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

Anju KOCHHAR, Jitendra P. KHURANA, and Akhilesh K. TYAGI*
Department of Plant Molecular Biology, University of Delhi South Campus, New Delhi-110021, INDIA
(Received 5 August 1996)

Abstract

The psbP gene encoding the precursor of 23-kDa polypeptide of the oxygen-evolving complex of photosystem II has been isolated from Arabidopsis thaliana genomic library and sequenced. The gene harbors three introns and encodes a mature polypeptide of 186 amino acid residues and a transit peptide of 77 amino acid residues. The deduced molecular mass of the mature polypeptide is 23.5-kDa and it contains 22.6% charged amino acid residues which may contribute to the hydrophilic nature of the protein. The transcript encoded by psbP gene of Arabidopsis is approximately 1.3-kb long. In wild-type Arabidopsis seedlings, its expression is organ-specific and is regulated by endogenous developmental cues, light and sucrose. In a constitutively photomorphogenic mutant of Arabidopsis, designated as phol, the psbP gene is partly derepressed in young, dark-grown seedlings, resulting in a slightly higher level of the transcript. Additionally, the phol mutant shows slow accumulation of psbP transcript upon illumination of young, dark-grown seedlings. However, the derepression is not markedly displayed on dark-adaptation of phol plants grown in continuous light. These studies, therefore, define the activity of at least one cellular effector involved in regulation of psbP expression.

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

1. Introduction

The oxygen-evolving complex associated with photosystem II of the chloroplast thylakoid membrane is a vital component of the photosynthetic machinery. This complex consists of three major polypeptides (33-, 23- and 16-kDa) encoded by nuclear genome, which are synthesized on cytoplasmic ribosomes and translocated into thylakoid lumen with the help of a bipartite transit peptide.\(^1\) Mutational analysis in Chlamydomonas has demonstrated that 23-kDa polypeptide is necessary for PSII to be fully functional \textit{in vivo}.\(^2\) Although the precise role of this polypeptide in oxygen evolution is not understood, \textit{in vitro} studies have shown that it modulates chloride requirement for maximal oxygen evolution by increasing the affinity of water-oxidation site for chloride.\(^3\) Additionally, it provides conditions for high-affinity binding of calcium.\(^4,5\) This is perhaps accomplished by serving as a physical barrier between a high-affinity calcium-binding site and the medium resulting in restricted calcium exchange with the environment and not by increasing its affinity for calcium \textit{per se}.\(^6\) Recently, genes for the polypeptide components of oxygen-evolving complex have also been investigated.\(^7\) Since the first report on the isolation of a cDNA clone for the 23-kDa polypeptide from spinach,\(^8\) cDNA clones from a few other plants (tomato,\(^9\) tobacco,\(^10-12\) wheat,\(^13\) pea,\(^14\) rice\(^15\)) and complete genomic counterpart from Sinapis alba\(^16\) and an incomplete gene from tobacco,\(^11\) have been isolated. The expression of the psbP gene (encoding 23-kDa polypeptide) has been shown to be regulated by phytochrome in \textit{S. alba}\(^17\) and the spatio-temporal pattern of its light-regulated expression has also been studied during tobacco seedling development.\(^18\)

Recently, as a result of extensive studies on a class of photomorphogenic mutants, it has become clear that light signal is transduced through a hierarchy of cellular effectors to finally affect various cellular processes and gene expression patterns.\(^19-21\) At least eleven such loci
designated as either det (de-etiolated), cop (constitutively
photomorphogenic) or fusca have been identified. The
J corresponding wild-type alleles defined by these loci en-
code putative repressors of photomorphogenesis, and a
mutation in any of them leads to manifestation of a
range of light-dependent processes in dark. A similar
class of mutants designated as phe (plumular hook
open, in dark) is being characterized by us. The phe
mutants display a number of photomorphogenic features
in dark such as plumular hook opening, cotyledon ex-
pansion, hypocotyl growth suppression, leaf differentiation,
bolting and floral bud formation. In addition, they undergo slow greening when transferred from dark
to light and flower earlier than wild-type when grown in
continuous light.

