Proteomics of Light-Harvesting Proteins in Different Plant Species. Analysis and Comparison by Liquid Chromatography-Electrospray Ionization Mass Spectrometry. Photosystem II[1]

Lello Zolla*, Anna-Maria Timperio, Wolfgang Walcher, and Christian G. Huber

Department of Environmental Sciences, University of Tuscia, 01100 Viterbo, Italy (L.Z., A.-M.T.); Institute of Analytical Chemistry and Radiochemistry, Leopold-Franzens-University, 6020 Innsbruck, Austria (W.W.); and Instrumental Analysis and Bioanalysis, Saarland University, 66123 Saarbrücken, Germany (C.G.H.)

An overview of the intact molecular masses and the hydrophobic properties of the photosystem II (PSII) light-harvesting proteins in 14 different plant species is presented. The protein separation and identification was achieved by means of reversed-phase high-performance liquid chromatography-electrospray ionization-mass spectrometry. The good correspondence of the molecular masses measured by reversed-phase high-performance liquid chromatography-electrospray ionization-mass spectrometry with those deduced from the DNA sequence (0.008%–0.016% relative deviation in Arabidopsis) enabled the identification of the different protein types. Utilizing this correlation, it was possible in several cases to spot a gene product for the previously cloned genes. In PSII, all antenna proteins show hydrophobic properties considerably different within the same as well as among various species, in contrast to observations made previously with PSI. These differences might reflect a tuning of protein-protein interactions that play a role in inducing different supramolecular organizations of PSII: within the same species as a consequence of short-term adaptations, and among species for seasonal species adaptation. The relative antenna stoichiometry was readily established on the basis of relative peak areas of the separated proteins in the ultraviolet chromatograms. The correspondence found between the high copy number of genes with the gene products reveals that the genes are not silent in their protein expression. Moreover, the high copy number of gene products as well as protein heterogeneity observed in PSII suggest a possible plant strategy to realize the high degree of organization and interconnection of the light-harvesting systems under any environmental conditions.

The antenna proteins of PSII diverge in structure and composition within the same phyla, and even species (Jansson, 1994). The minor antenna proteins include CP29 (Lhcb4), CP26 (Lhcb5), and CP24 (Lhcb6), whereas Lhcb1, Lhcb2, and Lhcb3 comprise the group of major antenna proteins. For both antennas, several genes have been cloned and sequenced and numerous protein bands have been resolved by gel electrophoresis in different species (Bassi and Dainese, 1992; Jansson, 1994). Although a large number of genes for antenna proteins has been identified in several plants, it is not clear at present how the expression of the gene family members correlates to the gene products observed in vivo in any plant species (Morishige and Thornber, 1994). For instance, it is not known if these genes are expressed differentially during development of the plant, under changing environmental conditions, and/or in different cell types (Falbel and Staehelin, 1992).

In general, assignment of a protein to its corresponding gene may be achieved by (partial) protein sequencing, which is laborious, time consuming, expensive, and sometimes impossible because of the presence of amino-terminal blocking (Michel et al., 1991). Alternatively, assignments can be made by comparison of the molecular masses derived from the DNA sequence and the apparent molecular masses measured by SDS-PAGE, which today represents the most widely used method for the separation and isolation of small amounts of chlorophyll-binding proteins (Machold, 1991; Bassi and Dainese, 1992). However, it is well known that molecular masses estimated by SDS-PAGE may significantly diverge from the molecular masses calculated for the individual proteins on the basis of their nucleotide-derived amino acid sequences (Matsuoka et al., 1987; Schwartz and Pichersky, 1990).

During the past years, reversed-phase HPLC with on-line detection by electrospray ionization (ESI)-mass spectrometry (MS) has evolved into a highly powerful tool for accurate mass measurements of proteins (Covey et al., 1988; Chait and Kent, 1992; Premstaller et al., 2001). Applications of reversed-phase HPLC-ESI-MS to the analysis of plant proteins

[1] This work was supported by the Ministero dell'Università e della Ricerca Scientifica Co-Finanziamento 2001, by the Austrian Science Fund (grant no. P–13442–PHY), and by the CE INCO-COPERNICUS Project (grant no. IC15 CT98–0126).

* Corresponding author; e-mail zolla@unitus.it; fax 0039–0761–357–630.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.012823.

include the characterization of PSII reaction center subunits (Sharma et al., 1997a, 1997b; Whitelegge et al., 1998), the analysis of antenna proteins of PSI (Damm and Green, 1994; Zolla and Timperio, 2000; Zolla et al., 2002) and PSII (Corradini et al., 2000), the study of isoforms of antenna proteins (Huber et al., 2001), and the study of the appressed membrane subfractions of thylakoids (Gómez et al., 2002). The measured intact molecular masses of the proteins may sometimes be correlated with molecular masses calculated from genomic sequences to identify the corresponding genes (Gómez et al., 2002). The success of such correlations is, however, frequently complicated because of incomplete knowledge of post-translational modifications, unknown processing of precursors to mature proteins, isomeric forms of the proteins, and DNA sequencing errors, which all result in significant deviations of measured molecular masses with those deduced from the DNA sequences.

In due consequence, we explored in this investigation as to what extent reversed-phase HPLC-ESI-MS is applicable to the study and identification of the antenna proteins of PSII in 14 monocot and dicot species with the aim of establishing a system for rapid and correct protein identification based on intact molecular masses. The generated data set of different protein masses should enable an overview on antenna protein heterogeneity and its possible implications on the supramolecular organization of PSII. Because complete sequence information is not available for all of the investigated species, identification of the proteins needs to be based upon a comparison of the measured intact molecular masses with the mass range expected from the known genes of other plant species. To validate the approach of protein identification by intact molecular masses, the antenna proteins from Arabidopsis, from which a complete, high-quality genome sequence is available, are used as reference components.

RESULTS

Validation of Intact Molecular Mass Measurements as Analytical Method for the Identification of Antenna Proteins

Before application of reversed-phase HPLC-ESI-MS to the identification of antenna proteins in various plant species, we evaluated the accuracy of mass measurements using the antenna proteins from Arabidopsis as reference compounds, for which all antenna genes have been recently cloned and sequenced with high accuracy (Jansson, 1999). The components of bands 2 and 3 from Suc gradient ultracentrifugation were chromatographically separated in a 250- × 4.6-mm i.d. C4 column (The Separation Group, Hesperia, CA) and mass spectra were extracted and deconvoluted from the reconstructed ion chromatograms. Table I summarizes the molecular mass data of the major antenna proteins of Arabidopsis obtained from 14 independent analyses of the same sample preparation. The identity of the antenna proteins was established by comparison of the measured molecular masses with the mass range predicted from the DNA sequences. It can be seen that the reproducibility of mass measurements is better than 0.01% and that the relative mass deviations ranged from 0.008% to 0.016%.

In a second step, we focused our attention on the reproducibility of molecular masses measured in different sample preparations extracted from different batches of plant material. For that purpose, seven batches of spinach (*Spinacia olearacea*) leaves were collected at different periods of the year and from various locations to minimize the influence of growing conditions as well as any seasonal or species-related effects. Moreover, the thylakoid membranes were extracted after storage of the leaves for at least 18 h in the dark with the aim to minimize any light effects. In seven to 19 measurements conducted over a period of 36 months, the molecular masses were reproduced with SDs of ±6 to 12 D, corresponding to relative SDs of only 0.02% to 0.04% (Table II). Given an inaccuracy of molecular masses of 0.008% to 0.016% within the same sample batch, this means that the uncertainty in molecular masses because of sample preparation and sample heterogeneity is in the range of 0.012% to 0.024% and, therefore, essentially equivalent to the mass accuracy characteristic for ESI-MS. The molecular mass values obtained for spinach correspond excellently with the values recently reported by Gómez et al. (2002).

Separation and Identification of the PSII Antenna Proteins by Reversed-Phase HPLC-ESI-MS in Various Plant Species

Figure 1 illustrates as an example the reconstructed ion chromatogram and the extracted raw spectra together with the deconvoluted mass spectra of band 2 from pea. The chromatographic peak eluting from the column at 21 min contained two proteins of molecular masses 22,842 and 22,907 D, corresponding to Lhcb6.1 and 6.2. The chromatographic peak eluting around 22 min contained two proteins with molecular masses of 24,836 and 24,961 D, respectively, whereas the one eluting at 22.5 min revealed only one protein of intact mass of 24,976 D. Because the literature reports three Lhcb1 genes, whose predicted molecular masses are 24,825, 24,956, and 24,972 D (including acetylation of the amino-terminal Arg), it is reasonable to assume that the proteins contained in the two peaks represent the Lhcb1 antenna proteins (Table I). Moreover, the fourth peak revealed a protein with a molecular mass of 24,835 D, coming close to the value expected for Lhcb2 of 24,837 D. Three well-defined chromatographic peaks, revealing molecular masses of 24,330 (Lhcb 3), 28,686 (Lhcb 4), and 26,556 D (Lhcb 5) contained only one protein. The

Table I. *Comparison of molecular masses determined by HPLC-ESI-MS with the protein masses expected from the DNA sequence in Arabidopsis, pea (Pisum sativum), and tomato (Lycopersicon esculentum)*

*, Amino-terminal acetylation. X, Unknown chemical group. ?, Unknown group blocking the N-terminal acid of Lhca3 does not allow determination of the mass deviation.

