J|A|C|S
ARTICLES
Published on Web 06/19/2004

Evidence for a Post-Translational Modification, Aspartyl Aldehyde, in a Photosynthetic Membrane Protein

Lorraine B. Anderson,[†,#] Anthony J. A. Ouellette,[†,||] Julian Eaton-Rye,[‡] Melissa Maderia,[†,⊥] Michael J. MacCoss,[§] John R. Yates, III,[§] and Bridgette A. Barry*,[†,#]

Contribution from the Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, St. Paul, Minnesota 55108, Department of Biochemistry, University of Otago, Dunedin, New Zealand, and Department of Cell Biology, The Scripps Research Institute, La Jolla, California 92037

Received April 13, 2004; E-mail: bridgette.barry@chemistry.gatech.edu

Abstract: In oxygenic photosynthesis, photosystem II (PSII) carries out the oxidation of water and reduction of plastoquinone. Three PSII subunits contain reactive groups that covalently bind amines and phenylhydrazine. It has been proposed that these reactive groups are carbonyl-containing, co- or post-translationally modified amino acids (Ouellette et al. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2204 and Anderson et al. *J. Biol. Chem.* **2000**, *275*, 4920). To identify modified amino acid residues in one of the PSII subunits (CP47), tandem mass spectrometry was performed. Modified residues were affinity-tagged with either biotin-LC-hydrazide or biocytin hydrazide, which are known to label carbonyl groups. The affinity-tagged subunit was isolated by denaturing gel electrophoresis, and tryptic peptides were then subjected to affinity purification and tandem mass spectrometry. This procedure identified a hydrazide-labeled peptide, which has the sequence XKEGR. This result is supported by quantitative results acquired from peptide mapping and methylamine labeling. The gene sequence and these tandem data predict that the first amino acid, X, which is labeled with the hydrazide reagent, is a modified form of aspartic acid. On the basis of these data, we propose that D348 of the CP47 subunit is post- or co-translationally modified to give a novel amino acid side chain, aspartyl aldehyde.

In plants, algae, and cyanobacteria, oxygenic photosynthesis converts solar energy into chemical energy. Photosystem II (PSII) is one of the two photosynthetic reaction centers that carry out this process. PSII catalyzes the light-driven oxidation of water and reduction of plastoquinone. Reduced plastoquinone dissociates from PSII to act as a proton and electron carrier in the thylakoid membrane, and released oxygen maintains the aerobic component of the atmosphere. The oxidation of water to oxygen occurs within the oxygen-evolving complex of PSII. Manganese, calcium, and chloride are essential cofactors for oxygen evolution.[1]

PSII contains multiple intrinsic subunits, which cross the thylakoid membrane. D1 and D2 (both intrinsic subunits that bind most of the prosthetic groups in photosystem II) form the heterodimeric core of PSII and bind the redox-active cofactors involved in electron transfer.[2] CP47 (chlorophyll binding protein of apparent molecular mass 47 000) and CP43 (chlorophyll binding protein of apparent molecular mass 43 000) bind chlorophyll *a* and serve as light-harvesting proteins. In addition, both subunits have been shown to interact with the oxygen-evolving site (reviewed in ref 3). Extrinsic polypeptides also associate with the oxygen-evolving complex and are important for active-site stabilization and inorganic cofactor retention.[4]

Biochemical evidence suggests that there are three carbonyl-containing, modified amino acids in the oxygen-evolving complex.[5,6] These studies were performed using reagents, such as phenylhydrazine, methylamine, or benzylamine, which react with activated carbonyl groups and form either hydrazones or Schiff base derivatives. These experiments revealed that the CP47, D2, and D1 subunits contain reactive groups, which covalently bind amines and phenylhydrazine. The addition of chloride inhibited binding; the functional dependence suggested that the binding reactions occurred at or near the water splitting

† University of Minnesota, St. Paul.
‡ University of Otago.
§ The Scripps Research Institute.
|| Present address: Department of Biology, SUNY-Oswego, Oswego, NY 13126.
⊥ Present address: Laboratory of Medicinal Chemistry, National Cancer Institute, NIH, NCI-FCRDC, Frederick, MD 21702.
New address: School of Chemistry and Biochemistry and Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA 30332.

(1) Yocum, C. F. In *Manganese Redox Enzymes*; Pecoraro, V. L., Ed.; VCH Publishers: New York, 1992; pp 71−83.
(2) Nanba, O.; Satoh, K. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 109−112.
(3) Bricker, T. M.; Frankel, L. K. *Photosynth. Res.* **2002**, *72*, 131−146.
(4) Bricker, T. M.; Ghanotakis, D. F. In *Oxygenic Photosynthesis: The Light Reactions*; Ort, D. R., Yocum, C. F., Eds.; Kluwer Academic Publishers: Boston, 1996; Vol. 4, pp 113−136.
(5) Ouellette, A. J. A.; Anderson, L. B.; Barry, B. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2204−2209.
(6) Anderson, L. B.; Ouellette, A. J. A.; Barry, B. A. *J. Biol. Chem.* **2000**, *275*, 4920−4927.

site. Both plant and cyanobacterial PSII exhibited this reactivity.[5,6] At least one of these groups is redox-active and oxidatively deaminates primary amines.[5]

Modified amino acids, which contain carbonyl groups, have been identified in other proteins. Examples of carbonyl-containing amino acids include the redox-active quinocofactors, trihydroxyphenylalaninequinone,[7] lysine tyrosylquinone,[8] tryptophan tryptophylquinone,[9] and cysteine tryptophylquinone.[10] The arylsulfatases contain formylglycine[11] at the active site, which is an aldehyde generated from a cysteine[12] or serine[13] side chain. The barley lipid transfer protein has a lipid molecule attached through an ester linkage to a modified aspartate.[14]

Tandem mass spectrometry (MS^2) can be used to sequence both unmodified and modified peptides.[15−18] In this study, tandem mass spectrometry was utilized to identify a novel amino acid residue in the CP47 subunit of PSII.

