Characterization of a variant of the spinach PSII type I light-harvesting protein using kinetically controlled digestion and RP-HPLC-ESI-MS

ARTICLE *in* ANALYTICAL CHEMISTRY · JANUARY 2004

Impact Factor: 5.64 · DOI: 10.1021/ac034866+ · Source: PubMed

4 AUTHORS, INCLUDING:

Anna Maria Timperio
Tuscia University
85 PUBLICATIONS **1,459** CITATIONS
SEE PROFILE

Christian G Huber
University of Salzburg
189 PUBLICATIONS **5,012** CITATIONS
SEE PROFILE

Anal. Chem. **2003,** *75,* 6775−6780

Characterization of a Variant of the Spinach PSII Type I Light-Harvesting Protein Using Kinetically Controlled Digestion and RP-HPLC-ESI-MS

Wolfgang Walcher,[†] Anna-Maria Timperio,[‡] Lello Zolla,*[,‡] and Christian G. Huber*[,§]

Institute of Analytical Chemistry and Radiochemistry, Leopold-Franzens-University, 6020 Innsbruck, Austria, Department of Environmental Sciences, University of Tuscia, 01100 Viterbo, Italy, and Instrumental Analysis and Bioanalysis, Saarland University, 66123 Saarbrücken, Germany

A previously unknown isoform of the type I major antenna protein of photosystem II of spinach was identified, and its amino-terminal sequence was characterized by a novel kinetic digestion approach, in which sequential tryptic digestion was followed by analysis of both released peptides and truncated proteins by reversed-phase high-performance liquid chromatography−electrospray ionization mass spectrometry. Using nonpolar, monolithic, 200-*μ*m-i.d. separation columns based on poly(styrene/divinylbenzene) copolymer and applying gradients of acetonitrile in 0.05% aqueous trifluoroacetic acid, released peptides and truncated proteins could be separated and mass analyzed in a single chromatographic run. This enabled a straightforward identification of the fragments removed from the amino-terminal ends of the protein, which was essential for the characterization of the antenna isomers showing the most significant sequence variation in the amino-terminal region. The sequences of the amino termini were derived from the differences in molecular mass between intact and truncated proteins and were corroborated by sequencing using tandem mass spectrometry and database searching. The sequence of the 23 amino-terminal residues of the previously unknown isoform differed from that of the other two known isoforms only in one and three amino acids, respectively. Such subtle changes in amino acid sequence are supposed to play an important role in the supramolecular organization of photosynthetic antenna proteins.

The presence of multigene families is common in the vegetable kingdom. The antenna proteins of photosystems I and II (PSI and PSII) are encoded by several genes that show a high homology and consequently encode proteins with similar molecular masses and molecular properties.[1] All the major antenna proteins of PSII (Lhcb 1−3) fold into three membrane-spanning helices, with an additional small amphipatic helix near the C-terminal end (Figure 1a). Two of the three α-helices are held together within the thylakoid membrane by ion-pair formation, and only the amino-terminal end of the proteins is exposed on the stromal surface of the membrane.[2,3] The sequences of the transmembrane domains are usually highly conserved among different isoforms and even different types of antenna proteins. The most significant differences in sequence are located in the amino terminus (Figure 1a),[4] which is supposed to play a substantial role in the subunit−subunit interactions responsible for the supramolecular organization and regulation of photosynthetic protein complexes.[5]

The different Lhcb isoforms have been revealed in a number of plant species by intact molecular mass determinations using reversed-phase high-performance liquid chromatography-electrospray ionization mass spectrometry (RP-HPLC-ESI-MS), which represents a useful identification tool in cases where the exact protein or DNA sequence and eventual posttranslational modifications are known.[5−7] Nevertheless, it is not possible to assign amino acid sequences for previously unknown isoforms. In such cases, amino acid sequencing is mandatory in order characterize the variations in protein structure, but conventional Edman sequencing is often hampered by amino-terminal acetylation of antenna proteins.[8] Amino acid sequencing by tandem mass spectrometry (MS/MS) on the other hand, in which sequence information is deduced from fragment ions obtained upon collision-induced dissociation of tryptic peptides,[9] represents a very powerful alternative sequencing method that is not impaired by amino-terminal blockage. After extensive proteolytic digestion of the antenna proteins, this approach has been successfully applied to the characterization of the amino-terminal sequences of two isoforms of the Lhcb 1 antenna protein of spinach.[10,11]

* Corresponding authors. E-mail: christian.huber@mx.uni-saarland.de. Tel.: +49 681 302 2433. E-mail: zolla@unitus.it. Fax: +49 681 302 2963. Tel.: +39 0761 357 100. Fax: +39 0761 357 630.

[†] Leopold-Franzens-University.

[‡] University of Tuscia.