Species	Measured Mass (±SD)[a]	Calculated Mass[b]	Mass Deviation[c]	Protein Identification[d]	Accession No.[e]
	D		%		
Arabidopsis					X03907
	24,902.2 ± 2.0	24,904.2*	0.008	Lhcb1.1-1.3	X03908
					X03909
		24,775.0*		Lhcb1.4	X64459
	24,912.6 ± 1.8	24,915.2*	0.011	Lhcb1.5	X64460
					AF134122
	24,940.62 ± 2.4	24,943.3*	0.011	Lhcb2.1-2.3	AF134123
					AF134124
	24,927.23 ± 2.3	24,931.3*	0.016	Lhcb2.4	AF134125
	24,277.64 ± 1.7	24,280	0.013	Lhcb3	AF134126
Pea	24,836 ± 2.9	24,825*	0.11	Lhcb1.1	P04159
	24,961 ± 3.2	24,955.1*	0.05	Lhcb1.2	P07371
	24,955.1 ± 2.5	24,971.3*	0.04	Lhcb1.3	P27490
	24,835 ± 2.7	24,837*	0.004	Lhcb2	P27520
	24,330 ± 1.5	24,331	0.01	Lhcb3	X69215
Tomato	24,880 ± 1.9	24,879*	0.004		P07369
	24,894 ± 3.5				
	24,958 ± 1.9			Lhcb1	
	24,976 ± 3.2				
	24,696 ± 2.1	24,692*	0.01		P07370
	24,866 ± 2.1	24,860*	0.02	Lhcb2	P14278
		24,788*			P14279
	24,300 ± 2.4	24,308	0.03	Lhcb3	P27489
	24,330 ± 3.1				
	22,830 ± 3.5	22,800	0.1	Lhcb6	P27525
	22,610 ± 2.7	22,610	0.0		P27527
	26,602 ± 2.9	26,581+X	?	Lhcb5	S16294

[a]The mean ± SD of at least three (13 or 14 for Arabidopsis) measurements is presented. [b]Calculated average mass of the uncharged assigned gene product. [c]Percent difference between expected and observed masses. [d]Protein IDs indicate assignments made by comparison between measured and calculated molecular masses. [e]Accession no. in GenBank, National Center for Biotechnology Information (NCBI), and SWISS-Prot databases.

measured molecular mass of Lhcb3 correlated excellently with the predicted mass of is 24,331 D. Interestingly, all six antenna found have molecular masses similar to the values recently measured by Gómez et al. (2002).

Table I also reports values found for antenna proteins in tomato, where most of the genes for major antenna proteins, as well as those for the two minor antenna proteins Lhcb5 and Lhcb6, have been sequenced. By comparison of the measured molecular masses with those deduced from DNA sequence (see Table I), the identification of each antenna protein was performed for this species, too. In the case of Lhcb5, the measured value is higher than the expected because of an unknown chemical group blocking the amino terminus.

Similarly, the major and minor antenna proteins of PSII were identified in the reconstructed ion chromatograms obtained from 14 different plant species. In cases where the molecular mass measured for a protein did not correspond well to mass values expected from the DNA sequence, or where the gene is not known, the protein types were assigned on the basis of the mass range calculated from known genes of other species.

Tables III through VI give an overview of the experimental molecular masses of the major and minor antenna proteins in dicots and monocots. The exper-

Table II. *Reproducibility of molecular masses of major and minor antenna proteins in antenna protein preparations from spinach prepared and analyzed over a period of 36 months*

Protein	Average Molecular Mass	SD	Relative SD	No. of Measurements
	D		%	
Lhcb1.1	24,933.7	5.7	0.023	19
Lhcb1.2	25,004.2	6.2	0.025	18
Lhcb2	24,758.4	8.6	0.035	19
Lhcb3	24,320.3	7.7	0.032	16
Lhcb4	28,072.5	11.4	0.041	11
Lhcb5	27,071.1	9.9	0.037	7
Lhcb6	22,812.6	8.8	0.039	12

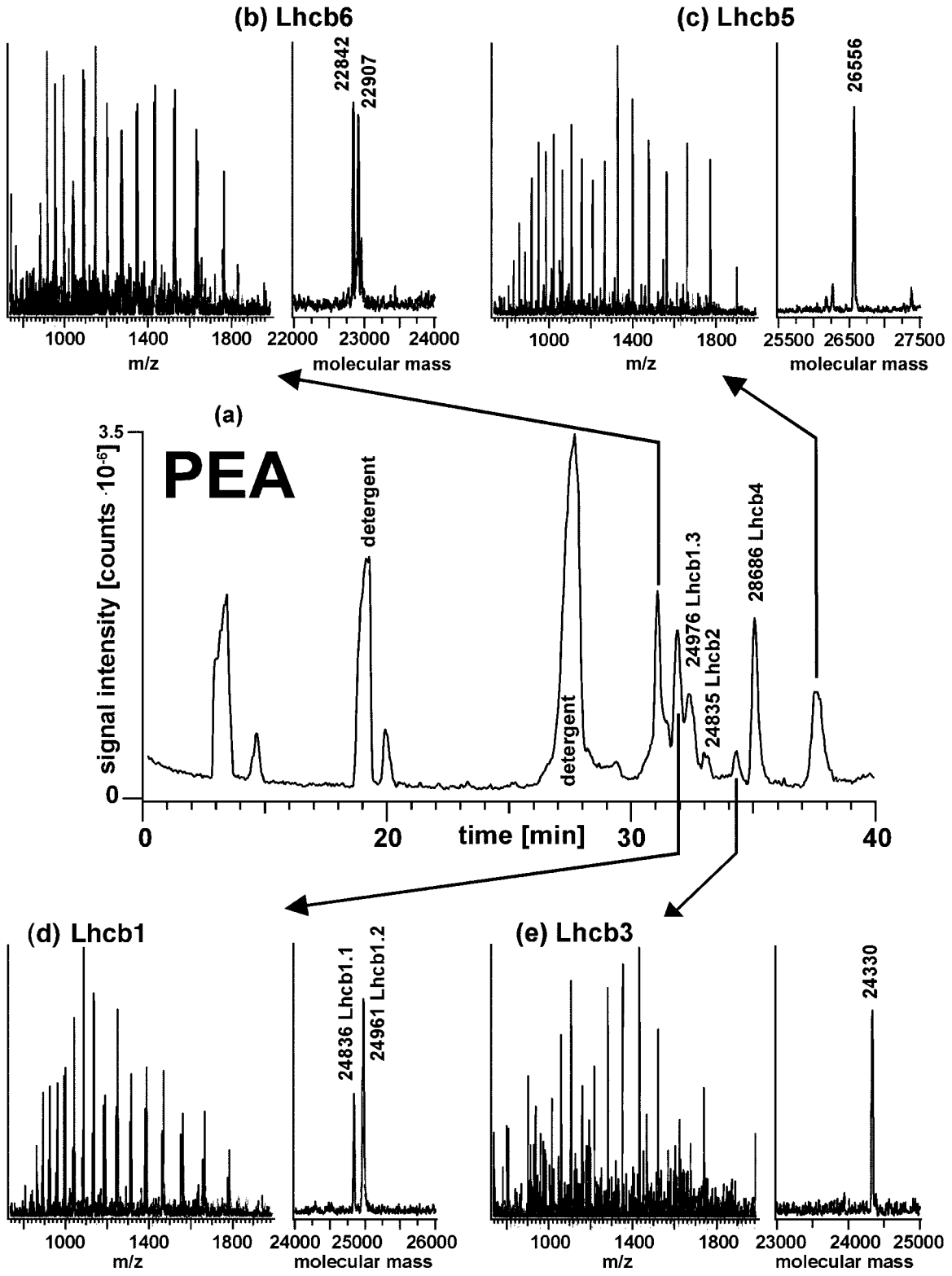

Figure 1. Identification of the protein components of the PSII major and minor antenna system from pea by reversed-phase HPLC-ESI-MS. Column, Vydac Protein C-4 (250- × 4.6-mm i.d.) packed with 5-μm butyl silica; mobile phase, 45-min linear gradient from 27.5% to 63.5% (v/v) acetonitrile in water containing 0.05% (v/v) trifluoroacetic acid (TFA); flow rate, 1.0 mL min^{-1}; flow of column effluent entering the mass spectrometer, 50 μL min^{-1}; detection, ESI-MS, 500 to 2,000 atomic mass units; and injection volume, 100 μL.

Table III. *Comparison of experimental and computed values of the molecular masses of the Lhcb1 major antenna protein of PSII from dicots*
The SDS of measured molecular masses range between ±1.5 and ±2.7 D.

Species	Lhcb1					
	Intact mass measured	Mass calculated[a]	Accession no.[b]	Gene name[c]	Mass apparent[d]	No. of genes by DNA probes[e]
	D					
Spinach						
1	24,936	–	–	–	28,400	4
2	25,014	25,021	P12333	Cab 21	27,800	–
Pea						
1	24,836	24,857	P04159	CabAB86	–	8
2	24,961	24,957	P07371	Cab	–	–
3	24,976	24,973	P27490	Cab 8	–	–
Tomato						
1	24,880	24,879	P07369	Cab 3C	28,400	8
2	24,894	–	–	–	28,200	–
3	24,958	–	–	–	27,700	–
4	24,976	24,692	P07370	Cab 1B	–	–
5	24,696	–	–	–	–	–
Petunia (*Petunia hybrida*)						
1	24,820	24,818	P12062	Cab 37	26,700	16
2	24,936	24,953	P04782	Cab 25	27,200	–
					27,900	
3	24,878	25,029	P04781	Cab 22R	28,200	–
4	24,924	25,009	P04783	Cab 91R	28,400	–
		25,036	P04780	Cab 22L	28,900	–
5	24,964	24,966	P04779	Cab 13	29,300	–
Cucumber (*Cucumis sativus*)						
1	24,690	24,712	P08221	Cab	–	2
2	24,688	–	–	–	–	–
3	24,674	–	–	–	–	–
4	24,726	–	–	–	–	–
Tobacco (*Nicotiana tabacum*)						
1	24,920	24,925	P27492	Cab 16	–	–
2	24,888	–	–	–	–	–
3	24,748	24,750	P27493	Cab 21	–	–
4	24,928	24,911	P27495	Cab 40	–	–
		24,945	P27491	Cab 7		
		24,987	P27496	Cab 50		
Vicia faba						
1	24,964	–	–	–	27,000	–
2	24,806	–	–	–	–	–
Populus albae						
1	24,654	–	–	–	–	–
2	24,772	–	–	–	–	–
3	24,964	–	–	–	–	–
Soybean (*Glycine max*)						
1	24,820	24,761	P09756	Cab 3	–	–
2	24,841	24,781	P12471	Cab	–	–

