Complete amino acid sequence of 33 kDa protein isolated from spinach photosystem II particles

H. Oh-oka, S. Tanaka, K. Wada°, T. Kuwabara+ and N. Murata*

Department of Biology, Faculty of Science, Osaka University, Toyonaka, Osaka 560, +Department of Chemistry, Faculty of Science, Toho University, Miyama, Funabashi, Chiba 274 and **Department of Regulation Biology, National Institute of Basic Biology, Okazaki, Aichi 444, Japan*

Received 12 December 1985

The amino acid sequence of the 33 kDa protein isolated as an extrinsic protein from spinach photosystem II particles was determined by using solid-phase sequencing and conventional procedures. The 33 kDa protein was found to be composed of 248 amino acid residues, to lack histidine and to have a molecular mass of 26 663 Da, which was considerably smaller than the value deduced from SDS-polycrylamide gel electrophoresis. The sequence of the 33 kDa protein was compared with those of the bacterial superoxide dismutases (SOD) with Mn atoms at an active site. A part of the sequence of the 33 kDa protein was similar to a region in Mn-SODs from *Bacillus stearothermophilus* and *Escherichia coli*, which was expected to be the Mn-binding site.

33 kDa protein *Photosynthesis* *Oxygen evolution* *Amino acid sequence* *Manganese-binding site*

1. INTRODUCTION

Since the photosystem II particles which retain a high oxygen-evolving activity were prepared with Triton X-100 treatment [1,2] of choloroplast membrane fragments, many experimental data were accumulated in order to elucidate the mechanism of oxygen evolution. Protein components in these particles are known to be the light harvesting-chlorophyll protein, reaction center antenna protein (43 kDa), reaction center core protein (47 kDa), cytochrome *b*-559, herbicide-binding protein (32 kDa) and 34 kDa protein (function unknown) as the intrinsic proteins and 3 extrinsic proteins with molecular masses of 18, 24 and 33 kDa [3]. Concentrated NaCl can dissociate the 18 and 24 kDa proteins from the particles with partial loss of oxygen-evolving activity [4]. When photosystem II particles were treated with concentrated Tris buffer or subjected to high pH, these 3 extrinsic proteins were released from the particles concomitantly with losses of Mn atoms and the oxygen-evolving activity [1,5,6], indicating that these extrinsic proteins were components of the oxygen-evolving complex. The photosystem II particles washed with $CaCl_2$ [7] or urea plus NaCl [8] removed the 33 kDa protein but preserved Mn atoms. The oxygen-evolving activity of the washed particles was appreciably restored by the addition of Cl^- [8] or rebinding of the 33 kDa protein [9,10]. The essential role of the 33 kDa protein is the preservation of Mn in the photosynthetic oxygen-evolving complex, and Cl^- can partially substitute the 33 kDa protein [8,10]. The 33 kDa protein was isolated first from spinach photosystem II particles by isoelectric focusing [11], and then from acetone powder of spinach thylakoids by column chromatography [12]. It has been well characterized chemically and physicochemically [11,12].

To clarify further the function of the 33 kDa protein in photosystem II particles, the amino acid sequence of this protein needs to be determined. This paper reports the complete amino acid se-

° To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

quence of the 33 kDa protein from spinach chloroplasts and a sequence of the 33 kDa protein from spinach chloroplasts and a sequence similarity to the Mn-binding site of Mn superoxide dismutases.

2. MATERIALS AND METHODS

All reagents and chemicals for determination of amino acid sequence are described in [13]. The 33 kDa protein was prepared from spinach chloroplast photosystem II particles as described [10]. The protein was carboxymethylated by the method of Crestfield et al. [14]. Sequence determination was carried out essentially as described in [13]. The digest by staphylococcal V8 protease (Miles Laboratories, Elkhart, IN) was fractionated by Toyopearl HW-50F (Toyo Soda, Tokyo, Japan) column chromatography. The peptide fractions were further purified by reverse-phase high performance liquid chromatography (Gilson HPLC system, Gilson France S.A., Villiers le Bel, France). Large peptides were further digested with TPCK-treated trypsin, chymotrypsin or thermolysin and purified as described above. Another lot of carboxymethyl (Cm)-33 kDa protein was digested with trypsin. The resultant peptides were purified and sequenced to overlap the peptides derived from the staphylococcal V8 protease digestion. To confirm the overlap of peptides, succinylated Cm-33 kDa protein was digested with trypsin and sequenced.

