

[Click here to view linked References](#)**Type of article: Regular Paper****Title**

Isolation of novel PSII-LHCII megacomplexes from pea plants characterized by a combination of proteomics and electron microscopy.

**Authors names and affiliation**

Pascal Albanese<sup>a,b</sup>, Jon Nield<sup>c</sup>, Jose Alejandro Muñoz Tabares<sup>d</sup>, Angelica Chiodoni<sup>d</sup>, Marcello Manfredi<sup>e,f</sup>, Fabio Gosetti<sup>f</sup>, Emilio Marengo<sup>f</sup>, Guido Saracco<sup>a</sup>, James Barber<sup>g</sup>, Cristina Pagliano<sup>a,\*</sup>

<sup>a</sup>Applied Science and Technology Department - BioSolar Lab, Politecnico di Torino, Viale T. Michel 5, 15121 Alessandria, Italy

<sup>b</sup>Department of Biology, University of Padova, Via Ugo Bassi 58 B, 35121 Padova, Italy

<sup>c</sup>School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4NS, United Kingdom

<sup>d</sup>Center for Space Human Robotics IIT@POLITO, Istituto Italiano di Tecnologia, Corso Trento 21, 10129 Torino, Italy

<sup>e</sup>ISALIT–Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy

<sup>f</sup>Department of Science and Technological Innovation, University of Eastern Piedmont, Viale T. Michel 11, 15121 Alessandria, Italy

<sup>g</sup>Department of Life Sciences, Faculty of Natural Sciences, Imperial College London, London SW7 2AZ, United Kingdom

**Corresponding author\***

Cristina Pagliano E-mail: [cristina.pagliano@polito.it](mailto:cristina.pagliano@polito.it); Tel. No. +39 131 229301; Fax No. +39 131 229344; Politecnico di Torino, Applied Science and Technology Department - BioSolar Lab, Viale T. Michel 5, 15121 Alessandria, Italy

## Abstract

In higher plants, Photosystem II (PSII) is a multisubunit pigment-protein complex embedded in the thylakoid membranes of chloroplasts, where it is present mostly in dimeric form within the grana. Its light harvesting antenna system, LHCII, is composed of trimeric and monomeric complexes, which can associate in variable number with the dimeric PSII core complex in order to form different types of PSII-LHCII supercomplexes. Moreover, PSII-LHCII supercomplexes can laterally associate within the thylakoid membrane plane, thus forming higher molecular mass complexes, termed PSII-LHCII megacomplexes (Boekema et al. 1999 *Biochemistry* 38:2233–2239; Boekema et al. 1999 *Eur J Biochem* 266:444–452).

In this study, pure PSII-LHCII megacomplexes were directly isolated from stacked pea thylakoid membranes by a rapid single-step solubilization, using the detergent n-dodecyl- $\alpha$ -D-maltoside, followed by sucrose gradient ultracentrifugation. The megacomplexes were subjected to biochemical and structural analyses. Transmission electron microscopy on negatively stained samples, followed by single particle analyses, revealed a novel form of PSII-LHCII megacomplexes, as compared to previous studies (Boekema et al. 1999 *Biochemistry* 38:2233–2239; Boekema et al. 1999 *Eur J Biochem* 266:444–452), consisting of two PSII-LHCII supercomplexes sitting side-by-side in the membrane plane, sandwiched together with a second copy. This second copy of the megacomplex is most likely derived from the opposite membrane of a granal stack. Two predominant forms of intact sandwiched megacomplexes were observed and termed, according to (Dekker and Boekema 2005 *Biochim Biophys Acta* 1706:12–39), as  $(C_2S_2)_4$  and  $(C_2S_2 + C_2S_2M_2)_2$  megacomplexes. By applying a gel-based proteomic approach, the protein composition of the isolated megacomplexes was fully characterized. In summary, the new structural forms of isolated megacomplexes and the related modelling performed provide novel insights into how PSII-LHCII supercomplexes may bind to each other, not only in the membrane-plane, but also between granal stacks within the chloroplast.

**Keywords:** photosystem II; proteomics; PSII-LHCII megacomplex; single particle analysis; transmission electron microscopy

**Abbreviations:** ATP-ase, ATP-synthase; Chl, chlorophyll; Cyt, cytochrome; 2D, two-dimensional; 3D, three-dimensional;  $\alpha$ -DDM, n-dodecyl- $\alpha$ -D-maltoside; HPLC, high performance liquid chromatography; LHC, light harvesting complex; IpBN-PAGE, large pore blue native polyacrylamide gel electrophoresis; MS, mass spectrometry; OEC, oxygen evolving complex; PS, photosystem; RC, reaction centre; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis; TEM, transmission electron microscopy

## Introduction

The thylakoid membranes of higher plants form a highly plastic and dynamic membrane system characterized by a peculiar feature, the presence of granal stacks, or grana (Nevo et al. 2009). Tight granal stacking is typically observed, with two grana membrane pairs being separated by ~3.5 nm on the stromal side and with a lumenal width of 4.5 nm (Daum et al. 2010; Kirchhoff et al. 2011). A consequence of grana formation is the structural sub-compartmentalization of the thylakoid membrane system into tightly stacked and exposed unstacked regions, and the subsequent lateral distribution of protein complexes between these sub-compartments. Photosystem I (PSI) (Amunts and Nelson 2009) and ATP-synthase (ATP-ase) (Groth and Pohl 2001) have bulky stromal protrusions and, mainly due to this, are confined to unstacked stroma lamellae and the granal margins, whereas Photosystem II (PSII) and its light harvesting complex (LHCII), with their relatively flat stromal surfaces, are concentrated in the grana membranes (Andersson and Anderson 1980; Albertsson 2001). It is widely believed that interactions between the stromal protein moieties of LHCII, localized in adjacent stacked membranes, are one of the main determinants for grana formation (Mullet 1983), although surface electrical charge properties play an important role (Barber 1982).

PSII is a multi-subunit pigment-protein complex that normally functions as a dimer within the grana of higher plants (Hankamer et al. 1997). Within the dimer, each of the two monomeric cores contains the integral subunits D1 and D2. These proteins bind most of the redox cofactors required

to propagate the photosynthetic electron transport chain and, together with the  $\alpha$  and  $\beta$  subunits of cytochrome b559 (Cyt b559) and the low molecular subunit PsbI, constitute the reaction centre (RC) of PSII, a term used in this case to define the minimum unit for sustaining primary charge separation (Nanba and Satoh 1987). The RC proteins are closely associated with the two inner antenna proteins CP47 and CP43 and bind additional low molecular mass membrane subunits, which play a key role in the assembly and stabilization of the overall PSII core complex (for details see the recent review Pagliano et al. 2013). This core complex, to be fully functional, requires the assembly of the oxygen-evolving complex (OEC), containing the  $Mn_4CaO_5$  cluster which, in higher plants, is stabilized by the extrinsic subunits PsbO, PsbP, PsbQ and PsbR. Up to now, the most highly resolved three-dimensional (3D) structure available for the plant PSII core complex has been obtained by electron crystallography at 8 Å, allowing for the position and helix organization of the major subunits to be revealed (Rhee et al. 1998; Hankamer et al. 2001). Moreover, crystal structures have been determined for the isolated extrinsic polypeptides PsbP (Ifuku et al. 2004) and PsbQ (Calderone et al. 2003; Balsera et al. 2005). In higher plants, the dimeric PSII RC core is serviced by a variable number of intrinsic light harvesting proteins, named Lhcb1-6, binding chlorophyll (Chl) *a* and Chl *b*. The Lhcb1-3 subunits typically occur in a ratio of about 8:3:1 (Jansson 1994; Jackowski et al. 2001; Caffarri et al. 2004) and form several types of major LHCII homo- and hetero-trimers, whose high resolution structures have been solved by X-ray crystallography (Liu et al. 2004; Standfuss et al. 2005). They interact specifically with the PSII core via the minor LHCII monomeric antenna proteins Lhcb4, Lhcb5 and Lhcb6. So far, among these minor LHCII antenna proteins, an atomically resolved 3D structure is available only for Lhcb4 (Pan et al. 2011).

A variable number of LHCII proteins can associate with the dimeric PSII core complex to form different types of PSII-LHCII supercomplexes, named according to their composition (Dekker and Boekema 2005). In terms of nomenclature, the so-called  $C_2S_2$  supercomplex, whose 3D structure at the highest resolution has been obtained at 17 Å from spinach by single particle cryo-transmission electron microscopy (cryo-TEM) (Nield et al. 2000; Nield and Barber 2006), consists of a dimeric PSII core complex ( $C_2$ ) that binds to two LHCII trimers in a relatively strong way (i.e.,

S-trimers) via two copies of the monomeric Lhcb4 and Lhcb5 subunits. Larger PSII-LHCII supercomplexes, containing two extra copies of the monomeric Lhcb6 with one or two additional LHCII trimers, moderately bound to the dimeric PSII core complex (i.e., M-trimers) via Lhcb4 and Lhcb6, are known as  $C_2S_2M_{1-2}$  (Dekker and Boekema 2005). For the  $C_2S_2M_2$  supercomplex, 2D projection maps and a low resolution (30 Å) 3D model have been obtained by TEM and single particle analysis of negatively stained particles isolated respectively from *Arabidopsis* (Yakushevskaya et al. 2001; Caffarri et al. 2009) and pea plants (Pagliano et al. 2014). Occasionally, but only in membranes partially solubilized with detergents, larger supercomplexes have been observed by TEM, with one or two additional LHCII trimers (i.e., L-trimer) even more loosely bound to the dimeric PSII core complex, via Lhcb6, known as  $C_2S_2M_{0-2}L_{1-2}$  (Boekema et al. 1999a). By using TEM and image analysis on negatively stained BBY membranes isolated from spinach, only partially solubilized with the mild detergent n-dodecyl- $\alpha$ -D-maltoside ( $\alpha$ -DDM), Boekema and colleagues observed even larger complexes, termed megacomplexes, identified as dimeric associations of PSII-LHCII supercomplexes interacting with each other, within the membrane, in three different binding sites (Boekema et al. 1999a; Boekema et al. 1999b). The type I megacomplex ( $C_4S_4M_{2-4}$ ), with a square-like shape, and type II megacomplex ( $C_4S_4M_{2-4}$ ), with a rectangular-like shape, consist of two  $C_2S_2M$  supercomplexes arranged in an antiparallel fashion with a translational shift of about 7.5 nm for one of the  $C_2S_2M$  units (Boekema et al. 1999a); type III megacomplex ( $C_4S_4$ ), with a windmill-like shape, is made up of two  $C_2S_2$  supercomplexes connected at their Lhcb5 tips (Boekema et al. 1999b).

In this study, PSII-LHCII megacomplexes were isolated by a quick, single-step, solubilization using  $\alpha$ -DDM on stacked thylakoid membranes extracted from pea plants, and their structures have been analyzed by TEM and subsequent image analysis of negatively stained single particles at a resolution of  $\sim 18$  Å. This structural study was integrated with biochemical information, obtained by applying a gel-based proteomic approach to the isolated PSII-LHCII megacomplexes, to analyze their protein composition. Overall, these data indicate that screening for the native association of membrane protein complexes by single particle TEM and mass spectrometry is a valuable

approach to fully characterize large megacomplexes and their supramolecular organization within the granal membranes of higher plant chloroplasts.