[a]Values reported were calculated assuming that a conserved Arg, not a Met, was the N-terminal residue of the mature polypeptide. The calculation also took into account the fact that an acetyl group blocks this Arg. [b]Accession no. in NCBI, SWISS-Prot, or Protein Information Resource (PIR) databases. [c]Gene names according to nomenclature reported by Jansson et al. (1992). [d]Values reported in literature: spinach, tomato, and petunia according to Sigrist and Staehelin (1992); *V. faba* according to Machold (1991). [e]Values reported in literature: spinach from Mason (1989), pea from Coruzzi et al. (1983), tomato from Pichersky et al. (1985), petunia from Dunsmuir (1985), and cucumber from Greenland et al. (1987).

imental masses were correlated with the protein types on the basis of DNA sequences available from the literature and from the SWISS-PROT database (http://www.ExPAsy.ch). Expected mass values were calculated assuming that a conserved Arg, not a Met, was the amino-terminal residue of the mature polypeptide, as indicated by tandem MS (Michel et al., 1991). The calculation also took into account the fact that an acetyl group blocks this Arg in both Lhcb1 and Lhcb2 (Jansson, 1994). The apparent molecular masses determined by SDS-PAGE are also reported in the tables with the aim to demonstrate

Table IV. *Comparison of experimental and computed values of the molecular masses of the Lhcb1 major antenna protein of PS II from monocots*

The SDS of measured mass range between ±1.5 and ±2.7.

Species	Lhcb1					
	Intact mass measured	Mass calculated[a]	Accession no.[b]	Gene name[c]	Mass apparent[d]	No. of genes by DNA probes[e]
	D					
Maize (*Zea mays*)						
1	24,901	24,901	P12329	Cab 1	29,500	12
2	24,883	24,881	P27497	CabM9	29,800	–
3	24,833	24,861	P06671	Cab	28,800	–
4	24,799	24,938	Q00827	Cab 48	–	–
Barley (*Hordeum Vulgare*)						
1	24,630	24,620	P08963	Cab 2	26,900	–
2	24,853	–	–	–	–	–
Wheat (*Triticum aestivum*)	24,827	24,910	P04784	Cab	–	–
Rice						
1	24,880	24,986	P12330	Cab 1	–	–
2	24,794	25,003	P12331	Cab 2	–	–
Rye (*Secale cereale*)						
1	24,668	–	–	–	–	–
2	24,830	–	–	–	–	–
3	24,924	–	–	–	–	–

[a]Values reported were calculated assuming that a conserved Arg, not a Met, was the N-terminal residue of the mature polypeptide. The calculation also took into account the fact that an acetyl group blocks this Arg. [b]Accession no. in NCBI, SWISS-Prot, or PIR databases. [c]Gene names according to nomenclature reported in Jansson et al. (1992). [d]Values reported in literature: maize according to Bassi et al. (1988), and barley according to Sigrist and Staehelin (1994). [e]Values reported in literature: maize according to Sheen and Bogorad (1986).

the relatively large discrepancy between molecular masses obtained for these hydrophobic proteins determined by SDS-PAGE and ESI-MS.

UV Chromatograms of PSII Protein Preparations from Different Plants and Reproducibility of Chromatographic Profiles

After identification by reversed-phase HPLC-ESI-MS, each preparation of antenna proteins isolated from each plant was analyzed by reversed-phase HPLC with UV detection in triplicate to obtain retention data and to evaluate the reproducibility of the chromatographic profiles and resolution of the antenna proteins. Because the retention times of major antenna proteins were not influenced by the presence of the minor antenna proteins also present in band 2 and vice versa, only the chromatograms of band 2 will be presented and discussed. Figures 2 through 5 give an overview of the chromatographic profiles of the protein components of band 2 from 12 different plant species. The reproducibility both of the chromatographic patterns and the retention times ($< 0.4\%$ relative SD) was equivalent to that already observed with the protein components of PSI (see Zolla et al., 2002), providing further evidence for the high suitability of reversed-phase HPLC for the separation and identification of antenna proteins.

Figures 2 through 4 refer to the dicots more commonly studied, namely spinach, petunia, pea, tomato, tobacco, cucumber, *P. albae*, *V. faba*, and soybean, whereas Figure 5 collects the chromatograms of the monocots maize, rice, and barley. Upon comparing the chromatograms in Figure 2 through 5, it was observed that the profiles varied characteristically and reproducibly from plant to plant. Moreover, the Lhcb1 retention times in all species ranged from 26 to 34 min, corresponding to an elution window of rather narrow acetonitrile concentration of only 3.9% (v/v). From this observation, it can be inferred that the antenna proteins from different species have similar hydrophobicities, although the differences in sequence are in many cases sufficient to allow their chromatographic separation. This is in accordance with variations in amino acid sequence deduced from the DNA sequence reported in the literature for other plants (Jansson, 1994) and with different band patterns obtained by denaturing SDS-PAGE.

A comparison of the retention times of the major antenna proteins revealed that Lhcb3 eluted after Lhcb1 and Lhcb2 in all species, whereas the retention time of Lhcb2 was more variable as its hydrophobic properties in relation to Lhcb1 varied from species to species. Regarding the minor antenna proteins, Lhcb6 eluted first as the most hydrophilic antenna protein in all species examined, in contrast to Lhcb5,

Table V. *Comparison of experimental and computed values of the molecular masses of the Lhcb2 and Lhcb3 major antenna proteins of PSII from different species*

The SDS of measured mass ranges between ±1.5 and ±2.7.

Species	*Lhcb2*			*Lhcb3*		
	Intact mass measured	Mass calculated[a]	Mass apparent[b]	Intact mass measured[c]	Mass calculated[d]	Mass apparent[e]
	D			*D*		
Dicots						
Spinach	24,761	–	26,600	24,323	–	–
Petunia						
1	24,790	–	26,700	24,280	–	–
2	–	–	26,400	24,232	–	–
Pea	24,835	24,838 (P27520)[c]	–	24,330	24,331 (Q05918)[c]	–
Tomato						
1	24,866	24,860 (P14278)[c]	26,800	24,300	24,308 (P27489)[c]	–
			24,100	24,330		
2	–	24,788 (P14279)[c]			–	–
Cucumber	24,702	–	–	24,252	–	–
Tobacco	24,840	24,825 (P27519)[c]	–	24,224	–	–
V. faba	24,740	–	25,000	24,338	–	–
P. albae	24,690	–	–	24,254	–	–
Soybean	–	–	–	24,322	–	–
Monocots						
Maize	24,721	24,576[d]	28,500	24,289	–	26,400
Barley	24,690	–	27,500	24,292	24,285 (P27523)[c]	26,000
Wheat	24,675	–	–	24,288	–	–
Rice	24,800	24,857 (P27519)[c]	–	24,350	–	–
Rye	24,624	–	–	24,290	–	–

[a]Values reported were calculated assuming that a conserved Arg, not a Met, was the N-terminal residue of the Lhcb2 mature polypeptide. The calculation also took into account the fact that an acetyl group blocks this Arg. In the case of Lhcb3, values were calculated according to Jansson (1994) without adding chemical modification. [b]Values reported in literature: spinach and barley according to Allen and Staehelin (1992), petunia and tomato according to Sigrist and Staehelin (1992), *V. faba* according to Machold (1991), and maize according to Bassi et al. (1998). [c]Accession no. in NCBI, SWISS-Prot, or PIR databases. [d]According to Viret et al. (1990).

which was the most hydrophobic in all cases. Based on its retention, Lhcb4 can be considered to be relatively hydrophobic compared with the other major antenna proteins in all species except in spinach, where it eluted before Lhcb1. In tomato and petunia, two Lhcb6 proteins were separable by reversed-phase HPLC, whereas only one single peak for Lhcb5 and Lhcb4 was observed. Finally, it was noticed that the relative abundance of minor antenna proteins in relation to the major antenna proteins was generally higher in dicots in comparison with monocots.

The proteins of the same type showed quite dissimilar retention times and relative abundances in different species, whereas they displayed quite similar electrophoretic patterns in SDS-PAGE (data not shown). These characteristic chromatographic patterns may serve as a highly confident and reproducible fingerprint for comparison of the antenna proteome within a single and among different species. Moreover, because the identity of the protein(s) contained in each HPLC peak has been previously identified by ESI-MS, the UV-chromatograms shown in Figures 2 through 5 can now be used as a reference for identification of protein types in a variety of plant species in future experiments without the necessity for expensive and comparably complex ESI-MS detection.

DISCUSSION

Protein Identification Based on Reversed-Phase HPLC-ESI-MS

The chromatographic separation of the antenna proteins is based on the different hydrophobicities of the analytes. Therefore, chromatographic retention times may serve as a first indication of the identity of a protein, although coincidental elution of other compounds at the same position in the chromatogram may result in false positive identification. As a consequence, identification needs to be corroborated by additional molecular properties of the analyte. The intact molecular mass represents an important intrinsic property of a protein molecule suitable for its characterization and identification (Nguyen et al., 1995). Although the molecular mass of a protein is primarily defined by its amino acid sequence that can be readily deduced from the DNA sequence, permanent or environmentally induced posttranslational modification or processing may alter its molecular mass. Moreover, partial proteolysis or other chemical reactions effective during sample preparation may change the molecular mass, resulting in failure to correctly identify the protein under investigation.

Table VI. *Comparison of experimental and computed values of the molecular masses of the minor (Lhcb6, Lhcb5, and Lhcb4) antenna protein of PSII from different species*

The SDS of measured mass range between ±1.5 and ±2.7.

