Primary Structure Characterization of the Photosystem II D1 and D2 Subunits*

Jyoti Sharma, Maria Panico, Catherine A. Shipton, Fredrick Nilsson, Howard R. Morris, and James Barber‡

From the Wolfson Laboratories, Department of Biochemistry, Imperial College, London, SW7 2AY, United Kingdom

Mass spectrometry techniques have been applied in a protein mapping strategy to elucidate the majority of the primary structures of the D1 and D2 proteins present in the photosystem II reaction center. Evidence verifying the post-translational processing of the initiating methionine residue and acetylation of the free amino group, similar to those reported for other higher plant species, are presented for the two subunits from pea plants (*Pisum sativum L.*). Further covalent modifications observed on the D1 protein include the COOH-terminal processing with a loss of nine amino acids and phosphorylation of Thr. In addition, the studies reported in this paper provide the first definitive characterization of oxidations on specific amino acids of the D1 and D2 proteins. We believe that these oxidations, and to a much lesser extent the phosphorylations, are major contributors to the heterogeneity observed during the electrospray analysis of the intact subunits reported in the accompanying paper (Sharma, J., Panico, M., Barber, J., and Morris, H. R. (1997) *J. Biol. Chem.* 272, 33153–33157). Significantly, all of the regions that have been identified as those particularly susceptible to oxidation are anticipated (from current models) to be in close proximity to the redox active components of the photosystem II complex.

The accurate molecular mass information afforded by techniques such as electrospray ionization (ESI)† mass spectrometry (MS) is particularly useful in cases where gene sequences are available; a correlation between experimental and expected measurements can be sufficient in confirming the identity and homogeneity of intact proteins. However, in situations where molecular weight discrepancies are detected or more detailed structural characterization is required, an MS peptide mass mapping strategy is usually applied (1). This procedure uses suitable MS techniques to analyze the peptides generated by enzymatic and/or chemical digestion of the protein under investigation. The molecular weights for the products must match within experimental error those predicted from the protein sequence; those that do not are then presumed to correspond to peptides containing sequence errors or post-translational modifications, or they may provide evidence for contaminating proteins.

Detailed structural information of proteins is of fundamental importance in developing an understanding of their biological activities. It therefore follows that analysis of the primary structural features that govern the higher orders of structures of proteins are of considerable interest. The application of a wide range of biochemical and molecular biological studies has identified several post-translational modifications on the D1 and D2 proteins of photosystem II (PSII) reaction centers, for example NH₄⁺ and COOH-terminal proteolytic processing (2, 3), acetylation (2), lipid attachment and acylation (4, 5), and phosphorylation (2, 6). Most of these covalent processes have been associated with the biosynthesis and assembly of active photosynthetic complexes. By contrast, details of the structural modifications that influence other aspects of PSII function, in particular the triggering and subsequent degradation of damaged PSII reaction center proteins that are observed during illumination conditions, remain to be resolved (7). The majority of experimental evidence obtained to date favors the possibility of oxidative damage resulting in conformational changes and eventually the degradation of the D1 subunit and to a lesser extent the D2 subunit (8).

In the accompanying paper (34) we described how we utilized the powerful ESI-MS technique to examine the D1 and D2 proteins that had been isolated from dark treated PSII reaction center preparations. The data obtained indicated the presence of molecular heterogeneity within the populations of both subunits. In the studies reported in this paper we have employed a protein mapping strategy to characterize the nature and sites of the structural modifications that result in the observed diversity.

**MATERIALS AND METHODS**

Isolation of PSII Reaction Centers—These procedures are detailed in the accompanying paper (34).

Preparative Polycrylamide Gel Electrophoresis and Gel Elution—Separation of the D1 and D2 protein bands to enable isolation of these subunits was performed by preparative polycrylamide gel electrophoresis using the LKB 2001 vertical electrophoresis apparatus. Linear 2-mm-thick 15% non–SDS-polyacrylamide gels containing 6 M urea were used for this procedure. The buffer system was 0.2 M glycine, 25 mM Tris-HCl, and 0.1% SDS. The reaction center samples were solubilized in an equal volume of 25 mM Tris-HCl, 4% SDS, and 2% β-mercaptoethanol (pH 8.9) before application to the gel (a maximum of 200 μg of chlorophyll was loaded on each gel). After the resolution of the D1 and D2 proteins by preparative polycrylamide gel electrophoresis, the proteins were identified by negative staining using a 5% (w/v) solution of copper chloride in 0.2 M glycine, 25 mM Tris-HCl. The relevant bands were cut out and washed in a solution of 0.2 M glycine, 25 mM Tris-HCl (elution buffer) to remove any excess copper solution. The gel slices were then reduced by incubation in a solution of 10% β-mercaptoethanol in elution buffer. Finally, the proteins were eluted from the gel slices into elution buffer using a Hoefer Scientific Instruments GT114 tube gel unit, running at 4 °C, 10 mA/tube for 4 h.