In the present investigation, the gene encoding the
23-kDa component of the oxygen-evolving complex in
Arabidopsis thaliana has been isolated, sequenced and
characterized for its expression in response to intrinsic
and extrinsic factors in the wild-type A. thaliana. Since
photomorphogenic mutant, phe, is likely to define an
important cellular effector mediating light-regulated gene
expression, the effect of this mutation on phe gene expression has also been examined.

2. Materials and Methods

2.1. Isolation of gene and sequencing

A genomic library of A. thaliana, ecotype Columbia, in
EMBL3 replacement vector, was screened with a heterol-
ogous cDNA clone from spinach at a density of 50,000
pfu per 22.5 × 22.5 cm² culture dish. Prehybridiza-
tion and hybridization of the replica filters were done
as described for analysis of eukaryotic genomic DNA in
Sambrook et al. Phage DNA was isolated from posi-
tive clones by procedure given by Santos. Restriction
analysis and cloning was done using standard protocols.
A genomic library of A. thaliana, ecotype Columbia, in
EMBL3 replacement vector, was screened with a heterol-
ogous cDNA clone from spinach at a density of 50,000
pfu per 22.5 × 22.5 cm² culture dish. Prehybridiza-
tion and hybridization of the replica filters were done
as described for analysis of eukaryotic genomic DNA in
Sambrook et al. Phage DNA was isolated from posi-
tive clones by procedure given by Santos. Restriction
analysis and cloning was done using standard protocols.

2.2. Plant material and growth conditions

Seeds of A. thaliana, ecotype Estland, and its pho-
tomorphogenic mutant, pho1, were treated with 5% sodium
hypochlorite containing 0.02% Triton X-100 for 5 min,
ashed thrice with sterile distilled water, followed by
three washings with mineral medium. Seeds were finally
suspended in 0.1% agar solution prepared in the mineral
medium and inoculated in 9.0 cm Petri dishes containing
gelled (0.8% agar) mineral medium. Sucrose was added at the 2% level only in the experiment involving the study
of its effect on phe gene expression. Seeds were sub-
jected to chilling treatment (at 4°C) for 24-48 h and then
illuminated with white light for 6–8 h. Following this
treatment for potentiating seed germination, Petri
dishes were either shifted to dark or maintained in con-
tinuous light depending upon the requirement. For study-
ing organ-specific expression and extracting total plant
DNA, plants were grown in clay pots containing ‘Soil-
rite mix’ (a mixture of vermiculite, perlite and Sphag-
num moss in 1:1:1 ratio; Kelperlite, Bangalore), satu-
rated with mineral medium, under continuous light and
at 23 ± 2°C.

2.3. DNA extraction and Southern analysis

Total plant DNA was extracted by the method of Del-
laporta et al. and digested separately with several re-
striction enzymes. Digested DNA samples (8 µg each)
were electrophoresed on a 1% agarose gel in 1X TAE
buffer. The gel was treated with depurination (0.15
N HCl; 15 min), denaturation (1.5 M NaCl, 0.5 M
NaOH; 45 min) and neutralization (1 M Tris-Cl, pH 8.0,
1.5 M NaCl; 45 min) solutions, respectively, and blot-
ted on Hybond-C membrane (Amersham International
Inc.). Prehybridization (in 50% formamide, 5X SSC, 5X
Denhardt’s solution, 5% dextran sulfate, 200
µg ml⁻¹ herring sperm DNA) and hybridization (in 50%
formamide, 5X SSC, 0.1% SDS, 5X Denhardt’s solution,
10% dextran sulfate, 1% SDS, 250 µg ml⁻¹ herring sperm
DNA) of the blot were performed at 37°C. The 6.9-kb
BamHI fragment from Arabidopsis was used as probe
after labelling with Multprime DNA Labelling System
(Amersham International Inc.). Subsequently, the blot
was washed, in sequence, with 50% formamide, 5X SSC,
0.1% SDS; 5X SSC, 0.1% SDS; 2X SSC, 0.1% SDS,
and 1X SSC, 0.1% SDS, each time for 10 min, at room tem-
perature and exposed to X-ray film for autoradiography.