## **Materials and methods**

### Isolation of PSII-LHCII megacomplexes

Pea plants were grown for three weeks under 8 h daylight at 20 °C, 60% humidity and 150  $\mu\text{mol m}^{-2} \text{s}^{-1}$  photons (conditions provided by the growth chamber SANYO MLR-351H with 15 Fluorescent lamps, of 40W each, switched on). Stacked thylakoid membranes were isolated from plants at the end of the daily dark phase of growth as described in detail earlier (Pagliano et al. 2012). Briefly, stacked thylakoid membranes, at a Chl concentration of 1 mg mL<sup>-1</sup>, were treated with 50 mM  $\alpha$ -DDM for 1 min at 4 °C in the dark according to Barera et al. (2012). Phenylmethylsulphonylfluoride (500  $\mu\text{M}$ ) was present during the solubilization to inhibit protease activity. After centrifugation, at 21,000  $g$  for 10 min at 4 °C, 700  $\mu\text{L}$  of supernatant was added to the top of a linear sucrose gradient, previously prepared by a freezing and thawing cycle applied to ultracentrifuge tubes filled with a buffer made of 0.65 M sucrose, 25 mM MES pH 5.7, 10 mM NaCl, 5 mM CaCl<sub>2</sub> and 0.03% (w/v)  $\alpha$ -DDM. Centrifugation was carried out at 100,000  $g$  for 12 h at 4 °C (Surespin 630 rotor, Thermo Scientific).

The sucrose band containing PSII-LHCII megacomplexes was carefully harvested using a syringe and, if necessary, concentrated by membrane filtration via Amicon Ultra 100 kDa cut-off devices (Millipore) and then flash frozen for storage at -80 °C.

### Spectroscopic analyses

The Chl concentration was determined spectrophotometrically after extraction in 80% (v/v) acetone according to (Arnon 1949). Absorption spectra in native conditions were recorded using a Lambda25 spectrophotometer (Perkin Elmer) at 12 °C. Low temperature (77 K) fluorescence

emission spectra were registered by a FL55 spectrofluorometer (Perkin Elmer), equipped with a red sensitive photomultiplier and a low temperature cuvette holder. Samples were excited at 436 nm. The spectral bandwidths were 7.5 nm (excitation) and 3 nm (emission). The Chl concentration was approximately  $0.5 \mu\text{g mL}^{-1}$  in 90% (v/v) glycerol and 10% (v/v) sucrose gradient buffer.

## Gel electrophoresis

Optimal separation of the thylakoid membrane protein complexes and PSII-LHCII megacomplexes was obtained as in Järvi et al. (2011), using large pore blue native polyacrylamide gel electrophoresis (IpBN-PAGE) on gels 20 cm x 16 cm, containing an acrylamide gradient of 3-12% (w/v) T and 3.33% (w/v) C in the resolving gel, and 3% (w/v) T and 20% (w/v) C in the stacking gel. Either solubilized pea thylakoids or the band from the sucrose density gradient were mixed to a final Chl concentration of  $300 \mu\text{g mL}^{-1}$  with BN-loading buffer (5% (w/v) Coomassie Blue G250, 750 mM  $\epsilon$ -amino caproic acid) at a volume ratio of 16:1. After incubation for 10 min on ice and a brief centrifugation at 20,000  $g$  for 10 min, the subsequent supernatants were loaded onto the gel. Electrophoresis was performed at a constant voltage of 70 V at 9 °C for 20 h. The anode buffer was made of 50 mM Bis-Tris-HCl pH 7.0, the cathode buffer of 50 mM Tricine, 15 mM Bis-Tris-HCl pH 7.0, and 0.02% (w/v) Coomassie G250; the latter being replaced, after two-thirds of the run, by a cathode buffer with the same composition but devoid of Coomassie G250. For molecular mass markers, a mixture of lyophilized standard proteins (Amersham, high molecular weight calibration kit code 17-0445-01, GE Healthcare) and blue dextran (code D575, Sigma-Aldrich) was used.

For two-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (2D SDS-PAGE), lanes from the IpBN-PAGE were excised and denatured in 66 mM  $\text{Na}_2\text{CO}_3$ , 0.66% (v/v)  $\beta$ -mercaptoethanol and 2% (w/v) SDS at room temperature for 40 min. After denaturation, proteins were resolved by SDS-PAGE according to the Laemmli's system (Laemmli 1970) on a 15% (w/v) polyacrylamide gel containing 6 M urea. Electrophoresis was carried out at a constant amperage of 10 mA for 17 h at 9 °C. Gels were stained using a mass spectrometry compatible silver staining protocol as described in Shevchenko et al. (1996).

2D SDS-PAGE and/or mono-dimensional SDS-PAGE, the latter being performed on a 12.5% (w/v) polyacrylamide gel containing 5 M urea using Laemmli's system (Laemmli 1970), were transferred onto nitro-cellulose membranes and further immunodetected with specific antisera against PsbO, PsbP, PsbQ and PsbR polypeptides (Agrisera codes AS05092, AS06167, AS06142-16, AS05059 respectively), by using the alkaline phosphatase conjugate method, with 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium as chromogenic substrates (Sigma-Aldrich).

#### Mass spectrometry and chromatographic conditions

For liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis, spots from the 2D SDS-PAGE were cut out and proteins digested in-gel with trypsin (code V5111, Promega), as described in Hellmann et al. (1995).

LC-MS/MS analyses were performed by a micro-LC Eksigent Technologies (Dublin, USA) system that included a micro LC200 Eksigent pump with flow module 5-50  $\mu\text{L}$  and a programmable autosampler CTC PAL with a Peltier unit (1.0-45.0  $^{\circ}\text{C}$ ). The stationary phase was a Halo Fused C18 column (0.5 x 100 mm, 2.7  $\mu\text{m}$ ; Eksigent Technologies Dublin, USA). The mobile phase was a mixture of 0.1% (v/v) formic acid in water (A) and 0.1% (v/v) formic acid in acetonitrile (B), eluting at a flow-rate of 15.0  $\text{mL min}^{-1}$  and at an increasing concentration of solvent B from 2% to 40% in 30 minutes. The injection volume was 4.0  $\mu\text{L}$ . The oven temperature was set at 40  $^{\circ}\text{C}$ .

The LC system was interfaced with a 5600<sup>+</sup> TripleTOF™ system (AB Sciex, Concord, Canada) equipped with DuoSpray™ Ion Source and CDS (Calibrant Delivery System). The mass spectrometer worked in information dependent acquisition (IDA) mode. Peptide profiling was performed using a mass range of 100-1600 Da (TOF scan with an accumulation time of 100.0 ms), followed by a MS/MS product ion scan from 200 to 1250 Da (accumulation time of 5.0 ms) with the abundance threshold set at 30 cps (35 candidate ions can be monitored per cycle). The ion source parameters in electrospray positive mode were set as follows: curtain gas ( $\text{N}_2$ ) at 25 psig, nebulizer gas GAS1 at 25 psig, and GAS2 at 20 psig, ionspray floating voltage (ISFV) at 5000 V, source temperature at 450  $^{\circ}\text{C}$  and declustering potential at 25 V.



MS data were acquired with Analyst TF 1.7 (AB SCIEX, Concord, Canada). Subsequently, PeakView version 2.1 (AB SCIEX, Concord, Canada) was used to generate Mascot (.mgf) files with peak lists from the Analyst software files (.wiff); the default parameters were used (<http://www.matrixscience.com>).

The MS/MS spectra obtained by digested samples were analyzed as Mascot generic files against all entries in the public NCBI nr database using the online Mascot server (Matrix Science: <http://www.matrixscience.com/>) without a taxonomy filter.

The principal parameter settings for the Mascot search were as follows: database NCBI nr (version 2015.04.25; containing 65519838 sequence entries); enzyme trypsin; allow up to one missed cleavage; possible variable modifications carbamidomethylation of cysteine (C), oxidation of methionine (M), deamidation of asparagine and glutamine (NQ); precursor ion mass and fragment masses tolerance of 60 ppm and 0.3 Da, respectively; instrument default; default charge state set to 2+, 3+, and 4+. Widely accepted positive identifications of proteins by MS/MS analysis require a minimum of two unique peptides with at least one peptide having a significant ion score ( $p \leq 0.05$ ). Considering that the genome of *Pisum sativum* is not fully sequenced, and that only some protein sequences of *Pisum sativum* are present in the database, we also accepted hits identified by at least one peptide with a significant ion score according to the MASCOT MS/MS ion search algorithm as being confident assignments.

#### Transmission electron microscopy and single particle image analysis

PSII-LHCII mega-complexes from the analyzed sucrose gradient band, immediately after isolation, were applied to glow discharged carbon-coated copper grids, washed with distilled water and negatively stained with 2% (w/v) uranyl acetate. After checking the reproducibility of the preparation by preliminary TEM analyses on five independent isolated samples, two datasets of images were collected from two individual samples chosen as representative of the preparation, and then subsequently merged. For acquisition of micrographs, a FEI Tecnai F20-ST transmission electron microscope, equipped with a field emission gun (FEG) operated at 200 keV, was used. A

total of 339 images were recorded at 29,000x on a Gatan Orius 4.0K x 2.7K CCD camera, leading to a final pixel size of 3.4 Å at the specimen level. Imaging conditions were optimised, in terms of defocus and astigmatism, to ensure the first minima of the Fourier power spectrum was within 15 Å and an underfocus range of 0.5 to 1.5 µm present. Particles were floated out into boxes using the boxer module of EMAN2 (Tang et al. 2007). All latter image processing steps were performed within the IMAGIC-V software environment (Imagic Science GmbH, Berlin, Germany) (van Heel et al. 1996), including initial reference-free alignment, multi-reference alignments, multivariate statistical analyses and classification, Fourier ring correlation estimation, and as further described in van Heel et al. (2000). Relevant crystallographic co-ordinate atom and chain data (PDB IDs: 3ARC, 2BHW, 3PL9) were overlaid onto characteristic views derived from the single particle sub-populations identified, using UCSF Chimera (Pettersen et al. 2004) modelling software.

## **Results and Discussion**

### **Isolation of PSII-LHCII megacomplexes**

In previous work (Pagliano et al. 2012), conditions for the direct solubilisation of stacked thylakoid membranes isolated from pea leaves were optimized in order to solubilize high molecular mass PSII-LHCII complexes. In so doing, it was found that after treating isolated stacked pea thylakoids, at a Chl concentration of 1 mg mL<sup>-1</sup>, with α-DDM concentrations ranging between 50 and 100 mM for 1 min, an almost complete solubilisation of the entire membrane compartments could be achieved and well defined high molecular mass PSII-LHCII complexes solubilized with high reproducibility (Pagliano et al. 2012). The amount and molecular mass of solubilized PSII-LHCII complexes did not change even when longer times of solubilisation (10 and 30 min) were tested (data not shown), attesting that the duration of the solubilisation, at least in these experimental conditions, does not lead to either aggregation or dissociation of PSII-LHCII complexes. By solubilizing stacked pea thylakoid membranes at a Chl concentration of 1 mg mL<sup>-1</sup>, for the shortest time tested (1 min) and with the lowest concentration of α-DDM (50 mM) effective in solubilizing

high molecular mass PSII-LHCII complexes, followed by a relatively quick (12 h) ultracentrifugation step on a sucrose density gradient, we isolated homogenous preparations of PSII-LHCII supercomplexes, subsequently used for structural studies (Barera et al. 2012; Pagliano et al. 2014). In detail, in the gradient five different green bands, termed  $\alpha 1$  to  $\alpha 5$ , were separated: band  $\alpha 1$  contained mostly LHCII trimers; band  $\alpha 2$  was composed of PSI-LHCI complexes co-migrating with ATP-ase; and bands  $\alpha 3$  to  $\alpha 5$  were observed to be composed predominantly of PSII-LHCII particles (Fig. 1). Band  $\alpha 3$  has been characterized previously (Barera et al. 2012) and corresponds to PSII-LHCII supercomplexes of type  $C_2S_2M_2$ . This latter preparation was also used as starting material for single particle electron microscopy studies that led to a three dimensional model of the supercomplex of type  $C_2S_2M_2$  from pea plants (Pagliano et al. 2014). The lower, highly reproducible bands of the gradient,  $\alpha 4$  and  $\alpha 5$ , given their higher apparent molecular mass, were expected to contain either PSII-LHCII supercomplexes of increasing size, or even larger forms of PSII-LHCII, termed megacomplexes. In the work presented here, therefore, a detailed biochemical and structural characterization of band  $\alpha 5$  was performed, this being the heaviest and most abundant of the two bands observed.