[§] Saarland University.

(1) Buetow, D. E.; Chen, H.; Erdös, G.; Yi, L. S. H. *Photosynth. Res.* **1988**, *18*, 61−97.

(2) Bürgi, R.; Suter, F.; Zuber, H. *Biochim. Biophys. Acta* **1987**, *890*, 346−51.

(3) Kühlbrandt, W.; Wang, D. N.; Fujiyoshi, Y. *Nature* **1994**, *367*, 614−21.

(4) Jansson, S. *Biochim. Biophys. Acta* **1994**, *1184*, 1−19.

(5) Huber, C. G.; Timperio, A.-M.; Zolla, L. *J. Biol. Chem.* **2001**, *276*, 45755−61.

(6) Gomez, S. M.; Nishio, J. N.; Faull, K. F.; Whitelegge, J. P. *Mol. Cell. Proteomics* **2002**, *1*, 46−59.

(7) Zolla, L.; Timperio, A.-M.; Walcher, W.; Huber, C. G. *Plant Physiol.* **2003**, *131*, 198−214.

(8) Zolla, L.; Timperio, A. M.; Testi, M. G.; Bianchetti, M.; Bassi, R.; Manera, F.; Corradini, D. *Photosynth. Res* **1999**, *61*, 281−90.

(9) Biemann, K. *Anal. Chem.* **1986**, *58*, 1288A−300A.

(10) Michel, H.; Griffin, P. R.; Shabanowitz, J.; Hunt, D. F.; Bennett, J. *J. Biol. Chem.* **1991**, *266*, 17584−91.

10.1021/ac034866+ CCC: $25.00
Published on Web 11/06/2003

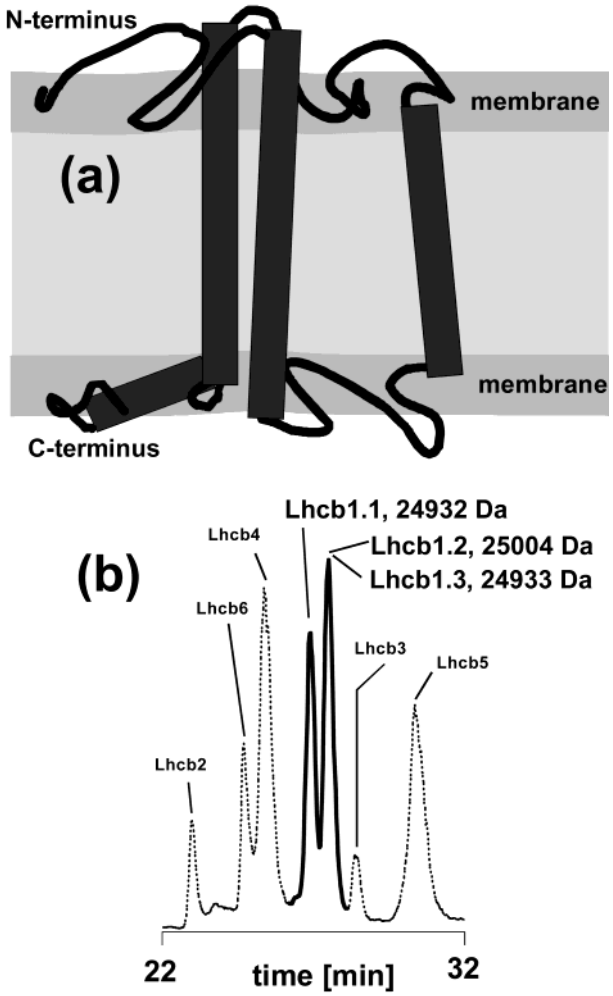

Figure 1. (a) Molecular structure of the pea Lhcb1 antenna protein showing the highly conserved transmembrane helices and the more variable amino terminus.[3] (b) Detection of Lhcb 1 isoforms of spinach (represented in the chromatogram by solid lines; dashed lines represent the other antenna proteins) by RP-HPLC-ESI-MS. Column, Vydac Protein C-4 (250 × 4.6 mm i.d.); mobile phase, (A) 0.10% trifluoroacetic acid in water, (B) 0.10% trifluoroacetic acid in acetonitrile; linear gradient 27.5−63.5% B in 45 min; flow rate 1.0 mL/min; detection, ES-MS, scan, 500−2000 amu.