Materials and Methods

Spinach photosystem II was prepared according to ref 19 with some modifications.[6] The rate of oxygen evolution in PSII samples was >700 μmol O_2/mg chl·hr, as measured with a Clark electrode.[20] Spinach PSII-3 (tris-washed PSII membranes) samples,[5,6] which lack extrinsic polypeptides and the manganese cluster,[21] were used for hydrazide binding. Cultures of the cyanobacterium, *Synechocystis sp*. PCC 6803, were grown under conditions previously described.[20] A site-directed mutation in the CP47 subunit (D348A) was generated by standard procedures.[22] PSII was purified from wild-type and mutant cyanobacterial cultures through the use of ion exchange chromatography.[23] The resulting PSII oxygen evolution rates were similar for wild-type and mutant cultures and were ∼1400 μmol O_2/mg chl·hr.

Hydrazide binding to PSII was performed at room temperature and under room light under the following conditions: 0.6 mg chl/mL PSII-3, 2.3 mM biotin-linked hydrazide (Pierce Chemical, Rockford, IL), 2.7 M urea, 2.1% sodium dodecyl sulfate (SDS), 52 mM Na_2CO_3.[6] After a 1 h incubation, the samples were subjected to electrophoresis using the modified SDS-polyacrylamide gel electrophoresis (PAGE) Neville method.[24] Gels were fixed, stained, and destained in an aqueous solution containing 45% methanol, 10% acetic acid, and ∼0.1% Coomassie Brilliant Blue R (Aldrich, Milwaukee, WI) for a total time less than 1.5 h. The gels were then washed in several changes of water over the course of 2−3.5 h.

In some experiments, PSII was incubated either with [^{14}C]methylamine (specific activity, 56mCi/mmol, Amersham Pharmacia Biotech, Piscataway, NJ), [^{14}C]phenylhydrazine (specific activity 11mCi/mmol, Morovek Biochemicals, Inc., Brea, CA), *t*-[^{14}C]- butylamine (specific activity, 55 mCi/mmol, Morovek Biochemicals), or biotin-linked hydrazides. This material was subjected to SDS-PAGE before transfer of the gel contents to Immobilon P^{SQ} membranes (Millipore Corporation, Bradford, MA) using a semi-dry transfer cell (Bio-Rad Laboratories, Inc., Hercules CA). Phosphorimages of ^{14}C binding were obtained as described previously.[6] Alkaline phosphatase-conjugated streptavidin (Jackson Immunoresearch Laboratories, Inc, West Grove, PA) and 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium were used to visualize hydrazide binding (Sigma-Aldrich, St. Louis, MO), as described.[25] Using these approaches, we have detected some degree of nonspecific binding with these commercially available hydrazides, which is due to an aldehyde contaminant (data not shown). Tandem mass spectrometry revealed that this aldehyde contaminant can bind to lysine residues in PSII subunits. However, the mass of this aldehyde contaminant is 30 atomic mass units smaller than the mass of the hydrazide compound, and binding of the contaminant is easily distinguished with mass spectrometry.

To quantitate the amount of methylamine binding to CP47, spinach PSII samples were treated with ^{14}C methylamine, SDS-PAGE was performed, and the gel was stained.[5,6] The band corresponding to CP47 was excised,[5] and the gel slice was treated with a solution containing cyanogen bromide (Fluka, Buchs, Switzerland), according to a procedure previously described.[26] After 4 h, the derived peptide mixture was subjected to another round of gel electrophoresis,[27] and the contents of the gel were transferred to Immobilon P^{SQ} membranes, as described above. Autoradiography and protein staining of the membranes were performed by published methods.[5,6] ^{14}C quantitation of the gel was performed by dissolving the gel and then scintillation counting.

For mass spectrometry, proteolytic treatment of the hydrazide-tagged, protein band corresponding to the CP47 subunit was performed as previously described.[28,29] Hydrazide-linked peptides were purified using an avidin column (Pierce). The peptide mixture (0.5 to 2.0 mL) was incubated overnight with 1 to 2 mL of monomeric avidin beads, which were then loaded into a disposable chromatography column. The column was washed with up to 14 mL of buffer containing 150 mM sodium chloride and 100 mM sodium phosphate, pH 7.2, to remove unbound peptides. The material bound to the avidin beads was then eluted with no more than 9 mL of 100 mM glycine-HCl, pH 2.8. The eluant was collected in ∼500 μL fractions.

These affinity purified peptides were analyzed by collision-induced dissociation (CID), generating a tandem mass spectrum (MS^2).[28,29] High-performance liquid chromatography (HPLC) solvents contained 0.2% acetic acid, and the peptides were eluted with an acetonitrile gradient directly into the LCQ ion trap mass spectrometer (Thermo-Finnigan, San Jose, CA). The flow rate was 300−700 nL/min using a MAGIC 2002 capillary HPLC and variable splitter (Michrom BioResources, Auburn, CA) or a Hewlett-Packard 1100 series HPLC and a homemade splitter. A PicoView (New Objective, Cambridge, MA) or a homemade source was utilized.