### **Spectroscopic characterization of PSII-LHCII megacomplexes**

The absorption spectrum of band  $\alpha 5$  is presented in Fig. 2a. The  $\alpha 5$  particles registered absorption peaks at 436, 469 and 677 nm, similar to values previously observed for PSII-LHCII supercomplexes of type  $C_2S_2M_2$  present in band  $\alpha 3$  of the sucrose gradient (see Barera et al. 2012). The relatively high intensity of the absorption in the Chl *b* region (470 and 650 nm) of band  $\alpha 5$ , together with its Chl *a/b* ratio between 2.4-2.6, is consistent with studies conducted on forms of PSII isolated directly from thylakoid membranes that are larger than the basic  $C_2S_2$  unit and enriched in the Lhcb6 antenna subunit (Eshaghi et al. 1999; Barera et al. 2012). It should be noted that the 2.4-2.6 ratio is lower than values typically observed for PSII-LHCII supercomplexes of type  $C_2S_2$  (Eshaghi et al. 1999; Barera et al. 2012).

The low temperature emission spectrum of the  $\alpha 5$  particles is shown in Fig. 2b. It reveals a single fluorescence emission peak at 683 nm, an intermediate value compared to that of PSII-LHCII supercomplexes of type  $C_2S_2$  (682 nm) and  $C_2S_2M_2$  (684 nm) (see Barera et al. 2012). Performing fluorescence studies at 77 K can give valuable information on the structural organization of the photosynthetic apparatus (van Grondelle et al. 1994). Detailed 77 K fluorescence studies have shown that PSII has two main emission bands at 685 and 695 nm, corresponding to two different Chl-protein complexes of PSII: the RC and core antenna complexes, respectively (Dekker et al. 1995). An additional peak at 680 nm has been ascribed to the LHCII originating from the red-most Chl *a* pigment in LHCII, absorbing at 676 nm (Hemelrijk et al. 1992). For the PSII-LHCII particles of band  $\alpha 5$  (see Fig. 2b), the main peak at 683 nm is red-shifted compared to the expected characteristics of the LHCII trimers alone, confirming that there is energy transfer between LHCII and the PSII core complex. Notably, there is no 77 K fluorescence peak at 735 nm, confirming the absence of any PSI contamination in the  $\alpha 5$  particles (Gobets and van Grondelle 2001).

### **Protein composition of PSII-LHCII megacomplexes**

In order to investigate the oligomerization state of the isolated high molecular mass PSII-LHCII containing particles present in band  $\alpha 5$ , and search for any counterpart within the original thylakoid membranes, large-pore blue native PAGE (lpBN-PAGE) was performed either on isolated particles or on thylakoids. In Fig. 3a the lpBN-PAGE profile of the band  $\alpha 5$  is shown and compared to that of thylakoid membranes solubilized with  $\alpha$ -DDM under the same conditions used for band  $\alpha 5$  isolation. In these lpBN-PAGE profiles, the high molecular mass region showed: 1) a variable number of bands with molecular masses falling within the region 669 to 2,000 kDa, with three such bands detectable only after Coomassie staining in the band  $\alpha 5$  and, intriguingly, four such bands being predominant in solubilized thylakoids; 2) three green bands with molecular masses above 2,000 kDa, predominant in the  $\alpha 5$  sample, and also detectable in solubilized thylakoids. The nature of these high molecular mass PSII-LHCII containing complexes was probed by performing the denaturing 2D SDS-PAGE of the native-PAGE lanes of the thylakoid membranes and of band  $\alpha 5$ ,

whose profiles, after silver staining, are shown in Fig. 3b and supplementary Fig. S1a, respectively. In Fig. 3c an enlargement of the region above 2,000 kDa of the denaturing 2D SDS-PAGE of the band  $\alpha 5$  of Fig. S1a is shown. The denatured second dimension profile of the IpBN-PAGE of solubilized thylakoids, interpreted according to previous reports (Aro et al. 2005; Granvogl et al. 2006; Pagliano et al. 2012; Pagliano et al. 2014), revealed that: 1) the four bands with molecular masses between 669 and 2,000 kDa were indeed PSII-LHCII supercomplexes of different sizes, one with a lower mass containing only LHCII S trimers; the further three, with increasing molecular masses, containing additional LHCII M trimers; 2) the bands with molecular masses above 2,000 kDa contained PSII-LHCII megacomplexes, given a protein composition similar to that displayed by PSII-LHCII supercomplexes, but a molecular mass over twice that of the supercomplex with smaller size. Hence, by comparing the denaturing second dimension profile of the IpBN-PAGE of band  $\alpha 5$  (Fig. S1a and 3c) to that of the solubilized thylakoid membranes (Fig. 3b), it was clear the analogous nature of PSII-LHCII supercomplexes for the faint bands lying between 669 and 2,000 kDa and that of PSII-LHCII megacomplexes for the predominant bands, in the region  $> 2,000$  kDa (hereafter referred to as  $\alpha 5$  PSII-LHCII mc region). The low amount of PSII-LHCII supercomplexes with respect to the predominant content of PSII-LHCII megacomplexes found in band  $\alpha 5$  might be taken as an indication of the relatively high molecular stability of the isolated PSII-LHCII megacomplexes. Moreover, the detection of their counterparts within the solubilized thylakoid membranes would suggest that the isolation step by sucrose gradient ultracentrifugation maintains the genuine nature of megacomplexes of these high molecular mass PSII-LHCII complexes.

After in-gel trypsin digestion of all the spots stained on the second dimension of the  $\alpha 5$  PSII-LHCII mc region, microLC-ESI-MS/MS analysis of the digested peptides revealed the presence of the RC core subunits D2 and D1, the inner antenna proteins CP47 and CP43, the six Lhcb outer antenna proteins (Lhcb1-6), and the OEC subunit PsbO (see supplementary Table S1), thus allowing their positions on the 2D SDS-PAGE map to be identified (Fig. 3c). The detection of the Lhcb3 and Lhb6 antennae subunits, the former being an exclusive component of LHCII M trimers (Boekema et al. 1999a), and the latter a linker for the LHCII M trimer to the  $C_2S_2$  supercomplex (Barera et al. 2012), confirmed the occurrence of LHCII M trimers in the preparation. In particular, the proteomic

analysis proved the existence of at least three distinguishable sub-populations of PSII-LHCII mega-complexes in the preparation, visible in the  $\alpha 5$  PSII-LHCII mc region (Fig. 3c), with different sizes and containing LHCII M trimers.

For the particles of band  $\alpha 5$ , western blot analyses for the OEC polypeptides, based on identical Chl loading (Fig. 4), showed the presence of a high quantity of PsbO, a low amount of PsbP, and the absence of PsbQ and PsbR. In band  $\alpha 5$  the presence of the subunit PsbO bound to the PSII-LHCII mega-complexes particles was confirmed by MS/MS analysis (see Table S1 and Fig. 3c), and that of PsbP by western blot analysis performed on the second dimension of the entire IpBN-PAGE lane of band  $\alpha 5$  (Fig. S1b, see the low amount in the PSII-LHCII mc region between 20-25 kDa).

### **Determination of the supramolecular structure of PSII-LHCII mega-complexes**

PSII-LHCII mega-complexes, immediately after isolation from band  $\alpha 5$  by sucrose density gradient ultracentrifugation, were adsorbed onto carbon film, washed with distilled water, and negatively stained using uranyl acetate. The reproducibility of the preparation was checked by preliminary TEM imaging upon five independent isolated samples. In Fig. 5 a micrograph, chosen as representative of the preparation, shows the distribution on the carbon support of the particles present in band  $\alpha 5$ . PSII-LHCII mega-complexes in top, side and end view orientations are clearly discernible; in addition many large particles are present in tilted view, most likely PSII-LHCII mega-complexes, with a low amount of  $C_2S_2$  PSII-LHCII supercomplexes in top view orientation and some small particles, possibly detached LHCII.

To compile the single particle image dataset, TEM images from two independent samples were recorded at a sampling frequency of 3.4 Å per pixel using a consistent  $\sim 1 \mu\text{m}$  underfocus, equating to the first minimum of the contrast transfer function being  $\sim 15 \text{ Å}$ , and then merged. All possible single particles, i.e. intact or degraded, were selected, resulting in a total of 8,963 available 2D projections of varying orientation. The dataset was processed by reference-free alignment and multivariate statistical analysis. Over five iterations of refinement,  $\sim 5\%$  of particles were set

inactive at each stage to ameliorate against the heterogeneous nature of the raw data, resulting in 1,792 particles being excluded and 600 class averages being requested from 7,171 particles (see supplementary Fig. S2). The averages displayed a greatly improved signal-to-noise ratio over the raw dataset of particle images and subjectively represented the most relevant spread of variation for observing characteristic views.

About one third of the entire dataset consisted of particles ascribed to top, side and end view orientations, the rest being considered tilted views of intermediate orientation (see Table 1). Among the particles seen in top view, ~14% were attributed to single, non-sandwich, PSII-LHCII supercomplexes of type  $C_2S_2$ ,  $C_2S_2M$  and  $C_2S_2M_2$  (Table 1). No single supercomplexes were observed in side view orientation, but the possibility does remain that further non-sandwich supercomplex projections may exist within the averages of other tilted orientations observed in the total classification (see Fig. S2). However, such a minute percentage is in agreement with the low quantity of PSII-LHCII supercomplexes revealed in band  $\alpha 5$  by electrophoretic methods (see Fig. 3a and Fig. S1a). Roughly 86% of the particles seen in top, side and end view were ascribable to PSII-LHCII megacomplexes (see Table 1), these resulting the predominant complexes within band  $\alpha 5$ . This would suggest that a strong supramolecular interaction exists between PSII-LHCII supercomplexes when they constitute larger complexes. Among these, the most intact and striking views were attributed to top, side and end views of the PSII-LHCII megacomplexes shown in Fig. 6a, specifically those classes numbered 1 and 4 for top views, class 2 for a side view, and class 3 for an end view. All averages shown are ~18 Å in resolution as estimated by Fourier ring correlation. The number of raw particles for these classes is shown in Table 1 and their domain arrangement is proposed with labels in white on Fig. 6a (according to Dekker and Boekema 2005), where C represents a PSII core monomer, S the strongly bound LHCII trimer and M the moderately bound LHCII trimer. No obvious views were observed that would suggest the presence of LHCII L trimers loosely bound to the PSII core.

