gene in the dark-grown wild-type and *pho2* mutant seedlings. RNA was isolated from seedlings grown in dark for 1–7 days and processed for northern analysis. The blot was hybridized with the 1.0-kb *psbO* gene-specific probe. The *rDNA* probe was used as a control to check the quality and quantity of RNA samples prepared.

further channeled into the thylakoid.[34] The signal peptide responsible for routing proteins to the thylakoid lumen are characterized by a short amino-terminal positively charged region, a central apolar region consisting of 7–15 residues and a carboxy-terminal region typically 5–6 residues long, with small, uncharged amino acids at positions −3 and −1, counting from the cleavage site.[35,36] The 33-kDa signal peptide fulfills all the above requirements; it has a charged amino-terminus region rich in serine and threonine and an apolar region comprising of AVKIAGFAFATSALVVS. The cleavage site in the 33-kDa precursor polypeptide is preceded by A-S-A, which agrees well with the consensus sequence (A-X-A).

The hydropathy profile generated for the 33-kDa protein indicates a predominantly hydrophilic protein,[9] which is confirmed by the frequency of charged residues (His, Lys, Asp, Glu and Arg) of 26.3%. The deduced sequence is 72% homologous to precursor of the 33-kDa polypeptide from spinach.[8] A 27-amino acid residue long calcium-binding domain similar to that found in pea and spinach 33-kDa polypeptides[8,12] is also present in the 33-kDa polypeptide of *Arabidopsis* (Fig. 2B). The exact function of this domain has not been discerned yet. Although Ca^{2+} is required for oxygen evolution, it has been suggested that the Ca^{2+} binding at the putative Ca^{2+}-binding domain in the 33-kDa polypeptide (as inferred from the sequence homologies to known Ca^{2+}-binding proteins[12,37,38]) may not be involved in oxygen evolution.[39] In comparison with the other known sequences for 33-kDa polypeptides (spinach[8] and pea[12]), a region speculated in binding Mn atoms has also been delineated in the *Arabidopsis* 33-kDa polypeptide (amino acid residues 103 to 120). The involvement of 33-kDa polypeptides in the stabilization of Mn atoms associated with the water-splitting complex has been suggested by several workers;[4,5] however, convincing evidence to confirm such a role is still lacking.

3.2. *Development-dependent expression*

The *psbO* gene encodes a transcript of about 1.2-kb in *A. thaliana*. To understand the regulation of *psbO* expression as influenced by endogenous developmental cues, its steady-state transcript levels were examined in 1- to 7-day-old seedlings grown in the dark. The accumulation of *psbO* transcript was found to be development-dependent and the transcript levels increased significantly with growth in the dark, attaining normally a peak in the 3-day-old seedlings (Fig. 3) followed by a steady decline. At the same time, no significant alteration in the control 25S rRNA was observed. The development-dependent accumulation of *psbO* transcripts was also checked in one of the constitutively photomorphogenic mutants, *pho2* (*p*leiomular *h*ook *o*pen 2), which displays light-grown phenotype in the dark. Although the steady-state transcript levels displayed a development-dependent accumulation pattern analogous to that of the wild-type, the levels were much higher in *pho2*, in comparison to the wild-type, and the difference was particularly striking in 3-day-old seedlings (Fig. 3). It thus appears that *PHO2* product plays an important role in regulating the transcript abundance of *psbO* gene in young dark-grown *Arabidopsis* seedlings. Studies from our laboratory have also shown that PHO2 (and PHO1) also regulates the transcript abundance of other genes encoding polypeptides of the photosynthetic apparatus, both of nuclear (e.g. *rbcS* and *cab*) and plastid (e.g. *psbA* and *rbcL*) origin.[18,19,20] In addition, several other cellular effectors (e.g. COP1 and DET1) have also been shown to cause severe derepression of photosynthesis-related genes in young dark-grown *Arabidopsis* seedlings.[20,40]

3.3. *Light-induced changes in transcript abundance*

The 4-day-old dark-grown seedlings were subjected to increasing durations of white light illumination and analyzed for *psbO* transcript abundance. In wild-type *Arabidopsis* seedlings, a prominent increase in transcript level was detectable with as little as 2 hr of illumination, and a steady-state was attained within 4–8 hr of illumination (Fig. 4). Thus, the light-regulated nature of *psbO* is evident by the increase in steady-state transcript levels during light-induced greening. The regulation of *psbO* transcript by light has also been reported

