Biochimica et Biophysica Acta 933 (1988) 501–505
Elsevier

BBA 40212

Partial amino acid sequences of two nuclear-encoded Photosystem I polypeptides from barley

Henrik Vibe Scheller [a], Peter Bordier Høj [a], Ib Svendsen [b] and Birger Lindberg Møller [a]

[a] *Department of Plant Physiology, Royal Veterinary and Agricultural University, Frederiksberg C, and* [b] *Department of Chemistry, Carlsberg Laboratory, Valby (Denmark)*

(Received 12 October 1987)

Key words: Photosystem I; Thylakoid polypeptide; Amino acid sequence homology; Nuclear encoding; (*H. vulgare*)

Partial amino acid sequences of two nuclear-encoded Photosystem I polypeptides of barley (*Hordeum vulgare* L.) have been determined. The sequence data include the N- and C-terminal parts of the polypeptides. The N-terminal parts of the polypeptides are homologous and rich in proline and alanine. In one polypeptide, the homologous region is repeated. The two polypeptides are peripheral membrane proteins.

Photosystem I (PS I) preparations contain the reaction center protein P-700-chlorophyll *a*-protein 1 and a varying number of polypeptides with lower molecular masses. Of the latter, those polypeptides of apparent molecular masses 18, 16, 14 and 9 kDa belong to the PS I core [1–3]. P-700-chlorophyll *a*-protein 1 and the 9 kDa polypeptide have been identified as iron sulfur proteins carrying the electron acceptors X, A and B [2–4]. The function of the additional polypeptides is largely unknown, although these polypeptides have been implicated as important in mediating reduction of $P\text{-}700^+$ by plastocyanin and in mediating electron transfer to soluble ferredoxin [5].

In order to elucidate the molecular structure and function of the PS I components, it is necessary to determine their primary structure. In this paper we report partial amino acid sequences of the two nuclear encoded 16 and 18 kDa polypeptides isolated from PS-I preparations of barley.

The 16 and 18 kDa polypeptides were isolated from urea-treated PS-I vesicles of barley (*Hordeum vulgare* L. cv. Svaløfs Bonus) as described previously [3]. Prior to amino acid analysis, sequencing, and fragmentation, the cysteine residue of the 18 kDa polypeptide was modified with 2-vinylpyridine essentially as reported [3]. Chemical cleavage with CNBr, dilute acid and hydroxylamine was performed essentially as in Refs. 6 and 7. The polypeptide fragments generated by these treatments were purified by gel filtration on a column (0.9 × 60 cm) of Bio-Gel P-30, 100–200 mesh, or Bio-Gel P60, 100–200 mesh (Bio-Rad Laboratories, Richmond, CA). Equilibration and elution were carried out using 30% (v/v) HOAc. The polypeptide content of the fractions was monitored by SDS-polyacrylamide gel electrophoresis [8].

The amino acid composition of the purified polypeptides and of isolated fragments was analyzed using a Durrum D500 amino acid analyzer after hydrolysis for 24 h in 6 M HCl at 110 °C in evacuated, sealed tubes. Cysteine con-

Abbreviation: PS I, Photosystem I.

Correspondence: H.V. Scheller, Department of Plant Physiology, Royal Veterinary and Agricultural University, 40 Thorvaldsensvej, DK-1871 Frederiksberg C, Denmark.

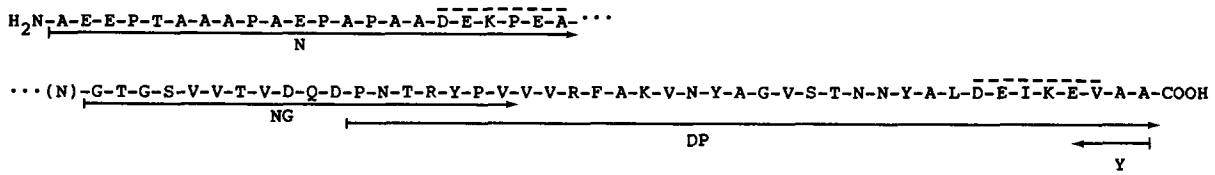

Fig. 1. Partial amino acid sequence of the 16 kDa polypeptide and of the fragments used to construct the sequence. N, N-terminal sequence obtained from the uncleaved polypeptide; DP, fragment obtained by limited acid hydrolysis; NG, fragment obtained by cleavage with hydroxylamine; Y, C-terminal sequence obtained by digestion with carboxypeptidase Y. The parentheses indicate tentative identification. The dashed overlines indicate amino acid sequences with little degeneration in their codons and are therefore suitable for the construction of oligonucleotide probes.

tent was determined as S-β-2-pyridylethylcysteine.