(7) Janes, S. M.; Mu, D.; Wemmer, D.; Smith, A. J.; Kaur, S.; Maltby, D.; Burlingame, A. L.; Klinman, J. P. *Science* **1990**, *248*, 981−987.
(8) Wang, S. X.; Mure, M.; Medzihradszky, K. F.; Burlingame, A. L.; Brown, D. E.; Dooley, D. M.; Smith, A. J.; Kagan, H. M.; Klinman, J. P. *Science* **1996**, *273*, 1078−1084.
(9) McIntire, W. S.; Wemmer, D. E.; Chistoserdov, A.; Lidstrom, M. E. *Science* **1991**, *252*, 817−824.
(10) Datta, S.; Mori, Y.; Takagi, K.; Kawaguchi, K.; Chen, Z. W.; Okajima, T.; Kuroda, S.; Ikeda, T.; Kano, K.; Tanizawa, K.; Mathews, F. S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14268−14273.
(11) von Figura, K.; Schmidt, B.; Selmer, T.; Dierks, T. *BioEssays* **1998**, *20*, 505−510.
(12) Schmidt, A.; Selmer, T.; Ingendoh, A.; von Figura, K. *Cell* **1995**, *82*, 271−278.
(13) Miech, C.; Dierks, T.; Selmer, T.; von Figura, K.; Schmidt, B. *J. Biol. Chem.* **1998**, *273*, 4835−4837.
(14) Lindorff-Larsen, K.; Lerche, M. H.; Poulsen, M. F.; Roepstorff, P.; Winther, J. R. *J. Biol. Chem.* **2001**, *276*, 33547−33553.
(15) Eng, J. K.; McCormack, A. L.; Yates, J. R., III. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976−989.
(16) Dongre, A. R.; Eng, J. K.; Yates, J. R., III. *Trends Biotechnol.* **1997**, *15*, 418−425.
(17) Aebersold, R.; Goodlett, D. R. *Chem. Rev.* **2001**, *101*, 269−295.
(18) Kinter, M.; Sherman, N. E. In *Protein Sequencing and Identification Using Tandem Mass Spectrometry*; Desiderios, D. M., Nibbering, N. M., Eds.; John Wiley: New York, 2000; pp 64−116.
(19) Berthold, D. A.; Babcock, G. T.; Yocum, C. F. *FEBS Lett.* **1981**, *134*, 231−234.
(20) Barry, B. A. *Methods Enzymol.* **1995**, *258*, 303−319.
(21) Yamomoto, Y.; Doi, M.; Tamura, N.; Nishimura, N. *FEBS Lett.* **1981**, *133*, 265−268.
(22) Eaton-Rye, J. J.; Vermaas, W. F. J. *Plant Mol. Biol.* **1991**, *17*, 1165−1177.
(23) Noren, G. H.; Boerner, R. J.; Barry, B. A. *Biochemistry* **1991**, *30*, 3943−3950.
(24) Piccioni, R.; Bellemare, G.; Chua, N.-H. In *Methods in Chloroplast Molecular Biology*; Edelman, H., Hallick, R. B., Chua, N.-H., Eds.; Elsevier Biomedical Press: Amsterdam, 1982; pp 985−1014.
(25) Harlow, E.; Lane, D. *Antibodies: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Plainview, NY, 1988; p 598.
(26) Nikodem, V.; Fresco, J. *Anal. Biochem.* **1979**, *97*, 382−386.
(27) Schagger, H.; van Jagow, G. *Anal. Biochem.* **1987**, *166*, 368−379.
(28) Anderson, L. B.; Maderia, M.; Ouellette, A. J. A.; Putnam-Evans, C.; Higgins, L.; Krick, T.; MacCoss, M. J.; Lim, H.; Yates, J. R., III; Barry, B. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14676−14681.
(29) Ouellette, A. J. A.; Barry, B. A. *Photosynth. Res.* **2002**, *72*, 159−173.

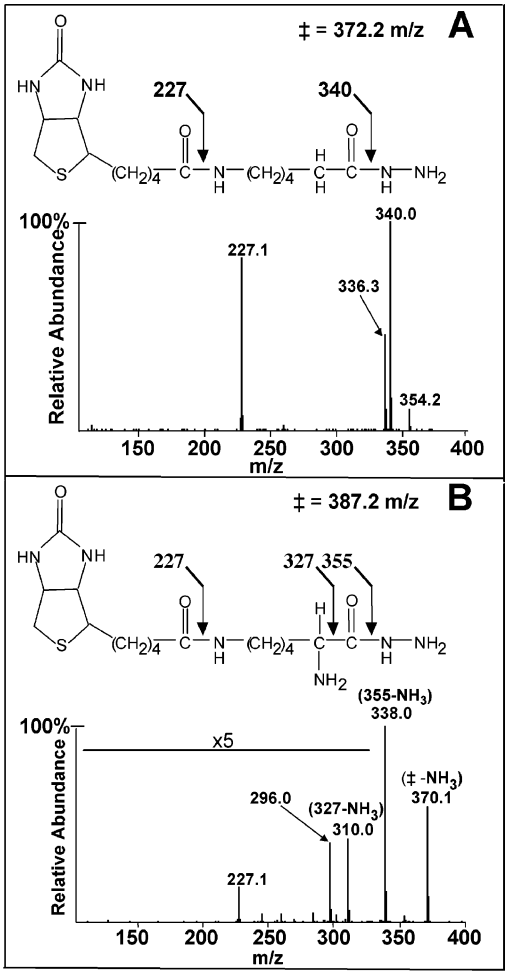

Figure 1. MS^2 of biotin-LC-hydrazide (A) and biocytin hydrazide (B). Arrows indicate some of the bonds that are broken to generate fragment ions. The fragment ions include portions of the molecules to the left of the arrows. ‡ = parent ion. Note that the *y*-axis has been expanded in one region of the spectrum in B.