A novel arrangement of four  $C_2S_2$  supercomplexes was observed, where two were bound side-by-side in the plane of the membrane and then further sandwiched together with a second copy; these high molecular mass complexes, with an estimated molecular mass of at least 4.4 MDa, were

termed  $(C_2S_2)_4$  megacomplexes. Such a macromolecular form was immediately obvious in both top and end view orientations (locations 1 and 3, Fig. 6a), in total representing 23.4% of the 1,883 particles, unequivocally attributable to megacomplexes present in the dataset (see Table 1). The percentage of this form of megacomplex may increase to ~70% if the side views of location 2 are considered exclusively as  $(C_2S_2)_4$  sandwich particles, but the presence of  $(C_2S_2)_2$  sandwich particles within these side view averages may not be excluded, since these would not be discernible in this orientation. The overall top view dimensions of the megacomplex were, at most, 410 Å in width (x) and 350 Å in length (y), with an internal corner-to-corner diagonal of 450 Å. Any height (z) of the side views must be inclusive of the lumenally-bound PSII subunits known to optimize the water-splitting environment, identified in this study as PsbO and PsbP (see Fig. 3c, Fig. 4, Fig. S1b and Table S1). These luminal protrusions are clearly present, extending out from the membrane domain of the side view and end view of locations 2 and 3 (Fig. 6a). The height (z) of the side view sandwich is 290 Å and, within it, each single side view, therein, being 145 Å (z), of which 60 to 65 Å may be attributed to the extrinsic lumenally-protruding subunits. Taken as a whole, these views of the new megacomplex form show that it is a sandwich, the top half viewed from the luminal surface sitting on top of an identical copy, viewed from underneath. The result is that the internal “diagonal” appearance of the single dimeric PSII core, typically being top left to bottom right when observed in top view (see Barera et al. 2012), is obscured to form a square, multi-domain arrangement when viewed in 2D projection from the luminal surface of a sandwich megacomplex (locations 1 and 4). To aid the reader further, four copies of the 3D map of a  $C_2S_2$  supercomplex, reported previously (Pagliano et al. 2014), were imported into a UCSF Chimera modelling environment (Fig. 6b) and overlaid onto locations 2 and 3. The color of these maps was chosen arbitrarily, but the two side-by-side  $C_2S_2$  supercomplex 3D maps of the top half of the sandwich are colored purple and cyan using a see-through mesh render, whereas the other half of the megacomplex sandwich, two side-by-side  $C_2S_2$  supercomplex 3D maps viewed from underneath, are orange and magenta solid surface representations. In addition to the central lumenally protruding domain, attributable to the PsbO subunit, the 3D maps (Fig. 6b) contain a second protruding luminal domain off the edge of each PSII monomer. Caution was urged



regarding the interpretation of the 3D maps presented in Pagliano et al. (2014), due to the relative lack of tilted views present within the negatively stained data used, however, one could speculate that this second protruding luminal domain might be a candidate for the position of PsbP, in accordance also with a similar positioning of this subunit within the PSII-LHCII supercomplex by Ido et al. (2014). The use of Protein Databank (.pdb) files for modeling purposes with the program UCSF Chimera was extended to include all-atom spheres for each PSII and LHCII subunit, assigned as in Pagliano et al. (2014) (Fig. 6b). The presence of sandwiched particles necessarily increases the complexity of any attempted domain assignment process for further novel characteristic views classified, but such an exercise aids in understanding the majority of views observed. The other most prevalent class of top view intact particles is shown in location 4 (Fig. 6a), representing 17% of the 1,883 particles univocally attributable to megacomplexes present in the dataset (see Table 1). From many models attempted (data not shown), it was concluded that location 4 represents a top view of a sandwiched megacomplex that must include two moderately bound LHCII M trimers within each half of the megacomplex sandwich, in addition to two  $C_2S_2$  units, thus forming a  $(C_2S_2 + C_2S_2M_2)_2$  megacomplex. Sandwich-form megacomplexes in three-dimensions, once projected down into two-dimensional views, will result in a complex overlapping arrangement of atoms. An expanded interpretation of the sandwich-form megacomplexes of locations 1 and 4 is offered in Fig. 6b, the former shown as a lumenally viewed  $(C_2S_2 + C_2S_2)$  megacomplex merged atop a stromally viewed  $(C_2S_2 + C_2S_2)$  megacomplex, and the latter as a lumenally viewed  $(C_2S_2 + C_2S_2M_2)$  megacomplex merged atop a stromally viewed  $(C_2S_2M_2 + C_2S_2)$  megacomplex. The 3D maps and schematic representations used in Fig. 6b aid in visualizing the interactions between the stromal surfaces of PSII-LHCII supercomplexes in the two halves of the sandwich, forming the stromal gap. Contacts between PSII-LHCII supercomplexes in adjacent membranes are mediated by the stromal flat surfaces of both the LHCS and the PSII reaction center dimers. The LHCII trimers at either end of one supercomplex sit atop the LHCII trimers of the supercomplex in the opposite membrane. The interaction of LHCII trimers with one another in grana thylakoids, like molecular Velcro, is supported either by the charge complementarity shown in the atomic structure of the trimeric LHCII proposed by Standfuss et al. (2005) or by the direct

tomographic observations in pea chloroplast membranes (Daum et al. 2010). The overall geometry of the interaction between the stromal surfaces of the PSII-LHCII supercomplexes in two adjacent membranes is similar to that observed in isolated supercomplex sandwich dimers by single-particle cryo-TEM (Nield et al. 2000), although different for a rotational offset of  $\sim 90^\circ$  from that displayed by PSII-LHCII supercomplexes that form ordered 2D arrays in the grana observed by tomography (Daum et al. 2010). It can not be excluded that the similarity with other isolated supercomplex sandwiches is related to the unavoidable use of detergents during the purification step. Nevertheless, two different detergents,  $\alpha$ -DDM (in this study) and  $\beta$ -DDM (Nield et al. 2000), have been used for comparable structural studies (Nield et al. 2000; Barera et al. 2012; Pagliano et al. 2014), the former considered milder and more effective than the latter in solubilizing larger PSII-LHCII complexes from stacked thylakoid membranes (see Pagliano et al. 2012). The isolated megacomplexes of this study derive from a complete solubilization of the thylakoid membranes of pea plants grown at moderate light (i.e.,  $150 \mu\text{mol m}^{-2} \text{s}^{-1}$  photons). Therefore, they can likely be considered possible representatives of the supramolecular arrangements of PSII-LHCII supercomplexes within thylakoids of plants acclimated to this light condition, and not exclusively within local patches of semi-crystalline 2D arrays that occur occasionally in grana membranes, at higher frequency in low light treated plants (Daum et al. 2010; Kouřil et al. 2013). This observation, together with the fact that the isolated megacomplexes displayed additional antenna proteins (i.e., LHCII M trimers present at least in one class of isolated megacomplexes, see Fig. 6 location 4) bound to the basic  $\text{C}_2\text{S}_2$  units, might explain the different geometry present at the conjoint stromal surfaces for PSII-LHCII supercomplexes in two adjacent membranes, differently to that observed in tomograms of 2D arrays of  $\text{C}_2\text{S}_2$  PSII-LHCII supercomplexes by Daum et al. (2010).

During the preparation of the TEM grid, the detachment of some LHCII (i.e., monomeric LHCII, and trimeric LHCII of type M and even S) from the band  $\alpha 5$  megacomplexes may not be excluded, since an amount of small particulate matter remains visible in the background of the TEM images (see Fig. 5) and the occurrence of megacomplexes with broken symmetry, likely of type  $\text{C}_2\text{S} + (\text{C}_2\text{S}_2)_2$ , was observed during the classification (see Table 1 and Fig. S2). The detachment of some LHCII from the megacomplexes might be partially induced by the

experimental technique of washing the samples on the grid with water, prior to the negative stain fixation. This technique is well known to improve the imaging by lowering the stain background, through the dilution of the high sucrose concentration present within sucrose gradients bands (>1 M in the isolated band  $\alpha 5$ ). Notwithstanding the low amount of detached LHCII and supercomplexes observed, the majority of the particles in this preparation were stable PSII-LHCII megacomplexes, in a sandwich form, characterized by two new types of symmetrical organization.

### **Concluding remarks**

A direct method for the isolation of pure PSII-LHCII megacomplexes has been demonstrated, consisting of a single rapid (1 min) solubilization of stacked pea thylakoid membranes with  $\alpha$ -DDM at experimental conditions that effectively solubilize entire thylakoids, followed by a relatively quick (12 h) ultracentrifugation step using continuous sucrose density gradients. Spectroscopic and proteomic analyses revealed that the heaviest band of the sucrose gradient contained pure PSII-LHCII megacomplexes. TEM analyses on negatively stained samples, and classification of the single particle views observed therein, revealed the presence of megacomplexes having a novel structure compared to that reported previously (Boekema et al 1999a; Boekema et al 1999b); whereby two PSII-LHCII supercomplexes sit side-by-side in the membrane plane, interacting with a second inverted copy of a similar megacomplex forming a sandwich organization. Two predominant forms of intact sandwiched megacomplexes were observed, a  $(C_2S_2)_4$  megacomplex and a  $(C_2S_2 + C_2S_2M_2)_2$  megacomplex. The  $(C_2S_2)_4$  megacomplex displayed a central arrangement of two  $C_2S_2$  supercomplexes, interacting with an intra-membrane-plane twofold symmetry; the organization of the second copy of two  $C_2S_2$  supercomplexes in the other appressed membrane is related to the first layer by a  $180^\circ$  rotation about an axis at a position half way between the two membranes. The  $(C_2S_2 + C_2S_2M_2)_2$  megacomplex differs from the  $(C_2S_2)_4$  megacomplex by having two additional moderately bound LHCII M trimers present within each half of the megacomplex sandwich, one peripheral and the other in a central position, making a bridge between the two supercomplexes in the same membrane plane. The occurrence of LHCII M trimers in this preparation was further confirmed by proteomic analyses.

For a full understanding of plant photosynthesis, it is necessary to establish how the PSII-LHCII supercomplexes are organized and interact with one another in the highly dynamic thylakoid membranes. These new structural forms of isolated megacomplexes, and the modelling proposed, provide intriguing insights into how PSII-LHCII supercomplexes may bind to each other, in register, both in the membrane-plane and between granal stacks within the chloroplast. We have provided evidence of different geometries for the interaction between the stromal surfaces of PSII-LHCII supercomplexes across two adjacent membranes, in addition to those proposed by Daum et al. (2010) for the case of thylakoid patches containing crystalline 2D arrays of C<sub>2</sub>S<sub>2</sub> supercomplexes. Noteworthy, however, is the fact that, despite the different starting material, either plant (i.e., spinach or pea) or preparation (i.e., native chloroplasts/thylakoids or isolated complexes; chloroplasts/thylakoids exposed to light or megacomplexes isolated in the dark during the purification step), the direct interaction between LHCII trimers of the supercomplexes in the two adjacent membranes is maintained. Therefore, it seems reasonable to conclude that the transmembrane interactions, holding two PSII-LHCII half-sandwiches together in the isolated megacomplexes, are very similar to those that govern granal stacking in PSII-LHCII 2D arrays in native chloroplasts, and that the electrostatic interactions between LHCII trimers are a key factor in this stacking.

In summary, the overall data collected on these isolated sandwiched megacomplexes will improve our understanding of the arrangement of PSII and its LHCII antenna system in chloroplast membranes of higher plants, complementing the structural work performed by cryo-electron tomography on native thylakoid membranes and the sectioning of intact chloroplasts (Daum et al. 2010).

### **Acknowledgements**

This work was supported by the Italian Ministry of Education, University and Research, “Futuro in Ricerca 2013” program RBFR1334SB to CP. We kindly thank Dr Alessandro Di Blasio (Politecnico di Torino, Italy) for helping in the glow discharge treatment of grids used for TEM analysis.

## References

- Albertsson PA (2001) A quantitative model of the domain structure of the photosynthetic membrane. *Trends Plant Sci* 6:349–354.
- Amunts A, Nelson N (2009) Plant photosystem I design in the light of evolution. *Structure* 17:37–650.
- Andersson B, Anderson JM (1980) Lateral heterogeneity in the distribution of chlorophyll-protein complexes of the thylakoid membranes of spinach. *Biochim Biophys Acta* 593:427–440.
- Arnon DJ (1949) Copper enzymes in isolated chloroplasts, polyphenoloxidase in *Beta vulgaris*. *Plant Physiol* 24:1–14.
- Aro EM, Suorsa M, Rokka A, Allahverdiyeva Y, Paakkarinen V, Saleem A, Battchikova N, Rintamaki E (2005) Dynamics of photosystem II: a proteomic approach to thylakoid protein complexes. *J Exp Bot* 56:347–356.
- Balsera M, Arellano JB, Revuelta JL, De Las Rivas J, Hermoso JA (2005) The 1.49 Å resolution crystal structure of PsbQ from photosystem II of *Spinacia oleracea* reveals a PPII structure in the N-terminal region. *J Mol Biol* 350:1051–1060.
- Barber J (1982) Influence of surface charges on thylakoid structure and function. *Ann Rev Plant Physiol* 33:261–295.
- Barera S, Pagliano C, Pape T, Saracco G, Barber J (2012) Characterization of PSII-LHCII supercomplexes isolated from pea thylakoid membrane by one-step treatment with  $\alpha$ - and  $\beta$ -dodecyl-D-maltoside. *Philos Trans R Soc Lond B Biol Sci* 367:3389–3399.