Amino acid compositions of proteins and peptides were determined with an IRICA amino acid ana-

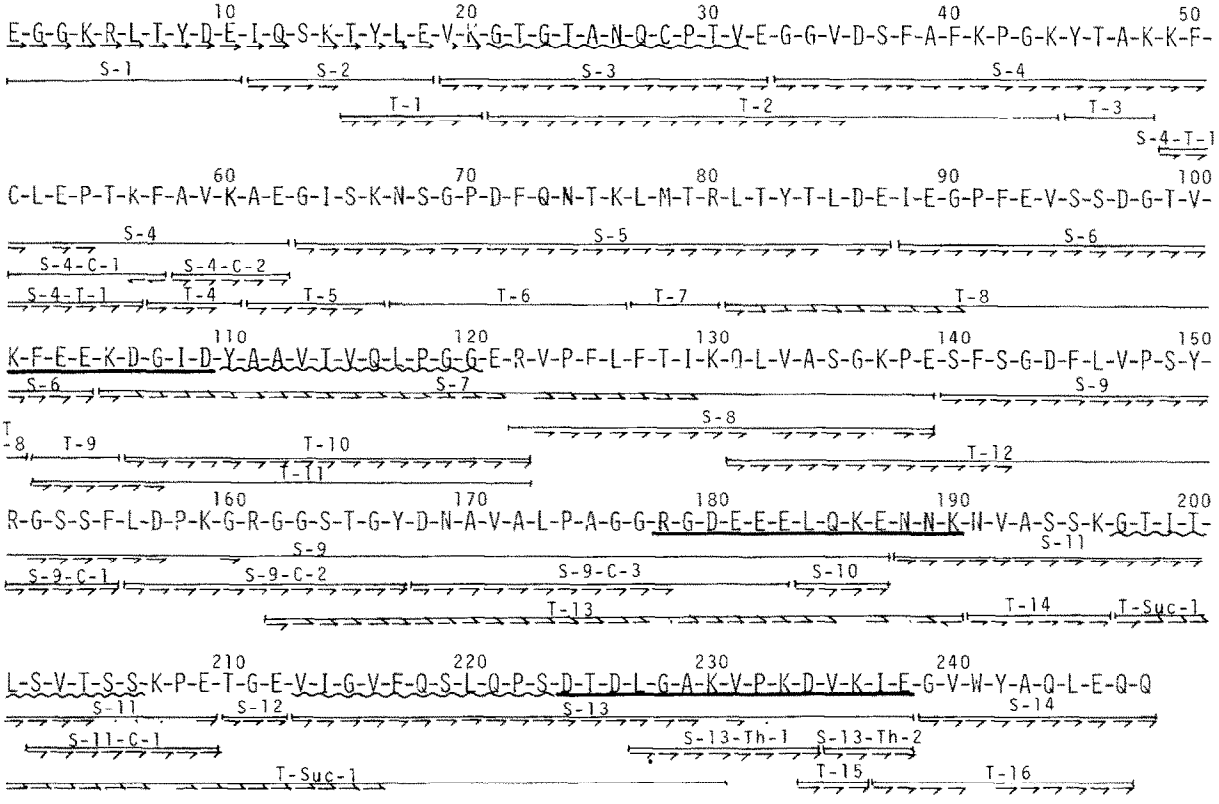

Fig.1. A summary of the sequence studies of spinach 33 kDa protein. (→, → and ←) Solid-phase, manual Edman degradation and carboxypeptidase A and/or B digestion, respectively. (→) N-terminal sequence analysis of the whole 33 kDa protein by manual Edman degradation. S-, T-, C- and Th-; staphylococcal V8, tryptic, chymotryptic and thermolytic peptides, respectively. T-Suc-, tryptic peptide derived from the succinylated Cm-33 kDa protein. The dotted arrows denote the ambiguous identifications of the amino acid derivatives. ~~~ and ▬ show the regions of the successive non-charged amino acids and the regions rich in the charged amino acids, respectively.

lyzer, model A-5500 (IRICA Instrument, Kyoto, Japan).