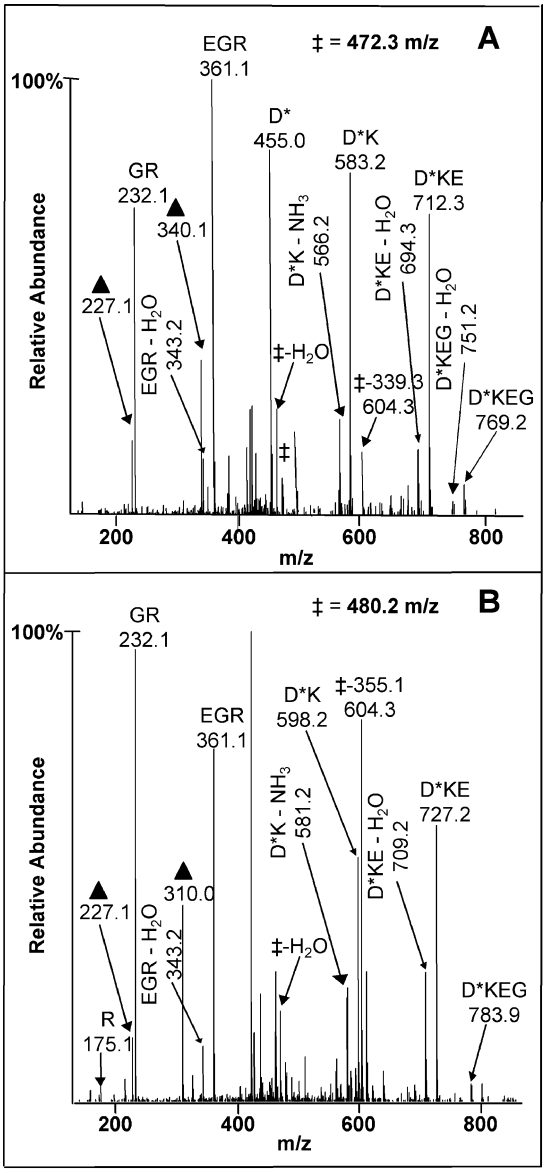

Figure 2. MS^2 interpreted as the D*KEGR peptide, labeled with either biotin-LC-hydrazide (A) or with biocytin hydrazide (B). ‡ = parent ion. The triangles denote ions that can be assigned to biotin-linked hydrazide fragments (biotin-LC-hydrazide in A; biocytin hydrazide in B). The ion at 604 *m/z* is consistent with loss of the biotin tag, but consistent with retention of the $-NHNH_2$ group on aspartyl aldehyde. In these ion trap experiments, low abundant fragment ions may have *m/z* values that differ by 1 *m/z* unit, when compared to the theoretical prediction (discussed in ref 6). For peaks with good ion statistics, the mass accuracy in these LCQ experiments is approximately ±0.3 *m/z*.

Results

The PSII subunits—CP47, D2, and D1—covalently bind ^{14}C-labeled methylamine, benzylamine, and phenylhydrazine, presumably through carbonyl-containing modifications of amino acid side chains.[5,6] Hydrazides also derivatize carbonyl groups and have been used to derivatize the carbonyl-containing trihydroxyphenylalanine quinone cofactor in copper amine oxidase.[30] In the present study, biotin-LC-hydrazide and biocytin hydrazide were used to label the carbonyl-containing amino acid in the CP47 subunit. This procedure then allowed the use of the biotin moiety for affinity purification. In addition, biotin-LC-hydrazide (Figure 1A) and biocytin hydrazide (Figure 1B) provide characteristic CID fragment ions (see below), because CID fragmentation occurs in a facile fashion at amide linkages.[16−18,31] This is shown in tandem mass spectra recorded from the reagents alone (Figure 1). Tandem mass spectra of the singly charged biotin-LC-hydrazide ion exhibit two major ions (227 and 340 *m/z*) caused by fragmentation at the amide bonds (Figure 1A). Tandem mass spectra of the singly charged biocytin hydrazide ion exhibit ions at 227 and 310 *m/z*, which result from amide bond fragmentation and loss of ammonia (Figure 1B). These fragment ions, produced from the biotin-linked hydrazides, were used as diagnostic ions in the PSII peptide mapping studies.

Affinity-purified, biotin-LC-hydrazide-labeled, CP47 peptides were separated by reversed-phase capillary HPLC, and peptides were eluted from the column directly into the mass spectrometer. The resulting ion chromatogram was searched for biotin-LC-hydrazide fragment ions, expected at 227 and 340 *m/z*. Representative data, derived from a doubly charged 472 *m/z* parent ion, are shown in Figure 2A. The spectrum can be assigned to the sequence, XKEGR, using both b-ions and y-ions

(30) Medda, R.; Padiglia, A.; Pedersen, J. Z.; Rotilio, G.; Agro, A. F.; Floris, G. *Biochemistry* **1995**, *34*, 16375−16381.
(31) Hunt, D. F.; Yates, J. R., III; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6233−6237.

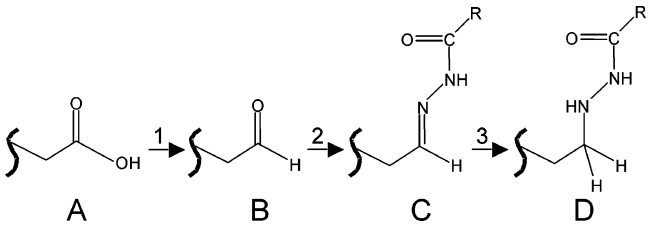

Figure 3. Structures (B−D) and reaction scheme (1−3) for aldehyde generation and subsequent modification by biotin-linked hydrazides. In (1), CP47 aspartic acid 348 (A) is reduced to aspartyl aldehyde (B). Incubation with hydrazides (2) and subsequent reduction (3) yields structures that are consistent with the mass spectral data.