Species	Lhcb6			Lhcb5			Lhcb4		
	Intact mass measured	Mass calculated[a]	Mass apparent[b]	Intact mass measured	Mass calculated[a]	Mass apparent[c]	Intact mass measured	Mass calculated[a]	Mass apparent[c]
		D			*D*			*D*	
Dicots									
Spinach	22,820	22,813 (P36494)[e]	21,000	27,091	–	27,500 27,800	28,076	–	29,500 30,000
Petunia									
1	22,688	–	–	27,926	–	–	28,354	–	–
2	22,750	–	–	–	–	–	–	–	–
3	22,814	–	–	–	–	–	–	–	–
4	22,862	–	–	–	–	–	–	–	–
Pea									
1	22,842	–	–	26,556	–	26,500	28,686	–	30,000
2	22,907	–	–	–	–	–	–	–	–
Tomato									
1	22,830	22,800 (P27525)[e]	–	26,602	26,581 (S16294)[e] (CAA4390)[e]	26,000	28,486	–	28,800
2	22,610	22,610 (P27527)[e]	–	–	–	–	–	–	–
Cucumber	22,758	–	–	26,722	–	–	28,194	–	–
V. faba	22,910	–	–	26,508	–	–	28,686	–	–
Tobacco									
1	22,614	–	–	–	–	–	28,420	–	–
2	22,652	–	–	–	–	–	–	–	–
P. albae	–	–	–	26,552	–	–	28,176	–	–
Soybean	–	–	–	26,620	–	–	28,212	–	–
Monocots									
Maize	22,714	22,809 (AAA64416)[e]	20,500	26,679	26,624 (T02251)[e]	–	28,613	28,558 (T02986)[e]	31,000
Barley	22,838	–	21,000	27,170	27,137 (S21386)[e]	–	28,388	27,374 (CAA44777)[e]	30,000
				27,180					31,000
Wheat	22,800	–	–	27,316	–	–	28,344	–	–
Rice									
1	22,576	–	–	26,712	–	–	–	–	–
2	22,644	–	–	–	–	–	–	–	–
Rye	22,822	–	–	26,738	–	–	28,354	–	–

[a]Values reported were calculated assuming a conserved Leu as starting amino acid for Lhcb5 and an LF as the first two amino acids for Lhcb4 of the mature polypeptide. The calculation does not take into account the fact that an unknown chemical residue blocks both Leu. Values for Lhcb6 were calculated assuming two amino acids as initial amino acids of the NH_2 terminal, without adding any chemical modification. [b]Values reported in literature: spinach according to Barbato et al. (1989), maize according to Bassi et al. (1988), and barley according to Morishige and Thornber (1992). [c]Values reported in literature: spinach and pea according to Allen and Staehelin (1992), tomato according to Falbel and Staehelin (1992), maize according to Bassi and Simpson (1987), and barley according to Peter and Thornber (1991). [e]Accession no.

Nevertheless, our approach to protein identification by intact molecular masses is justified by the data from Arabidopsis, pea, and tomato confirming that the measured molecular masses represent the real molecular masses of the proteins, which are in excellent correspondence with DNA sequence data. However, in Tables III through VI, significant deviations between the experimental and the expected molecular masses can be found that are far beyond the experimental error of mass measurements. Hence, there has to be some chemical difference between the actual protein structure and the structure predicted from the DNA sequence, which can only be explained by incorrect DNA sequences (probably a consequence of the chemical sequencing methods common more than 10 years ago), different precursor processing, or unknown posttranslational modifications. The molecular masses determined for Lhcb3 and Lhcb6 by ESI-MS are in excellent agreement with the masses expected from the protein sequence, which could be obtained because of the absence of amino-terminal acetylation. Moreover, the correct assignment of major and minor antenna proteins in spinach by intact molecular mass measurements (Corradini et al., 2000) could be also confirmed by immunoblotting and amino-terminal microsequence

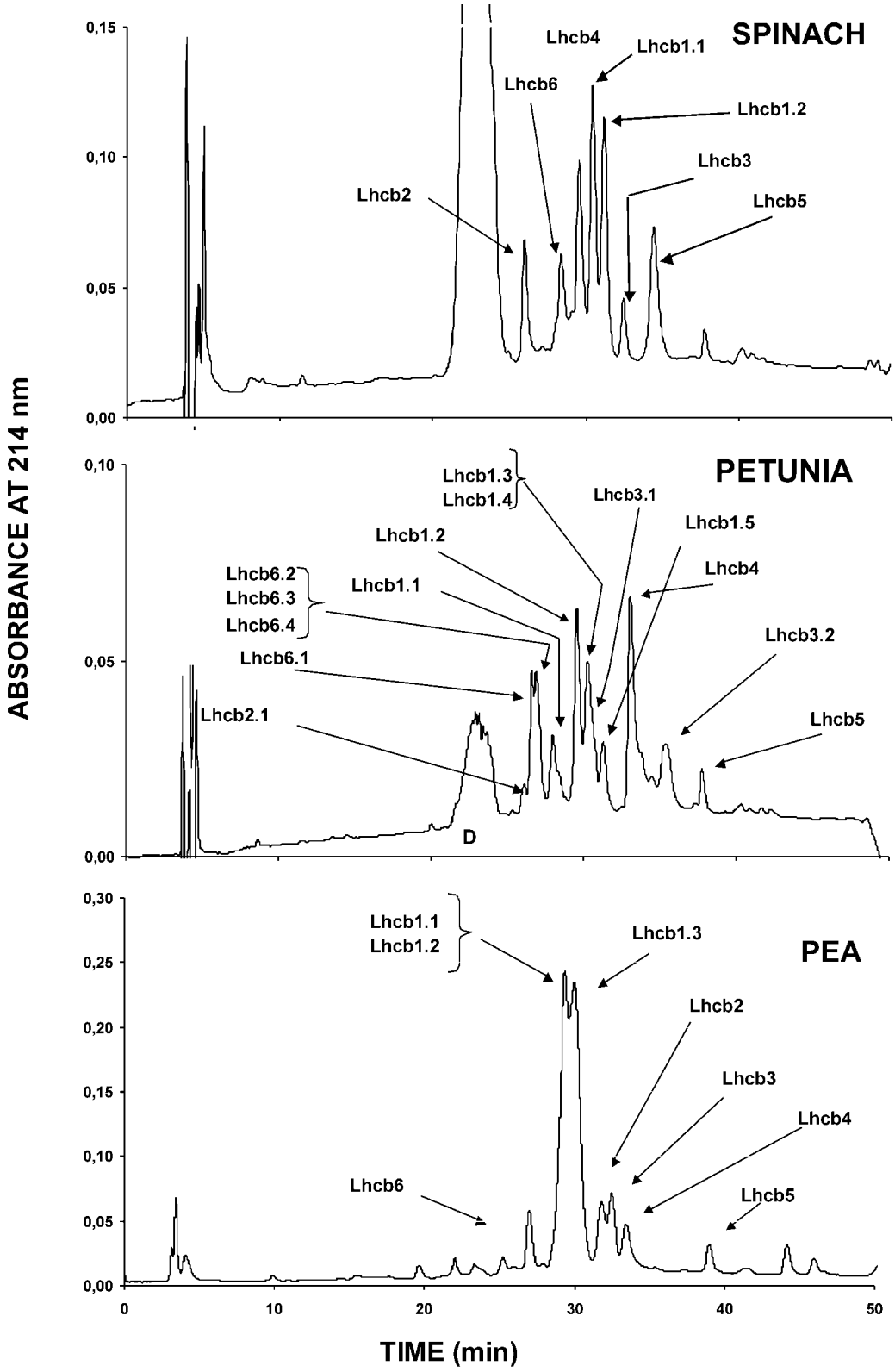

Figure 2. Chromatographic fingerprints of the PSII antenna proteins of band 2 from Suc gradient separation of BBY preparations from the following dicot plant species: a, spinach; b, petunia; and c, pea. Column, Vydac Protein C-4 (250- × 4.6-mm i.d.) packed with 5-μm butyl silica; mobile phase, 45-min linear gradient from 27.5% to 63.5% (v/v) acetonitrile in water containing 0.1% (v/v) TFA; flow rate, 1.0 mL min; UV detection, 214 nm; injection volume, 50 μL; and sample size, approximately 40 μg of total protein. D, Peak containing detergent and free pigments.