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 JF, 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.

Caffarri S, Croce R, Cattivelli L, Bassi R (2004) A look within LHCII: differential analysis of the Lhcb1-3 complexes building the major trimeric antenna complex of higher-plant photosynthesis. *Biochemistry* 43:9467–9476.

Caffarri S, Kouřil R, Kerešič S, Boekema EJ, Croce R (2009) Functional architecture of higher plant photosystem II supercomplexes. *EMBO J* 28:3052–3063.

Calderone V, Trabucco M, Vujčić A, Battistutta R, Giacometti GM, Andreucci F, Barbato R, Zanotti G (2003) Crystal structure of the PsbQ protein of photosystem II from higher plants. *EMBO Rep* 4:900–905.

Daum B, Nicastro D, Austin JII, McIntosh R, Kühlbrandt W (2010) Arrangement of photosystem II and ATP synthase in chloroplast membranes of spinach and pea. *Plant Cell* 22:1299–1312.

Dekker JP, Hassoldt A, Pettersson A, van Roon H, Groot ML, van Grondelle R (1995) On the nature of the F695 and F685 emission of Photosystem II. In: Mathis P (eds) *Photosynthesis: from light to biosphere*. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 53–56.

Dekker JP, Boekema EJ (2005) Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim Biophys Acta* 1706:12–39.

Eshaghi S, Andersson B, Barber J (1999) Isolation of a highly active PSII-LHCII supercomplex from thylakoid membranes by a direct method. *FEBS Lett* 446:23–26.

Gobets B, van Grondelle R (2001) Energy transfer and trapping in photosystem I. *Biochim Biophys Acta* 1507:80–99.

Granvogl B, Reisinger V, Eichacker LA (2006) Mapping the proteome of thylakoid membranes by de novo sequencing of intermembrane peptide domains. *Proteomics* 6:3681–3695.

Groth G, Pohl E (2001) The structure of the chloroplast F1-ATPase at 3.2 Å resolution. *J Biol Chem* 276:1345–1352.

Hankamer B, Barber J, Boekema E (1997) Structure and membrane organization of photosystem II in green plants. *Ann Rev Plant Physiol Plant Mol Biol* 48:641–671.

Hankamer B, Morris EP, Nield J, Gerle C, Barber J (2001) Three dimensional structure of the photosystem II core dimer of higher plants determined by electron microscopy. *J Struct Biol* 135:262–269.

Hellmann U, Wernstedt C, Genez J, Heldin CH (1995) Improvement of an “In-Gel” digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. *Anal Biochem* 224:451–455.

Hemelrijk PW, Kwa SLS, van Grondelle R, Dekker JP (1992) Spectroscopic properties of LHCII, the main light-harvesting chlorophyll a/b protein complex from chloroplast membranes. *Biochim Biophys Acta* 1098:159–166.

Ido K, Nield J, Fukao Y, Nishimura T, Sato F, Ifuku K (2014) Cross-linking evidence for multiple interactions of the PsbP and PsbQ proteins in a higher plant photosystem II supercomplex. *J Biol Chem* 289:20150–20157.

Ifuku K, Nakatsu T, Kato H, Sato F (2004) Crystal structure of the PsbP protein of photosystem II from *Nicotiana tabacum*. *EMBO Rep* 5:362–367.

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 (LHCII). *Biochim Biophys Acta* 1504:340–345.

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

Järvi S, Suorsa M, Paakkarinen V, Aro EM (2011) Optimized native gel systems for separation of thylakoid protein complexes: novel super- and mega-complexes. *Biochem J* 439:207–214.

Kirchhoff H, Hall C, Wood M, Herbstová M, Tsabari O, Nevo R, Charuvi D, Shimoni E, Reich Z (2011) Dynamic control of protein diffusion within the granal thylakoid lumen. *Proc Natl Acad Sci USA* 108:20248–20253.

Kouřil R, Wientjes E, Bultema JB, Croce R, Boekema EJ (2013) High-light vs. low-light: effect of light acclimation on photosystem II composition and organization in *Arabidopsis thaliana*. *Biochim Biophys Acta* 1827:411–419.

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685.



Liu ZF, Yan HC, Wang KB, Kuang TY, Zhang JP, Gui LL, An XM, Chang WR (2004) Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature* 428:287–292.

Mullet JE (1983) The amino acid sequence of the polypeptide segment which regulates membrane adhesion (grana stacking) in chloroplasts. *J Biol Chem* 258:9941–9948.

Nanba O, Satoh K (1987) Isolation of a photosystem II reaction center consisting of D-1 and D-2 polypeptides and cytochrome b-559. *Proc Natl Acad Sci USA* 84:109–112.

Nevo R, Chuartzman SG, Tsabari O, Reich Z (2009) Architecture and plasticity of thylakoid membrane networks. In: Wada H, Murata N (eds) *Lipids in photosynthesis*, Springer, Berlin, Germany, pp 295–328.

Nield J, Orlova EV, Morris EP, Gowen B, van Heel M, Barber J (2000) 3D map of the plant photosystem two supercomplex obtained by cryoelectron microscopy and single particle analysis. *Nat Struct Biol* 7:44–47.

Nield J, Barber J (2006) Refinement of the structural model for the photosystem II supercomplex of higher plants. *Biochim Biophys Acta* 1757:353–361.

Pagliano C, Chimirri F, Saracco G, Marsano F, Barber J (2011) One-step isolation and biochemical characterization of a highly active plant PSII monomeric core. *Photosynth Res* 108:33–46.

Pagliano C, Barera S, Chimirri F, Saracco G, Barber J (2012) Comparison of the  $\alpha$  and  $\beta$  isomeric forms of the detergent n-dodecyl-D-maltoside for solubilizing photosynthetic complexes from pea thylakoid membranes. *Biochim Biophys Acta* 1817:1506–1515.

Pagliano C, Saracco G, Barber J (2013) Structural, functional and auxiliary proteins of photosystem II. *Photosynth Res* 116:167–188.

Pagliano C, Nield J, Marsano F, Pape T, Barera S, Saracco G, Barber J (2014) Proteomic characterization and three-dimensional electron microscopy study of PSII-LHCII supercomplexes from higher plants. *Biochim Biophys Acta* 1837:1454–1462.

Pan X, Li M, Wan T, Wang L, Jia C, Hou Z, Zhao X, Zhang J, Chang W (2011) Structural insights into energy regulation of light-harvesting complex CP29 from spinach. *Nat Struct Mol Biol* 18:309–316.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.

Rhee KH, Morris EP, Barber J, Kühlbrandt W (1998) Three-dimensional structure of the plant photosystem II reaction centre at 8 Å resolution. *Nature* 396:283–286.

Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* 68:850–858.

Standfuss J, van Scheltinga ACT, Lamborghini M, Kühlbrandt W (2005) Mechanisms of photoprotection and nonphotochemical quenching in pea light harvesting complex at 2.5 Å resolution. *EMBO J*. 24:918–928.

Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, Ludtke SJ (2007) EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* 157:38–46.

Umena Y, Kawakami K, Shen JR, Kamiya N (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* 473:55–60.

van Grondelle R, Dekker JP, Gillbro T, Sundstrom V (1994) Energy transfer and trapping in photosynthesis. *Biochim Biophys Acta* 1187:1–65.

van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M (1996) A new generation of the IMAGIC image processing system. *J Struct Biol* 116:17–24.

van Heel M, Gowen B, Matadeen R, Orlova EV, Finn R, Pape T, Cohen D, Stark H, Schmidt R, Schatz M, Patwardhan A (2000) Single-particle electron cryo-microscopy: towards atomic resolution. *Q Rev Biophys* 33:307–369.

Yakushevskaya AE, Jensen PE, Keegstra W, van Roon H, Scheller HV, Boekema EJ, Dekker JP (2001) Supermolecular organization of photosystem II and its associated light-harvesting antenna in *Arabidopsis thaliana*. *Eur J Biochem* 268:6020–6028.

## Figure legends

**Fig. 1** Isolation of PSII-LHCII megacomplexes by sucrose density gradient ultracentrifugation of pea thylakoid membranes solubilized with  $\alpha$ -DDM.

**Fig. 2** Absorption spectrum normalized to the maximum in the red region (**a**) and low temperature (77 K) fluorescence emission spectrum normalized to the maximum value (**b**) obtained from sucrose gradient band  $\alpha 5$ .

**Fig. 3** Electrophoretic separation of pea thylakoid membranes solubilized with  $\alpha$ -DDM and sucrose gradient band  $\alpha 5$ . (**a**) IpBN-PAGE of solubilized thylakoids (25  $\mu$ g Chl) and of band  $\alpha 5$  (6  $\mu$ g Chl),

Coomassie stained and non Coomassie stained. Lane M, mixture of native high molecular weight marker (GE Healthcare) and blue dextran (Sigma-Aldrich). **(b-c)** Silver stained second dimension SDS-PAGEs of solubilized thylakoids and of the PSII-LHCII mc region of band  $\alpha 5$ , respectively. Labels on the left indicate the molecular weight marker positions (Bio-Rad precision plus). In **(b)** identification of the macromolecular protein complexes of thylakoid membranes is given on the top of the gel. Isolated complexes were indexed as follows: megacomplex (mc), supercomplex (sc), trimer (trim), dimer (dim), monomer (mon). In **(c)** identification of the polypeptide composition of the PSII-LHCII mc region of band  $\alpha 5$  by MS/MS spectrometry (see also Table S1).

**Fig. 4** Western blot with antibodies against the extrinsic subunits PsbO, PsbP, PsbQ and PsbR of sucrose gradient band  $\alpha 5$  and of thylakoid membranes. The same amount of Chl (4  $\mu\text{g}$ ) was loaded on each lane.

**Fig. 5** Electron micrograph of sucrose gradient band  $\alpha 5$  particles, negatively stained with 2% uranyl acetate. Top-view, side-view and end-view projections of the megacomplexes are indicated by white rectangles, ovals and circles, respectively. The scale bar represents 100 nm.

**Fig. 6** Characteristic views of the largest megacomplex sub-populations observed in band  $\alpha 5$ . See text for detail. **(a)** Four averages of the most intact megacomplexes in top, side and end view, representing 1,639 analyzed particles out of a total of 1,883 megacomplexes observed with these orientations. Nomenclature: C represents a PSII core monomer, S the strongly bound LHCII trimer and M the moderately bound LHCII trimer. Bar represents 10 nm; all views to the same scale. **(b)** Interpretation of the domains within averages numbered 2 and 3, using a  $\text{C}_2\text{S}_2$  3D map and color co-ordinated all atom sphere overlays as reported previously (Pagliano et al. 2014). Here, modelled high-resolution X-ray structures of the PSII dimeric core from cyanobacteria (Umena et al. 2011) (PDB ID: 3ARC; subunits D1, D2, CP47, CP43 and PsbO are in yellow, orange, red, sandy brown and purple, other subunits in grey, respectively; PsbU and PsbV omitted), the LHCII trimer (Standfuss et al. 2005) (PDB ID: 2BHW; in blue) and Lhcb4 (Pan et al. 2011) (PDB ID:

3PL9; in pale green - PDB ID also used for the fitting of Lhcb5 and Lhcb6 in this study), the latter two from higher plants. Interpretation of averages numbered 1 and 4; the former shown as a lumenally viewed ( $C_2S_2 + C_2S_2$ ) megacomplex (top layer) merged atop a stromally viewed ( $C_2S_2 + C_2S_2$ ) megacomplex (bottom layer), and the latter as a lumenally viewed ( $C_2S_2 + C_2S_2M_2$ ) megacomplex (top layer) merged atop a stromally viewed ( $C_2S_2M_2 + C_2S_2$ ) megacomplex (bottom layer) (PSII dimeric core in red; LHCII trimer S in green; LHCII trimer M in yellow; Lhcb4 in blue; Lhcb5 in green; Lhcb6 in yellow; PDB IDs as above; percentage of transparency top layer:bottom layer 55:45).