(Table 1; Figure 2A). The sequence, DKEGR, is present in lumenal loop E of the CP47 subunit and corresponds to residues 348−352.[32] Subtraction of the calculated KEGR $(M + H)^+$ mass (489.6 Da) from the mass of the parent ion predicts that X will have a mass of 454.0 Da. An $(M + H)^+$ ion at 455 *m*/*z* is observed in the spectrum (Figure 2A). Therefore, the 455 *m*/*z* ion is tentatively assigned to the b_1-ion (X).

To identify X, the primary sequence of CP47, PIFRDKEGR, was considered. No satisfactory assignment for the 455 *m*/*z* ion could be made, considering only an unmodified peptide. For example, calculated *m*/*z* values for the singly charged ions, corresponding to DKEGR (604), RDKEGR (761), FRDKEGR (908), or IFRDKEGR (1021), were either too small or too large to account for the data. Note that hydrazides do not react with carboxylic acids[33] and that the acyl form of the hydrazide derivative (retaining an oxygen) is not consistent with the observed *m*/*z* value. Also, reaction of the (−30 amu) aldehyde contaminant with the lysine residue in DKEGR does not give a self-consistent interpretation of the data, and ions consistent with loss of ammonia are observed in the MS^2 data (Figure 2). These observations are inconsistent with an assignment of the lysine as the modified amino acid. Therefore, to explain our results, we must propose that aspartate 348 is post-translationally modified. An interpretation (Figure 2A), which can explain the data, is that position 348 contains a novel amino acid that is modified by the hydrazide (D*). A structure that is consistent with measured *m*/*z* value and the chemistry of hydrazides is shown in Figure 3D.

To test this interpretation, PSII was labeled with biocytin hydrazide (Figure 1B), and experiments were carried out as described for the biotin-LC-hydrazide samples. If the D*KEGR assignment is correct, then the *m*/*z* value for singly charged X and X-containing ions will increase by 15. Furthermore, fragment ions from the biocytin hydrazide (Figure 1B) should be present in the tandem mass spectrum. Representative data, derived from a doubly charged 480 *m*/*z* parent ion, are shown in Figure 2B. The ions at 227 *m*/*z* and 310 *m*/*z* (Figure 2B) are assigned to biocytin hydrazide fragments (see Figure 1B). Also, b-ions, which were predicted to contain D*, shift by 15 *m*/*z* (compare Figure 2B to 2A; see Table 1), whereas y-ions, which were not predicted to contain D*, do not shift in *m*/*z* value. Therefore, this experiment supports the D*KEGR assignment and supports the interpretation that hydrazide binding occurs at residue 348 of CP47.

To obtain further information concerning the identities of fragment ions, tandem mass spectrometry of an MS^2 fragment ion (MS^3) was conducted both on biotin-LC-hydrazide-labeled material and on biocytin-labeled material. In this technique, a fragment ion from the MS^2 experiment is isolated in the ion trap, and another round of CID is performed. MS^3 was conducted on the 583.2 *m*/*z* ion, assigned as D*K, from a biotin-LC-hydrazide labeling experiment (Figure 4A). In the MS^3 data, biotin-LC-hydrazide fragment ions are again observed at 227 and 340 *m*/*z*, supporting the assignment of the 583 *m*/*z* ion to D*K.

This approach was also used with biocytin-labeled material. MS^3 was performed on the 598 *m*/*z* ion, assigned as D*K (Figure 4B). A 310 *m*/*z* ion is observed, consistent with the presence of biocytin hydrazide in the original fragment ion. Figure 4C shows an MS^3 spectrum derived from the 727.2 *m*/*z* ion, assigned as D*KE. The 310 *m*/*z* ion, observed in the MS^3 data, is consistent with the release of a biocytin hydrazide fragment. Also, ions consistent with production of the dipeptides, KE and D*K, are observed in this spectrum.

These mass spectrometry experiments do not provide quantitative information concerning labeling at D348. To rule out the possibility that affinity purification is selecting for a rare subset of labeled molecules, radioactive methylamine was employed; a radioactive form of the biotin hydrazides is not commercially available. The results of these peptide mapping experiments are shown in Figure 5. CP47 was digested with cyanogen bromide, and the peptide mixture was subjected to SDS-PAGE on 2-amino-2-hydroxymethyl-1,3 propanediol (Tris)-tricine gels. The results of peptide staining with Coomassie are shown in Figure 5, lane 1. Eleven peptides are predicted to result from complete cyanogen bromide cleavage of CP47; the predicted masses are 0.7, 0.8, 1.2, 2.0, 2.5, 2.8, 3.3, 3.8, 8.5, 14.1, and 16.7 kDa. The stained lane (Figure 5, lane 1) shows many peptides with apparent masses larger than expected from complete cleavage. The autoradiogram (Figure 5, lane 2) reveals one major radioactive band with an apparent molecular mass

Table 1. Calculated and Measured Ions for the D*KEGR Peptide Labeled with Biotin LC Hydrazide or Biocytin Hydrazide

	b-ions					y-ions			
	biotin-LC-hydrazide (D +339 Da)		biocytin hydrazide (D +354 Da)			biotin-LC-hydrazide (D +339 Da)		biocytin hydrazide (D +354 Da)	
residue	calculated	measured	calculated	measured	residue	calculated	measured	calculated	measured
D*[a]	455.1	455.0	470.1	—[b]	D*KEGR[a]	NA[c]	NA[c]	NA[c]	NA[c]
D*K	583.3	583.2	598.3	598.2	KEGR	489.6	—[b]	489.6	—[b]
D*KE	712.4	712.3	727.4	727.2	EGR	361.4	361.1	361.4	361.1
D*KEG	769.4	769.2	784.4	783.9	GR	232.3	232.1	232.3	232.1
D*KEGR	NA[c]	NA[c]	NA[c]	NA[c]	R	175.2	—[b]	175.2	175.1

[a] Biotin-labeled aspartyl aldehyde, which is +339 (biotin-LC-hydrazide labeled) or +354 Da (biocytin hydrazide labeled), as compared to unmodified aspartic acid. Calculated ions are for singly charged ions using the average mass. Measured ions come from the spectra in Figure 2. [b] Ions that are not above background. [c] Not applicable.

