Volume 250, number 2, 575–579 FEB 07342 July 1989

A cDNA clone encoding the precursor for a 10.2 kDa photosystem I polypeptide of barley

Jens Sigurd Okkels, Henrik Vibe Scheller, Lars Bæk Jepsen and Birger Lindberg Møller

Plant Biochemistry Laboratory, Department of Plant Physiology, Royal Veterinary and Agricultural University, Frederiksberg C, Denmark

Received 28 March 1989; revised version received 12 May 1989

Two cDNA clones for the barley photosystem I polypeptide which migrates with an apparent molecular mass of 9.5 kDa on SDS-polyacrylamide gels have been isolated using antibodies and an oligonucleotide probe. The determined N-terminal amino acid sequence for the mature polypeptide confirms the identification of the clones. The 644 base-pair sequence of one of the clones contains one large open reading frame coding for a 14 882 Da precursor polypeptide. The molecular mass of the mature polypeptide is 10 193 Da. The hydropathy plot of the polypeptide shows one membrane-spanning region with a predicted α-helix secondary structure. The gene for the 9.5 kDa polypeptide has been designated *PsaH*.

Photosystem I; Polypeptide, 9.5 kDa; cDNA sequence; Transit peptide; Gene, *PsaH*; (Barley)

1. INTRODUCTION

Photosystem (PS) I preparations from barley are composed of polypeptides with apparent molecular masses of 83, 82, 18, 16, 14, 9.5, 9, 4 and 1.5 kDa [1,2]. One copy of each of these polypeptides is present per P700 reaction center [2]. The PS I preparation binds the photoreducible electron acceptors A_0, A_1, X, A and B [3]. These electron acceptors are bound to the 83, 82 and 9 kDa polypeptides [2–6]. The genes for these polypeptides are located in the chloroplast genome [1,5,7,8].

The role of the remaining polypeptides of PS I is less clear. Experiments with chemical cross-linkers have shown that a 20–22 kDa polypeptide from spinach can be cross-linked to ferredoxin [9,10]. Similarly, a 19 kDa polypeptide of PS I from spinach can be cross-linked to plastocyanin [11]. The presence of these additional polypeptides may therefore be important for the efficiency of the PS I-mediated electron transport from plastocyanin to ferredoxin. One approach for characterization of these polypeptides is to determine their primary structure. cDNA clones encoding four such polypeptides have so far been isolated from various plants [12–16]. We present here the nucleotide sequence of a fifth cDNA clone encoding the precursor for the 9.5 kDa polypeptide of barley along with the partial amino acid sequences for the N-terminal of the mature polypeptide.

Correspondence address: J.S. Okkels, Department of Plant Physiology, Royal Veterinary and Agricultural University, 40 Thorvaldsensvej, DK-1871 Frederiksberg C, Denmark

2. MATERIALS AND METHODS

Total RNA was isolated from 5-day-old etiolated barley seedlings after prior exposure to light for 5.5 h. The total RNA was prepared using a modification of the techniques of Martin and Northcote [17] as reported by Collinge et al. [18]. Poly(A)$^+$ mRNAs were purified on oligo(dT) Sepharose 4B (Pharmacia LKB Biotechnology, Uppsala) [17] and used for construction of a λgt11 cDNA library (Clontech Labs, Palo Alto, CA).

PS I particles and the 9.5 kDa polypeptide were isolated as in [1,2]. Amino acid sequencing was carried out as in [1].