Amino acid sequencing was carried out as in Ref. 3. The C-terminal amino acid residues of the 16 and 18 kDa polypeptides were determined by digestion with carboxypeptidase Y [3,9].

The partial amino acid sequences obtained for the 16 and 18 kDa polypeptides are shown in Figs. 1 and 2. For the 16 kDa polypeptide, an N-terminal sequence of 22 residues and a C-terminal sequence of 45 residues were obtained. Sequencing of the intact 18 kDa polypeptide and fragments of this polypeptide yielded an N-terminal sequence of 40 residues plus two internal sequence fragments. Carboxypeptidase Y digestion enabled determination of the three amino acids constituting the C-terminal (Fig. 3).

The two polypeptides were assigned apparent molecular masses of 16 and 18 kDa on the basis of their electrophoretic mobilities on denaturing SDS-polyacrylamide gels [3]. From these masses, it can be estimated that the number of sequenced residues constitute about 45 and 60%, respectively, of the total sequences. The amino acid composition of the 16 and 18 kDa polypeptides and the composition of the sequenced portions are shown in Table I. Amino acid analysis of the 16 kDa polypeptide indicated a small content of methionine (Table I), but experiments with ^{35}S-labeled plants clearly demonstrated that the 16 kDa polypeptide does not contain sulfur amino acids [3].

The 16 kDa polypeptide contains no histidine residues and the 18 kDa polypeptide contains only one. Since histidine residues have been implicated in the non-covalent binding of chlorophyll to the protein backbone of chlorophyll proteins [10], it is unlikely that the 16 and 18 kDa polypeptides represent labile chlorophyll proteins.

The polarity index of the 16 and 18 kDa polypeptides calculated according to Capaldi and Vanderkooi [11] is 44 and 40%, respectively. Thus, the 16 kDa polypeptide appears to be more hydrophilic than the 18 kDa polypeptide. Both polypeptides contain slightly hydrophobic amino acid segments (VVVRFAKV for the 16 kDa polypeptide and APAGFVP, PIFGGG and VFPIEV for the 18 kDa polypeptide), but none of these segments are sufficiently long to represent membrane-spanning regions. Of the polypeptides constituting the PS-I core, the 16 and 18 kDa poly-

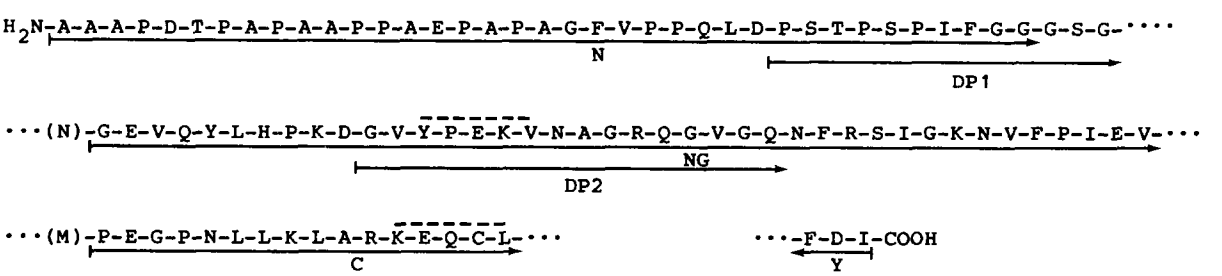

Fig. 2. Partial amino acid sequence of the 18 kDa polypeptide and of the fragments used to construct the sequence. DP1 and DP2, fragments obtained by acid hydrolysis; C, fragment obtained by cleavage with CNBr. Other symbols: see legend to Fig. 1. The position of fragment C relative to fragments NG and DP2 is arbitrary.

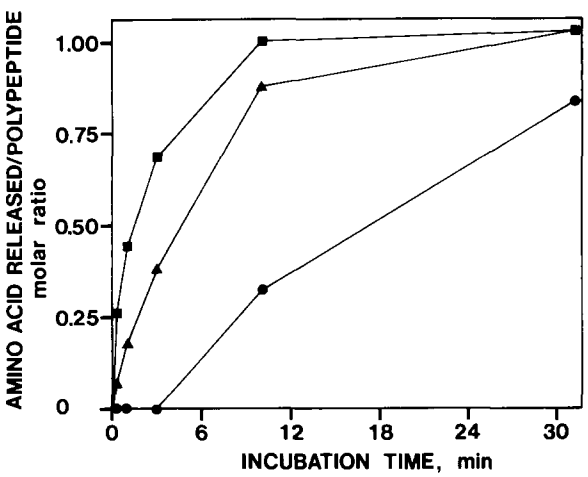

Fig. 3. Determination of C-terminal amino acid in the 18 kDa polypeptide by digestion with carboxypeptidase Y. Purified polypeptide (22 nmol) was incubated with carboxypeptidase Y (0.55 μg). Norleucine (15 nmol) was added as an internal standard. Release of other amino acids than isoleucine (■), asparctic acid (▲), and phenylalanine (●) was not detected.