The efficacy of proteases in cleaving proteins is dependent upon the accessibility to open stretches of primary amino acid sequence. In the case of membrane proteins, the native three-dimensional structure of the substrate protein may block access to many sites, resulting in the release of only a few peptides.[12] To derive the highly variable amino-terminal sequences, we evaluate here a new approach of kinetic digestion of photosynthetic membrane proteins, in which only the most accessible amino-terminal peptides are cleaved during the first few minutes of tryptic digestion of photosynthetic membrane proteins. Both truncated proteins and cleaved peptides are subsequently analyzed by RP-HPLC-ESI-MS(MS) using monolithic separation columns, which facilitates an assignment of the formed peptides to the corresponding proteins based on differences in molecular mass between intact and truncated proteins. Moreover, sequences of the released peptides can be determined by MSMS. This approach is utilized to determine the amino-terminal sequence of a previously unknown isoform of the Lhcb 1 antenna protein.

EXPERIMENTAL SECTION

Chemicals and Samples. Acetonitrile (HPLC gradient grade) was obtained from Merck (Darmstadt, Germany). Water was purified by a NANOpure-Infinity water deionization system from Barnstead (Dubuque, IA). 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) buffer (pH 7.3), urea (analytical reagent grade), bradykinin, and carbonic anhydrase were obtained from Sigma (St. Louis, MO). Trifluoroacetic acid (TFA, for protein sequence analysis) was purchased from Fluka (Buchs, Switzerland), and trypsin (sequencing grade modified) from Promega (Madison, WI).

Chloroplast thylakoid membranes (PSII membranes) were isolated from spinach leaves (*Spinacia olearacea*) according to the method described in ref 13 with the modification reported elsewhere.[8] Leaves were collected at night in the dark and at different periods of the year in order to minimize any seasonal and light effects. The light-harvesting complex was isolated from the PS II membranes as previously described[14] with the following modifications: PSII membranes were pelleted by centrifugation at 10000*g* for 5.0 min at 4 °C, suspended in 50 mM MES buffer, pH 6.3, containing 15 mM sodium chloride, 5 mM magnesium chloride, and 1.0 mg/mg chlorophyll, and then solubilized by adding 1% (w/v) *n*-dodecyl *β*-D-maltoside. Unsolubilized material was removed by centrifugation at 10000*g* for 10 min. The supernatant was rapidly loaded onto a 0.1−1.0 M sucrose gradient containing 50 mM MES buffer, pH 6.3, 15 mM sodium chloride, 5 mM magnesium chloride, and 5.0 mM *n*-dodecyl *β*-D-maltoside. The gradient was then spun on a Kontron model Centricon T-1080 ultracentrifuge equipped with a model TST 41.14 rotor at 39 000 rpm for 18 h at 4 °C. Green bands were harvested with a syringe. The SDS−PAGE analysis of these green bands revealed that band 2 contained a mixture of the protein components of the major and minor PSII antenna systems, whereas band 3 essentially contained the protein components of the major PSII antenna system, as previously reported.[14]

High-Performance Liquid Chromatography and Electrospray Ionization Mass Spectrometry. Monolithic capillary columns (60 × 0.20 mm i.d. and 60 × 0.10 mm i.d.) were prepared according to the published protocol.[15] The 50 × 0.20 mm i.d. monolithic capillary columns are commercially available from LC-Packings-A Dionex Co. (Monoliths, Amsterdam, The Netherlands). The Ultimate capillary HPLC system (LC-Packings) was used for HPLC-ESI-MS(MS) experiments. ESI-MS was performed on a quadrupole ion trap mass spectrometer (LCQ, Thermo Finnigan, San Jose, CA) equipped with a triaxial electrospray ion source. Fine-tuning for ESI-MS of peptides and proteins using bradykinin and carbonic anhydrase as tuning compounds was performed as described previously.[16]

Tryptic Digestion and Micropreparative Fractionation of Lhcb Proteins from Spinach. Sample manipulations were

(11) Ouellette, A. J. A.; Barry, B. A. *Photosynth. Res.* **2002**, *72*, 159−73.
(12) Aebersold, R.; Goodlett, D. R. *Chem. Rev.* **2001**, *101*, 269−95.
(13) Berthold, D. A.; Babcock, G. T.; Yocum, C. F *FEBS Lett* **1981**, *134*, 231−34.
(14) Bassi, R.; Dainese, P. *Eur. J. Biochem.* **1992**, *204*, 317−26.
(15) Premstaller, A.; Oberacher, H.; Huber, C. G. *Anal. Chem.* **2000**, *72*, 4386−93.
(16) Oberacher, H.; Walcher, W.; Huber, C. G. *J. Mass Spectrom.* **2003**, *38*, 108−16.

preformed in the dark in order to avoid photodegradation of the proteins. Trypsin (20 μg) was dissolved in 20 μL of acetate buffer (Promega, Madison, WI) and heated for 30 min at 37 °C to activate trypsin. After activation, the trypsin solution and 30 μL of HEPES buffer were added to 50 μL of band 3 obtained from sucrose gradient ultracentrifugation of PSII of spinach. Digestion was stopped by the addition of trifluoroacetic acid to a total concentration of 0.05%.