* Cyanogen Bromide (CNBr) and Trypsin Digestion of the D1 and D2

This paper is available on line at http://www.jbc.org
Proteins for Mapping—The D1 and D2 proteins purified by the gel elution technique were dialyzed against 2 × 2000 ml of 20% aqueous formic acid at 4 °C for 6 h. These samples were then made up to a concentration of 70% aqueous formic acid before the digestion. A few crystals of CNBr were added to the sample solutions, and the reaction mixture was incubated in the dark for 5 h at room temperature. The reaction mixture was diluted with 5 volumes of water and dialyzed against 2 × 2000 ml of 50 mM ammonium bicarbonate (pH 8.5) at 4 °C for 6 h. The appropriate amount of trypsin (1.50 w/w, enzyme:substrate) was added and incubated for 6 h at 37 °C. Digestion was terminated by freezing followed by lyophilization.

Reverse Phase High Performance Liquid Chromatography (HPLC) for the Separation of Peptide Mixtures—Partial separation of the polypeptides produced by enzymatic digestion was accomplished using an Aquapore OD-300 (220 × 4.6 mm) column fitted on a Kontron HPLC system (Datsystem 450, HPLC pump 420, detector 430, and mixer M500). The samples were dissolved in 0.1% trifluoroacetic acid (buffer A) for injection on the column. The column was held for 5 min at 0% B followed by a linear increase to 100% B (90% acetonitrile in aqueous 0.1% trifluoroacetic acid) over 60 min at a flow rate of 1.0 ml min⁻¹. Elution was monitored at 214 and 280 nm, and fractions were collected at 1-min intervals.

Manual Edman Degradation—The samples for manual sequencing were dried down in a glass tube and redissolved in 100 µl of Milli-Q water (Millipore system). 100 µl of a 5% solution of phenylisothiocyanate in pyridine was added to each sample, and the coupling reaction was flushed with nitrogen before its incubation at 45 °C. After 1 h this reaction was terminated by removing excess reagents under a stream of nitrogen. To each of the dried samples 100 µl of trifluoroacetic acid was added, and the mixtures were incubated for a further 10 min at 45 °C, causing the cleavage of the phenylthio- carbamyl-labeled terminal amino acids to form the thiodycyanate de- rivative. The products of the remaining truncated peptides were subsequently dried under nitrogen and analyzed by MS without the removal of phenylthiodicyanate derivatives.

FAB-MS Analysis—FAB mass spectra were acquired using a Fisons VG ZAB 2 SEFPD mass spectrometer fitted with a cesium ion gun operating at 30 kV. The instrument was calibrated using cesium iodide clusters. Data acquisition and processing were performed using VG Analytical Opus software. The samples were dissolved either in aqueous 5% acetic acid or in 40% propan-1-ol in 5% aqueous 5% acetic acid (for the more hydrophobic peptides), and 1-µl aliquots were loaded onto the probe that had been treated previously with a 1:1 mixture of glycerol:monothioglycerol.

ESI-MS Analysis—These procedures are detailed in the accompanying paper (34).

MALDI-MS Analysis—Matrix-assisted laser desorption ionization (MALDI) mass spectra were acquired using a Fisons VG ZAB 2 SEFPD mass spectrometer fitted with a UV laser (337 nm). The instrument was calibrated using cesium iodide clusters. Data acquisition and processing were performed using VG Analytical Opus software. The samples were dissolved in 60% propan-1-ol in 5% aqueous, and 1-µl aliquots were loaded onto the probe that had been treated with a saturated solution of 2,5-dihydroxybenzoic acid made up in a 70:30 v/v solution of 90% acetonitrile in aqueous 0.1% trifluoroacetic acid and aqueous 0.1% trifluoroacetic acid, respectively, followed by air-drying.

RESULTS

Purification of the D1 and D2 Proteins from PSII Reaction Center Preparations—The MS mapping strategy used in these studies to analyze the primary structure of the D1 and D2 proteins is facilitated greatly by the availability of pure protein. Thus an alternative purification protocol enabling better separation of these PSII subunits than that reported by our HPLC purification system (described in Ref. 34) was required. The D1 and D2 proteins share sequence homology and are of approximately the same mass, consequently distinct separation of these subunits is an extremely difficult achievement. One technique by which it has been possible to segregate these proteins is SDS-polyacrylamide gel electrophoresis; 15% acrylamide gels containing 6 M urea have been found to give a good separation of the two subunits. We have developed an isolation procedure based on this electrophoretic system followed by electroelution of proteins from the gel. This purification protocol was found to be far more efficient than the electroblotting procedure that we had originally developed for use in these mapping studies.

The main problem encountered during the development of our purification technique was that the use of SDS had to be minimized to enable MS analysis. Fortunately, the resolution of the D1 and D2 proteins on polyacrylamide gels that did not contain SDS (in the gel matrix) was found to be similar to those with SDS (Fig. 1), with the only obvious difference being that the former had a larger proportion of aggregated products. By contrast, analysis of the products after the gel elution procedure indicated that the presence of SDS had a marked effect on this step; in fact it was found that protein elution from gels that did not utilize SDS in the gel matrix was less efficient than that of the standard SDS-containing system. However, by incubating the non–SDS-containing gel slices in a buffer containing mercaptoethanol before electroelution we were able to increase the purification yields significantly.

The pure forms of the D1 and D2 proteins, isolated using the gel elution technique, were extremely hydrophobic, and even slight changes in the procedure led to their precipitation. Moreover, even the propan-1-ol and 5% acetic acid solvent system, which had enabled partial purification of the D1 and D2 for ESI-MS analysis (reported in Ref. 34), failed to elute these gel-purified proteins from the reverse phase HPLC column. Thus, special care has to be taken in handling these proteins after their isolation. We found that an effective method for mapping the D1 and D2 proteins avoided any drying or concentrating steps and instead utilized a dialysis strategy to transfer buffers. Application of two consecutive digestions, CNBr followed by trypsin digestion leading to the cleavage of all accessible methionine, arginine, and lysine residues, was found to give the greatest peptide yields.