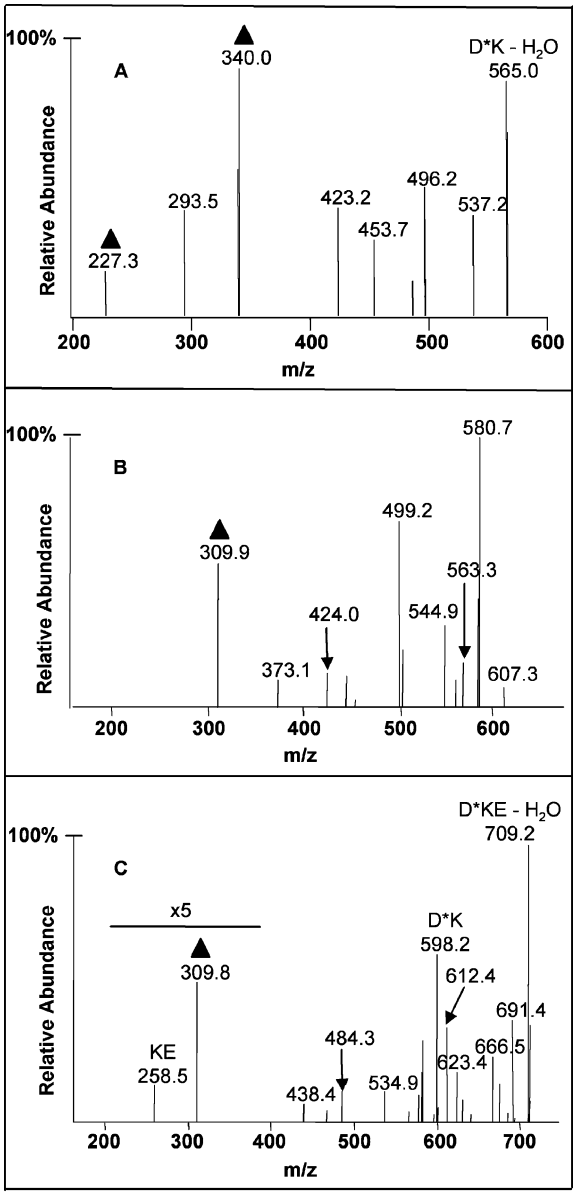

Figure 4. MS^3 spectra from fragment ions in the MS^2 experiments shown in Figure 2. (A) MS^3 spectrum of 583.2 *m/z* ion interpreted as D*K labeled with biotin-LC-hydrazide. (B) MS^3 spectrum of 598.2 *m/z* ion interpreted as D*K labeled with biocytin hydrazide. (C) MS^3 spectrum of 727.2 *m/z* ion interpreted as D*KE labeled with biocytin hydrazide. The triangles denote ions that can be assigned to biotin-linked hydrazide fragments (biotin-LC-hydrazide in (A) and to biocytin hydrazide in (B) and (C)). Note that the *y*-axis has been expanded in one region of the spectrum in (C).

at approximately 17 kDa. The amount of labeling retained on this 17 kDa band was 12% of the amount of radioactivity retained on CP47 immediately after cyanogen bromide treatment. Another 19% of the label was retained at the top of the Tris-tricine gel; this material most likely corresponds to undigested protein. Retention of a total of only ~30% of the starting radioactivity on the Tris-tricine gel is most likely due to the oxidative loss of the reductively trapped amine adduct.

Many peptides derived from membrane proteins are blocked for N-terminal sequencing, and membrane proteins can run anomalously on SDS-PAGE. Therefore, mass spectrometry

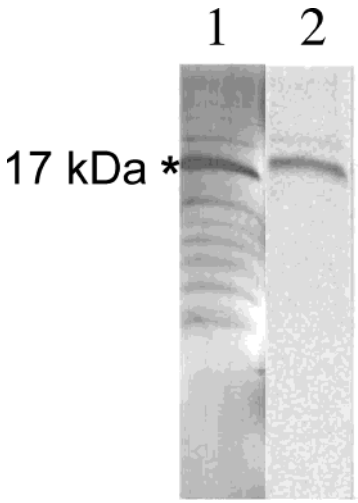

Figure 5. Cyanogen bromide cleavage of ^{14}C-methylamine-labeled, plant CP47. CP47 was gel-purified, cleaved with cyanogen bromide, subjected to SDS-PAGE, and transferred to a membrane. In lane 1, the blot was stained for protein; in lane 2, an autoradiogram was obtained.

experiments were performed after tryptic digest of the apparent 17 kDa, ^{14}C-labeled band. These experiments detected peptides before and after D^{348}KEGR^{352} in the primary sequence of the 17 Da band (data not shown). The methylamine-labeled D*KEGR tryptic peptide was not detected, perhaps because tryptic digestion generated an undetectable dipeptide, D*K (see Discussion section). A DKEGR-containing peptide of 17.3 kDa could result from incomplete cyanogen bromide digestion and concomitant acid-induced cleavage at one of the Asp-Pro linkages near the CP47 carboxyl terminus.[34] Incomplete cleavage at the Asp-Pro amide bond could account for the faint, higher molecular weight band in the autoradiogram (Figure 5, lane 2). It should be noted that the 17.3 kDa band does overlap other peptides, such as the 16.7 kDa peptide predicted from complete cyanogen bromide cleavage.[35,36] However, we found no evidence for label incorporation into the 16.7 kDa peptide. Therefore, this peptide-mapping experiment is consistent with the D^{348}KEGR^{352} CP47 sequence retaining a substantial fraction of the methylamine-derived ^{14}C label after SDS-PAGE.