### Table title

**Table 1** Relative occurrence of the various PSII-LHCII supercomplexes and megacomplexes in the sucrose gradient band  $\alpha 5$  of Fig. 1. Values are based on two datasets (see text) of 8,963 particles collected in total from two independently isolated bands  $\alpha 5$ .

### Legends for Supplementary data

**Fig. S1** (a) Silver stained second dimension SDS-PAGE of the entire IpBN-PAGE lane of band  $\alpha 5$ . Labels on the left indicate the molecular weight marker positions (Bio-Rad precision plus). (b) Western blot with the antibody against PsbP, performed on the second dimension of the entire IpBN-PAGE lane of band  $\alpha 5$ .

**Fig. S2** The single particle image classification as a visual whole; the entire 7,171 particle dataset (see Table 1) is shown as 600 characteristic views (2D averaged projections), revealing not only the views of the major sub-populations chosen as being representative of the megacomplex observed (see Fig. 6, Table 1), but also the remaining tilted orientations more difficult to ascribe. The scale bar represents 50 nm.

**Table S1** List of integral PSII core subunits, extrinsic polypeptides, and LHCII proteins identified by LC-MS/MS present in the PSII-LHCII mc region of band  $\alpha 5$ , shown in Fig. 3c. The table reports: sequences of peptides obtained by LC-MS/MS (second column) with their corresponding precursor ion mass (third column); for each identified protein (first column), the calculated molecular mass (MW, fourth column), the accession number and the database in which the protein was found (fifth column), and the percentage of residue identities with *Pisum sativum*, when available, or the homolog *Arabidopsis thaliana* (sixth column). Underlined amino acid residues (second column) indicate modifications such as carbamidomethylation of cysteine (C), oxidation of methionine (M), deamidation of asparagine and glutamine (N, Q).



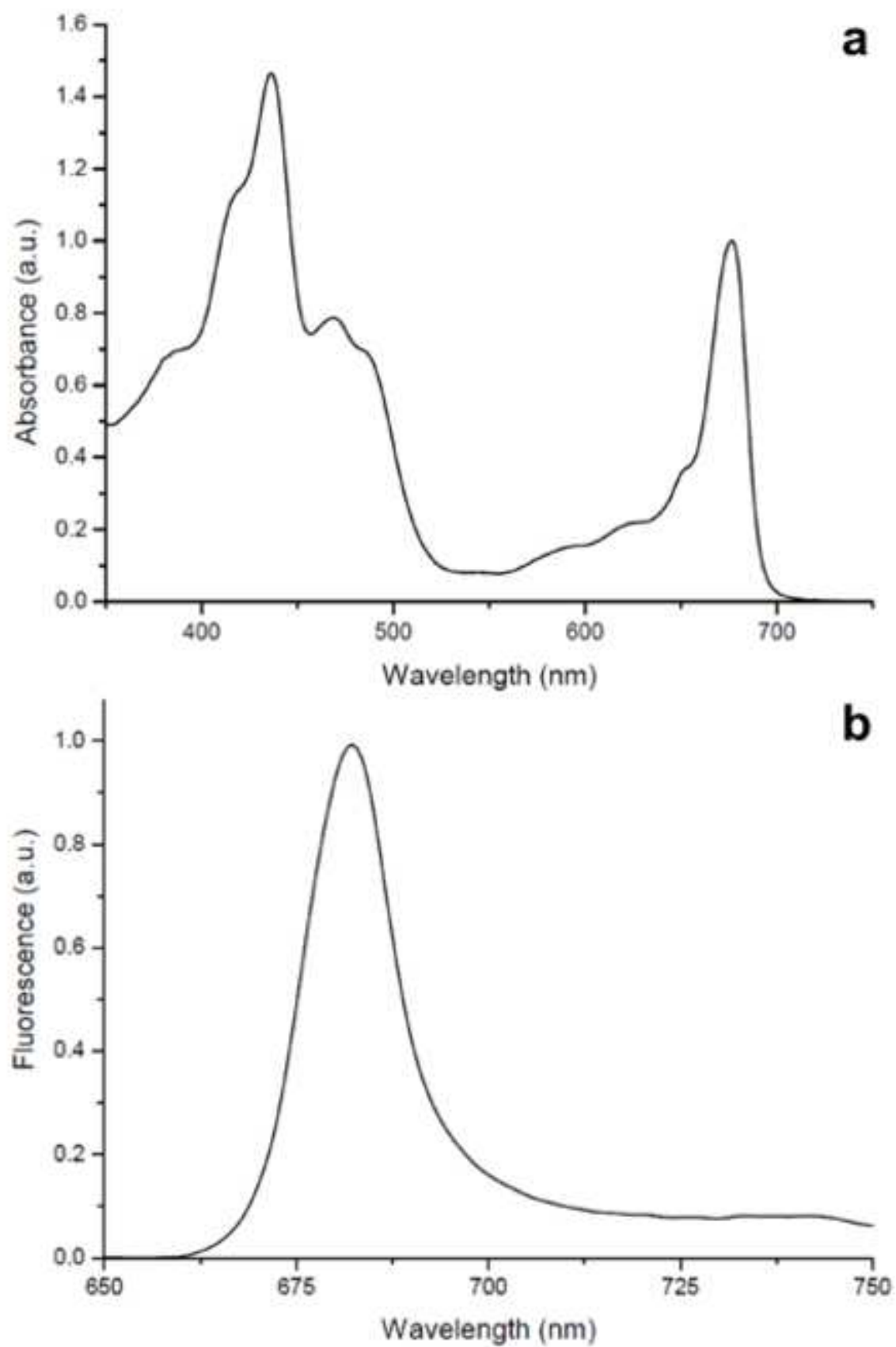
|  **$\alpha$ 1:** LHCII trimers

|  **$\alpha$ 2:** PSI-LHCI complexes/ATP-ase

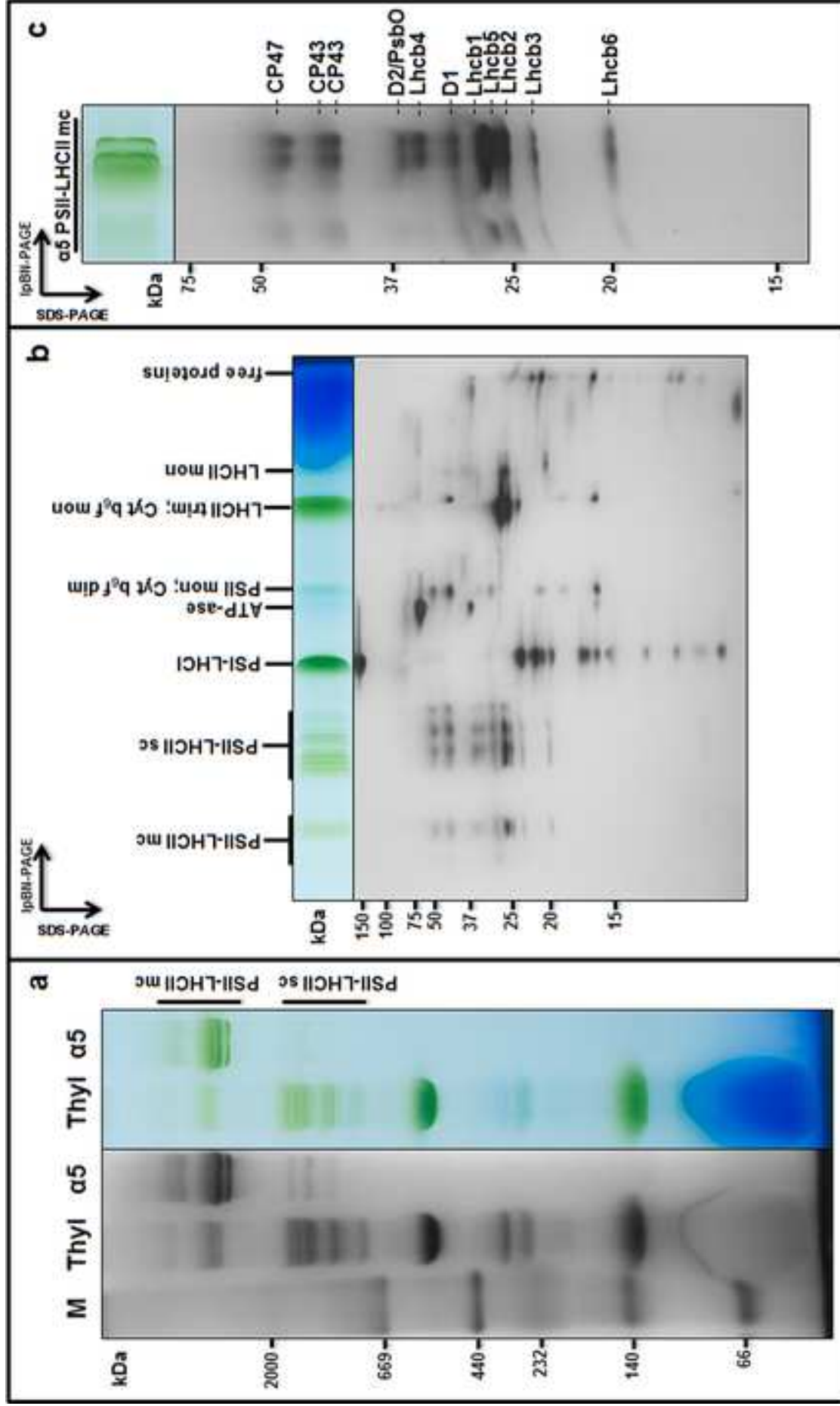
|  **$\alpha$ 3:** PSII-LHCII supercomplexes