Structural Characterization of the Gel-purified D1 Protein Obtained from PSII Reaction Center Preparations—The spectrum presented in Fig. 2 was obtained by direct FAB-MS analysis of an aliquot taken from the sucrose CNBr and tryptic digestion mixture of the D1 protein purified by the gel elution technique. Most of the observed quasimolecular ions can be assigned to expected D1 peptides (calculated from the protein sequence derived from pea psba gene; see Ref. 9), and some important information on the primary structure can also be obtained from this spectrum. For example, as can be seen, the quasimolecular ions detected at m/z 744.4, 824.4, 900.4, and

![Fig. 1. Polypeptide profile of the PSII reaction center preparation, resolved on 15% polyacrylamide gels containing 6 M urea.](image-url)
The quasimolecular ions \( m/z \) 744.4 and 900.4 differ by a mass equivalent to an arginine residue (156 atomic mass units), and their molecular weights correspond to that expected for the NH\(_2\)-terminal processed (des-Met) and acetylated peptides (2), acetyl-Thr\(^2\)-Arg\(^7\) and acetyl-Thr\(^2\)-Arg\(^8\), respectively. By analogy, the quasimolecular ions \( m/z \) 824.4 and 980.4 also differ by 156 atomic mass units; these signals are both 80 atomic mass units larger than the 744 and 900 components, respectively. The 80-atomic mass unit increment can be ascribed to a phosphoryl group (2), and these components are likely to represent the NH\(_2\)-terminal phosphorylated peptides, acetyl/phosphoryl-Thr\(^2\)-Arg\(^7\) and acetyl/phosphoryl-Thr\(^2\)-Arg\(^8\), respectively. Furthermore, although the spectrum is only semiquantitative, it is clear that in these samples the nonphosphorylated form predominates. Verification of these assignments was obtained after purification of the individual components.

Direct FAB-MS analysis of complicated mixtures, such as the one under investigation here, often leads to suppression effects (10). This phenomenon is usually overcome by partial separation of the digestion products; in this study we used reverse phase HPLC. All fractions correlating with regions of UV absorbance were screened by FAB-MS and ESI-MS. A large majority of the molecular ions detected during this process could be assigned to expected peptides (Table I; note that the methionine cleavage by CNBr results in the formation of homoserine, which is detected as the lactone in these experiments). These assignments have been verified by manual Edman reaction followed by MS analysis of the truncated peptides. The mass shift observed on all of these assigned components corresponds to the molecular weight of the expected NH\(_2\)-terminal residue (from the nucleotide-derived protein sequence) except on the NH\(_2\)-terminal tryptic peptides (Thr\(^2\)-Arg\(^7\) and Thr\(^2\)-Arg\(^8\)) where no mass shift was detected because the D1 protein is blocked at its NH\(_2\) terminus.

In fractions 27–29 the four molecular ions attributed to the processed NH\(_2\)-terminal peptides, which had been detected during examination of the digestion mixture, were identified. By contrast, the molecular ions representing the unprocessed NH\(_2\)-terminal peptide of the D1 protein were not observed in any of the UV-absorbing fractions. The NH\(_2\)-terminal processing event characterized for the pea D1 protein entails removal of the initiating formylmethionine residue followed by N-acetylation of the second residue, Thr\(^2\). This process is therefore similar to that reported for the D1 protein of spinach (2). In addition, our studies have also provided direct evidence for the presence of an NH\(_2\)-terminal phosphorylation site on the D1 subunits of pea plants, which is also analogous to that observed in its spinach counterpart (2). Even though it is difficult to quantitate exactly, careful analysis of these samples indicates that the phosphorylated form occurs in approximately one-third of the D1 population. This modification would thus contribute to the heterogeneity observed by the ESI-MS analysis of the intact protein (34).

The COOH terminus of the D1 protein in most photosynthetic organisms (except Euglena) is thought to be processed to the mature form (11). In spinach plants this event was found to result in the removal of nine amino acid residues 335–344 (3); however, as yet there has been no direct evidence confirming the occurrence of this modification in pea plants (the organism under investigation in these studies). During our analyses of the D1 protein we did not detect the quasimolecular ion corresponding to the unprocessed COOH-terminal peptide, \( m/z \) 1679.8, Asp\(^{335}\)-Gly\(^{353}\), although the quasimolecular ion at \( m/z \)
1111.6 which corresponds to the processed COOH-terminal component, Asn\(^{335}\)-Ala\(^{344}\), was observed readily (Fig. 2). These findings not only provide direct evidence verifying that the D1 protein in pea plants is processed proteolytically in a way similar to that reported for spinach and other plants, but it also indicates that at least in the granal regions of thylakoid membranes (the regions from which these PSII reaction centers were isolated) there is no heterogeneity at this modification site.