peptides are less tightly bound as judged from their preferential dissociation from the PS-I vesicles upon urea treatment and chromatofocusing [3]. Treatment of the PS-I vesicles with proteinase K caused the 16 kDa polypeptide to be degraded much faster than the 18 kDa polypeptide (data not shown). These observations support the assignment of the 16 kDa polypeptide as a more peripherally located polypeptide compared to the 18 kDa polypeptide. It remains to be seen whether the unsequenced parts of the two polypeptides contain membrane spanning hydrophobic segments which serve to anchor the polypeptides or whether these polypeptides are bound to the other subunits of the PS-I complex solely by

TABLE I

AMINO ACID COMPOSITION OF THE 16 kDa AND 18 kDa POLYPEPTIDES

A: Average values of three different preparations. Each analysis was made with 1–2 nmol polypeptide. A total of 149 and 165 residues (not counting possible Trp residues) was assumed for the 16 and 18 kDa polypeptide, respectively. Based on the amino acid composition reported, the calculated molecular masses would be 15.7 and 17.7 kDa. B: Nearest integer values. n.d., not determined.

	Number of residues					
	16 kDa			18 kDa		
	total		sequenced	total		sequenced
	A	B		A	B	
Asp	13.35	13	4	11.80	12	4
Asn [a]	–	–	4	–	–	4
Thr	9.95	10	5	7.48	7	2
Ser	6.40	6	2	8.53	9	4
Glu	16.17	16	7	18.10	18	6
Gln [a]	–	–	1	–	–	5
Pro	18.44	18	7	17.76	18	17
Gly	9.81	10	3	17.91	18	12
Ala	25.68	26	14	17.23	17	11
Cys	0.06 [b]	0	0	0.86 [b]	1	1
Val	12.37	13	9	10.65	11	7
Met	0.35	0	0	1.61	2	0
Ile	2.79	3	1	6.49	6	4
Leu	3.84	4	1	11.71	12	6
Tyr	6.68	7	3	5.26	5	2
Phe	2.02	2	1	8.61	9	5
His	0.29	0	0	0.56	1	1
Lys	11.57	12	3	9.90	10	5
Arg	9.23	9	2	8.65	9	3
Trp	n.d.	n.d.	0	n.d.	n.d.	0
Total	149.00	149	67	163.11	165	99

[a] Asn and Gln total number are included in the values for Asp and Glu, respectively.

[b] Determined as 2-pyridylethyl-cysteine.

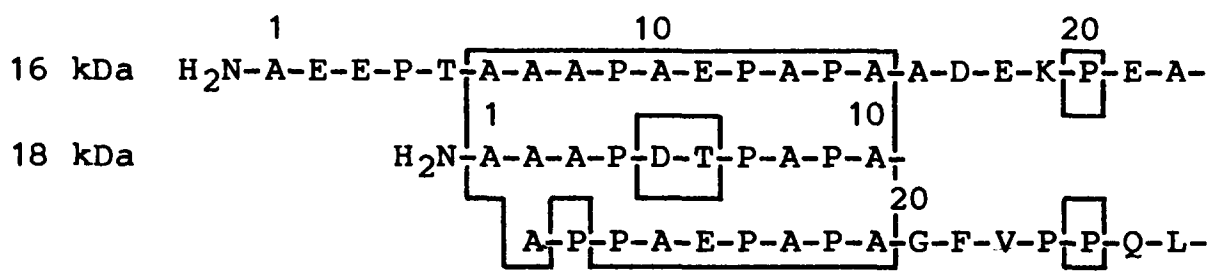

Fig. 4. Intermolecular homology in the N-terminal amino acid sequences of the 16 and 18 kDa polypeptides and intramolecular homology in the N-terminal sequence of the 18 kDa polypeptide. Residues that are identical in the two polypeptides are boxed.

hydrophobic interaction and/or electrostatic forces. Nearly all the basic residues of the 16 kDa polypeptide are located in the unsequenced region.

The N-terminal ends of the 16 and 18 kDa polypeptides reveal a high content of proline and alanine. These N-terminal ends are homologous to each other (Fig. 4). Additionally, the homologous sequence is repeated in the 18 kDa polypeptide (Fig. 4). The other regions of the polypeptides in which the sequence was determined show no convincing inter- or intramolecular homology. Therefore, the 5′ regions of the genes for the two polypeptides may be derived from a common ancestral gene or gene fragment which was added onto the unrelated 3′ regions. It shall be interesting to see whether the 5′ homology extends beyond the nucleotide and amino acid sequences corresponding to the mature polypeptides.

