Isolation and characterization of a large Photosystem I-Light Harvesting complex II supercomplex with an additional Lhca1-a4 dimer in Arabidopsis

Aurelie Crepin, Zuzana Kučerová, Artemis Kosta, Eric Durand, Stefano Caffarri

Isolation and characterization of a large Photosystem I-Light Harvesting complex II supercomplex with an additional Lhca1-a4 dimer in Arabidopsis

Aurélie Crepin[1,2], Zuzana Kučerová[1,3], Artemis Kosta[4], Eric Durand[5], Stefano Caffarri[1#]

[1]Aix Marseille Université, CEA, CNRS, Biosciences and Biotechnologies Institute of Aix-Marseille (BIAM), Equipe de Luminy de Génétique et Biophysique des Plantes, 13009 Marseille, France.
[2]Centre Algatech, Institute of Microbiology of the Czech Academy of Sciences, Opatovický mlýn, 379 81 Třeboň, Czech Republic
[3]Department of Biophysics, Centre of the Region Haná for Biotechnological and Agricultural Research, Faculty of Science, Palacký University, Šlechtitelů 27, 78371 Olomouc, Czech Republic
[4]Microscopy Core Facility, Institut de Microbiologie de la Méditerranée (IMM), FR3479, CNRS, Aix Marseille University, Marseille, France.
[5]Laboratoire d'Ingénierie des Systèmes Macromoléculaires (LISM), Institut de Microbiologie de la Méditerranée (IMM), Aix-Marseille Université, CNRS, UMR 7255, 13402 Marseille Cedex 20, France.

[#]for correspondence (email: stefano.caffarri@univ-amu.fr)

Keywords: Photosynthesis, Photosystem I, LHCII, Light harvesting complex, Lhca, *Arabidopsis thaliana*

SUMMARY

The biological conversion of light energy into chemical energy is performed by a flexible photosynthetic machinery located in the thylakoid membranes. Photosystem I and II (PSI and PSII) are the two complexes able to harvest light. PSI is the last complex of the electron transport chain and is composed of multiple subunits: the proteins building the catalytic core complex that are well conserved between oxygenic photosynthetic organisms, and, in green organisms, the membrane Light harvesting complexes (Lhc) necessary to increase light absorption. In plants, four Lhca proteins (Lhca1-4) make up the antenna system of PSI, which can be further extended to optimize photosynthesis by reversible binding of LHCII, the main antenna complex of photosystem II. Here, we used biochemistry and electron microscopy in Arabidopsis to reveal a previously unknown supercomplex of PSI with LHCII that contains an additional Lhca1-a4 dimer bound on the PsaB-PsaI-PsaH side of the complex. This finding contradicts recent structural studies suggesting that the presence of an Lhca dimer at this position is an exclusive feature of algal PSI. We discuss the features of the additional Lhca dimer in the large plant PSI-LHCII supercomplex and the differences with the algal PSI. Our work provides further insights into the intricate structural plasticity of photosystems.

INTRODUCTION

Photosynthesis is the mechanism by which organisms use light energy to form carbohydrates. In plants, light is absorbed at the level of two photosystems (PS), which function in series to drive an electron transport chain in the thylakoid membrane of chloroplasts. In order to keep the chain running with maximal efficiency and avoid potentially damaging saturation, the excitation of the two PS must be carefully balanced. One mechanism for balancing PS excitation is through the fine-tuning of the composition of antenna system composition.

The first complex involved in the photosynthetic electron transport chain and the one responsible for water-splitting is called PSII. In green organisms the main PSII antenna is composed of LHCII trimers. These complexes are composed of different combinations of Lhcbm (in algae and mosses) or Lhcb1-3 (in vascular plants) proteins. All these isoforms present a highly similar sequence and structure (Crepin and Caffarri, 2018; Pan *et al.*, 2019): they have three transmembrane domains, bind ~14 chlorophylls (Chls) and have a low Chl a/b ratio of ~1.3/1.4.

In plant PSI-LHCI (hereafter PSI), the antenna system is composed of four Lhca proteins organized in two heterodimers, Lhca1-a4 and Lhca2-a3 (Wientjes and Croce, 2011) arranged as a single row adjacent to the core proteins PsaG, PsaF, PsaJ and PsaK (Qin *et al.*, 2015; Mazor *et al.*, 2017; Pan *et al.*, 2018; Pan *et al.*, 2019). Lhca monomers present a structure similar to Lhcb, and also bind a high number of Chls, although with a Chls a/b ratio that is higher than that found in LHCII. Lhca complexes contain red forms (Chls a that absorb light above 700 nm) (Wientjes and Croce, 2011; Morosinotto *et al.*, 2003), which are important for energy trapping and transfer in PSI. Lhca3 and Lhca4 have the most red-shifted maxima.