3. RESULTS AND DISCUSSION

The N-terminal sequence of the 33 kDa protein was determined by the manual Edman degradation method up to the 20th step as shown in fig.1. This sequence was different from that reported by Wölter et al. [15] in 5 out of the 16 N-terminal residues compared. Reasons for this discrepancy are not clear at present.

Fourteen peptides (S-1 to S-14) derived from the straphylococcal V8 protease digestion covered the whole sequence of this protein. Tryptic peptides gave overlap for straphylococcal peptides except for peptides S-11, S-12 and S-13. An overlapping peptide (T-Suc-1) for these 3 was obtained from the tryptic digest of the Cm-33 kDa protein insufficiently succinylated. The C-terminal amino acid of the protein was identified as glutamine, asparagine or threonine by carboxypeptidase A digestion. All the straphylococcal protease peptides had glutamic acid as the C-terminus except for peptide S-14, which terminated in glutamine of the sequence -Leu-Glu-Gln-Gln and therefore, S-14 was concluded to be the C-terminal peptide. However, the tryptic peptide T-16, which also lacked lysine or arginine, terminated in Leu-Glu-Gln. These results suggested that the C-terminus of the 33 kDa protein was heterogeneous.

A summary of the sequence study of the 33 kDa protein from spinach thylakoid membrane is given in fig.1. This protein contains 2 cysteine residues at the 28th and 51st position and one each of methionine and tryptophan. An apparent intramolecular duplication was found in the sequence: the sequence from Arg-5 to Gln-12 was quite similar to the sequence from Arg-80 to Glu-88 with insertions of Thr-83 and Leu-84. The 33 kDa protein was fairly hydrophilic [12] and comprised 64 charged amino acid residues. There are 3 regions rich in the charged amino acids: Lys-101 to Asp-109, Arg-178 to Lys-190 and Asp-224 to Glu-238. The second region contained 8 charged residues in 13 residues. In contrast, 4 regions comprised 10–11 successive non-charged amino acids: Gly-21 to Val-31, Tyr-110 to Gly-120, Gly-197 to Ser-206 and Val-213 to Ser-223, respectively.

The total number of amino acid residues of this protein was 248 giving a molecular mass of 26663 Da. This value disagrees with that deduced (33 kDa) from SDS-polyacrylamide gel electrophoresis. Amino acid compositions calculated from the sequence data and amino acid analysis in this experiment are shown in table 1. Our independent analysis of the 33 kDa protein (table 1, column (2)) agreed with that of [12] (column (1)), both of which were calculated for the molecular mass of 33 kDa. However, if these values were adjusted to take account of 27 kDa obtained from the sequence, the composition recalculated (column (3)) agreed very well with that deduced from the sequence, indicating that the sequence presented here is probably correct.

In general a glycoprotein behaves on SDS-polyacrylamide gel electrophoresis as if it had a larger molecular mass. After methanolysis and trimethylsilylation of this protein, the sugar content was

Table 1

Amino acid compositions of 33 kDa protein

	(1)	(2)		(3)		(4)
Cmc	2	2.11	2	1.71	2	2
Asp	26	27.0	27	21.9	22	21
Thr	24	24.3	24	19.7	20	20
Ser	24	24.2	24	19.6	20	20
Glu	39	41.4	41	33.7	34	32
Pro	18	18.4	18	14.9	15	14
Gly	36	36.2	36	29.4	29	29
Ala	18	17.7	18	14.4	14	14
Val	23	23.8	24	18.9	19	19
Met	1	1.21	1	0.98	1	1
Ile	10	9.73	10	7.71	8	8
Leu	21	21.7	22	17.2	17	17
Tyr	10	9.51	10	7.72	8	8
Phe	16	15.8	16	12.8	13	13
Lys	29	27.0	27	21.9	22	23
His	0	0	0	0	0	0
Arg	8	7.04	7	5.71	6	6
Trp	(1)*		(1)*		(1)*	1
Total	306		308		251	248
M_r			33082		27023	26663

(1) From Kuwabara and Murata [12]; (2) and (3) were calculated from the data in this experiment based on 2 different molecular masses of 33 and 27 kDa, and (4) from sequence analysis. *Determined from the spectrophotometric analyses