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

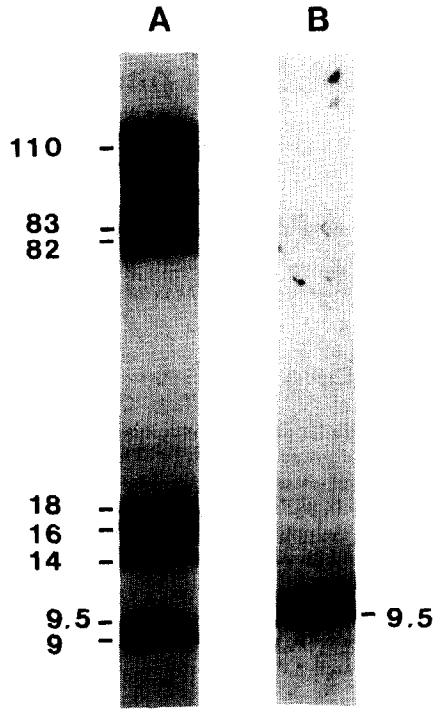

Fig.1. Western blotting analysis of the polyclonal antibody against the isolated 9.5 kDa polypeptide. (A) Polypeptide composition of the PS I preparation as analyzed by SDS-PAGE [20] and staining with Coomassie brilliant blue. The 4 and 1.5 kDa polypeptides are not detected in this gel system. (B) Reaction of the 9.5 kDa antibody with the PS I preparation as monitored by Western blotting and use of an alkaline phosphatase conjugated secondary antibody. The 9.5 kDa antibody did not cross-react with the isolated 9 kDa polypeptide (not shown).

Polyclonal antibodies towards the PS I preparation and the isolated 9.5 kDa polypeptides were raised in rabbits. The immunoglobulin fraction of the antiserum was purified according to Harboe and Ingild [19]. The specificity of the antibody against the 9.5 kDa polypeptide was tested by Western blotting analysis (fig.1).

The immunological screening of the library was conducted as described by Mierendorf et al. [21]. The fusion proteins were immobilized on nitrocellulose filters (Schleicher & Schuell, Keene, NH). Due to unspecific binding of the PS I antibody to λgt11 plaques, this antibody was preincubated with saturating amounts of wild-type λgt11 plaques immobilized on nitrocellulose filters. Antigen-antibody complexes were detected enzymatically with alkaline phosphatase-conjugated antibodies (Dakopatts a/s, Glostrup, Denmark).

Based on the partial amino acid sequence data, an oligonucleotide probe specifying the antisense strand of the N-terminal region of the polypeptide, H_2N-KYGEK, was synthesized: 5′-TT$^{T}_{C}$TCNCC$^{A}_{G}$TA$^{T}_{C}$TT-3′ (32 different oligonucleotides, N = T, C, G, A).

The positive clones obtained by the immunological screening were screened with the 5′-end-labeled [13] probe using nitrocellulose filters [22]. Southern blotting and hybridization was performed using Zetaprobe membranes according to the manufacturer (Bio-Rad, Richmond, CA).

DNA from λ phages was prepared according to Grossberger [23]. Inserts from λ phages were subcloned in the pTZ18/19 plasmids [24]. Single-stranded DNA was obtained from the pTZ18/19 plasmids using helper phage M13K07 [25]. DNA sequencing was carried out by the dideoxy chain termination method of Sanger et al. [26], using [α-^{32}P]dATP (Amersham, Bucks), and a Sequenase kit (US Biochem., Cleveland, OH). A *Pst*I restriction enzyme site in the middle of clone 9.5-1 was used for subcloning in the pTZ18U plasmid to obtain the sequence of both DNA strands (fig.2).

Hydropathy plot and secondary structure prediction was carried out with the Seqanal program package (A.R. Crofts, Biotechnology Center, University of Illinois, IL).