2.4. RNA extraction and northern analysis

RNA was extracted from the tissue following the pro-
cedure of Nagy et al. and quantified spectrophotomet-
ically at 260 nm. Denatured RNA (10–20 µg) was
electrophoresed on 1.2% agarose-formaldehyde gel us-
ing 1X MOPS buffer. Capillary blotting was done on
Hybond-C membrane using 2X SSC. Prehybridization
was performed in a solution containing 50% formamide,
5X SSC, 10X Denhardt’s solution, 50 mM sodium phos-
phate buffer (pH 6.5) and 250 µg ml⁻¹ herring sperm
DNA, at 37°C. The 6.9-kb BamHI fragment from Ara-
bidopsis, as above, was used for hybridization in a so-
lution containing 10% dextran sulfate in addition to all
other components of the prehybridization solution. The
hybridization was performed for 48 h at 37°C. Washing of
Figure 1. Southern blot analysis of Arabidopsis genomic DNA digested with SaI (S), HindIII (H), EcoRI (E) and BamHI (B) using a 6.9-kb BamHI fragment that harbors complete psbP gene of Arabidopsis along with extensive 3' neighboring sequences. The genomic fragments which bear complete or a part of the psbP gene, are marked by arrowheads. Other fragments appear due to homology with the region downstream of the gene. The uppermost band in lane H is the extended genomic version of the 1.6-kb fragment present on the 5' end of the probe and contains N-terminus coding region of the gene and more than 2.0-kb of the upstream region (see text). The small HindIII fragments of 0.3-kb and 0.15-kb became apparent only after very long exposure and were not visible in this autoradiograph. Numerals on left indicate size markers (ADNA HindIII digest).

blots was done under the conditions described above for washing Southern blot. The blots were also hybridized under similar conditions with a cDNA clone encoding 25S ribosomal RNA30 to ensure equal loading and integrity of the RNA samples being used.

3. Results and Discussion

3.1. Isolation and characterization of the gene

A genomic library of Arabidopsis DNA in EMBL3 vector was screened using a cDNA clone encoding a 23-kDa polypeptide component of the oxygen-evolving complex of photosystem II from spinach.8 Twenty plaques giving a signal were replated, out of which only 8 yielded overlapping signals in the secondary screening. The number of positive clones was reduced to only two after tertiary screening and their purity was confirmed in quaternary screening. Phage DNA isolated from the two purified clones (designated as 'A' and 'B') was characterized by restriction digestion and Southern analysis using the heterologous probe from spinach. On digestion with SaI, clones 'A' and 'B' revealed single inserts of ca. 13- and 15-kb, respectively, which hybridize with the heterologous spinach probe. In EcoRI-digested clones, a 7-kb fragment binds the probe. The smallest single hybridizing fragment (ca. 6.9-kb) was obtained with BamHI in both the clones. This is also obvious in genomic Southern blots hybridized with the 6.9-kb BamHI fragment of clone 'A', where signals with SaI-, HindIII-, EcoRI-, and BamHI-restricted DNA reflect the presence of single or identical gene(s) within the limit of detection, for 23-kDa polypeptide (Fig. 1). A single psbP gene has also been reported earlier in pea14 and spinach,8 but multigene families are reportedly present in Sinapis (4 genes16) and tobacco (4 or more genes12). The 6.9-kb BamHI fragment from clone 'A' was subcloned in pBluescript KS" vector for further characterization. Restriction and Southern analysis of this recombinant clone revealed that three fragments (approximate sizes 0.3-kb, 1.6-kb and 2.7-kb) obtained after digestion with HindIII give positive signals. These three fragments were further subcloned in pBluescript KS" vector and sequenced. Another HindIII fragment of 152 bp was revealed during sequencing of the overlapping region between the 1.6- and 0.3-kb fragments with the help of a oligonucleotide primer. For obtaining complete sequence information of the gene, cloning and sequencing of Sau3AI subfragments of the 1.6-kb fragment were also undertaken. Thus, the C-terminus coding region was located on the 2.7-kb fragment, the N-terminus was located on the 1.6-kb fragment, while the 0.15- and 0.3-kb fragments encoded the middle region of the protein.