Micropreparative fractionation of major antenna proteins using a Vydac C4 column (250 × 4.6 mm i.d. plus 10 × 4.6 mm precolumn, The Separation Group, Hesperia, CA) was accomplished with a gradient pump (model Rheos 2000, Flux Instruments, Basel, Switzerland), a vacuum degasser (ERC 3215, Ercatech, Bern, Switzerland), an injector (Rheodyne, model 7125, Cotati, CA) equipped with a 100-μL sample loop, and a UV detector (UVD320, Gynkotek, Germering, Germany) set to 215 nm. PSII major antenna proteins from band 3 of sucrose gradient ultracentrifugation were eluted by a linear gradient of 42−64% acetonitrile in 0.1% aqueous TFA at ambient temperature. Collected fractions were lyophilized, resuspended in 50 μL of HEPES buffer (pH 7.3), and digested as described above.

RESULTS AND DISCUSSION

Identification of Protein Isoforms by Intact Molecular Mass Determination. So far, only two isoforms of the Lhcb 1 antenna protein of spinach have been described in the literature.[10] In our RP-HPLC-ESI-MS analyses, however, we were able to detect a total of three different isoforms of Lhcb1 in spinach having intact molecular masses of 24 932, 24 933, and 25 004. Formerly, we interpreted the presence of a protein of M_r 24 933 in the second peak containing a protein of M_r 25 004 as an impurity carried over from the first peak.[7,17] However, detailed analysis of selected mass traces and mass spectra corresponding to the M_r 24 932 and 24 933 species in this study clearly revealed two separated peaks that represent two different proteins having more or less identical molecular masses. It is interesting to note that two of the isoforms of essentially identical molecular masses were almost separated to baseline, while two isoforms with a difference of 71 in mass coeluted into one peak (Figure 1b). Because none of the common posttranslational modifications matched with this mass difference, we assumed that the proteins represent sequence isoforms.

Kinetically Controlled Digestion of Antenna Proteins. In an attempt to characterize and assign the amino-terminal sequence of the two known and the previously unknown isoforms, a preparation of the major antenna proteins (band 3 of sucrose gradient ultracentrifugation) was treated with trypsin before analysis by RP-HPLC-ESI-MS. Digestion experiments were conducted under nondenaturing conditions (HEPES buffer, pH 7.2) and revealed that upon short incubation times only the amino-terminal ends of the antenna proteins were digested. After removal of a few peptides, digestion stopped at truncated proteins having molecular masses around 22 400, most probably due to restricted access of the enzyme to the stable three-dimensional structure of the transmembrane part of the proteins persisting even after removal of the thylakoid membranes (Figure 1a).

The chromatograms shown in Figure 2 illustrate the analysis of the initial and final stages of such a digestion experiment. This RP-HPLC-ESI-MS analysis is remarkable because both peptides and membrane proteins could be separated, detected, and mass