During the screening of the isolated products we observed that several molecular ions that corresponded to peptides of the D1 protein had additional satellite components at multiples of 16-atomic mass unit increments associated with them. The 16-atomic mass unit increment can be ascribed to an oxygen atom; thus it seems that the molecular ions with these mass increments are oxidized products. The D1 peptides found to carry this modification are presented in Table I. In some cases, for example, Glu\(^{130}\)-Arg\(^{136}\) (966.4), Gly\(^{128}\)-Arg\(^{136}\) (1179.6), Ala\(^{324}\)-Arg\(^{334}\) (1286.6), and Ile\(^{184}\)-Hse\(^{199}\) (1916.9), the quasimolecular ions corresponding to the oxidized peptides were quite prominent. On peptides Glu\(^{130}\)-Arg\(^{136}\) and Gly\(^{128}\)-Arg\(^{136}\), the quasimolecular ions \(m/z\) 982.4 and 1195.6, which are both 16 atomic mass units larger than the signals expected for these molecules, are the principal components detected. Similarly, for the peptide Ala\(^{324}\)-Arg\(^{334}\) the main quasimolecular ion observed, at \(m/z\) 1933.1, is 32 atomic mass units heavier (addition of two oxygen atoms) than that of its calculated mass, whereas on the peptide Ile\(^{184}\)-Hse\(^{199}\) quasimolecular ions at \(m/z\) 1318.6, 1949.1, and 1965.2 corresponding to additions of 16, 32, and 48 atomic mass units (one, two, and three oxygen atoms, respectively) were detected. Collectively, these results, which show the absence of oxidations on some peptides while others have up to three oxygen additions, indicate that there is a degree of specificity of this modification. According to the folding model of the D1 protein (Fig. 3), which has been developed by analogy to the L-subunit of the corresponding bacterial photosynthetic unit (28), the observed oxidation sites appear to be located primarily in two luminal regions, the COOH-terminal region and the beginning of the fourth transmembrane segment, and one stromal region, the end of the II transmembrane segment. This group of modifications exemplifies a further source of heterogeneity which may also explain the ESI-MS data obtained for the intact D1 protein presented in the accompanying paper (34).

At the time of transferring the CNBr digestion products into the buffer required for tryptic digestion we noted a slight precipitate. This product, which could probably be attributed to the poor solubility of some extremely hydrophobic peptides, remained insoluble in the solvent conditions used to initiate the HPLC separation. To complete the analysis of the D1 protein we analyzed this precipitate by MALDI-MS, a technique that is anticipated to be more amenable to hydrophobic molecules. The D1 protein digestion mixture was centrifuged and subsequently separated into soluble and precipitated phases. Direct MALDI-MS analysis of the precipitate identified (among other signals) the quasimolecular ion, \(m/z\) 2827.3 (Fig. 4). This com-

---

**TABLE I**

**MS mapping of the CNBr/tryptic peptides from D1 protein**

The expected \([M+H]^+\) were calculated by MacBioSpec. ppt corresponds to the precipitate fraction present in the CNBr/trypsin digestion mixture.

| Peptide | Fraction | Expected \([M+H]^+\) | Measured \([M+H]^+\) | Total + 16 atomic mass units observed |
|---------|----------|----------------------|----------------------|-------------------------------------|
| AcThr\(^2\)-Arg\(^7\) | 28-29 | 744.4 | 744.4 | 32 |
| AcThr\(^2\)-Arg\(^8\) | 28-29 | 744.4 | 744.4 | 32 |
| Ac/Glu\(^4\)-Arg\(^9\) | 27 | 900.5 | 900.5 | 28 |
| Ac/Glu\(^4\)-Arg\(^9\) | 27-28 | 900.4 | 900.4 | 28 |
| Asp\(^9\)-Arg\(^16\) | 26 | 976.4 | 976.5 | 24 |
| Ile\(^{35}\)-Arg\(^{64}\) | ppt | 2826.9 | 2827.3 | 23 |
| Gly\(^{126}\)-Arg\(^{136}\) | 29 | 1179.6 | 1179.6 | 23 |
| Glu\(^{128}\)-Arg\(^{136}\) | 29 | 1179.6 | 1179.6 | 23 |
| Arg\(^{148}\)-Hse\(^{172}\) | 41 | 1338.9 | 1338.7 | 18 |
| Pro\(^{172}\)-Hse\(^{183}\) | 32-34 | 1315.4 | 1315.4 | 15 |
| Ile\(^{184}\)-Hse\(^{194}\) | 30 | 1266.6 | 1266.6 | 14 |
| Ile\(^{184}\)-Hse\(^{194}\) | 30 | 1266.6 | 1266.6 | 14 |
| Leu\(^{205}\)-Hse\(^{214}\) | 36-40 | 1364.7 | 1364.6 | 10 |
| His\(^{215}\)-Arg\(^{235}\) | 26-29 | 1169.6 | 1169.6 | 12 |
| Glu\(^{226}\)-Arg\(^{238}\) | 22-23 | 1499.6 | 1499.6 | 8 |
| Phe\(^{339}\)-Arg\(^{357}\) | 31 | 2188.5 | 2188.4 | 6 |
| Leu\(^{356}\)-Arg\(^{369}\) | 31 | 1459.7 | 1459.6 | 5 |
| Ala\(^{391}\)-Arg\(^{412}\) | 32 | 2114.2 | 2114.2 | 3 |
| Val\(^{391}\)-Arg\(^{412}\) | 32 | 1314.8 | 1314.9 | 3 |
| Ala\(^{324}\)-Arg\(^{334}\) | 24 | 1286.6 | 1286.6 | 3 |
| His\(^{392}\)-Ala\(^{394}\) | 30 | 1533.6 | 1533.4 | 3 |
| Asn\(^{335}\)-Ala\(^{344}\) | 31-32 | 1111.6 | 1111.6 | 3 |

---

**Fig. 3.** Protein sequence of the D1 subunit. The circled residues are those that have been mapped, and the shaded regions correspond to the peptides with oxidized amino acids.
ponent can be assigned to peptide Ile\textsuperscript{38}-Arg\textsuperscript{64}, which had not been observed previously. The quasimolecular ion at \( m/z \) 2615.3, also present in this spectrum, is yet to be assigned.