The nucleotide sequence of the gene along with 289 bases of the 5' region and 146 bases of the 3' untranslated region is shown in Fig. 2. The gene consists of four exons and three introns as is the case with psbP genes of Sinapis16 and tobacco.11 Length of introns and their sequences show considerable variation in these species with tobacco having particularly longer introns I (516 bases) and II (487 bases) but they share a common feature of all plant introns in being AT-rich.31 Introns I (111 bases), II (90 bases) and III (255 bases) of Arabidopsis psbP gene have 73%, 62.2% and 66.2% of AT base complement, respectively. The GT-AG consensus, found at the borders of most introns33 has helped in defining the boundaries of all three introns. Also, certain other features of yeast and vertebrate introns31 such as a branch point sequence (CYRAY, Y = pyrimidine and R = purine), nearly 30 bases upstream from the 3' splice site, is present in in-
Figure 2. Nucleotide and deduced amino acid sequence of the *Arabidopsis* psbP gene encoding precursor of 23-kDa polypeptide of oxygen-evolving complex associated with photosystem II. Nucleotides are numbered on the left. The branch point sequences in introns II and III are underlined. Amino acid residues implicated in forming β-pleated sheets, hydrophobic domain and terminal processing site are shown in bold. The asterisk marks the site of terminal processing of the precursor polypeptide.
tron II and intron III, and a 3' polypyrimidine tract is present in intron I of the Arabidopsis psbP gene. The intron/exon boundaries have been further confirmed by comparison of the Arabidopsis psbP gene with the psbP gene sequence from Sinapis alba, a member of the same family (Cruciferae), which share 88.8% homology in the coding region.

The psbP gene of Arabidopsis codes for a precursor polypeptide of 263 amino acid residues comprising a 77-amino acid-long transit peptide and a 186-amino acid-long mature polypeptide (Fig. 2). Genomic and cDNA sequences of this gene known from other plants also encode a mature polypeptide of almost the same length and length of the transit peptide varies from 72 in tomato9 to 82 amino acid residues in tobacco.11'12 The transit peptide (77 amino acid residues) encoded by the Arabidopsis psbP gene is three amino acids longer than that from Sinapis,16 both sharing 85.7% homology. Despite differences in the primary sequence of the transit peptide of the 23-kDa polypeptide from divergent species, certain features suggested to be important for the import process of thylakoid lumen proteins are clearly present. These include a preponderance of hydroxylated and charged amino acids, a β-sheet-forming region (amino acids 40–43), a hydrophobic domain (amino acids 57–69) and presence of short chain residues at −3 and −1 positions with respect to the terminal cleavage site. Homology at the mature polypeptide level between Arabidopsis (present investigation) and Sinapis16 is 92.5%. The deduced size of the mature polypeptide encoded by the Arabidopsis psbP gene is 23.5-kDa, while that of the transit peptide is 9.3-kDa. The presence of 22.6% charged residues in the mature polypeptide suggests that the polypeptide is hydrophilic in nature. The context of the AUG start codon consists of an A at −3 position and a G at +4 position and thus fulfills the requirement for the optimal context suggested by Kozak, which is found only in ~ 35% of eukaryotic mRNA species. Additional AUG codons in the optimal context are also present at nucleotide positions 5 and 38 but are followed by in-frame stop codons after a few bases.