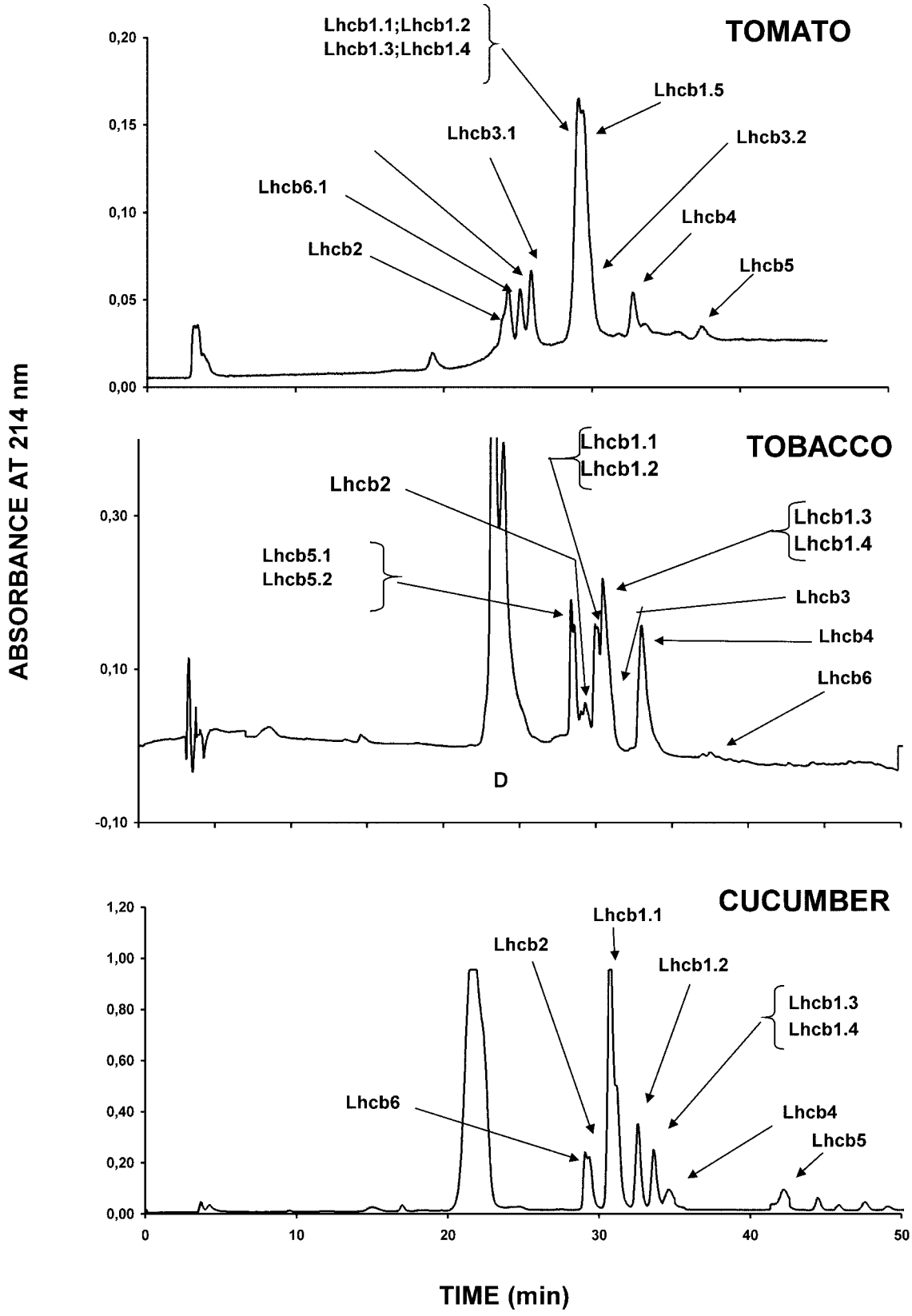

Figure 3. Chromatographic fingerprints of the PSII antenna proteins of band 2 from Suc gradient separation of BBY preparations from the following dicot plant species: a, tomato; b, tobacco; and c, cucumber. Conditions as in Figure 2.

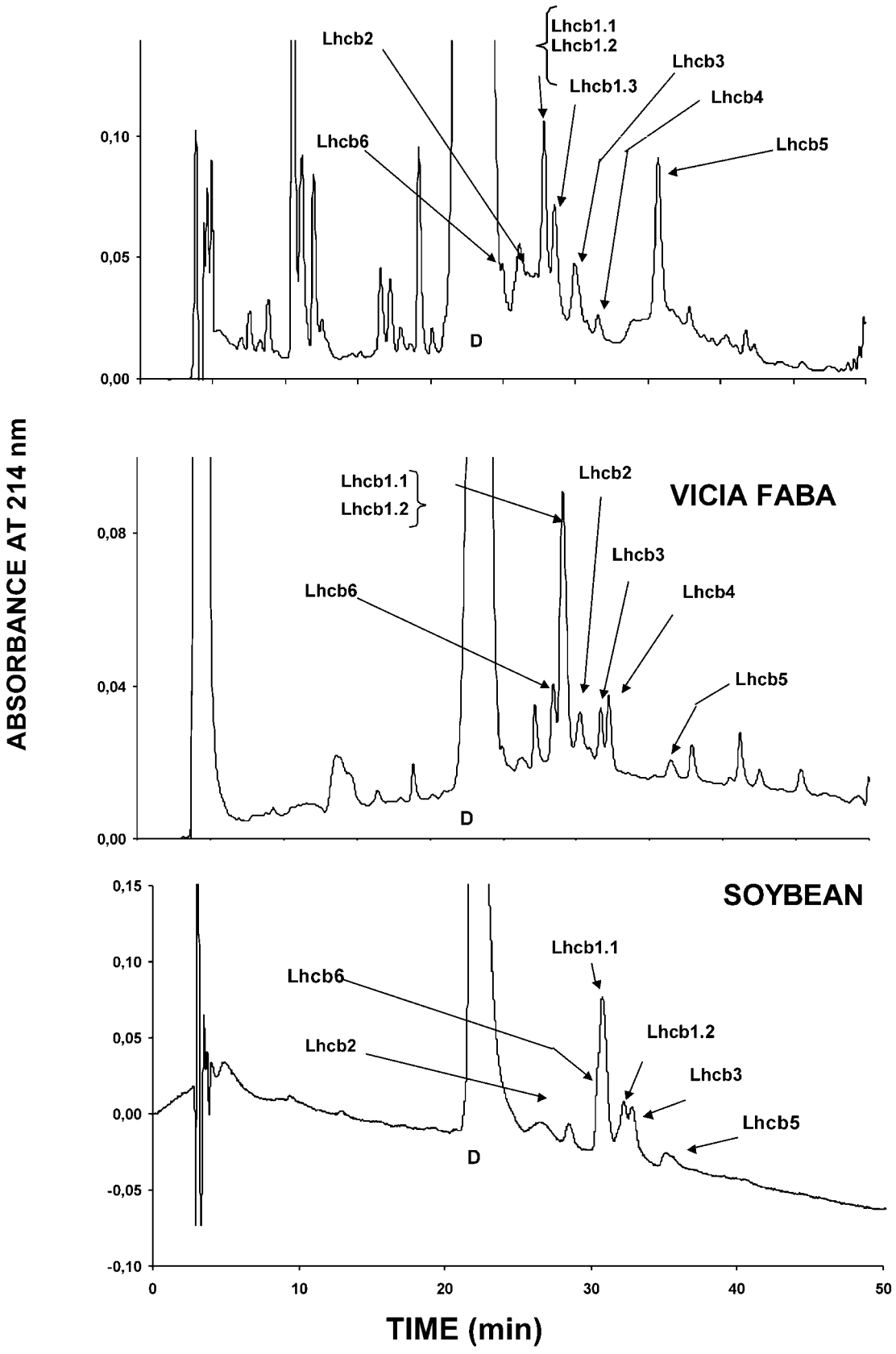

Figure 4. Chromatographic fingerprints of the PSII antenna proteins of band 2 from Suc gradient separation of BBY preparations from the following dicot plant species: a, *P. albae*; b, *V. faba*; and c, soybean. Conditions as in Figure 2.

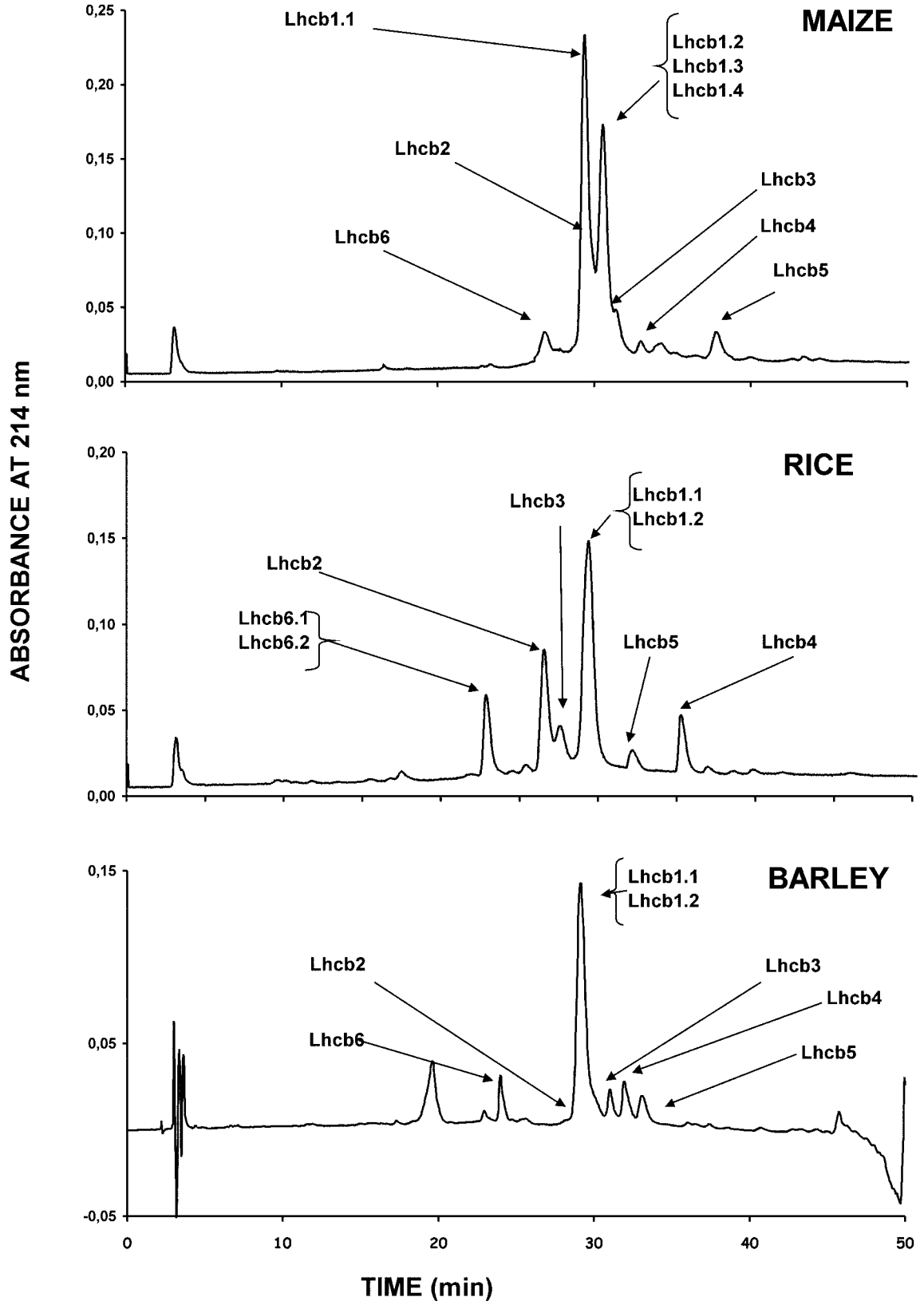

Figure 5. Chromatographic fingerprints of the PSII antenna proteins of band 2 from Suc gradient separation of BBY preparations from the following monocot plant species: a, maize; b, rice; and c, barley. Conditions as in Figure 2.

analysis (Zolla et al., 1999), although the measured masses sometimes differed significantly from the molecular masses derived from the DNA sequence.