The PSI antenna size can be increased in non-saturating light conditions through a regulatory mechanism called state transitions (Rochaix, 2014): in the dark or in a light that preferentially excites PSI, the un-phosphorylated LHCII trimers are bound and transfer absorbed energy mainly to PSII (state 1). However, in case of preferential PSII excitation, the high reduction of the plastoquinone pool induces the activation of the STN7 kinase (Bellaflore *et al.*, 2005), the phosphorylation of part of the LHCII and its movement and binding to PSI next to the PsaH, PsaK, PsaL and PsaO subunits (state 2) (Jensen *et al.*, 2007; Pan *et al.*, 2018). There, the bound LHCII trimer can transfer absorbed energy with a high efficiency to the PSI core (Galka *et al.*, 2012; Santabarbara *et al.*, 2017). Other studies have also reported the binding of additional LHCII trimers to PSI, though little is known on their exact binding sites (Benson *et al.*, 2015; Bell *et al.*, 2015; Bos *et al.*, 2017; Yadav *et al.*, 2017).

In other organisms of the green lineage, the PSI antenna can also be enlarged by additional Lhca proteins. In the moss *Physcomitrella patens*, for instance, two teams have described a larger complex including a second row of Lhca dimers below the first one, although these additional Lhca complexes are shifted towards PsaK and seem to also bind to the LHCII trimer found on this side of the complex (Iwai *et al.*, 2018; Pinnola *et al.*, 2018). In one case, another complex was also described that has only two additional unidentified Lhca proteins whose binding site is adjacent to the first Lhca1-Lhca4 complex, and which seems involved in the docking of an LHCII trimer at this position (Pinnola *et al.*, 2018). The formation of these complexes, however, seems to depend on the *Physcomitrella*-specific Lhcb9 protein.

In the green alga *Chlamydomonas reinhardtii* the LHCI antenna, which contains 9 different Lhca proteins (Lhca1-9) (Busch and Hippler, 2011), is organized as a double row, with the

more internal one in a similar position as Lhca1-4 in plants. Besides these Lhca, the binding of an additional dimer was recently described next to PsaH, PsaI and PsaB, i.e. the side of the core complex opposite to the LHCII binding site (Ozawa *et al.*, 2018; Suga *et al.*, 2019; Su *et al.*, 2019). A similar organization was also described in the green alga *Bryopsis corticulans* (Qin *et al.*, 2019) as well as in the red alga *Cyanidioschyzon merolae* (Pi *et al.*, 2018), suggesting that the binding site of this additional dimer is conserved among several photosynthetic organisms. The binding site of the Lhca dimer in green algae, though, is closer to PsaH compared to the one in red algae. This difference was suggested to be due to differences between PSI cores, and especially to the loss of PsaM and the gain of PsaH and PsaG (Suga *et al.*, 2019) in the PSI of green algae. To this day, however, no such complex has been described in plants, leading to the belief that this binding site was lost in plants.

In this article, we describe a previously unknown PSI supercomplex in *Arabidopsis thaliana* that contains an additional Lhca dimer on the PsaI-PsaH side of the core as well as an LHCII trimer bound to PsaH-PsaK. This complex demonstrates that the Lhca-binding domain on on the PsaI-PsaH side of PSI is also conserved in plants. However, compared to green algae the plant Lhca dimer binding site is shifted towards PsaH, and the presence of LHCII seems necessary for the proper binding.