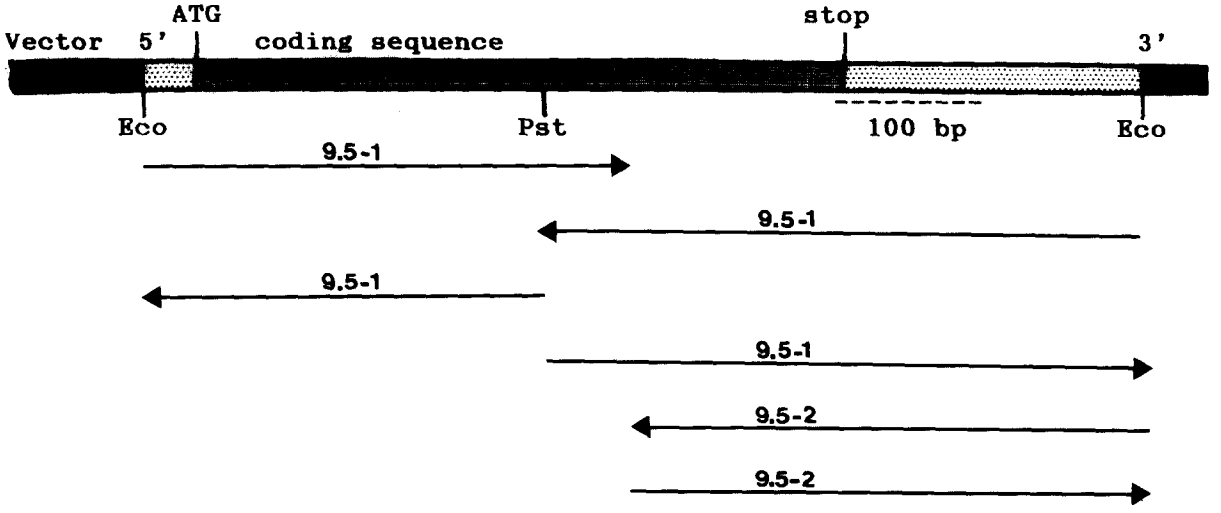

Fig.2. Sequencing strategy and the different sets of overlapping subclones.

3. RESULTS AND DISCUSSION

Immunological screening of the cDNA library with the PS I antibody gave 14 positive clones from approx. 2×10^4 recombinants. The PS I antibody was used in the primary screening to obtain a number of cDNA clones for the nuclear encoded polypeptides of PS I. Two of the 14 clones were positive with a polyclonal antibody towards the 9.5 kDa polypeptide. One of these clones was positive with the oligonucleotide probe on Southern blotting analysis of phage inserts. The two clones were sequenced in both directions (fig.2). The nucleotide sequence of clone 9.5-1 is shown in fig.3A. Clone 9.5-2 is a partial cDNA clone starting with the 5′-end at nucleotide 316 of clone 9.5-1 and with identical sequence except for an extra 9 bp at the poly(A) tail (fig.3A).

Analysis of clone 9.5-1 reveals one large open reading frame of 429 bp. The deduced amino acid sequence corresponding to bp 145–195 is identical to the N-terminal amino acid sequence determined

A:

```
                                                        -27                -1
                                                          CTCCTCC CACCGCACAG CTCGACCGCA

1                                                                          60
ATG GCG TCG CTC GTC GCC GTC CAG CCG GCC GCC GTG AAG GGC CTC TCC GGC AGC TCC ATC
Met Ala Ser Leu Val Ala Val Gln Pro Ala Ala Val Lys Gly Leu Ser Gly Ser Ser Ile

                                                                          120
TCC GGC CGC AAG CTC GCC GTC CGG CCC TCC TCC GCC GCC GTC TCC CGC TCC ACC CGC AGG
Ser Gly Arg Lys Leu Ala Val Arg Pro Ser Ser Ala Ala Val Ser Arg Ser Thr Arg Arg

                                                                          180
GCT CGC GGG GCG GCG GTG GTG GCC▼AAG TAC GGG GAG AAG AGC GTG TAC TTC GAC CTG GAC
Ala Arg Gly Ala Ala Val Val Ala Lys Tyr Gly Glu Lys Ser Val Tyr Phe Asp Leu Asp

                                                                          240
GAC ATC GCC AAC ACG ACG GGG CAG TGG GAC CTG TAC GGC AGC GAC GCG CCG TCG CCC TAC
Asp Ile Ala Asn Thr Thr Gly Gln Trp Asp Leu Tyr Gly Ser Asp Ala Pro Ser Pro Tyr

                                                                          300
AAC GGC CTG CAG AGC AAG TTC TTC AAC ACC TTC GCC GCG CCC TTC ACC AAG CGC GGC CTG
Asn Gly Leu Gln Ser Lys Phe Phe Asn Thr Phe Ala Ala Pro Phe Thr Lys Arg Gly Leu

                                                                          360
CTG CTC AAG TTC CTG CTC ATC GGC GGC GGC TCC CTG GTC GCC TAC GTC AGC GCA TCC GCC
Leu Leu Lys Phe Leu Leu Ile Gly Gly Gly Ser Leu Val Ala Tyr Val Ser Ala Ser Ala

                                                                          420
TCC CCC GAC CTG CTC CCC ATC AAG AAG GGG CCC CAG CTG CCG CCC ACC CCG GGA CCA CGC
Ser Pro Asp Leu Leu Pro Ile Lys Lys Gly Pro Gln Leu Pro Pro Thr Pro Gly Pro Arg

                                                                          490
GGC AAG ATC TAA ATCTAATC CTCGCCTCGC CTCGCCTCGT GCTTGTGTAC CATGGATCCC GTCTGATCTC
Gly Lys Ile STOP
                                                                          560
ATGGATCGGA ATCGGATCCC CCGTATCCAT GTCTTTCTGT TGATACTTGA TGCTGTACCT TGTTGCCTAC

GTGATGCACA CAATTACTGC TAGATC(A)31 ←———— Clone 9.5-1

/--------- ---------- ------ATTTGCTCT(A)62 ← Clone 9.5-2 (starts at bp 316)
```