In our opinion, false positive identification of a protein as a certain type of antenna protein by retention time plus molecular mass is very unlikely because the antenna proteins are the most abundant in the thylakoid membranes; they have been fractionated in several dimensions using selective extraction, ultracentrifugation, and reversed-phase HPLC; and their hydrophobicities and molecular mass ranges differ significantly. In due consequence, it is reasonable to suggest an assignment of an experimental molecular mass to one of the different types of antenna proteins, if it fits into the range of mass values expected from the DNA sequence.

The major advantage of intact molecular mass measurements by reversed-phase HPLC-ESI-MS relies on the fact that identification can be performed with reasonable effort and instrumentation in acceptable time compared with identification by antibody or protein sequencing. Such rapid analysis is indispensable for the comparison of the antenna proteomes in several plant species. The availability of more DNA sequence data as well as the characterization of posttranslational modifications in the near future will allow to find more and more exact correlations and it may be anticipated that some of our identifications will have to be corrected because of such new data. In this respect, the large body of molecular mass data presented in Tables III through VI may serve as a reference for future investigations of antenna proteins and as groundwork for revealing the differences observed between sequence data and the gene products.

Utility of UV Chromatograms for Protein Identification

Each species exhibits a unique chromatographic pattern that reflects the varying hydrophobicity and stoichiometry of the light-harvesting proteins in all species examined. Compared with the antenna proteins of PSI, the observed differences in PSII are more pronounced (see also Zolla et al., 2002). Reversed-phase HPLC holds the advantage over SDS-PAGE of being generally capable of the fractionation of all antenna types within less than 1 h without the necessity for individual optimization of the separation conditions. Although proteins differing in only two amino acids could be separated chromatographically (see Lhcb 1 in tomato), co-elution of different antenna protein types was observed in very few cases (e.g. Lhcb1 and Lhcb6 co-elute in soybean), and sometimes in the case of protein isoforms. In such situations, only MS is capable of identifying the co-eluting compounds.

PSII Major Antenna Proteins

The Lhcb1 protein, as expected, usually represents the most abundant peak in the HPLC profile. Although the Lhcb1 elution times were in a narrow time window, small but significant differences among different species were observed: Lhcb1 from tomato was more hydrophilic than that from pea and spinach. The measured molecular masses for Lhcb1 ranged from 24,630 in barley to 25,014 D in spinach (Tables III and IV). Moreover, it was found that more than one and up to five Lhcb1 isoforms could be distinguished (Huber et al., 2001). From a comparison of molecular masses measured for monocots and dicots, it is interesting to note that in monocots, both Lhcb1 and Lhcb2 show smaller masses than in dicots, suggesting a different evolution of the antenna proteins in dicots and monocots. However, the mass of Lhcb1 was always higher than that of Lhcb2, both in monocots and dicots. In tomato, two proteins of molecular masses of 24,880 and 24,696 D were measured, as expected from the gene sequences. Thus, the cab3C protein, which was expected to be two amino acids longer and less hydrophobic than cab1B (Green et al., 1992), corresponds to our Lhcb1.1.

The molecular masses of Lhcb2 ranged from 24,624 D in rye to 24,866 D in tomato (Table V). The Lhcb2 of cucumber had a mass very similar to Lhcb1, making it difficult to discriminate between Lhcb1 and Lhcb2 in this species. Only one Lhcb2 protein was revealed by reversed-phase HPLC-ESI-MS in all species examined, in agreement with one gene copy reported for these species in the literature. An exception is the two proteins found in Arabidopsis, for which three very similar *Lhcb2* genes have been reported (Jansson, 1999). Lhcb2 elution times were quite diverse compared with the relatively constant hydrophobicity observed for Lhcb1 and Lhcb3. The observed variability supports the hypothesis that the Lhcb2 protein plays an important role in modulating different aggregation states in different species. This is consistent with the "outer mobile" role of this major antenna because in most models of antenna conformation, it is located on the periphery of the PSII (Barber et al., 1997) and it has been shown to have the highest ability to migrate from grana into stroma (Drepper et al., 1993).

In all species examined, Lhcb3 is the most hydrophobic major antenna protein. As a consequence, it is reasonable to assume that Lhcb3 is tightly bound to the reaction center of PSII (Harrison et al., 1993) and moderately bound to the peripheral, trimeric light-harvesting complex (Boekema et al., 1999b). The molecular masses of Lhcb3 ranged between 24,292 and 24,330 D for barley and tomato, respectively (Table V).

In all species examined, the ratio of peak areas in the chromatograms is approximately 10:3:1 for Lhcb1:Lhcb2:Lhc3, which is consistent with that revealed by immunological analysis (Peter and Thornber, 1991), by the ratio of genes identified (30 Lhcb1 versus 10 Lhcb2 CAB sequences; Jansson et al., 1992), and the ratio of mRNA in Arabidopsis (Jansson, 1999). This good agreement between the ratio of

genes and gene products determined by HPLC again supports the assumption that all genes are active at the same time.

PSII Minor Antenna Proteins

Lhcb6 was usually the first peak eluting from the column, exhibiting the lowest hydrophobicity. The narrow range of molecular masses between 22,610 and 22,910 D (Table VI) is in agreement with the high sequence homology attributed to this protein (Schwartz and Pichersky, 1990). Interestingly, more than one Lhcb6 protein was found in many species, which is in contrast with only one Lhcb6 gene product expected in Arabidopsis (Jansson, 1999). In tomato, two Lhcb6 proteins having a mass of 22,610 and 22,832 D were found, which are in good agreement with the genes cloned (Schwartz and Pichersky, 1990).

The other two minor antenna proteins, Lhcb5 and Lhcb4, which were matter of ambiguity including differences in migration order in certain species (Falbel and Staehelin, 1992), were well differentiated by reversed-phase HPLC in all species examined. Lhcb4 is the largest antenna protein with a molecular mass ranging from 28,076 to 28,696 D (Table VI). In maize, comparison of the measured masses with that expected (Bergantino et al., 1998) indicates that the mature protein comprises 262 amino acids. Only one Lhcb4 protein was found in all species, except in Arabidopsis (Jansson, 1999), where two proteins were found although three genes have been reported. However, two of the three genes (Lhcb4.1 and Lhcb4.2) in Arabidopsis have the same level of expression, whereas the third (Lhcb4.3) significantly differs from Lhcb4.1 and Lhcb4.2 and it is expressed only at a low level, and, therefore, escapes detection (Jansson, 1999).

Lhcb5 is the last antenna protein eluting from the column, which indicates that it is the most hydrophobic antenna protein. This is also consistent with the finding that it is strongly associated with the core complex (Falbel and Staehelin, 1992), being more often present in super-complexes obtained from oxygen-evolving PSII membranes than the other minor antenna proteins (Boekema et al., 1999a). The molecular masses of Lhcb5 ranged from 26,508 D in *V. faba* to 27,316 D in wheat (Table VI). A good correspondence is found in maize, which has been sequenced recently and where the measured molecular mass of 26,679 D (including the unknown chemical group blocking the NH_2 terminal amino acid) is close to that deduced from the DNA, namely 26,624 D. From this information, we can conclude that in maize, the peptide cleavage of the precursor most probably takes place at position 36 to form the mature protein, with a Leu as starting amino acid, as observed for Arabidopsis (Table I; Jansson, 1999). As a consequence, in maize, a mature polypeptide of 247 amino acids is present, which comes close to the 243 found for Arabidopsis. Similarly in barley, assuming a cleavage of the Lhcb5 precursor (Morishige and Thornber, 1992) at residue number 36, the first three amino acids are LFD. In fact, the expected mass of the mature protein of 27,127 D is close to a mass of 27,179 D determined by ESI-MS (which includes the mass of the unknown chemical modification of the amino-terminal Leu).

Physiological Implications of Antenna Proteomes

Besides the improved chemical analysis of the PSII antenna proteins obtained by reversed-phase HPLC-ESI-MS and reversed-phase HPLC with UV detection, the analytical data enable valuable insights from a physiological point of view. The analyses allowed us: (a) to identify the protein(s) eluting in each chromatographic peak; (b) to estimate the relative stoichiometry of antenna proteins on the basis of peak areas in the UV chromatograms; (c) to reveal that more than one Lhcb1, Lhcb3, or Lhcb4 protein isoform exist in most species examined; (d) to realize that the number of found proteins and cloned genes match pretty closely, whereas the number of genes detected by DNA-probes is an overestimation; (e) to report the molecular masses of some antenna proteins the genes of which have not yet been sequenced; and (f) to compare the measured molecular masses with those expected to allow an assignment of most proteins to their respective gene families.

The largest peak(s) in most of the chromatograms, which is (are) most likely related to Lhcb1 because of its high abundance, usually contained two or more proteins as evidenced by partly separated chromatographic peaks and/or different masses measured by ESI-MS. These proteins were considered as heterogeneous forms of Lhcb1 (Huber et al., 2001), and indicated in the chromatograms by using the indexed labels Lhcb1.1, Lhcb1.2, etc., according to the nomenclature proposed by Jansson (1999) to identify the numerous Arabidopsis genes.