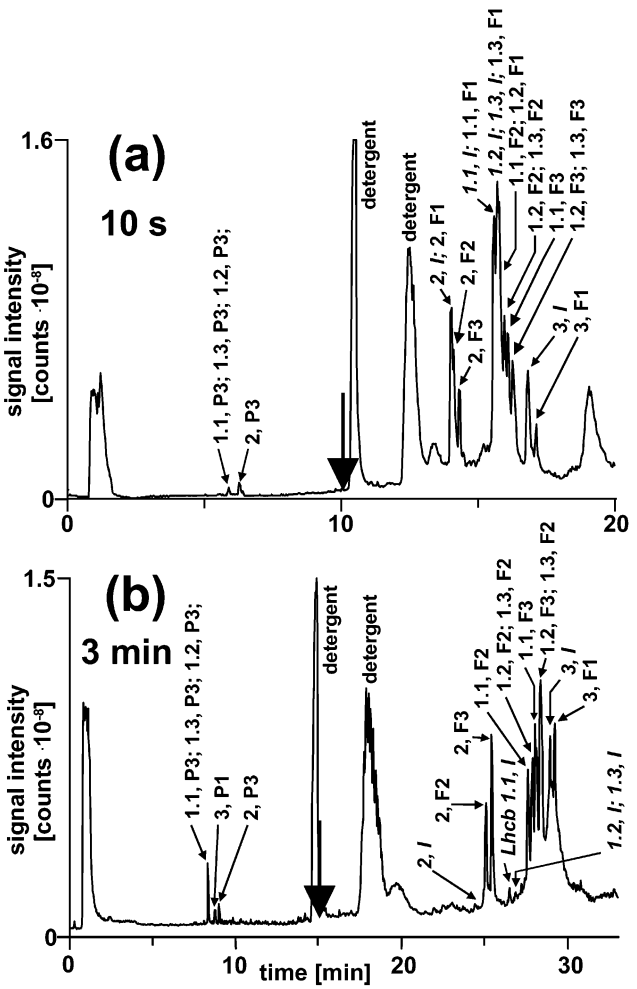

Figure 2. Simultaneous analysis of intact proteins (identified by index I), truncated proteins (identified by indexes F1−F3), and amino-terminal peptides (identified by indexes P1−P3) after trypsin treatment of the PSII major antenna system of spinach. Column, PS−DVB monolith, 60 × 0.20 mm i.d.; mobile phase, 0.050% TFA in water, (B) 0.050% TFA in acetonitrile; linear gradient, (a) 0−40% B in 10 min, 40−52% B in 20 min, (b) 0−36% B in 15 min, 36−48% B in 20 min; flow rate, 2.0 μL/min; temperature, 60 °C; detection, ESI-MS, scan, 500−2000 amu; electrospray voltage, 2.0 kV; sheath gas, nitrogen, 100 arbitrary units; sample, (a) 10- and (b) 180-s tryptic digest of band 3 of spinach PSII. Peak identification in Table 1; the arrow indicates the change from bradikinin to carbonic anhydrase tuning parameters.

analyzed in a single run utilizing monolithic, micropellicular PS−DVB columns. Such analyses cannot be obtained with comparable performances in a single column packed with conventional, porous, alkylated, silica-based separation columns, because the hydrophobic membrane proteins are difficult to elute from the octadecyl-silica stationary phases that are usually employed for peptide separations. Tryptic peptides, on the other hand, are poorly retained on butyl-silica columns, which are commonly used for HPLC of membrane proteins. Moreover, the optimal pore diameters of conventional particle-based stationary phases are different for peptides and proteins, namely, 10−15 nm for the former and of 20−100 nm for the latter, which prevents the use of the same type of stationary phase for both types of analytes. The micropellicular configuration of the monolithic PS−DVB stationary phase used in this study, however, comprises only macropores

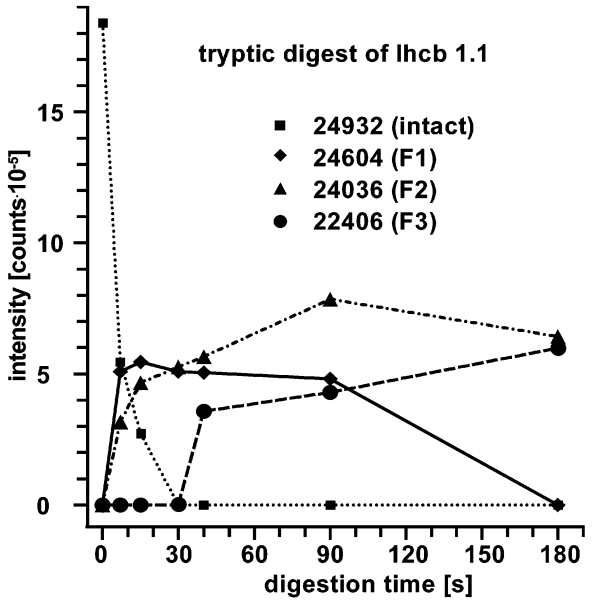

Figure 3. Kinetics of tryptic digestion of the major antenna proteins of spinach. Samples were taken after 7, 15, 30, 40, 90, and 180 s of digestion and analyzed under the conditions described in Figure 2.

for convective mass transport and lacks any micropores accessible to the analytes,[18] resulting in high column performance for both peptides and proteins despite their considerably different molecular size and diffusion coefficients.[19] Figure 2 clearly demonstrates that peptides and proteins can be eluted from the PS−DVB monoliths with high separation efficiency as evidenced by peak widths at half-height of less than 8 s.

The peptides eluted in Figure 2b during the first 15-min segment of the gradient at acetonitrile concentrations up to 36%, whereas membrane proteins eluted at acetonitrile concentrations higher than 40% in a second gradient segment. Detergents, necessary to solubilize the membrane proteins, eluted between the peptides and proteins and did not interfere with their analysis. To achieve sensitive detection for both peptides and proteins, tuning parameters were optimized independently by direct infusion using bradykinin and carbonic anhydrase as tuning compounds[16] and the parameters were changed automatically in the course of the chromatographic analysis, as indicated by an arrow in the chromatograms of Figure 2. Using this optimized analytical system, a total of 18 different intact proteins (identified in Figure 2 by italic letters and the index "I") and protein fragments (identified in Figure 2 by the indexes "F1−F3") as well as 4 different tryptic peptides (identified in Figure 2 by the indexes "P1 and P3") were detected in the chromatograms.