3.2. Development-dependent expression

The psbP gene encodes a transcript of 1.3-kb in A. thaliana. Its expression was examined in 1- to 7-day-old seedlings grown in continuous light or darkness to understand how the gene is regulated by light and endogenous developmental cues. Surprisingly, the pattern of psbP gene expression is similar in both dark- and light-grown seedlings although the steady-state transcript level is always higher in light-grown plants (Fig. 3A, B). The level of transcript gradually increases with development and attains a maximum in 3-day-old seedlings, whether grown in dark or light, and thereafter it declines remarkably in dark-grown seedlings, but is maintained at a higher steady-state level in light-grown seedlings. The expression pattern can be correlated with the early seedling development as the third day marks the time when hypocotyl elongates considerably (in dark) and cotyledon expansion also takes place in seedlings exposed to light. Therefore, it is the time when photosynthetic machinery is being established. A certain steady-state level of psbP transcript maintained in dark-grown seedlings perhaps allows rapid attainment of photosynthetic competence with the onset of light. The development-dependent accumulation of psbP transcript was also checked in the pho1 mutant displaying light-grown phenotype in dark. In comparison to wild-type, slightly higher level of psbP transcript was recorded in 2- to 3-day-old pho1 seedlings thereby indicating that the product of the wild-type gene (PHO1) defined by this mutation is involved in repression of the psbP gene during very early stages of seedling growth.

3.3. Light-induced changes in transcript abundance

The seedlings grown in dark for 4 days were illuminated with white light for different durations and analyzed for psbP transcript abundance. In the case of wild-type Arabidopsis seedlings, a prominent increase in transcript level was detectable with even 4 h of illumination, and a steady rise continued thereafter for at least 24 h, before stabilizing at a high steady-state level (Fig. 3C). However, pho1 seedlings displayed a delay in accumulation of psbP gene transcript, as did several other photosynthesis-related nuclear- as well as chloroplast-encoded genes; the significant increase in psbP transcript level was apparent only with 24 h of illumination, and thereafter increased rather slowly, leading eventually to a considerable delay in attainment of levels comparable to wild-type. This delayed accumulation of photosynthesis-related gene transcripts is paralleled by a delayed accumulation of chlorophyll and carotenoid pigments in similar sets of experiments with pho1 mutant (data not shown). The data indicate that mutation in the PHO1 gene causes a general delay in the attainment of photosynthetic competence when dark-grown seedlings are exposed to light and, therefore, defines at least one component of the light signal transduction chain which modulates the chloroplast biogenesis by altering expression levels of genes encoding chloroplast proteins.

3.4. Dark-adaptation and reinduction kinetics

Two-week-old plants of Arabidopsis grown in continuous light were dark-adapted for various durations and the transcript levels for psbP gene were estimated by northern hybridization. Figure 3D shows that the level of psbP transcript is reduced significantly after 2 h of dark-adaptation and is virtually undetectable after 24 h in
Figure 3. Regulated expression of psbP gene (upper row in each set) in Arabidopsis thaliana wild-type and a photomorphogenic mutant phol in relation to the control gene for 25S ribosomal RNA (lower row in each set). A, B, Effect of temporal development of the seedling in dark (A) or light (B). The numerals represent age of seedlings in days. C, Inducibility of gene by light after transfer of 4-day-old dark-grown seedlings (D) to light. Numerals represent hours of illumination. D, Gene expression during dark-adaptation (DA) of 14-day-old light-grown plants (14 dL) for various durations (h). After dark-adaptation for 48 h, plants were again exposed to light, DA + L (h). 16 dL and 17 dL represent controls grown in light for 16 and 17 days. E, Sucrose-dependent regulation of gene expression in seedlings grown in dark for varying days (see numerals), S for sucrose. F, Organ-specific expression of psbP gene in root (R), stem (S), leaf (L), and inflorescence (I) of 36-day-old plants grown in continuous light. Variation in transcript levels for similar treatments in different vertical panels is due to development of autoradiographs for different period in order to show the influence of most effective treatment.
dark. Re-irradiating these plants with white light causes an increased accumulation of psbP transcript within 8 h and the steady-state transcript level, as observed in continuous light-grown plants of the same age (17-day-light), is attained within 24 h. The pho1 plants when examined in this perspective show similar pattern of psbP transcript accumulation except that the transcript level in general is relatively lower than wild-type which could be due to reduced greening of leaves of the mutant. Another interesting observation is that the level of psbP transcript after 24 h of illumination of the dark-adapted plants is higher than the continuous light-grown plants of the same age (17-day-light). This could perhaps be to enable the plants to quickly regain its optimal photosynthetic ability. The results of the dark-adaptation experiment also suggest that the PHO1 product may not be important in down-regulation of psbP expression in dark-adapted plants.