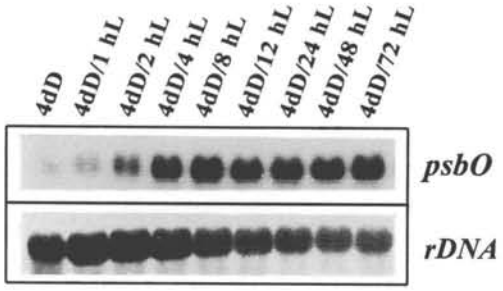


Figure 4. Light-dependent accumulation of *psbO* gene transcripts in young *Arabidopsis* seedlings. The 4-day-old dark-grown (4 dD) seedlings raised on MS medium (without sucrose) were irradiated with white light for various durations (hL, hours light). Total RNA (10 μg) from each sample was loaded per lane of the gel and subjected to northern analysis. The blot was hybridized with the 1.0-kb *psbO* gene-specific probe. The *rDNA* probe was used as a control to check the quality of RNA samples.

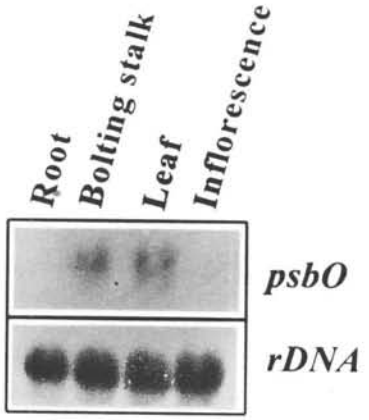


Figure 5. Organ-specific expression of *psbO* gene of *Arabidopsis*. Mature (36 days old) plants were used for isolating RNA from roots, leaves, bolting stalks and inflorescence. For each sample, 15 μg of RNA was processed for northern analysis, using a 1.0-kb fragment of the *psbO* gene. The *rDNA* probe was used as a control to check the quality of RNA samples.

earlier in tomato and *Arabidopsis*.[9] Similar observations have earlier been made in etiolated maize seedlings, where both *psbO* transcript and the corresponding protein levels increased several-fold when subjected to light treatment.[16,17]

3.4. Organ-specific expression

The distribution of 33-kDa gene transcripts was also examined amongst different organs of the *Arabidopsis* plant. Total RNA was isolated from roots, bolting stalks (without the cauline leaves), rosette leaves and inflorescence of 36-day-old wild-type plants grown under continuous light conditions, and analyzed for *psbO* transcript levels. Northern analysis revealed an organ-specific pattern of *psbO* transcript accumulation, with the transcript being most abundant in leaves, followed by bolting stalks, and was not detected in inflorescence and roots (Fig. 5). The differential expression of *psbO* has also been reported in various organs of the tomato plant.[9]

In conclusion, the *A. thaliana psbO* gene isolated and sequenced in this investigation shows near-complete homology of exons with the cDNA clone[9] and has been found to contain two introns. The expression of the *psbO* gene is organ-specific and is regulated by intrinsic and extrinsic factors. In this context, the influence of PHO2 (one of the cellular effectors of the PHO class[18,20]) on the expression of *psbO* gene deserves consideration. Knocking out the function of this effector, as in the *pho2* mutant, leads to de-repression of gene expression, thereby indicating direct involvement of PHO2 in repression of *psbO* expression in young, dark-grown *Arabidopsis* seedlings. With this basic and essential information available about the *psbO* gene, efforts can now be directed towards sequencing the upstream promoter region. Analysis for delineating specific regulatory elements can also be carried out.

Acknowledgements: We thank Dr. R. Oelmuller for providing us the genomic library of *Arabidopsis thaliana*, ecotype Columbia, and Dr. Subba Rao, In-charge of Bioinformatics Centre, JNU, for help in analysis of the gene sequence. This research was funded by the Departments of Science and Technology, and Biotechnology, of the Government of India, and the University Grants Commission, New Delhi. PKJ and AK acknowledge the award of Senior Research Fellowship by the Council of Scientific and Industrial Research, New Delhi.