B:

```
Barley: KYGEKSVYFDLDDIANTTGQWDLYGSDAPSPYNGLQ/
        |||·|||||||·||·||||||||||||||||·  ||
Pea:    KYGDKSVYFDLEDIGNTTGQWDLYGSDAPSPYSXLQ/
```

Fig.3. (A) DNA sequence of the 9.5-1 cDNA clone. Only the 3′-end of clone 9.5-2 is shown. The nucleotide sequence is written in the mRNA sense. The deduced amino acid residues are written below the nucleotide sequence. The amino acid sequence obtained by sequencing of the isolated 9.5 kDa polypeptide is underlined. The maturation site (▼), polyadenylation site (···), and predicted membrane-spanning α-helix (~~~) are indicated. (B) Comparison of the N-terminal amino acid sequence from pea [27] and the amino acid sequence derived from the barley cDNA clone. A vertical line indicates an amino acid identity and a dot represents a conservative replacement.

by sequencing of the mature 9.5 kDa polypeptide (fig.3A). The N-terminal residues determined by amino acid sequencing are underlined in fig.3A. The open reading frame codes for a precursor protein with a calculated molecular mass of 14882 Da. Presuming that the mature polypeptide is not C-terminally processed, the molecular mass of the mature polypeptide and the N-terminal transit peptide can be calculated to be 10193 and 4689 Da, respectively. The calculated molecular mass of the mature polypeptide is in agreement with the predicted value of 9.5 kDa based on migration in SDS-polyacrylamide gels.

Comparison of the amino acid sequence of the 9.5 kDa polypeptide of barley with a reported N-terminal sequence (35 amino acid residues) of an 11 kDa polypeptide from pea [27] reveals that these two polypeptides are homologous (fig.3B). The highly conserved amino acid sequence between the barley and pea polypeptides suggests an important function for the polypeptide in PS I. Presumably this polypeptide corresponds to subunit VI in the PS I preparations described by Nechushtai et al. [28]. The seven previous genes for PS I polypeptides have been denoted psa followed by an alphabetically ordered suffix. We have consequently named the gene for the 10.2 kDa polypeptide *PsaH*. A capital P has been used to indicate the location of the gene in the nuclear genome. No obvious sequence homology to other proteins was found by searching the National Biomedical Research Foundation protein bank.