An extrinsic or peripheral 18 kDa thylakoid polypeptide associated with photosynthetic oxygen evolution also possesses an N-terminus rich in proline [12], but is not homologous to the N-termini of the 16 and 18 kDa polypeptides of PS-I presented in this study. No convincing homologies to the partial sequences reported here were found in the data base provided by the National Biomedical Research Foundation.

The easiest way to determine the remaining amino acid sequences of the 16 and 18 kDa polypeptides will be from the DNA sequence of their cloned genes. The partial amino acid sequences obtained for the 16 and 18 kDa polypeptides reveal no homology with those deduced from the open reading frames listed for the chloroplast genomes of tobacco [13] and the liverwort *Marcantia polymorpha* [14]. In agreement with results obtained by in vivo labeling [15], in vitro translation [16] and inhibitor studies [17], it is concluded that the 16 and 18 kDa polypeptides are nuclear encoded. The partial amino acid sequences for both polypeptides contain segments with little degeneration in the corresponding codons (Figs. 1 and 2), permitting the construction of oligonucleotide probes. Thus, the partial sequences reported here should enable other workers in the field to initiate cloning experiments. Nuclear-encoded thylakoid polypeptides are known to be synthesized as precursors, containing N-terminal transit sequences specifying import into and routing within the chloroplast [18]. Thus, it has been shown that the 16 kDa polypeptide has a pre-sequence of approx. 4 kDa [16]. These pre-sequences are removed by specific peptidases and the N-terminal amino acid segments reported here will, when the corresponding genes have been cloned, also permit determination of the final maturation site.

This work was supported in part by grants from the Danish Natural Science Research Council, The Danish Agricultural Research Council, Dansk Investeringsfond, Thomas B. Thriges Foundation, The Carlsberg Foundation, The Tuborg Foundation, Stiftelsen Hofmansgave, and by a Niels Bohr grant from the Royal Danish Academy of Sciences and Letters.

Hanne Linde Nielsen, Inga Olsen, Bodil Corneliussen, Pia Breddam and Lone Sørensen are thanked for skillful technical assistance. Drs. Barbara Ann Halkier and David Simpson are thanked for many helpful discussions and Dr. Torben Graves Pedersen for help with computer search for homologous proteins.

References

1 Lagoutte, B., Sétif, P. and Duranton, J. (1984) FEBS Lett. 174, 24–29.
2 Høj, P.B. and Møller, B.L. (1986) J. Biol. Chem. 261, 14292–14300.
3 Høj, P.B., Svendsen, I., Scheller, H.V. and Møller, B.L. (1987) J. Biol. Chem. 262, 12676–12684.
4 Golbeck, J.H. and Cornelius, J.M. (1986) Biochim. Biophys. Acta 849, 16–24.
5 Bengis, C. and Nelson, N. (1977) J. Biol. Chem. 252, 4564–4569.
6 Kamp, R.M. (1986) in: Advanced Methods in Protein Microsequence Analysis (Wittmann-Liebold, B., Salnikow, J. and Erdmann, V.A., eds.), pp. 8-20, Springer-Verlag, Berlin.
7 Allen, G. (1981) Sequencing of Proteins and Peptides, pp. 43–71, Elsevier/North-Holland, Amsterdam.
8 Fling, S.P. and Gregorson, D.S. (1986) Anal. Biochem. 155, 83–86.
9 Martin, B., Svendsen, I. and Ottesen, M. (1977) Carlsberg Res. Commun. 42, 99–102.
10 Zuber, H. (1985) Photochem. Photobiol. 42, 821–844.
11 Capaldi, R.A. and Vanderkooi, G. (1972) Proc. Natl. Acad. Sci. USA 69, 930–932.
12 Kuwabara, T., Murata, T., Miyao, M. and Murata, N. (1987) in Progress in Photosynthesis Research (Biggins, J., ed.), Vol. I, pp. 705–708, Martinus Nijhoff, Dordrecht.

13 Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, T., Kato, A., Tohdoh, N., Shimada, H. and Sugiura,. M. (1986) EMBO J. 5, 2043–2049.
14 Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S.-i., Inokuchi, H. and Ozeki, H. (1986) Nature 322, 572–574.
15 Obokata, J. (1986) Plant Physiol. 81, 705–707.
16 Høyer-Hansen, G., Hønberg, L.S. and Høj, P.B. (1985) Carlsberg Res. Commun. 50, 211–221.
17 Nechushtai, R. and Nelson, N. (1981) J. Biol. Chem. 256, 11624–11628.
18 Colman, A. and Robinson, C. (1986) Cell 46, 321–322.