Figure 3 illustrates in more detail the progress of the digestion utilizing Lhcb 1.1 as an example. To follow the different stages of digestion, the enzyme was inactivated after time periods of 7, 15, 30, 40, 90, and 180 s by the addition of strong acid, and the products were analyzed by RP-HPLC-ESI-MS as shown in Figure 2. The relative concentrations of the different species were deduced from the peak heights of the protein signals in the deconvoluted mass spectra that were extracted from the corresponding peaks in the total ion current chromatograms. It can be seen that the intact molecule disappeared entirely after 30 s of digestion. As the intact molecule was hydrolyzed, fragment F1 was formed and reached its maximum concentration already after 10−20 s, after which its concentration continuously declined to background level at 180 s. The second fragment F2 was formed comparatively slowly with a maximum concentration found at ∼90 s. The final fragment F3 was first detected after 30 s of digestion, when its concentration increased relatively fast, whereas after 40 s the increase in concentration was only moderate. From the data, we conclude that all peptides were clipped off the proteins in a strictly sequential manner. The first two, very small peptides were removed very rapidly from the proteins while the third, comparatively large peptide was hydrolyzed at a slower rate. Under the chosen, nondenaturing conditions, digestion stopped at the truncated protein having a molecular mass of 22 406.

Interpretation of the complex elution patterns of peptides, intact proteins, and truncated proteins was assisted by analysis of the products after different stages of digestion and the fact that removal of the hydrophilic, amino-terminal peptides generally resulted in a slight increase in retention time. In the initial stages of digestion, intact proteins as well as fragments lacking the first and sometimes the second amino-terminal tryptic peptides were detectable (Figure 2a). After 3 min of digestion, intact Lhcb 1.1/1.2/1.3 and Lhcb 2 as well as their fragments lacking the first amino-terminal peptide had almost disappeared. Lhcb 3 lost only one amino-terminal peptide, and digestion ceased at fragment F1. Table 1 collects the molecular masses of the observed proteins, protein fragments, and peptides together with the mass differences between corresponding species.

Identification of the Amino-Terminal Sequences by Mass Spectrometry. Three different protein fragments were observed for Lhcb 1.1/1.2/1.3 and Lhcb 2. The first two peptides hydrolyzed from these proteins could not be directly detected in the chromatograms, not even using a scan range starting from m/z 100, because they eluted in the void volume due to their small size and high polarity, in which ion suppression by eluting salts and other low molecular mass compounds inhibited the detection of the small peptides. Nevertheless, their identity was readily inferable from the mass differences ΔM between the intact and truncated proteins and comparison with published amino-terminal sequences of the Lhcb 1 proteins[10,11] (Table 1). The mass differences for the two peptides of the third Lhcb 1 isoform matched very closely the molecular masses of acetyl-RK and TAGKPK, leading to the conclusion that Lhcb 1.2 and Lhcb 1.3 share common sequences in this part of their amino terminus (Table 2). Larger peptides could be identified either directly by intact molecular mass determination or indirectly by the mass differences between the different protein species. The peptide removed from Lhcb 3 had a molecular mass of 1419.0, which corresponded very well with the theoretical molecular mass of the published sequence (Table 1). The three other peptides having molecular masses of 1648.2, 1706.2, and 1601.1 were consistent with the known amino-terminal sequences of Lhcb 1.1, Lhcb 1.2, and Lhcb 2. The observation of only three peptides for four proteins suggested that the previously unknown Lhcb isoform shares a common partial sequence with one of the known

(17) Corradini, D.; Huber, C. G.; Timperio, A.-M.; Zolla, L. *J. Chromatogr., A* **2000**, *886*, 111−21.
(18) Oberacher, H.; Huber, C. G. *TrAC, Trends Anal. Chem.* **2002**, *21*, 166−73.
(19) Walcher, W.; Oberacher, H.; Troiani, S.; Hölzl, G.; Oefner, P.; Zolla, L.; Huber, C. G. *J. Chromatogr., B* **2002**, *782*, 111−25.

Table 1. Molecular Masses of Intact Proteins, Protein Fragments, and Peptides Obtained by Partial Tryptic Digestion of PSII Major Antenna Proteins from Spinach

	molecular mass or mass difference (Da)					
protein	intact[a]	protein fragments[a]		ΔM[b]	peptide[c]	sequence[d]
Lhcb 1.1	24 931.7	1	24 604.1	344.8	nd[e]	AcO−RK (343.2)
		2	24 036.9	584.2	nd	SAGKPK (586.4)
		3	22 406.0	1647.9	1648.2	NVSSGSPWYGPDRVK (1647.9)
Lhcb 1.2	25 003.5	1	24 676.8	343.8	nd	AcO−RK (343.2)
		2	24 093.0	600.7	nd	TAGKPK (600.4)
		3	22 401.0	1709.0	1706.2	TVQSSSPWYGPDRVK (1706.0)
Lhcb 1.3	24 933.3	1	24 605.2	345.1	nd	unknown
		2	24 025.8	596.4	nd	unknown
		3	22 395.2	1647.6	1648.2	unknown
Lhcb 2	24 757.5	1	24 402.0	372.5	nd	AcO−RR (371.2)
		2	24 073.7	345.4	nd	TVK (346.2)
		3	22 489.9	1600.8	1601.1	SAPQSIWYGPDRPK (1600.9)
Lhcb 3	24 319.4	1	22 918.0	1418.4	1419.0	GNDLWYGPDRVK (1419.6)