To ascertain whether D348 is conserved, the spinach sequence W257 → W450, containing loop E,[32] was subjected to a BLASTP search against the nonredundant database.[37] CP47 sequences for 72 different organisms were returned. Five entries contain a residue other than aspartate in the homologous position. *Arabidopsis* and *Thermosynechococcus* contain asparagine; *Chlorella* and *Heterocapsa* contain a glutamate. *Amphidinium* has an isoleucine, but has a glutamine at the adjacent position.

In *Synechocystis sp*. PCC 6803, D348 is conserved. A cyanobacterial mutant was generated, in which D348 was changed to alanine. This D348A mutant had oxygen evolution activity similar to wild type. Also, we found that substitution of an alanine for the encoded aspartic acid did not disrupt

(32) Bricker, T. M. *Photosynth. Res.* **1990**, *24*, 1−13.
(33) McMurry, J. *Organic Chemistry*, 3rd ed.; Brooks/Cole Publishing Company: Pacific Grove, CA, 1992.
(34) Allen, G. In *Laboratory Techniques in Biochemistry and Molecular Biology*; Burdon, R. H., van Knippenberg, P. H., Eds.; Elsevier: Amsterdam, 1989; Vol. 9, p 81.
(35) Bricker, T. M.; Odom, W. R.; Queirolo, C. B. *FEBS Lett.* **1988**, *231*, 111−117.
(36) Bricker, T. M.; Frankel, L. K. *Arch. Biochem. Biophys.* **1987**, *256*, 295−301.
(37) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res.* **1997**, *25*, 3389−3402.

methylamine, phenylhydrazine, or *tert*-butylamine binding to the *Synechocystis* CP47 subunit (data not shown). In mass spectrometry experiments on cyanobacterial PSII, we also found no evidence for a hydrazide-tagged, DKEGR sequence. This result will be discussed in more detail below.

Discussion

Our previous work employed reductive trapping of methylamine and benzylamine and binding of *tert*-butylamine and phenylhydrazine. These experiments provided support for the presence of an activated carbonyl group in the CP47 subunit.[5,6] To determine the identity of the reactive site in the plant CP47 subunit, PSII was labeled with biotin-linked hydrazides, CP47 tryptic peptides were generated, and the hydrazide-tagged peptide was purified using affinity purification. LC−MS/MS experiments indicate that the CP47 tryptic peptide, XKEGR, is modified with each of two commercially available hydrazide reagents. These MS/MS experiments also suggest that residue 348 is the covalently modified site. This position is encoded as an aspartate in the CP47. Therefore, we propose that the side chain of D348 is post-translationally modified.

Considering the masses of aspartate and the biotin-linked hydrazides, a reduced hydrazone structure is consistent with the mass of the hydrazide-labeled D*. A plausible structure for the nonhydrazide-tagged side chain is an aldehyde (Figure 3B). Reaction of a biotin-linked hydrazide with this aldehyde (Figure 3, step 2), and subsequent reduction (step 3), would give the structure in Figure 3D and account for the observed, measured masses. Endogenous reducing equivalents are present under the conditions used for labeling.[6] While an ester group would also be chemically facile for hydrazide/phenylhydrazine/amine labeling, the product from the reaction with hydrazides would be inconsistent with the measured masses.

Despite much effort, aspartyl aldehyde, without the hydrazide tag, has not been detected. We attribute this to the expected chemical reactivity of the aldehyde moiety, which can react with lysine side chains. Another possible issue is the presence of an internal lysine in the DKEGR peptide, which generates a potential trypsin cleavage site if the peptide is not protected with a bulky labeling reagent, such as hydrazide. Trypsin cleavage after the internal lysine would generate an undetectable dipeptide.

Post-translational modifications of carboxylic acid side chains have been described in other proteins, and examples of modified aspartates are known. For example, *Escherichia coli* ribosomal proteins contain β-carboxyaspartic acid[38] and ribosomal protein S12 contains β-methylthio-aspartic acid.[39] Vitamin K-dependent proteins contain β-hydroxy aspartic acid.[40−44] Aspartates and glutamates can be modified to give a covalent ester linkage to a heme[45,46] or to a lipid.[14] Blood coagulation depends on the carboxylation of glutamate side chains to form γ-carboxyglutamate residues.[47] Histone proteins are ADP-ribosylated via ester linkages on glutamates.[48,49] Reversible carboxymethylation of glutamic acids in transmembrane sensory receptors are important in chemotaxis.[50−52]

Aspartyl aldehyde is a novel aspartate modification. However, the aryl sulfatases contain a reactive aldehyde[11−13,53] modified from a cysteine[12] or serine,[13] and therefore, the generation of an aldehyde functional group from an amino acid side chain has precedents. The mechanism of aldehyde generation from carboxylate in PSII is not clear at this time; however, biological systems are known to reduce carboxylic acids to aldehydes.[54−56] In the biosynthesis of threonine, methionine, and lysine, aspartic acid is converted to aspartate semialdehyde, and the synthesis of proline from glutamate involves a glutamate semialdehyde intermediate.[56] The formation of these semialdehydes requires an aspartyl- or glutamyl-phosphate intermediate. The direct reduction of carboxylic acids to aldehydes, without the requirement of an activated phosphoryl intermediate, is catalyzed by carboxylic acid reductase in *Clostridium thermoaceticum*.[54] Aromatic carboxylic acids can also be reduced to aldehydes (for examples, see ref 55 and references within). On the basis of these precedents, we propose that aspartyl aldehyde is generated in CP47 post-translationally by an enzymatic mechanism, with a source of electrons that remains to be identified. The modification could be generated in the chloroplast stromal compartment during assembly of PSII.