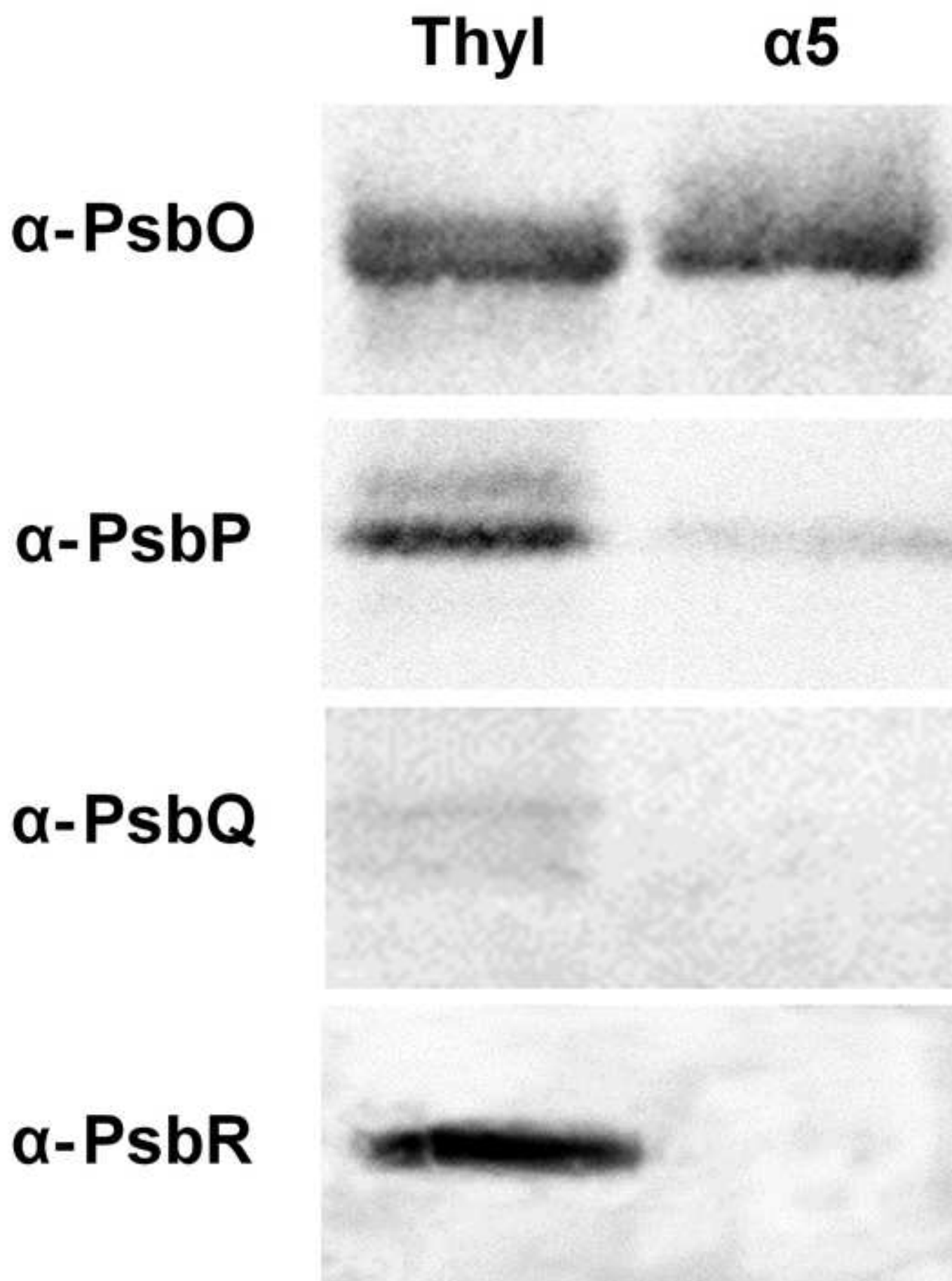
|  **$\alpha$ 4:** PSII-LHCII megacomplexes

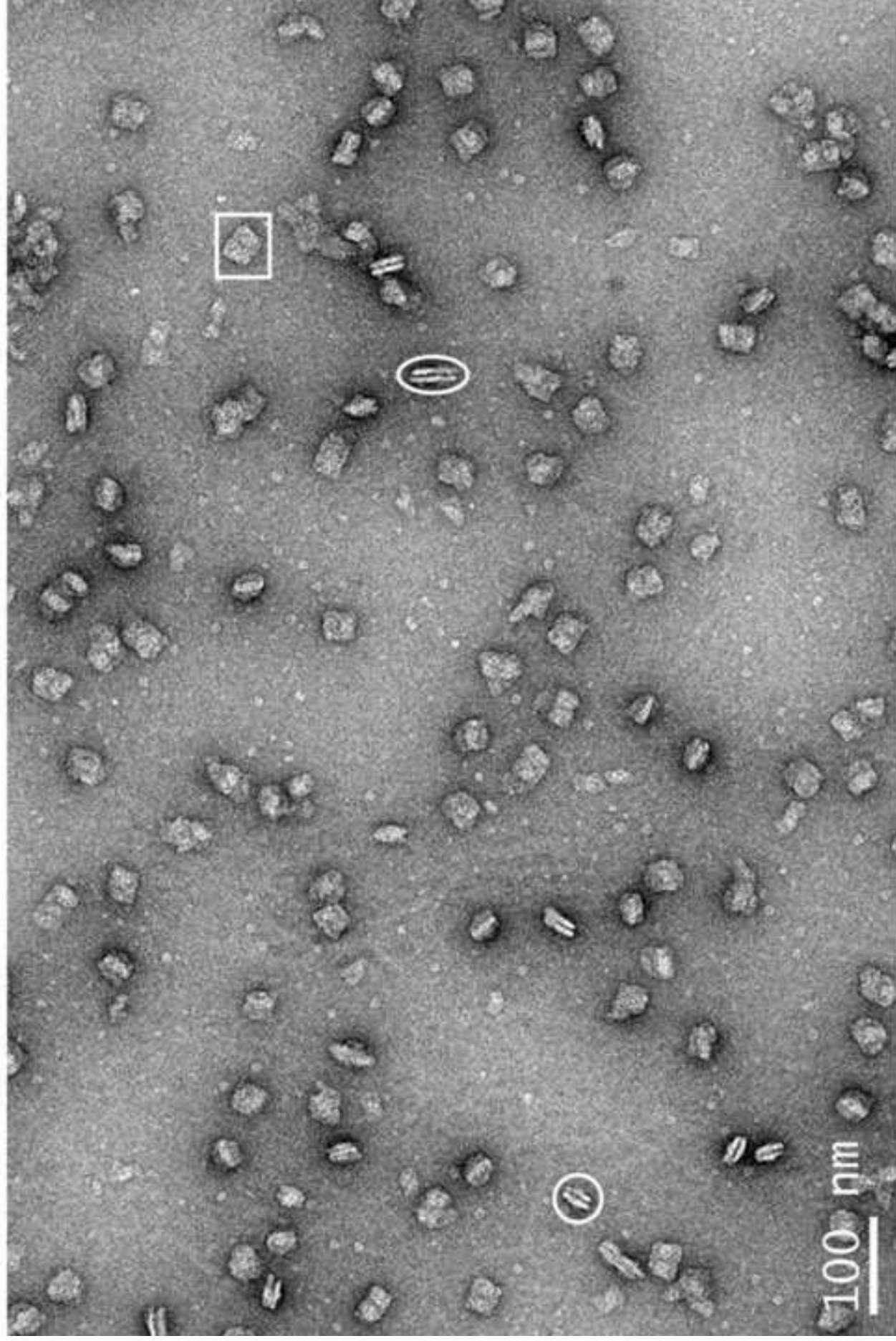
|  **$\alpha$ 5:** PSII-LHCII megacomplexes