It is quite clear from the above observations that the extreme hydrophobic properties coupled with the structural complexity of the D1 protein have made it a particularly difficult molecule to study; nevertheless the application of several complementary techniques has enabled a significant portion of this subunit to be mapped successfully and has also allowed characterization of several post-translational modifications. The overall map of proven MS assignments is shown in Fig. 3.

Structural Characterization of the Gel-purified D2 Protein Obtained from PSII Reaction Centers—Direct FAB-MS analysis of an aliquot from the successive CNBr and tryptic digestion experiment on the gel-purified D2 (Fig. 5) depicts a host of quasimolecular ions, most of which can be attributed to expected D2 peptides (derived from the pea \textit{psbD} sequence; see Ref. 12). Verification of these assignments has been carried out after purification of the individual components. One of the regions mapped by this process is the NH\textsubscript{2}-terminal tryptic peptide, acetyl-Thr\textsubscript{2}-Lys\textsubscript{7} (\( m/z \) 644.4); however, unlike the D1 protein, the quasimolecular ion with an 80-atomic mass unit increment (\( m/z \) 724.4), corresponding to the phosphorylated form of this peptide, was not immediately apparent in the FAB map spectrum. Subsequently, a more rigorous analysis of the D2 protein was carried out on the partially purified digestion products obtained by reverse phase HPLC.

All fractions correlating with regions of UV absorbance on the HPLC chromatogram were screened by FAB-MS and ESI-MS. A significant number of the molecular ions observed during this process could be assigned to expected D2 peptides (Table II). These assignments have been confirmed by manual Edman degradation followed by MS analysis. Collectively, the data obtained in this study have enabled a large proportion of the D2 protein to be mapped and have also revealed interesting structural information.

The D2 proteins of spinach plants have been shown previously to be NH\textsubscript{2}-terminally processed (2). This modification is similar to that reported for the D1 protein and entails the removal of the initiating methionine residue followed by N-acetylation. Michel \textit{et al.} (2) also found that the NH\textsubscript{2}-terminal regions of both the D1 and D2 were phosphorylated. In this study on the D2 protein isolated from pea plants, we did not detect the quasimolecular ion corresponding to the unprocessed NH\textsubscript{2}-terminal peptide, \( m/z \) 733.4 (Met\textsuperscript{1}-Lys\textsubscript{7}), although the quasimolecular ion at \( m/z \) 644.4, which corresponds to the NH\textsubscript{2}-terminal processed tryptic peptide acetyl-Thr\textsubscript{2}-Lys\textsubscript{7}, was observed in fraction 22. Careful analysis of the spectra obtained from the other fractions in this region did not reveal the presence of an NH\textsubscript{2}-terminal quasimolecular ion with an 80-atomic mass unit increment which would correspond to the phosphorylated form. The anticipated quasimolecular ion of the NH\textsubscript{2}-terminal phosphorylated peptide was also not present in the direct FAB analysis of the digestion mixture. There is no reason to suppose that the quasimolecular ion from the phosphorylated peptide would not have been detected, if present, since it was clearly seen in the D1 protein. Thus, it seems that the D2 protein that we purified from pea PSII reaction center preparations is NH\textsubscript{2}-terminally processed but not phosphorylated.

Several of the molecular ions that have been assigned to particular D2 peptides appear to have associated satellite signals at 16-atomic mass unit mass increments. The quasimolecular ion at \( m/z \) 2133.2 attributed to the peptide Arg\textsuperscript{266}-
Hse\textsuperscript{282} has only a single 16-atomic mass unit (one oxygen) satellite associated with it, whereas on the quasimolecular ion observed at $m/z$ 2623.1, which corresponds to peptide Ala\textsuperscript{328}-Arg\textsuperscript{349}, additions of 32 atomic mass units (two oxygens) and 48 atomic mass units (three oxygens) were observed (Table II and Fig. 6). In addition, during the screening of the HPLC fractions two quasimolecular ions, at $m/z$ 2396.6 and 2412.6, which also differed by 16 atomic mass units, were detected. These quasimolecular ions can be ascribed to the singly and doubly oxidized forms of the partial digestion peptide Phe\textsuperscript{183}-Hse\textsuperscript{200} ($m/z$ 2380.7), respectively. In this case the quasimolecular ion of the unmodified form of this peptide was only detected as a very minor component. As presented above, oxidation modifications were also observed for the D1 protein, although the level of oxidation of the D2 subunit did not appear to be as extensive as that found for the D1 protein. Based on the model of the D2 subunit (28) the oxidations observed would be located at two lumenal regions: one near the COOH terminus (Ala\textsuperscript{328}-Arg\textsuperscript{349}) of the protein and the other located on the helix connecting transmembrane regions III and IV (Phe\textsuperscript{183}-Hse\textsuperscript{200}), and one stromal region, on the V transmembrane segment (Arg\textsuperscript{266}-Hse\textsuperscript{282}). These types of modification are most probably one of the sources of heterogeneity attributed to the D2 protein from ESI-MS analysis (34).