Lhcb2 eluted as a single HPLC peak in all species examined, whereas Lhcb3 was resolved into two chromatographic peaks in tomato and petunia (Lhcb3.1 and Lhcb3.2). Because approximately 80% of the sequence of the major antenna proteins, especially in the three transmembrane helices, are highly conserved among different plant species (Jansson, 1994), the observed difference in chromatographic retention must be attributed to variations in the rather hydrophilic amino-terminal regions of the proteins, which can play a major role in protein-protein interaction and supramolecular organization of PSII. In fact, more than one type of trimeric populations have been postulated to exist (Jackowski et al., 2001).

In fact, in PSII, a number of gene products were revealed, corresponding to the multitude of genes cloned, indicating that all the light-harvesting genes

cloned are not silent in their protein expression (Walling et al., 1988). Until now, the reason for the existence of several genes encoding the light-harvesting proteins is not yet well understood, but the data reported allow us to assume that the genome organization in multiple gene families of the antenna proteins may partially reflect the high degree of organization and interconnection of the light-harvesting system within the thylakoid membrane complexes. Thus, because recent studies reported a specific and time-dependent protein reaction in forming supramolecular organization of PSII, especially for the minor antenna (Boekema et al., 1999b; Wollman et al., 1999), it is reasonable suppose that the observed differences in hydrophobicity of the antenna proteins, within the same species as well as among various species, used here to explain the elution times observed in the chromatograms, may reflect functional differences in the protein-protein interaction.

This hypothesis, together with the observation that the high copy number of more or less identical genes and gene products is presumably a consequence of the need for a high rate of protein synthesis in the photosynthetic apparatus of plants, could be a possible explanation of the biological significance of the numerous multigene families reported for the antenna proteins as well as for other plant genes (Theologis et al., 2000).

Conclusions and Future Perspectives

In all species examined, each antenna protein type showed molecular masses that fell within a relatively narrow range of mass values. In accordance, it is reasonable to suggest an assignment of an experimental molecular mass to one of the six types of antenna proteins, if it fits into these narrow ranges. In future investigations, we are planning to study the environmental effects on the photosynthetic apparatus directly in the BBY preparation, or even better, in thylakoid membranes, avoiding the separation step by Suc gradient ultracentrifugation. In fact, direct injection of thylakoids onto the HPLC column will make it possible to determine the quantitative relationship between chlorophyll *a*/*b*-binding proteins present in both PSs, which represents a great advantage of chromatographic analysis. Moreover, tandem MS of trypsin-digested antenna proteins will allow one to get partial sequence information that is suitable to unequivocally assign each protein to its respective gene, as previously performed for core proteins (Sharma et al., 1997b; Ouellette and Barry, 2002). The detection and study of posttranslational modifications of the antenna proteins, either by intact molecular mass measurements or tandem MS, will allow the gathering of insights into the molecular mechanism by which the chloroplast modulates the adaptation of photosynthetic apparatus to environmental changes. This is particularly attractive in Arabidopsis, where the entire genome is known. However, for the study of Arabidopsis having very small leaves, a further step of miniaturization of the analytical method will be mandatory, a concept that can be successfully realized by the use of monolithic capillary columns (Premstaller et al., 2001).

MATERIALS AND METHODS

Chemicals

Reagent grade phosphoric acid, magnesium chloride, sodium chloride, silver nitrate, sodium carbonate, TFA, methanol, ethanol, formamide, as well as HPLC grade water and acetonitrile, were obtained from Carlo Erba (Milan). Acrylamide, *N*,*N*′ methylene-bis-acrylamide, and all other reagents for SDS-PAGE were purchased from Bio-Rad (Segrate, Italy). Suc, Tricine, TRIS, *n*-octyl β-D glucopyranoside, *n*-dodecyl β-D maltoside (DM), chlorophyll *a* and *b*, and MES were obtained from Sigma (Milan). Triton X-100 and *n*-octyl-Suc were purchased from Calbiochem (San Diego).

Isolation of the Major and Minor Antenna Systems by Suc Gradient Ultracentrifugation

Chloroplast thylakoid membranes (PSII membranes) were isolated from the following dicot and monocot leaves: spinach (*Spinacia olearacea*), petunia (*Petunia hybrida*), pea (*Pisum sativum*), tomato (*Lycopersicon esculentum*), tobacco (*Nicotiana tabacum*), cucumber (*Cucumis sativus*), soybean (*Glycine max*), *Vicia faba*, *Populus albae*, maize (*Zea mays*), rice (*Oryza sativa*), barley (*Hordeum vulgare*), rye (*Secale cereale*), and wheat (*Triticum aestivum*), according to the method of Berthold et al. (1981) with the modification reported elsewhere (Zolla et al., 1999). Leaves were collected at night in the dark and at different periods of the year to minimize any seasonal and light effects. To ensure comparability of the results, the experimental conditions for extraction were kept constant for all species examined. Only in the case of petunia and soybean, the ratio of detergent to protein needed to be raised to completely solubilize the thylakoid membranes. The light-harvesting complex was isolated from the PSII membranes as previously described (Bassi and Dainese, 1992) with the following modifications: PSII membranes were pelleted by centrifugation at 10,000*g* for 5.0 min at 4°C, suspended in B3 buffer at 1.0 mg mL^{-1} chlorophyll, and then solubilized by adding 1% (w/v) DM. Unsolubilized material was removed by centrifugation at 10,000*g* for 10 min. The supernatant was rapidly loaded onto a 0.1 to 1.0 M Suc gradient containing B3 buffer (containing 1.5 mM NaCl, 5 mM $MgCl_2$, and 50 mM MES [pH 6.3]) and 5.0 mM DM. The gradient was then spun on a Centricon T-1080 ultracentrifuge equipped with a model TST 41.14 rotor (Kontron Instruments, Watford, Herts, UK) 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 (Bassi and Dainese, 1992). These bands were used for SDS-PAGE or HPLC analysis without any further treatment. Control analyses of the components of bands 2 and 3 were performed by denaturating SDS-PAGE. In the gels, all species displayed quite similar electrophoretic patterns, showing the three main bands of molecular mass ranging from 25 to 28 kD, which were identified as Lhcb1, Lhcb2, and Lhcb3 (Jansson and Gustafsson, 1991; Peter and Thornber, 1991). Small differences were observed for minor antenna proteins.

HPLC Separations and Hyphenation to ESI-MS

The HPLC separations were carried out on a model 200 C system having a model 785 A UV detector, and a model LC 240 fluorescence detector connected in series (PerkinElmer, Norwalk, CT). Samples were loaded onto the column by a model 7125NS-005 sample injection valve (Rheodyne, Cotati, CA) with a 50- or 100-μL sample loop. The analyses were performed using Vydac Protein C-4 columns of either 250- × 4.6-mm i.d. or 250- × 10-mm i.d., both containing 5-μm porous butyl silica. All solutions were

filtered through a Millipore (Milan) type FH 0.5-μm membrane filter and degassed by bubbling with helium before use. Antenna proteins were eluted with linear gradients of acetonitrile in 0.1% (w/v; with UV or fluorescence detection) or 0.05% (w/v; with ESI-MS detection) aqueous TFA at ambient temperature. The HPLC-ESI-MS experiments were carried out with a model Rheos 2000 low-pressure gradient pump (Flux Instruments, Basel), a degasser (ERC 3215, Ercatech, Bern, Switzerland), a Rheodyne model 7125 injector equipped with a 100-μL sample loop, and a TSQ 7000 triple quadrupole mass spectrometer (Finnigan MAT, San Jose, CA or by ion trap Esquire 3000 plus (Bruker, Daltonik, Germany). Details of instrumental setup and tuning are given in Corradini et al. (2000).

ACKNOWLEDGMENTS

We are grateful to Dr. Sara Rinalducci and Dr. Sonia Troiani (University of Viterbo, Italy) for technical assistance. The authors wish to thank Dr. Bart Thomma (Centre of Microbial and Plant Genetics, Katholieke Universiteit, Heverlee-Leuven, Belgium) for his generous gift of Arabidopsis plants. We also acknowledge Dr. Jaqueline Scarpa for manuscript revision.

Received August 10, 2002; returned for revision September 3, 2002; accepted September 24, 2002.

LITERATURE CITED

Allen KD, Staehelin LA (1992) Biochemical characterization of PS II antenna polypeptides in grana and stroma membranes of spinach. Plant Physiol **100:** 1517–1526

Barbato R, Rigoni F, Giardi MT, Giacometti GM (1989) The minor antenna complexes of an oxygen evolving photosystem II preparation: purification and stoichiometry. FEBS Lett **251:** 147–154

Barber J, Nield J, Morris EP, Zheleva D, Hankamer B (1997) The structure function and dynamics of photosystem two. Physiol Plant **100:** 817–827

Bassi R, Dainese P (1992) A supramolecular light-harvesting complex from chloroplast photosystem II membranes. Eur J Biochem **204:** 317–326

Bassi R, Rigoni F, Barbato R, Giacometti GM (1988) Light-harvesting chlorophyll a/b proteins (LHCII) populations in phosphorylated membranes. Biochim Biophys Acta **936:** 29–38

Bassi R, Simpson D (1987) Chlorophyll-protein complexes of barley photosystem I. Eur J Biochem **163:** 221–230

Bergantino E, Sandona D, Cugini D, Bassi R (1998) The photosystem II subunit Lhcb4 can be phosphorylated in both C3 and C4 plants as suggested by sequence analysis. Plant Mol Biol **36:** 11–22

Berthold DA, Babcock GT, Yocum CF (1981) A highly resolved, oxygen-evolving photosystem II preparation from spinach thylakoid membranes. FEBS Lett **134:** 231–234