[a] Average of between 5 (Lhcb3, F3) and 20 (Lhcb2, F2) mass measurements by RP-HPLC-ESI-MS, Standard deviations ranged between 1.2 and 3.3 Da, average 1.9 Da. [b] Mass difference + 17.0 (for OH) between intact protein and truncated proteins or between truncated proteins. [c] Masses obtained by direct mass measurement in the RP-HPLC-ESI-MS analyses, average of five measurements. [d] Sequences from refs 10 and 11. Numbers in parentheses give the molecular masses calculated from the sequence. [e] nd, peptide was not detected because of elution in the void volume.

Table 2. Amino-Terminal Sequences of the Five Major Antenna Proteins from Spinach Determined by RP-HPLC-ESI-MS(MS)

	sequences		
protein	peptide 1	peptide 2	peptide 3
Lhcb 1.1	AcO−RK	**S**AGKPK	NVSSGSPWYGPDRVK
Lhcb 1.2	AcO−RK	TAGKPK	**TVQSS**SPWYGPDRVK
Lhcb 1.3	AcO−RK	**T**AGKPK	**N**V**SSG**SPWYGPDRVK
Lhcb 2	AcO−RR	TVK	SAPQSIWYGPDRPK
Lhcb 3	GNDLWYG-PDRVK		

isoforms. Moreover, a mass difference of 1647.6 Da makes the sequence of peptide P3 of Lhcb 1.1 a strong candidate for the sequence of peptide 3 of the new Lhcb 1 isoform.

Sequence Confirmation by Tandem Mass Spectrometry. De novo sequencing of the m/z 1648.2 precursor ion eluting at 8.2 min from the monolithic column by RP-HPLC-ESI-MSMS clearly corroborated the sequence assigned to this peptide (Figure 4). To prove that peptide P3 originated from both Lhcb 1.1 and Lhcb 1.3, the proteins were separated by micropreparative RP-HPLC in an analytical butyl-silica column (comparable to the chromatogram shown in Figure 1b). Two fractions, the first containing Lhcb 1.1 and the second containing Lhcb 1.2 and 1.3 were collected and lyophilized. Following reconstitution, the two fractions were treated with trypsin for 3 min under denaturing conditions (HEPES buffer, pH 7.2, 1 mol/L urea) and the products were analyzed by RP-HPLC-ESI-MSMS (see Figure 5). Analysis of the data using the Sequest algorithm for fully automated peptide identification revealed a peptide of sequence NVSSGSPWYGPDR (from AAA50310, *Prunus persica*) in both fractions, which is in full accordance with the sequence of peptide P3 lacking the two carboxy-terminal amino acids (VK), which were obviously cleaved under denaturing conditions but not under nondenaturing conditions.

Differences in Amino-Terminal Sequences of Lhcb 1.1, 1.2, and 1.3. The amino-terminal sequence of the newly identified protein isoform Lhcb 1.3 differs only in one amino acid

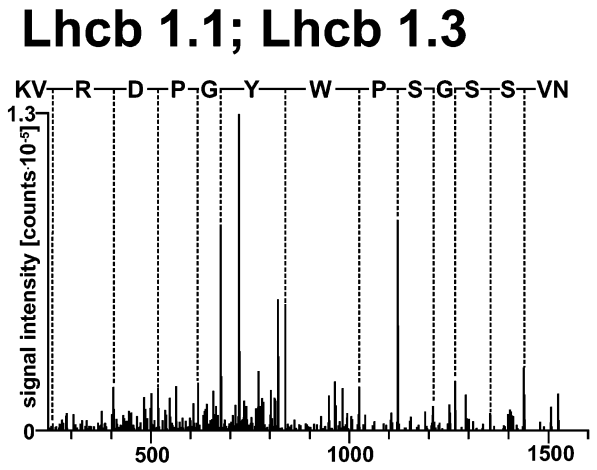

Figure 4. Tandem mass spectrum of the amino-terminal peptide P3 of Lhcb 1.1 and 1.3. Linear gradient, 40−52% B in 20 min; detection, MS/MS, scan, 500−2000 amu, relative collision energy, 36%; electrospray voltage, 1.6 kV; sample, tryptic digest of band 3 of spinach PSII. Other conditions as in Figure 2. Y-fragments are indicated by dotted lines.