As reported with the D1 protein, a slight precipitate appeared during the transfer of the D2 protein CNBr digestion products into buffer required for tryptic digestion. This precipitate was not soluble in the conditions used at the beginning of HPLC separation. Thus, after the digestions the sample mixture was fractionated by centrifugation to isolate the soluble and precipitated peptides, and the latter phase was then analyzed directly by magnetic sector MALDI-MS. By this procedure it was possible to detect the quasimolecular ion, $m/z$ 7242.0, which can be assigned to the peptide Ser\textsuperscript{136}-Hse\textsuperscript{199} ($m/z$ 7241.6) and covers a region that had not been mapped previously (Fig. 7). As can be seen there also appears to be a quasimolecular ion ($m/z$ 7274.4) which is at a 32-atomic mass unit

### Table II

| Peptide               | Fraction | Expected [M+H]\textsuperscript{+} | Measured [M+H]\textsuperscript{+} | Total +16 atomic mass units observed |
|-----------------------|----------|----------------------------------|-----------------------------------|-----------------------------------|
| AcThr\textsuperscript{2}-Lys\textsuperscript{7} | 22       | 644.4                            | 644.4                             | 22                                |
| AcThr\textsuperscript{2}-Lys\textsuperscript{10} | 26       | 1020.6                           | 1020.5                            | 26                                |
| Asp\textsuperscript{11}-Arg\textsuperscript{24} | 34       | 1818.5                           | 1818.6                            | 34                                |
| Trp\textsuperscript{105}-Arg\textsuperscript{135} | 40       | 3540.4                           | 3539.8                            | 40                                |
| Leu\textsuperscript{128}-Arg\textsuperscript{135} | 24       | 1032.6                           | 1032.6                            | 24                                |
| Gln\textsuperscript{130}-Arg\textsuperscript{135} | 23       | 763.4                            | 763.4                             | 23                                |
| Ser\textsuperscript{136}-Hse\textsuperscript{180} | ppt      | 7241.6                           | 7242.0                            | 22                                |
| Phe\textsuperscript{183}-Hse\textsuperscript{200} | 35       | 2380.7                           | 2380.7                            | 22                                |
| Gly\textsuperscript{201}-Arg\textsuperscript{234} | 37–39    | 3381.4                           | 3381.4                            | 37–39                             |
| Ala\textsuperscript{235}-Hse\textsuperscript{247} | 23       | 1440.6                           | 1440.6                            | 23                                |
| Phe\textsuperscript{253}-Lys\textsuperscript{255} | 32       | 1530.7                           | 1530.7                            | 32                                |
| Arg\textsuperscript{266}-Hse\textsuperscript{282} | 34–35    | 2133.3                           | 2133.2                            | 34–35                             |
| Leu\textsuperscript{273}-Arg\textsuperscript{295} | 38       | 2428.0                           | 2428.4                            | 38                                |
| Ser\textsuperscript{283}-Arg\textsuperscript{295} | 38–39    | 2492.5                           | 2492.8                            | 38–39                             |
| Ala\textsuperscript{296}-Arg\textsuperscript{305} | 25       | 1227.6                           | 1227.5                            | 25                                |
| Ala\textsuperscript{296}-Lys\textsuperscript{318} | 26–27    | 1547.5                           | 1547.6                            | 26–27                             |
| Asp\textsuperscript{319}-Arg\textsuperscript{327} | 25       | 1041.6                           | 1041.6                            | 25                                |
| Ala\textsuperscript{328}-Arg\textsuperscript{349} | 29       | 2622.9                           | 2623.1                            | 29                                |
| Ala\textsuperscript{331}-Arg\textsuperscript{349} | 28       | 2234.1                           | 2234.1                            | 28                                |
| Ala\textsuperscript{331}-Leu\textsuperscript{333} | ppt      | 2589.7                           | 2589.4                            | 22                                |
increment to the 7242.0 signal, indicating that the peptide Ser<sup>136</sup>-Hse<sup>199</sup> may have two oxidation modifications. A further quasimolecular ion detected by this procedure which has provided important structural information was the signal observed at m/z 2589.4. This component is attributed to the peptide Ala<sup>331</sup>-Leu<sup>353</sup> (m/z 2589.7), on the basis of mass and Edman data. The detection of this peptide not only confirms the identity of the COOH terminus of the D2 protein but also indicates that D2 is not processed COOH-terminally in the manner reported for the D1 subunit. All regions of the D2 protein which have been characterized successfully by these MS mapping studies are shown in Fig. 6.

**DISCUSSION**

MS combined with separation techniques provides a powerful analytical system for the structural study of proteins, and developments in this area remain of considerable importance. In particular, there has been a continued interest in procedures

---

**Fig. 6. Protein sequence of the D2 subunit.** The circled residues are those that have been mapped, and the shaded regions correspond to the peptides with oxidized amino acids.