The region around the translation initiation codon shows a good similarity to the consensus sequence AACAAUGGC for translation initiation codon regions in plants [29]. The sequence for the 48 N-terminal amino acid residues representing the transit peptide is very rich in positively charged amino acids (8 residues) and hydroxylated amino acids (10 residues) and lacks acidic residues. These properties are found in most other chloroplast transit peptides [30]. Weak similarity to the primary sequence of the three homology blocks on the transit peptides of the chloroplast suggested by Karlin-Neumann and Tobin [31] can be found. It has, however, been postulated that the secondary structure of transit peptides is the essential feature for their function [32–34]. Except for the 10.8 kDa polypeptide of barley [13], the transit peptides of PS I subunits have been reported to contain a highly hydrophilic segment [12,15,16]. A similar hydrophilic segment is contained in the transit peptide of the 9.5 kDa polypeptide (fig.4). The significance of this hydrophilic segment, which is generally not found in other chloroplast transit peptides, is presently unknown. The transit peptide of the 9.5 kDa polypeptide does not contain the hydrophobic thylakoid transfer domain which characterizes polypeptides directed to the thylakoid lumen [35,36]. Therefore, a stromal or transmembrane location of the polypeptide is more likely. Some similarity to the transit peptide of a 10.8 kDa polypeptide of PS II can be observed [37]. This polypeptide has a hydrophobic region of 20 amino acids at the C-terminal, suggesting a transmembrane location of the polypeptide [37]. A transmembrane location of the 9.5 kDa polypeptide is also suggested from the hydropathy plot (fig.4) which shows a hydrophobic area of 19 amino acids (residues 103–121) sufficiently long to represent a membrane-spanning region. Using the parameters for membrane-buried helices given by Rao and Argos [38], the hydrophobic region was predicted to have an α-helix conformation typical of membrane-spanning segments. The region contains several glycine residues which are strong helix breakers in soluble proteins [38]. However, glycine residues are neutral in membrane helices and tend to concentrate near the predicted helix midpoint

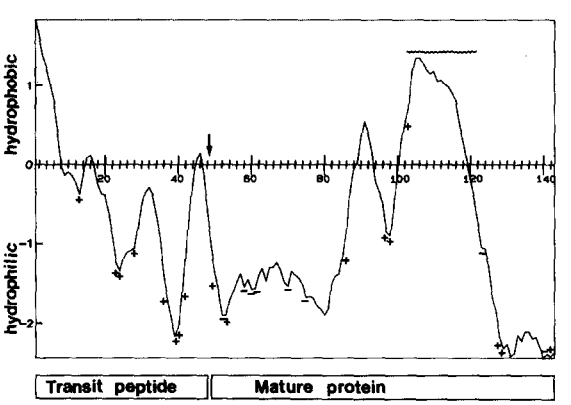

Fig.4. Hydropathy plot of the 14.9 kDa precursor polypeptide calculated with an averaging window of 7 amino acid residues and using membrane helix parameters according to Rao and Argos [38]. The distribution of positively (+) and negatively (−) charged residues is indicated. The predicted membrane-spanning α-helix (~~~) and the maturation site (↓) are indicated.

[38] as also observed for the 9.5 kDa polypeptide. Determination of the stroma/lumen location of the N- and C-terminal of the polypeptide has to await further analysis.

Acknowledgements: Dr Ib Svendsen is thanked for the determination of the N-terminal sequence and Dr Jan Lehmbeck for searching for homologous proteins. Inga Olsen and Hanne Linde Nielsen are thanked for skillful technical assistance. This work was supported in part by grants from the Danish Governmental Program for Biotechnology Research and from the Danish Agricultural and Veterinary Research Council.