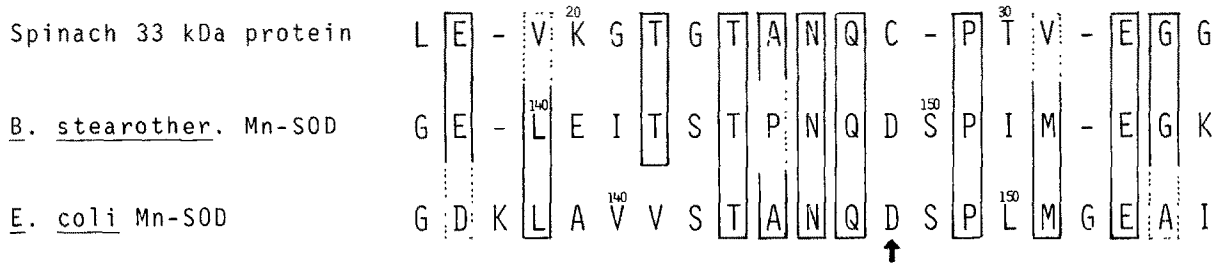

Fig.2. Comparison of similar parts of sequences of the 33 kDa protein and bacterial Mn superoxide dismutases. The common and similar amino acids are framed with solid and dotted lines, respectively. The arrow shows the site of one of the ligands to the Mn atom [18].

measured by using gas chromatography, but no sugar was found as a component. Therefore, the discrepancy between the molecular masses obtained by sequence analysis and SDS electrophoresis remain to be solved by future experiments.

The sequence of the 33 kDa protein was compared with those of bacterial Mn-superoxide dismutases [16,17]. As shown in fig.2, we found a considerable similarity between them. The aspartic acid shown with an arrow is considered to be a candidate for the ligands to the Mn atom [18]. Although the aspartic acid residue is substituted by a cysteine residue in the 33 kDa protein, these similarities suggest that this part of the sequence in the 33 kDa protein may play an important role in retaining Mn atom(s) situated between the 33 kDa protein and the reaction center core protein [3].

ACKNOWLEDGEMENTS

This work was supported in part by a grant for scientific researches from the Japanese Ministry of Education. The authors wish to thank Professor H. Matsubara for his encouragement during the course of this work and Mr Y. Yao for his thoughtful advice on sequencing.

REFERENCES

[1] Kuwabara, T. and Murata, N. (1982) Plant Cell Physiol. 23, 533–539.
[2] Berthold, D.A., Babcock, G.T. and Yocum, C.F. (1981) FEBS Lett. 134, 321–324.
[3] Murata, N. and Miyao, M. (1985) Trends Biochem. Sci. 10, 122–124.
[4] Kuwabara, T. and Murata, N. (1983) Plant Cell Physiol. 24, 741–747.
[5] Åkerlund, H.-E. and Jansson, C. (1981) FEBS Lett. 124, 229–232.
[6] Yamamoto, Y., Doi, M., Tamura, N. and Nishimura, M. (1981) FEBS Lett. 133, 265–268.
[7] Ono, T. and Inoue, Y. (1984) FEBS Lett. 164, 255–260.
[8] Miyao, M. and Murata, N. (1984) FEBS Lett. 170, 350–354.
[9] Ono, T. and Inoue, Y. (1984) FEBS Lett. 166, 381–384.
[10] Kuwabara, T., Miyao, M., Murata, T. and Murata, N. (1985) Biochim. Biophys. Acta 806, 283–289.
[11] Kuwabara, T. and Murata, N. (1979) Biochim. Biophys. Acta 581, 228–236.
[12] Kuwabara, T. and Murata, N. (1982) Biochim. Biophys. Acta 680, 210–215.
[13] Wakabayashi, S., Matsubara, H., Kim, C.H. and King, T.E. (1982) J. Biol. Chem. 257, 9335–9344.
[14] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–627.
[15] Wölter, F.P., Schmitt, J.M., Bohnert, H.J. and Tsugita, A. (1984) Plant Sci. Lett. 34, 323–334.
[16] Brook, C.J. and Walker, J.E. (1980) Biochemistry 19, 2873–2882.
[17] Steinman, H.M. (1978) J. Biol. Chem. 253, 8708–8720.
[18] Marres, C.A., Van Loon, A.P.G.M., Oudshoorn, R., Van Steeg, H., Grivell, L.A. and Slater, E.C. (1985) Eur. J. Biochem. 147, 153–161.