**Fig. 7. MALDI mass spectrum of the precipitate, present in the CNBr/trypsin digestion mixture of the D2 protein.**
for coupling gel electrophoresis with MS analysis. Such a system would combine the highly effective separation and characterization technique afforded by gel electrophoresis with the ability to determine accurate molecular weights independent of the structural and charge features that often compromise electrophoretic mobility. Unfortunately, analyses of intact proteins directly from acrylamide gel systems have proved to be technically difficult, and reports on successful applications are few (13). By contrast some significant achievements have been reported using in-gel digest (14) and electroblotting systems (15). In this study we report the development of an electroelution purification system that is effective in isolating extremely hydrophobic membrane proteins with relatively good yields, free of detergent. We believe that this purification system could be widely applicable, and we are currently using it in the characterization of other membrane proteins of PSII, including the specific light-induced degradation fragments of the D1 and D2 proteins for which only limited structural information is available.

Structural analysis of the electroeluted D1 and D2 proteins presented in this paper not only confirms the majority of their amino acid sequences but also identifies a variety of post-translational modifications, the characterized features are summarized in Fig. 3 and 6, respectively. As can be seen our analyses have confirmed that the D1 and D2 proteins present in the isolated PSII reaction center of pea are NH₂-terminally processed. This modification involves the removal of the initiating formylmethionine residue followed by N-acetylation of the threonine residue (Thr²). Peptides corresponding to unmodified or partially modified products were not observed during these analyses, thus suggesting a consistency of this processing event in mature PSII complexes located in the granal regions. Previously, an analogous NH₂-terminal modification on the D1 and D2 proteins isolated from spinach chloroplasts was reported (2). The presence of the post-translational processing event in both of these higher plant species indicates that this is a common modification that most probably occurs on D1 and D2 proteins of all higher plants. No specific function has yet been attributed to this modification, although there have been suggestions that the amino terminus may be of significance in controlling stability and organization of membrane proteins (16, 17).

In addition to the processing event, the D1 protein was also found to have a further modification at its NH₂ terminus, the covalent attachment of a phosphoryl group. This modification appears to be present on only approximately one-third of the D1 protein population. It was anticipated that, as with its spinach counterpart, the modification would be located on Thr² D1 protein population. It was anticipated that, as with its covalent attachment of a phosphoryl group. This modification is likely to contribute significantly to the molecular heterogeneity that was observed during characterization of the intact D1 and D2 proteins in our PSI-MS studies (see Ref. 34).

Previously, an analogous NH₂-terminal modification on the D1 and D2 proteins isolated from spinach chloroplasts was reported (2). The presence of the post-translational processing event in both of these higher plant species indicates that this is a common modification that most probably occurs on D1 and D2 proteins of all higher plants. No specific function has yet been attributed to this modification, although there have been suggestions that the amino terminus may be of significance in controlling stability and organization of membrane proteins (16, 17).

In addition to the processing event, the D1 protein was also found to have a further modification at its NH₂ terminus, the covalent attachment of a phosphoryl group. This modification appears to be present on only approximately one-third of the D1 protein population. It was anticipated that, as with its spinach counterpart, the modification would be located on Thr² D1 protein population. It was anticipated that, as with its covalent attachment of a phosphoryl group. This modification is likely to contribute significantly to the molecular heterogeneity that was observed during characterization of the intact D1 and D2 proteins in our PSI-MS studies (see Ref. 34). Having mapped the majority of the D1 and D2 proteins, we have found that in both subunits the oxidized amino acid residues are located on peptides in similar luminal regions (one is at the COOH terminus of the protein, and the other is on a segment around the beginning of the IV transmembrane helix) but differing stromal positions (on D1 the region at the end of transmembrane helix II is oxidized, whereas on D2 the modification is on the segment located on the beginning of the V transmembrane helix). Interestingly, all of the regions that appear to be particularly vulnerable to oxidation are positioned in close proximity to the potential binding sites of the redox active cofactors involved in the photochemistry of PSII (Figs. 3 and 6).

The functional domains of the D1 and D2 proteins have been identified on the basis of their respective structural analogy to the L and M subunits of the crystallized bacterial reaction center (28) and several mutagenesis experiments (for review, see Ref. 29). Thus, it is anticipated that the region around the COOH terminus and the beginning of the IV transmembrane helix on both the D1 and D2 subunits would be located near the binding sites of the P680 chlorophylls and manganese cluster. The stromal region located at the beginning of the V transmembrane helix on the D2 protein would be close to the Qₐ binding niche, whereas on the D1 protein one of the residues (Glu¹⁵⁰) positioned at the end of the second transmembrane segment has been shown to be near the phoehytin molecule (30).