Boekema EJ, van Roon H, Calkoen F, Bassi R, Dekker JP (1999a) Multiple types of association of photosystem II and its light-harvesting antenna in partially solubilized photosystem II membranes. Biochemistry **38:** 2233–2239

Boekema EJ, van Roon H, van Breemen JFL, Dekker JP (1999b) Supramolecular organization of photosystem II and its light-harvesting antenna in partially solubilized photosystem II membranes. Eur J Biochem **266:** 444–452

Chait BT, Kent SB (1992) Weighing naked proteins: practical high-accuracy mass measurement of peptides and proteins. Science **257:** 1885–1894

Corradini D, Huber CG, Timperio AM, Zolla L (2000) Resolution and identification of the protein components of the photosystem II antenna system of higher plants by reversed-phase liquid chromatography with electrospray-mass spectrometric detection. J Chromatogr A **886:** 111–121

Coruzzi G, Broglie R, Cashmore AR, Chua NH (1983) Nucleotide sequences of two pea cDNA clones encoding the small subunit of ribulose-1,5-bisphosphate carboxylase and the major chlorophyll a/b-binding thylakoid polypeptide. J Biol Chem **258:** 1399–1402

Covey TR, Bonner RF, Shushan BI, Henion J (1988) The determination of protein, oligonucleotide, and peptide molecular weights by ion-spray mass spectrometry. Rapid Commun Mass Spectrom **2:** 249–256

Damm I, Green BR (1994) Separation of closely related intrinsic membrane polypeptides of the photosystem II light-harvesting complex (LHC II) by reversed-phase high-performance liquid chromatography on a poly(styrene-divinylbenzene) column. J Chromatogr A **664:** 33–38

Drepper F, Carlberg I, Andersson B, Haehnel W (1993) Lateral diffusion of an integral membrane protein: Monte Carlo analysis of the migration of phosphorylated light-harvesting complex II in the thylakoid membrane. Biochemistry **32:** 11915–11922

Dunsmuir P (1985) The petunia chlorophyll a/b binding protein genes: a comparison of Cab genes from different gene families. Nucleic Acids Res **13:** 2503–2516

Falbel TGS, Staehelin LA (1992) Species related differences in the electrophoretic behavior of CP 29 and CP 26: an immunochemical analysis. Photosynth Res **34:** 249–262

Gómez SM, Nishio JN, Faull KF, Whitelegge JP (2002) The chloroplast grana proteome defined by intact mass measurements from liquid chromatography mass spectrometry. Mol Cell Proteomics **1:** 46–59

Green BR, Shen D, Aebersold R, Pichersky E (1992) Identification of the polypeptides of the major light-harvesting complex of photosystem II (LHCII) with their genes in tomato. FEBS Lett **305**: 18–22

Greenland AJ, Thomas MV, Walden RM (1987) Expression of two nuclear genes encoding chloroplast proteins during early development of cucumber seedlings. Planta **170:** 99–110

Harrison MA, Nemson JA, Melis A (1993) Assembly and composition of the chlorophyll a/b light-harvesting complex of barley (*Hordeum vulgare* L.): immunochemical analysis of chlorophyll b-less and chlorophyll b-deficient mutants. Photosynth Research **38:** 141–151

Huber CG, Timperio AM, Zolla L (2001) Isoforms of photosystem II antenna proteins in different plant species revealed by liquid chromatography-electrospray ionization mass spectrometry. J Biol Chem **276:** 45755–45761

Jackowski G, Kacprzak K, Jansson S (2001) Identification of Lhcb1/Lhcb2/Lhcb3 heterotrimers of the main light-harvesting chlorophyll a/b-protein complex of photosystem II (LHC II). Biochim Biophys Acta **1504:** 340–345

Jansson S, Gustafsson P (1991) Evolutionary conservation of the chlorophyll a/b-binding proteins: cDNAs encoding type I, II and III LHC I polypeptides from the gymnosperm Scots pine. Mol Gen Genet **229:** 67–76

Jansson S, Pichersky E, Bassi R, Green BR, Ikeuchi M, Melis A, Simpson DJ, Spangfort M, Staehelin LA, Thornber JP (1992) A nomenclature for the Genes Encoding the chlorophyll a/b binding proteins of higher plants. Plant Mol Biol Rep **10:** 242–253

Jansson S (1994) The light-harvesting chlorophyll a/b-binding proteins. Biochim Biophys Acta **1184:** 1–19

Jansson S (1999) A guide to the *Lhc* genes and their relatives in *Arabidopsis*. Trends Plant Sci **4:** 236–240

Machold O (1991) The structure of light-harvesting complex II as deduced from its polypeptide composition and stoichiometry I. Studies with *Vicia faba*. J Plant Physiol **138:** 678–684

Mason JG (1989) Nucleotide sequence of a cDNA encoding the light-harvesting chlorophyll a/b binding protein from spinach. Nucleic Acids Res **17:** 5387–5394

Matsuoka M, Kano-Murakami Y, Yamamoto N (1987) Nucleotide sequence of cDNA encoding the light-harvesting chlorophyll a/b binding protein from maize. Nucleic Acids Res **15:** 6302–6310

Michel H, Griffin PR, Shabanowitz J, Hunt DF, Bennett J (1991) Tandem mass spectrometry identifies sites of three post-translational modifications of spinach light-harvesting chlorophyll protein II. J Biol Chem **266:** 17584–17591

Morishige DT, Thornber JP (1992) Correlation of apoproteins with the genes of the major chlorophyll a/b binding protein of photosystem II in *Arabidopsis thaliana*. Confirmation for the presence of a third member of the LHC IIb gene family. FEBS Lett **293:** 183–187

Morishige DT, Thornber JP (1994) Identification of a novel light-harvesting complex II protein (LHCII). Photosynth Res **39:** 33–38

Nguyen DN, Becker GW, Riggin RM (1995) Protein mass spectrometry: applications to analytical biotechnology. J Chromatogr A **705:** 21–45

Ouellette AJA, Barry AB (2002) Tandem mass spectrometry identification of spinach photosystem II light harvesting components. Photosynth Res **72:** 159–173

Peter GF, Thornber JP (1991) Biochemical composition and organization of higher plant photosystem II light-harvesting pigment-proteins. J Biol Chem **266:** 16745–16754

Pichersky E, Bernatzky R, Tanksley SD, Breidenbach RB, Kausch AP, Cashmore AR (1985) Molecular characterization and genetic mapping of

two clusters of genes encoding chlorophyll a/b-binding proteins in *Lycopersicon esculentum* (tomato).Gene **40:** 247–258

Premstaller A, Oberacher H, Walcher W, Timperio A-M, Zolla L, Chervet J-P, Cavusoglu N, Van Dorsellaer A, Huber CG (2001) High-performance liquid chromatography-electrospray ionization mass spectrometry using monolithic capillary columns for proteomic studies. Anal Chem **73:** 2390–2396

Schwartz E, Pichersky E (1990) Sequence of two tomato nuclear genes encoding chlorophyll a/b-binding proteins of LHCB6, a PSII antenna component. Plant Mol Biol **15:** 157–160

Sharma J, Panico M, Barber J, Morris HR (1997a) Characterization of the low molecular weight photosystem II reaction center subunits and their light-induced modifications by mass spectrometry. J Biol Chem **272:** 3935–3943

Sharma J, Panico M, Shipton CA, Nilsson F, Morris HR, Barber J (1997b) Primary structure characterization of the photosystem II D1 and D2 subunits. J Biol Chem **272:** 33158–33166

Sheen JY, Bogorad L (1986) Differential expression of six light-harvesting chlorophyll a/b binding protein genes in maize leaf cell types. Proc Natl Acad Sci USA **83:** 7811–7815

Sigrist M, Staehelin LA (1992) Identification of type 1 and type 2 light-harvesting chlorophyll a/b-binding proteins using monospecific antibodies. Biochim Biophys Acta **1098:** 191–200

Sigrist M, Staehelin LA (1994) Appearance of type 1, 2, and 3 light-harvesting complex II and light-harvesting complex I proteins during light-induced greening of barley (*Hordeum vulgare*) etioplasts. Plant Physiol **104**: 135–145

Theologis A, Ecker JR, Palm CJ, Federspiel NA, Kaul S, White O, Alonso J, Altafi H, Araujo R, Bowman CL (2000) Sequence and analysis of chromosome I of the plant *Arabidopsis thaliana*. Nature **408:** 816–820

Viret JF, Schantz ML, Schantz R (1990) Nucleotide sequence of maize cDNA coding for a Light Harvesting chlorophyll *a/b* binding protein of photosystem II. Nucleic Acids Res **16:** 71–79

Walling LL, Chang YC, Demmin DS, Holzer FM (1988) Isolation characterization and evolutionary relatedness of three members from the soybean multigene family encoding chlorophyll a/b-binding proteins. Nucleic Acids Res **16:** 10477–10492

Whitelegge JP, Gundersen CB, Faull KF (1998) Electrospray-ionization mass spectrometry of intact intrinsic membrane proteins. Protein Sci **7:** 1423–1430

Wollman FA, Minai L, Nechushtai R (1999) The biogenesis and assembly of photosynthetic proteins in thylakoid membranes. Biochim Biophys Acta **1411:** 21–85

Zolla L, Rinalducci S, Timperio AM, Huber CG (2002) PSI proteomics of light-harvesting proteins in different plant species: analysis and comparison by liquid chromatography-electrospray ionization mass spectrometry. Photosystem I. Plant Physiol **130:** 1938–1950

Zolla L, Timperio AM (2000) High performance liquid chromatography-electrospray mass spectrometry for the simultaneous resolution and identification of intrinsic thylakoid membrane. Proteins **41:** 398–406

Zolla L, Timperio AM, Testi MG, Bianchetti M, Bassi R, Manera F, Corradini D (1999) Isolation and characterization of chloroplasts photosystem II antenna of spinach by reversed-phase liquid chromatograph. Photosynth Res **61:** 281–290