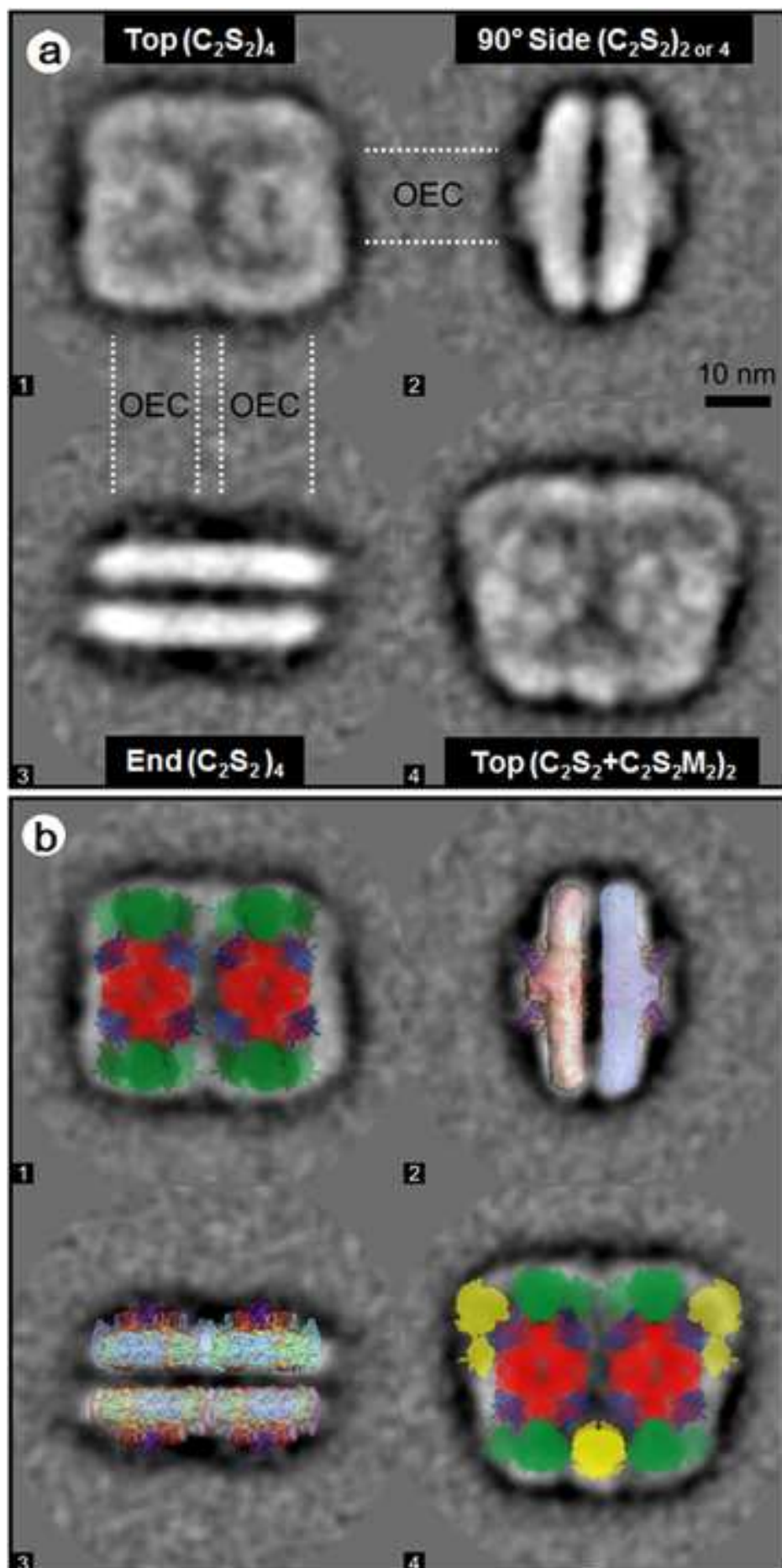




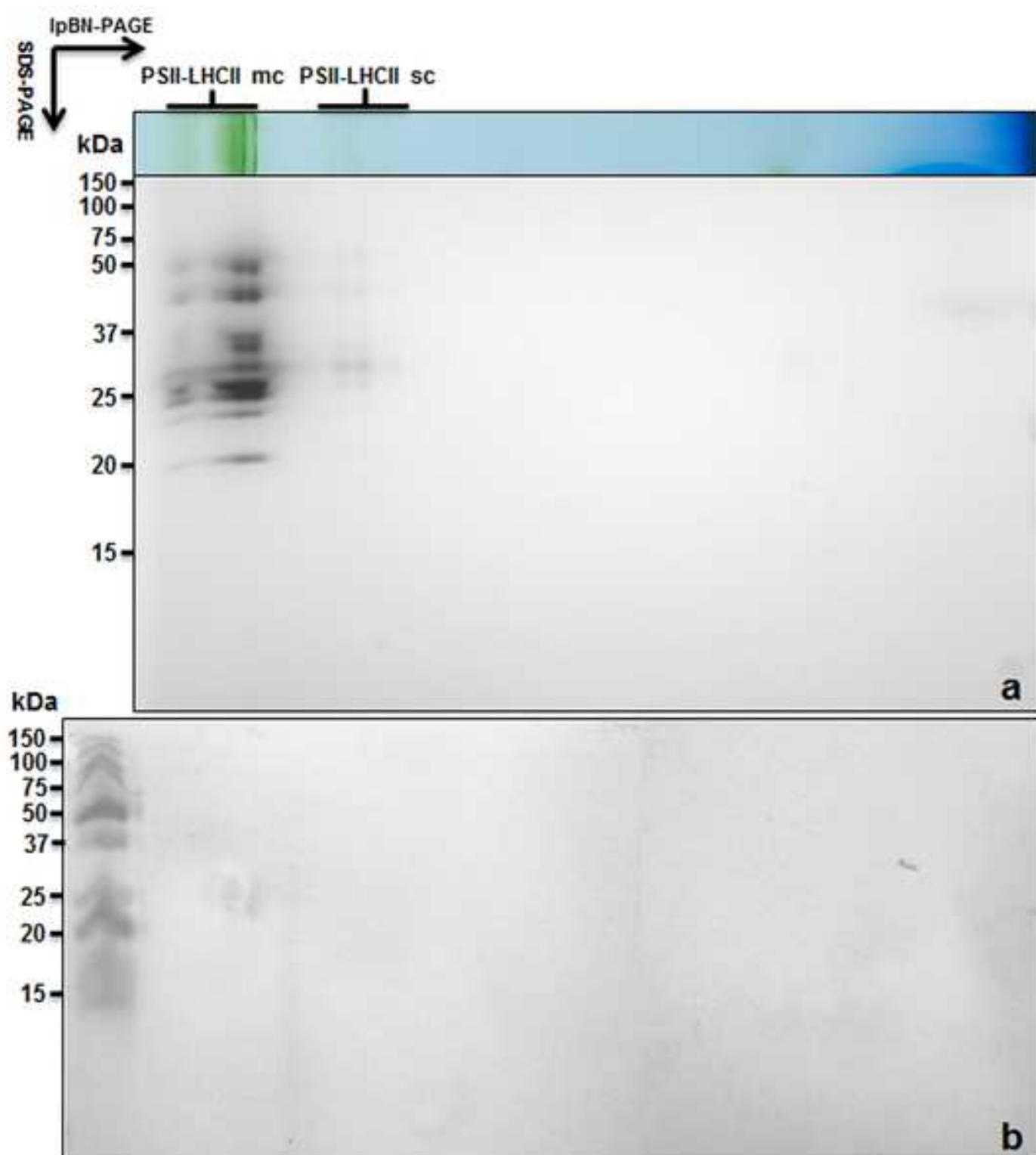


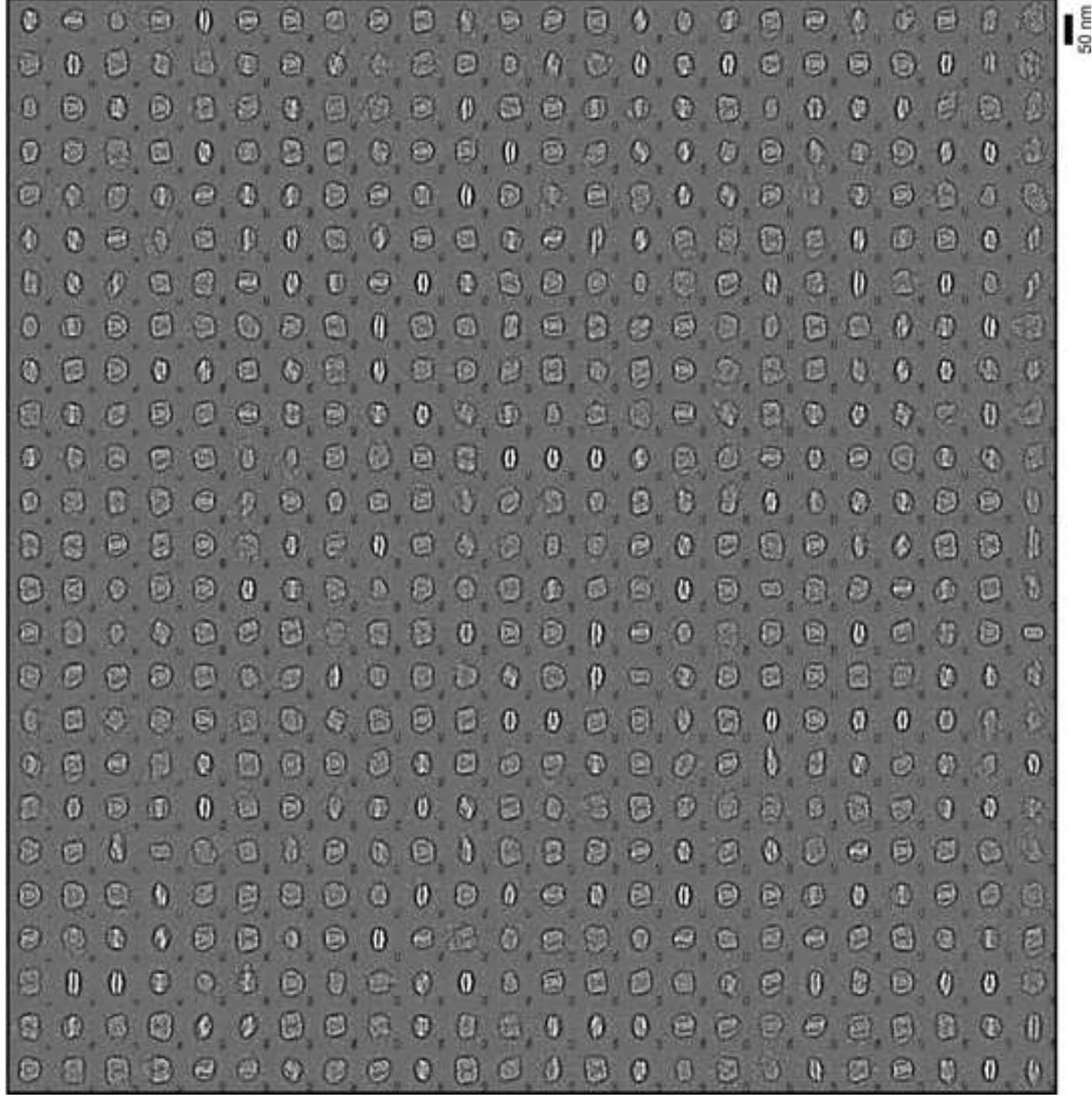






	<b>Number of particles observed</b>	<b>% total particles analysed</b>
Initial dataset of picked particles	8,963	-
Junk particles removed by processing	1,792	-
Total image-processed dataset (see Fig. S2)	7,171	100
Tilted views of intermediate orientation	4,987	69.5
Top/side/end views	2,184	30.5
<b>Complex</b> - nomenclature according to Dekker and Boekema (2005), [Schönflies notation point group symmetry]	<b>Number of particles observed</b>	<b>% total top/side/end views</b>
Supercomplex $C_2S_2$ top views [C2, single PSII dimeric core] (see Fig. S2)	35	1.6
Supercomplex $C_2S_2M$ top views [C2, single PSII dimeric core] (see Fig. S2)	167	7.7
Supercomplex $C_2S_2M_2$ top views [C2, single PSII dimeric core] (see Fig. S2)	99	4.5
Total supercomplexes	301	13.8
Megacomplex $(C_2S_2)_4$ top views [intact D2, four copies of PSII dimeric core] (see Fig. 6, location 1)	403	18.5
Megacomplex $(C_2S_2)_4$ end views [intact D2, four copies of PSII dimeric core] (see Fig. 6, location 3)	38	1.7
Megacomplex $(C_2S_2)_{2 \text{ or } 4}$ side views [intact D2, four copies of PSII dimeric core] (see Fig. 6, location 2)	878	40.2
Megacomplex $(C_2S_2+C_2S_2M_2)_2$ top views [intact D2, four copies of PSII dimeric core] (see Fig. 6, location 4)	320	14.7
Megacomplex $C_2S+(C_2S_2)_2$ top views [broken symmetry, three copies of PSII dimeric core] (see Fig. S2)	244	11.1
Total megacomplexes	1,883	86.2
Total supercomplexes + megacomplexes	2,184	100





Protein	Peptide sequence	Precursor ion mass m/z	MW (Da)	AC number (GI/NCBI) reference organism	% Identity with <i>P. sativum</i> or <i>A. thaliana</i> (sequence coverage)
	VHTVLNDPGR	1205,6537			
	TGKPSLDLPK	1054,6054			
	YQWDQGYFQQEIVR	1922,8596			
	VGGGLVENQSLSEAWSK	1759,8751			
	LAFYDIYGINNPAK	1484,7284			
CP47	AGSMNDNGDAVAVGWLGHPIFR	2170,0294	55988	gi 295137030 ( <i>Pisum sativum</i> )	100% gi 295137030 <i>P. sativum</i> (28%)
	MPTFEFTFPVVLVDGDIVR	2238,1441			
	DVFAQIDPDLDAQVEFGAFQK	2281,0917			
	RAQLGEIFEELDR	1445,7645			
	LGANVGSAAQPTGLGK	1425,7624			
CP43	SPTGEVIFGGETMR	1479,7042	51955	gi 295136994 ( <i>Pisum sativum</i> )	100% gi 295136994 <i>P. sativum</i> (19%)
	GPNGLDLSR	928,4619			
	DFEPLSMITPLN	1361,6560			
	AYDFVSOEIR	1226,5927			
D2	AAEDPEEFTTYTK	1546,6806	35386	gi 27435890 ( <i>Pisum sativum</i> )	100% gi 27435890 <i>P. sativum</i> (14%)
	AWMATQDDQPHENLIFPEEVLPR	2620,2732			
D1	ANLGMENVHER	1285,5906	38962	gi 131252 ( <i>Pisum sativum</i> )	100% gi 131252 <i>P. sativum</i> (6%)
	NAHNIFPLDIA	1110,5452			
	NTPLAFQNTK	1132,5858			
PsbO	LTYTLDIEIGPFEVSADGSVK	2269,1016	34872	gi 131384 ( <i>Pisum sativum</i> )	100% gi 131384 <i>P. sativum</i> (25%)
	DGIDYAAVTVQLPGER	1759,8728			
	GASTGYDNAAVALPAGGR	1575,7595			
	VASSGSPWYGPDR	1377,6311			
Lhcb1	ELEVIHSR	981,5235	28635	gi 115788 ( <i>Pisum sativum</i> )	100% gi 115788 <i>P. sativum</i> (23%)
	FGEAVVFK	982,4918			
	IAGGPLGEVVNPLYPGGSFDPQLGLADDPFAELK	3525,7641			
Lhcb2	SAPESIWYGFDRPK	1601,7863	28866	gi 115797 ( <i>Pisum sativum</i> )	100% gi 115797 <i>P. sativum</i> (25%)
	ELEVIHSR	981,5282			
	FGEAVVFK	982,4925			
	VGGGPLGEGLDPLYPGGAFDPLGLADDPDSFAELK	3455,6889			
	ALEVIHGR	893,5109			
Lhcb3	VDFKEPWFK	1293,6783	28710	gi 20671 ( <i>Pisum sativum</i> )	100% gi 20671 <i>P. sativum</i> (7%)
	ATLQLAEIK	985,5782			
Lhcb4	FDDPLGLADPEKK	1546,8051	31290	gi 346987811 ( <i>Dimocarpus longan</i> )	83% gi 15231990 <i>A. thaliana</i> (8%)
	SEIPEYL TGEVPGDYGYDPFGLSK	2632,2313			
Lhcb5	LTNGLDLEDK	1116,5676	30855	gi 470129540 ( <i>Fragaria vesca</i> subsp. <i>Vesca</i> )	81% gi 15235029 <i>A. thaliana</i> (12%)
	TAENFNNSITGLQGYPGGK	1854,8398			
Lhcb6			23518	gi 168009690 ( <i>Physcomitrella patens</i> )	82% gi 4741960 <i>A. thaliana</i> (7%)





Click here to access/download  
**Supplementary material**  
alpha3 cryo-TEM image.tif