There is now a consensus that PSII is a primary target for photodamage. The vulnerability of this complex has been linked to its photochemical reactions, and so far two main mechanisms for photoactivation (one induced by the photochemical effects of the acceptor side and the other by the donor side) have been identified (7). Both mechanisms of photoinhibition are believed to involve oxidative damage and subsequently result in the degradation of the D1 and, to a somewhat lesser extent, the D2 proteins. It has been shown that impairment of secondary electron transport at the acceptor side of PSII leads to recombination between the radical pair state P680⁺'Phe°⁻ and formation of the triplet state of P680 (31). Under aerobic conditions this triplet signal is quenched by oxygen, giving rise to singlet oxygen (32, 33). Alternatively,
inactivation of the donor side of PSII can lead to accumulation of long lived, highly oxidizing radicals such as P680$^+$ and Tyr$^\bullet$; these species can in turn cause the oxidation of nearby pigments, redox components, and amino acids. It seems therefore that the oxidations we have identified in this paper (all of which appear to be located on regions of the D1 and D2 proteins which may be close to the redox active components of PSII) are likely to have been generated as a direct result of the photochemistry within the intact leaf before isolation of the reaction center. Moreover, we have observed a considerable enhancement of oxidations on peptides of the D1 and D2 subunits isolated from light-treated reaction center complexes.\(^2\) It is believed that further work in characterizing the exact positions of these modifications could reveal additional information on the molecular process that underlies the vulnerability of D1 and D2 proteins to photodamage and is currently under way. Furthermore, the higher level of oxidation we observed on the D1 protein compared with the D2 subunit may be an important characteristic that could explain the faster turnover rate of the former component during illumination.

REFERENCES
1. Morris, H. R., Panico, M., and Taylor, G. W. (1983) Biochem. Biophys. Res. Commun. 30, 299–305
2. Michel, H., Hunt, D. F., Shabanowitz, J., and Bennett, J. (1988) J. Biol. Chem. 263, 1123–1130
3. Takahashi, M., Shiraiishi, T., and Asada, K. (1988) FERS Lett. 420, 6–8
4. Mattoo, A. K., and Edelman, M. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 1497–1501
5. Kruse, O., Radunz, A., and Schmid, G. H. (1994) Z. Naturforsch. 49c, 9651–9654
6. Owen, G. C., and Ohad, I. (1982) J. Cell Biol. 93, 712–718
7. Barber, J., and Andersson, B. (1992) Trends Biochem. Sci. 17, 61–66
8. Aro, E.-M., Virgin, I., and Andersson, B. (1993) Biochim. Biophys. Acta 1143, 113–118
9. Shapiro, D. R., and Tewari, K. K. (1986) Plant Mol. Biol. 6, 1–12
10. Morris, H. R., Panico, M., Barber, M., Sedwick, R. D., and Tyler, A. N. (1981) Biochem. Biophys. Res. Commun. 101, 623–631
11. Keller, M., and Stutz E. (1986) FEBS Lett. 175, 173–177
12. Rasmussen, O. F., Broom, G. Stuttman, B. M., and Henningsen, K. W. (1984) Plant Mol. Biol. 3, 191–199
13. Haelb, S., Jensen, C., Andersen, S. D., and Roestorff, P. (1995) Protein Sci. 4, 394–404
14. Wilm, M., Shevchenko, A., Houthaave, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) Nature 379, 466–469
15. Strupat, K., Karas, M., Hillenkamp, F., Eckerskon, C., and Lottspeich, F. (1994) Anal. Chem. 66, 464–470
16. Wold, F. (1992) FASEB J. 6, 404
17. Blenis, J., and Rosh, M. D. (1993) Curr. Opin. Cell Biol. 5, 984–989
18. Delpeleaire, P. (1984) EMBO J. 3, 701–706
19. Millner, P. A., Marder, J. B., Gounarias, K., and Barber, J. (1986) Biochim. Biophys. Acta 852, 30–37
20. Telfer, A., Marder, J. B., and Barber, J. (1987) Biochim. Biophys. Acta 893, 557–560
21. Eich, T. D., Edelman, M., and Mattoo, A. K. (1992) J. Biol. Chem. 267, 3523–3529
22. Ebert, V., and Godde, D. (1994) Biochem. Biophys. Acta 1187, 335–346
23. Kruse, O., Zheleva, D., and Barber, J. (1997) FEBS Lett. 408, 276–280
24. Callahan, F. E., Ghirardi, M. L., Sopory, S. K., Mehta, A. M., Edelman, M., and Mattoo, A. K. (1980) J. Biol. Chem. 265, 15357–15365
25. Kettunen, R., Tytjtarvi, E., and Aro, E. M. (1991) FERS Lett. 290, 153–156
26. Diner, B. A., Ries, D. F., Cohen, B. N., and Metz, J. G. (1988) J. Biol. Chem. 263, 8972–8980
27. Nixon, P. J., and Diner, B. A. (1992) in Plant Protein Engineering (Shewry, P. R., and Gutteridge, S., eds) Vol. 2, pp. 93–141, Cambridge University Press, Cambridge, U. K.
28. Michel, H., and Deisenhofer, J. (1988) Biochemistry 27, 1–7
29. Nixon, P. J., and Diner, B. A. (1994) Biochem. Soc. Trans. 22, 338–343
30. Giorgi, L. B., Nixon, P. J., Merry, S. A. P., Joseph, D. M., Ebbert, V., and Godde, D. (1992) Biochemistry 31, 468–474
31. Heide, E., and Andersen, B. (1994) Photosynth. Res. 39, 191–199
32. Telfer, A., Bishop, S. M., Phillips, D., and Barber, J. (1994) J. Biol. Chem. 269, 13244–13253
33. Sharma, J., Panico, M., and Morris, H. R. (1997) J. Biol. Chem. 272, 33153–33157

\(^2\) J. Sharma, M. Panico, H. R. Morris, and J. Barber, unpublished work.