REFERENCES

[1] Høj, P.B., Svendsen, I., Scheller, H.V. and Møller, B.L. (1987) J. Biol. Chem. 262, 12676–12684.
[2] Scheller, H.V., Svendsen, I. and Møller, B.L. (1989) J. Biol. Chem., in press.
[3] Andréasson, L.-E. and Vänngård, T. (1988) Annu. Rev. Plant Physiol. Plant Mol. Biol. 39, 379–411.
[4] Høj, P.B. and Møller, B.L. (1986) J. Biol. Chem. 261, 14292–14300.
[5] Fish, L.E., Kück, U. and Bogorad, L. (1985) J. Biol. Chem. 260, 1413–1421.
[6] Golbeck, J.H. and Cornelius, J.M. (1986) Biochim. Biophys. Acta 849, 16–24.
[7] Hayashida, N., Matsubayashi, T., Shinozaki, K., Sugiura, M., Inoue, K. and Hiyama, T. (1987) Curr. Genet. 12, 247–250.
[8] Oh-oka, H., Takahashi, Y., Kuriyama, K., Saeki, K. and Matsubara, H. (1988) J. Biochem. 103, 962–968.
[9] Zanetti, G. and Merati, G. (1987) Eur. J. Biochem. 169, 143–146.
[10] Zilber, A.L. and Malkin, R. (1988) Plant Physiol. 88, 810–814.
[11] Wynn, R.M. and Malkin, R. (1988) Biochemistry 27, 5863–5869.
[12] Lagoutte, B. (1988) FEBS Lett. 232, 275–280.
[13] Okkels, J.S., Jepsen, L.B., Hønberg, L.S., Lehmbeck, J., Scheller, H.V., Brandt, P., Høyer-Hansen, G., Stummann, B., Henningsen, K.W., Von Wettstein, D. and Møller, B.L. (1988) FEBS Lett. 237, 108–112.
[14] Hoffman, N.A., Pichersky, E., Malik, V.S., Ko, K. and Cashmore, A.R. (1988) Plant Mol. Biol. 10, 435–445.
[15] Steppuhn, J., Hermans, J., Nechushtai, R., Ljungberg, U., Thümmler, F., Lottspeich, F. and Herrmann, R.G. (1988) FEBS Lett. 237, 218–224.
[16] Münch, S., Ljungberg, U., Steppuhn, J., Schneiderbauer, A., Nechushtai, R., Beyreuther, K. and Herrmann, R.G. (1988) Curr. Genet. 14, 511–518.
[17] Martin, C. and Northcote, D.H. (1981) Planta 151, 189–197.
[18] Collinge, D.B., Milligan, D.E., Maxwell Dow, J., Scofield, G. and Daniels, M.J. (1987) Plant Mol. Biol. 8, 405–414.
[19] Harboe, N. and Ingild, A. (1973) in: A Manual of Quantitative Immunoelectrophoresis Methods and Applications (Axelsen, N.H. et al. eds) pp.161–164, Universitetsforlaget, Oslo.
[20] Fling, S.P. and Gregerson, D.S. (1986) Anal. Biochem. 155, 83–88.
[21] Mierendorf, R.C., Percy, C. and Young, R.A. (1987) Methods Enzymol. 152, 458–469.
[22] Vogeli, G. and Kaytes, P.S. (1987) Methods Enzymol. 152, 407–415.
[23] Grossberger, D. (1987) Nucleic Acids Res. 15, 6737.
[24] Mead, D.A., Szczesna-Skorupa, E. and Kemper, B. (1986) Protein Eng. 1, 67–74.
[25] Vieira, J. and Messing, J. (1988) Methods Enzymol. 153, 3–11.
[26] Sanger, F., Niklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[27] Dunn, P.P.J., Packman, L.C., Pappin, D. and Gray, J.C. (1988) FEBS Lett. 228, 157–161.
[28] Nechushtai, R., Nelson, N., Mattoo, A.K. and Edelman, M. (1981) FEBS Lett. 125, 115–119.
[29] Lütcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F. and Scheele, G.A. (1987) EMBO J. 6, 43–48.
[30] Schmidt, G.W. and Mishkind, M.L. (1986) Annu. Rev. Biochem. 55, 879–912.
[31] Karlin-Neumann, G.A. and Tobin, E.M. (1986) EMBO J. 5, 9–13.
[32] Roise, D. and Schatz, G. (1988) J. Biol. Chem. 263, 4509–4511.
[33] Keegstra, K. (1989) Cell 56, 247–253.
[34] Keegstra, K. and Olsen, L.J. (1989) Annu. Rev. Plant Physiol. Plant Mol. Biol. 40, 471–501.
[35] Smeekens, S., Bauerle, C., Hageman, J., Keegstra, K. and Weisbeek, P. (1986) Cell 46, 365–375.
[36] Smeekens, S. and Weisbeek, P. (1988) Photosynth. Res. 16, 177–186.
[37] Lautner, A., Klein, R., Ljungberg, U., Reiländer, H., Bartling, D., Andersson, B., Reinke, H., Beyreuther, K. and Herrmann, R.G. (1988) J. Biol. Chem. 263, 10077–10081.
[38] Rao, J.K. and Argos, P. (1986) Biochim. Biophys. Acta 869, 197–214.