References

1. Herrmann, R. G. 1996, photosynthesis research: aspects and perspectives, In: Andersson, B., Salter, A. H. and Barber, J. (eds), Molecular Genetics of photosynthesis, Oxford University Press, New York, pp. 1–44.
2. Ljungberg, U., Akerlund, H.-E., and Andersson, B. 1986, Isolation and characterization of the 10-kDa and 22-kDa polypeptides of higher plant photosystem 2, *Eur. J. Biochem.*, **158**, 477–482.
3. Ghanotakis, D. F., Topper, J. N., Babcock, G. T., and Yocum, C. F. 1984, Water soluble 17 and 23 kDa polypeptides restore oxygen evolution activity by creating a high-affinity binding site for Ca^{2+} on the oxidizing site of photosystem II, *FEBS Lett.*, **170**, 169–173.
4. Ono, T.-A. and Inoue, Y. 1983, Mn-preserving extraction of 33-, 24- and 16-kDa proteins from O_2-evolving PSII particles by divalent salt-washing, *FEBS Lett.*, **164**, 255–260.
5. Yamamoto, Y., Shinaki, H., Isogai, Y., Matsura, K., and Nishimura, M. 1984, Isolation of an Mn-carrying 33 kDa protein from an oxygen-evolving photosystem II preparation by phase partitioning with butanol, *FEBS Lett.*, **175**, 429–432.
6. Mayfield, S. P., Bennoun, P., and Rochaix, J. D. 1987,