While an aldehyde is generated in plant CP47 at position D348, we conclude that the aldehyde modification is not generated at position 348 in cyanobacterial CP47. Because cyanobacterial CP47 shows reaction with amine, hydrazine, and hydrazide reagents, an activated carbonyl binding site for amines is present in cyanobacterial CP47, but at a different location in the primary sequence. There are other differences in structure and function when cyanobacteria are compared to plants. These differences in structure−function include a difference in subunit composition, the degree of accessibility to the water splitting apparatus, and, possibly, the functional role of calcium.[57−59] For example, the complement of lumenal extrinsic proteins is known to be different in cyanobacteria and plants.[57] We therefore speculate and propose a function for aspartyl aldehyde in a plant-specific function, for example, as part of a binding site for a plant-specific intrinsic subunit or as a ligand to calcium. Alternatively, plant aspartyl aldehyde may be involved in a

(38) Christy, M. R.; Barkley, R. M.; Koch, T. H.; Van Buskirk, J. J.; Kirsch, W. M. *J. Am. Chem. Soc.* **1981**, *103*, 3935−3937.
(39) Kowalak, J. A.; Walsh, K. A. *Protein Sci.* **1996**, *5*, 1625−1632.
(40) Drakenberg, T.; Fernlund, P.; Roepstorff, V. P.; Stenflo, J. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 1802−1806.
(41) Fernlund, P.; Stenflo, J. *J. Biol. Chem.* **1983**, *258*, 12509−12512.
(42) McMullen, B. A.; Fujikawa, K.; Kisiel, W.; Sasagawa, T.; Howald, W. N.; Kwa, E. Y.; Weinstein, B. *Biochemistry* **1983**, *22*, 2875−2884.
(43) Sugo, T.; Fernlund, P.; Stenflo, J. *FEBS Lett.* **1984**, *165*, 102−106.
(44) Krishna, R. G.; Wold, F. In *Proteins: Analysis and Design*; Angeletti, R. H., Ed.; Academic Press: San Diego, CA, 1998; pp 121−207.
(45) Taylor, K. L.; Strobel, F.; Yue, K. T.; Ram, P.; Pohl, J.; Woods, A. S.; Kinkade, J. M. *Arch. Biochem. Biophys.* **1995**, *316*, 635−642.
(46) Fenna, R.; Zeng, J.; Davey, C. *Arch. Biochem. Biophys.* **1995**, *316*, 653−656.
(47) Nelsestuen, G. L.; Shah, A. M.; Harvey, S. B. In *Vitamins and Hormones: Advances in Research and Applications*; Litwack, G., Ed.; Academic Press: New York, 2000; Vol. 58, pp 355−389.
(48) Ogata, N.; Ueda, K.; Hayaishi, O. *J. Biol. Chem.* **1980**, *255*, 7610−7615.
(49) Ogata, N.; Ueda, K.; Kagamiyama, H.; Hayaishi, O. *J. Biol. Chem.* **1980**, *255*, 7616−7620.
(50) Kehry, M. R.; Doak, T. G.; Dahlquist, F. W. *J. Biol. Chem.* **1984**, *259*, 11828−11835.
(51) Dunten, P.; Koshland, D. E., Jr. *J. Biol. Chem.* **1991**, *266*, 1491−1496.
(52) Borkovich, K. A.; Alex, L. A.; Simon, M. I. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6756−6760.
(53) Lukatela, G.; Krauss, N.; Theis, K.; Selmer, T.; Gieselmann, V.; von Figura, K.; Saenger, W. *Biochemistry* **1998**, *37*, 3654−3664.
(54) White, H.; Strobl, G.; Feicht, R.; Simon, H. *Eur. J. Biochem.* **1989**, *184*, 89−96.
(55) Li, T.; Rosazza, J. P. *J. Bacteriol.* **1997**, *179*, 3482−3487.
(56) Heldt, H.-W. *Plant Biochemistry and Molecular Biology*; Oxford University Press: Oxford, New York, 1997; pp 261−263.
(57) Barry, B. A.; Boerner, R. J.; de Paula, J. C. In *The Molecular Biology of the Cyanobacteria*; Bryant, D., Ed.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1994; Vol. 1, pp 215−257.
(58) Tang, X.; Diner, B. A. *Biochemistry* **1994**, *33*, 4594−4603.
(59) Burnap, R. L.; Qian, M.; Pierce, C. *Biochemistry* **1996**, *35*, 874−882.

regulatory function, which is not necessary in the prokaryotic cyanobacteria. While X-ray structures of cyano-bacterial PSII are available,[60−62] there is currently no published structure of a higher plant PSII reaction center.

To summarize, the experiments described here provide evidence for a novel protein modification, aspartyl aldehyde, in the CP47 subunit of plant PSII. The modified residue, D348, is contained in the CP47 lumenal loop E and provides a hydrazide binding site in spinach PSII.

Acknowledgment. We thank Dr. LeeAnn Higgins and Thomas Krick for helpful discussions concerning mass spectrometry and the University of Minnesota Mass Spectrometry Consortium for the Life Sciences, where some of these experiments were conducted. We are also grateful to Dr. Hanjo Lim and Dr. Hayes MacDonald for experimental assistance. Supported by NIH GM43273 (B.A.B.) and F32DK59731 (M.J.M.).

JA0478781

(60) Zouni, A.; Witt, H.-T.; Kern, J.; Fromme, P.; Krauss, N.; Saenger, W.; Orth, P. *Nature* **2001**, *409*, 739−743.
(61) Kamiya, N.; Shen, J.-R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 98−103.
(62) Ferreira, K. N.; Iverson, T. M.; Maghlaoui, K.; Barber, J.; Iwata, S. *Science* **2004**, *303*, 1831−1837.