from Lhcb 1.1 whereas three amino acids are different in Lhcb 1.2 (Table 2). Lhcb 1.1 and Lhcb 1.3 share almost identical amino-terminal sequences in which only one S is mutated to T at position 3 from the amino terminus. The mass shift expected for an S to T mutation is 14 mass units; however, the measured mass difference between Lhcb 1.1 and 1.3 was only 1.6 mass units, so additional sequence variation has to be present in the part of the protein that was not accessed by kinetic digestion. The small difference in amino-terminal sequence significantly alters the hydrophobicity of the proteins, which facilitates their separation by RP-HPLC, while the difference is not sufficient for discrimination by a specific antibody.[20]

Although the individual sequences of all amino-terminal peptides were known before, only the application of a kinetically controlled protein digestion method in combination with simultaneous analysis of the resulting peptides by RP-HPLC-ESI-MS-

(20) Sigrist, M.; Staehelin, L. A. *Biochim. Biophys. Acta* **1992**, *1098*, 191−200.

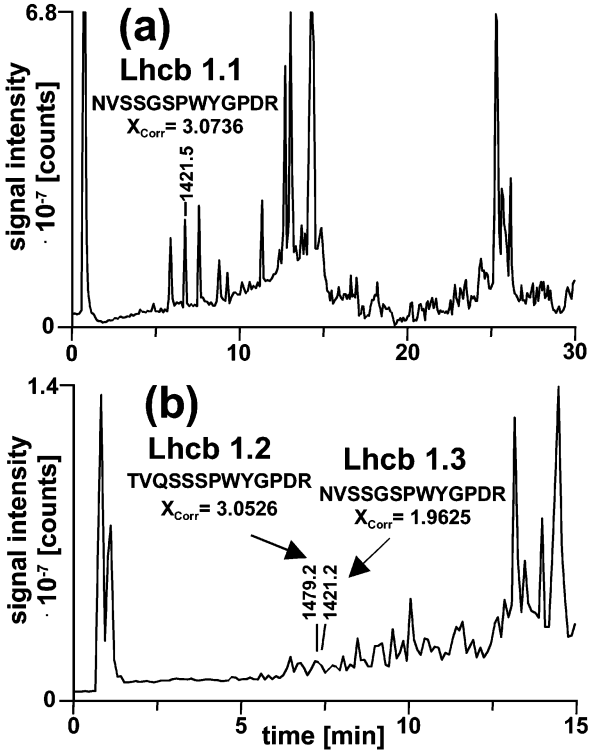

Figure 5. Identification of amino-terminal peptides in RP-HPLC fractions of Lhcb 1.1 and Lhcb 1.2/1.3. Linear gradient, (a) 0−40% B in 15 min, 40−52% B 20 min, (b) 0−36% B in 15 min; scan, (a) 0−15 min, data-dependent MS/MS, 15−30 min, full-scan MS, 500−2000 amu, (b) data-dependent MS/MS; sample, tryptic digest of the major antenna proteins of spinach. Other conditions as in Figure 2.

(MS) enabled the identification and characterization of the new isoform of Lhcb 1. This strategy can be applied to any mixture of antenna proteins. Hence, the method may be applied to other plant species with the final aim to verify whether the number of proteins expressed matches with the total number of cloned genes. Additionally, as in the case of spinach, this should result in attempts to isolate new polymorphic genes that are difficult to clone.

CONCLUSIONS

Because of their distinguished chemical and physical properties, membrane proteins remain difficult to analyze with conventional analytical methods. We demonstrated that kinetic digestion of membrane proteins followed by analysis of truncated proteins and released peptides by RP-HPLC-ESI-MS represents an effective method for the study of sequence variation in the extramembrane parts of the proteins, which are known to possess high functional relevance. Separation and ESI-MS detection of both peptides and hydrophobic membrane proteins by RP-HPLC-ESI-MS in a single analysis was essential and could be conveniently accomplished with monolithic separation columns based on PS−DVB copolymer. Although the applicability of the method was exemplified through the identification of isoforms of photosynthetic proteins, it is generally implementable to the study of sequence variation in the extramembrane domains of hydrophobic membrane proteins.

ACKNOWLEDGMENT

This work was supported by a grant from the Austrian Science Fund (P-13442-PHY) and Progetto Strategico DM 10/5/2000 (MUIR). We are grateful to Jean-Pierre Chervet of LC-Packings—A Dionex Co.—for generously providing a loan UltiMate capillary HPLC system. The 250 × 4.5 mm i.d. Vydac C4 column was generously provided by G. VanEssche (Grace Vydac).

Received for review July 28, 2003. Accepted September 30, 2003.

AC034866+