Expression of the nuclear-encoded OEE1 protein is required for oxygen evolution and stability in photosystem II particles in *C. reinhardtii, EMBO J.*, **6**, 313–318.
7. Rochaix, J.-D. and Erickson, J. 1988, Function and assembly of photosystem II: genetic and molecular analysis, *Trends Biochem. Sci.*, **13**, 56–59.
8. Tyagi, A., Hermans, J., Steppuhn, J., Jansson, Ch., Vater, F., and Herrmann, R. G. 1987, 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.
9. Ko, K., Granell, A., Bennett, J., and Cashmore, A. R. 1990, Isolation and characterization of cDNAs from *Lycopersicon esculentum* and *Arabidopsis thaliana* encoding the 33 kDa protein of the photosystem II-associated oxygen-evolving complex, *Plant Mol. Biol.*, **14**, 217–227.
10. van Spanje, M., Dirkse, W. G., Nap, J.-P., and Stiekema, W. J. 1991, Isolation and analysis of cDNA encoding the 33 kDa precursor protein of the oxygen-evolving complex of potato, *Plant Mol. Biol.*, **17**, 157–160.
11. Meadows, J. W., Hulford, A., Raines, C. A., and Robinson, C. 1991, Nucleotide sequence of a cDNA clone encoding the precursor of the 33 kDa protein of the oxygen-evolving complex from wheat, *Plant Mol. Biol.*, **16**, 1085–1087.
12. Wales, R., Newman, B. J., Pappin, D., and Gray, J. C. 1989, 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.
13. Goerlach, J., Schmid, J., and Amrhein, N. 1993, The 33 kDa protein of the oxygen-evolving complex: a multigene family in tomato, *Plant Cell Physiol.*, **34**, 479–501.
14. Palomares, H., Herrmann, R. G., and Oelmuller, R. 1993, Post-transcriptional and post-translational regulatory steps are crucial in controlling the appearance and stability of thylakoid polypeptides during the transition of etiolated tobacco seedlings to white light, *Eur. J. Biochem.*, **217**, 345–352.
15. Philbrick, J. B. and Zilinskas, B. A. 1988, 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.
16. Sheen, J., Sayre, R. T., and Bogorad, L. 1987, Differential expression of oxygen-evolving polypeptide genes in maize leaf cell types, *Plant Mol. Biol.*, **9**, 217–226.
17. Sutton, A., Sieburth, L. E., and Bennett, J. 1987, Light-dependent accumulation and localization of photosystem II proteins in maize, *Eur. J. Biochem.*, **164**, 571–578.
18. Khurana, J. P., Kochhar, A., and Jain, P. K. 1996, Genetic and molecular analysis of light-regulated plant development, *Genetica*, **97**, 349–361.
19. Kochhar, A., Khurana, J. P., and Tyagi, A. K. 1996, Nucleotide sequence of the *psbP* gene encoding precursor of 23-kDa polypeptide of oxygen-evolving complex in *Arabidopsis thaliana* and its expression in the wild-type and a constitutively photomorphogenic mutant, *DNA Res.*, **3**, 277–285.
20. Khurana, J. P., Kochhar, A., and Tyagi, A. K. 1998, photosensory perception and signal transduction in higher plants — molecular genetic analysis, *Crit. Rev. Plant. Sci.*, **38**, (In Press).
21. Hou, Y., von Arnim, A. G., and Deng, X.-W. 1993, A new class of *Arabidopsis* constitutive photomorphogenic genes involved in regulating cotyledon development, *Plant Cell*, **5**, 329–339.
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
23. Santos, M. A. 1991, An improved method for small scale preparation of bacteriophage DNA based on phage precipitation by zinc chloride, *Nucl. Acids Res.*, **19**, 5442.
24. Dellaporta, S. L., Wood, J., and Hicks, J. B. 1983, A plant DNA minipreparation: version II, *Plant Mol. Biol. Rep.*, **1**, 19–21.
25. Nagy, F., Kay, S. A., and Chua, N.-H. 1988, Analysis of gene expression in transgenic plants, In: Gelvin, S. B., Schilperoort, R. A. and Verma, D. P. S. (eds), Plant Molecular Biology Manual, Kluwer Academic Publishers, Dordrecht, pp. B4/1–B4/29.
26. Breathnach, R. and Chambon, P. 1981, Organization and expression of eukaryotic split genes coding for proteins, *Ann. Rev. Biochem.* **50**, 349–383.
27. Frendewey, D. and Keller, W. 1985, Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences, *Cell*, **42**, 355–367.
28. Brown, J. W. S. 1986, A catalogue of splice junction and putative branch point sequences from plant introns, *Nucl. Acids Res.*, **14**, 9549–9559.
29. Goodall, G. J. and Filipowicz, W. 1989, The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing, *Cell*, **58**, 473–483.
30. Goodall, G. J., Kiss, T., and Filipowicz, W. 1991, Nuclear RNA splicing and small nuclear RNAs and their genes in higher plants, *Oxford Survey Plant Mol. Cell Biol.*, **7**, 255–296.
31. McCullough, A. J., Lou, H., and Schuler, M. A. 1993, Factors affecting authentic 5′ splice site selection in plant nuclei, *Mol. Cell Biol.*, **13**, 1323–1331.
32. Simpson, C. G., Clark, G., Davidson, D., Smith, P., and Brown, J. W. S. 1996, Mutation of putative branchpoint consensus sequences in plant introns reduces splicing efficiency, *Plant J.*, **9**, 369–380.
33. Kozak, M. 1986, Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes, *Cell*, **44**, 283–292.
34. Smeekens, S., Bauerle, C., Hageman, J., Keegstra, K., and Weisbeek, P. 1986, The role of the transit peptide in the routing of precursors toward different chloroplast compartments, *Cell*, **46**, 356–375.
35. von Heijne, G., Steppuhn, J., and Herrmann, R. G. 1989, Domain structure of mitochondrial and chloroplast targeting peptides, *Eur. J. Biochem.*, **180**, 535–541.
36. Robinson, C. and Knott, T. G. 1996, Importing, sorting and assembly of photosynthetic proteins in higher plant chloroplasts, In: Andersson, B., Salter, A. H. and Barber, J. (eds), Molecular Genetics of photosynthesis, Oxford University Press, New York, pp. 145–159.

37. Coleman, W. J. and Govindjee 1987, A model for the mechanism of chloride activation of oxygen evolution in photosystem II, *Photosynth. Res.*, **13**, 199–223.
38. Yocum, C. F. 1991, Calcium activation of photosynthetic water oxidation, *Biochim. Biophys. Acta*, **1059**, 1–15.
39. Seidler, A. and Rutherford, A. W. 1996, The role of the extrinsic 33 kDa protein in Ca^{2+} binding in photosystem II, *Biochemistry*, **35**, 12104–12110.
40. von Arnim, A. G. and Deng, X.-W. 1996, Light control of seedling development, *Ann. Rev. Plant Physiol. Plant Mol. Biol.*, **47**, 215–243.