3.5. **Effect of sucrose**

Metabolizable sugars like sucrose and glucose have been shown to repress expression of photosynthetic genes including cab and rbcS. This has been suggested as the mechanism for feedback regulation of photosynthesis. The effect of sucrose on *Arabidopsis* psbP gene was analyzed in 2-, 5- and 9-day-old wild-type and pho1 seedlings grown in dark with or without sucrose. As shown in Fig. 3E, sucrose does repress distinctly the accumulation of psbP transcript in 2-day-old wild-type as well as pho1 seedlings. A similar effect has also been reported earlier for cab gene of *Arabidopsis*. Despite repression by sucrose, the level of psbP transcript is higher (though marginally) in 2-day-old pho1 seedlings, thereby indicating that the feedback regulation by this metabolite occurs by a pathway independent of the one which represses photomorphogenic development in the dark-grown seedlings.

3.6. **Organ-specific expression**

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 analysed for steady-state level of psbP transcript. The highest expression was detectable in leaves, followed by bolting stalks and inflorescence, respectively (Fig. 3F). The psbP transcript was almost undetectable in roots. The expression can thus be correlated with the photosynthetic competence of the tissue; in the case of inflorescence, sepals and pedicels fit into this criteria. This kind of expression pattern has also been reported for psbP gene of pea.14

In conclusion, *A. thaliana* psbP gene isolated and sequenced in this investigation shows an intron/exon arrangement similar to the genes from *Sinapis* and *tobacco*.11 Several important features of the gene and its product have also been described. The expression of *psbP* gene of *Arabidopsis* is organ-specific and is regulated by several intrinsic and extrinsic factors. Therefore, it provides a tool to study how different regulatory signals act together to affect the expression of a particular gene. In this context, we have studied the expression of this gene in a photomorphogenic mutant of *Arabidopsis* and the studies reveal the activity of at least one cellular effector, PHO1, which is partly responsible for dark-repression of *psbP* gene in 2- to 3-day-old seedlings and also participates in the rapid induction of this gene (and several other photosynthesis-related genes) on illumination of young dark-grown seedlings. It is worth mentioning here that genetic allelism has been found between a constitutively photomorphogenic mutant *cop* and *pho1* but differences are still evident at the phenotypic level, particularly in flowering time, which could be due to mutations at different sites on the gene. Moreover, the effect of *cop3* mutation on photosynthetic gene expression has not been reported in detail and differences between *pho1* and *cop3* may be revealed in this regard as well on further investigation. This approach, if extended to other possible combinations of mutants and the responses evoked by the affected pathways, will help resolve the components of intricate molecular circuitry driving temporal/spatial/inducible expression of various genes as well as other downstream responses.

**Acknowledgements:** We thank Professor Dr. R. G. Herrmann and Dr. R. Oelmuller (Munich) for kindly providing us cDNA probe from spinach and the genomic library of *Arabidopsis thaliana*, ecotype Columbia, respectively, and Prof. S. C. Maheshwari for encouragement and keen interest in this work. This research was funded by the Department of Science and Technology and the University Grants Commission, New Delhi. AK acknowledges the award of Senior research fellowship by the Council of Scientific and Industrial Research, New Delhi.

**References**


3. Andersson, B., Critchley, C., Ryrie, I. J., Jansson, C., Larsson, C., and Anderson, J. M. 1984, Modification of


32. Breathnach, R. and Chambon, P. 1981, Organization and expression of eucaryotic split genes coding for proteins,