RESULTS

Isolation of a larger PSI-LHCII supercomplex

In plants, the transition from state 1 to state 2 is characterized by the formation of a characteristic PSI-LHCII supercomplex, clearly visible on CN-PAGE (Pesaresi *et al.*, 2002; Järvi *et al.*, 2011; Galka *et al.*, 2012). However, the existence of larger complexes has been suggested, for instance other PSI supercomplexes with an even larger antenna, mostly by the binding of additional LHCII trimers (Yadav *et al.*, 2017; Benson *et al.*, 2015; Bell *et al.*, 2015; Bos *et al.*, 2017). To determine if such uncharacterized supercomplexes can be isolated from plants in either state, we separated solubilized thylakoid membranes from plants in state 1 and state 2 (WT and the mutant *pph1* blocked in state 2) on a non-denaturing CN-PAGE gel. We then used a delicate solubilization protocol developed by our team in the past (Galka *et al.*, 2012), based on a mix of digitonin and n-dodecyl-α-D-maltoside (α-DDM) detergents. Combining the two detergents allows for very efficient membrane solubilization with a final result almost identical to that obtained with much higher concentrations of the very mild digitonin as the only detergent. The detergent mix also avoids the negative effects of α-DDM on weak protein interactions. This type of solubilization already proved capable of yielding photosynthetic complexes that remain stable under many conditions (Galka *et al.*, 2012; Crepin *et al.*, 2016). In (Crepin *et al.*, 2016) further discussion is provided about the effect of several detergents on complex stability.

After solubilization, the complexes are separated by migration on a large pore native gel (Järvi *et al.*, 2011). Separation of the complexes on gel was further improved by adjusting the concentration of the detergents in the cathode buffer (see Experimental Procedures and Figure S2). As expected, the PSI-LHCII supercomplex is visible on this type of gel as a green band of low fluorescence above the PSI and this complex is only visible in state 2 (Figure 1a). The gel also showed a thinner additional band just above the PSI-LHCII supercomplex. This band shows a low fluorescence and is situated just below a PSII band containing C_2S particles

(Caffarri *et al.*, 2009). The low fluorescence of the band suggested a yet undescribed PSI supercomplex. This was confirmed by migration in a second dimension under denaturing conditions (Figure 1b and Figure S1 for a 2D gel of a whole lane). The additional band indeed contained PSI as well as LHCII subunits, indicating that this unknown PSI complex is a larger PSI-LHCII supercomplex. Though the band position, between PSI-LHCII (~700 kDa) and the PSII-C_2S supercomplex (~880 kDa) (Caffarri *et al.*, 2009), implied the addition of either a large subunit or few small ones, no additional spots were visible on the 2D gel. The absence of new subunits therefore suggests the presence of proteins already present in the PSI-LHCII supercomplex, possibly additional antenna proteins.

LHCII phosphorylation of the complex

As the PSI supercomplex of interest also contains LHCII, we investigated its phosphorylation to determine if it is indeed a classic PSI-LHCII containing additional proteins or an unknown supercomplex with different characteristics.

To quantify the phosphorylation, the complexes isolated from the CN-PAGE gel were separated on SDS-PAGE gels in the presence of Phos-tag, a molecule able to slow down phosphorylated proteins. Indeed, a band specific of phosphorylated LHCII was visible on the gel (Figure 2a). The quantification of this slowly-migrating, phosphorylated LHCII compared to the bulk, non-phosphorylated band indicated that 33.4 ± 2.9% of the total LHCII is phosphorylated in the large PSI-LHCII complex (SE, n=4), and 38.5 ± 1.2 % in the classic PSI-LHCII, a value in the range of what we obtained before (Crepin and Caffarri, 2015). These values also point to the phosphorylation of around one monomer per LHCII trimer, as is the case for the well characterized PSI-LHCII complex (Crepin and Caffarri, 2015).

To further characterize this phosphorylation, we performed immunoblots with antibodies specifically designed against phosphorylated Lhcb1 and Lhcb2 (Leoni *et al.*, 2013) and compared the results obtained between PSI-LHCII and the larger PSI-LHCII supercomplex (Figure 2b).

Considering that the immunoblot signal of P-Lhcb2 and P-Lhcb1 in the PSI-LHCII corresponds respectively to ~90% and ~10% of the total P-LHCII (Crepin and Caffarri, 2015), we estimated that P-Lhcb2 was very slightly decreased in the larger PSI-LHCII complex to 87% ± 1.6%, while on the contrary P-Lhcb1 increased to 12.6% ± 1.4%. Considering the prevalence of P-Lhcb2 in the complex, the difference between the samples is therefore very minor and the small increase in Lhcb1 phosphorylation can be explained by the slight contamination by the close-migrating PSII band, which is less phosphorylated and has a prevalence of P-Lhcb1 instead (Crepin and Caffarri, 2015).

Altogether, these values indicate a very similar phosphorylation of both PSI-LHCII supercomplexes, indicating that the undescribed band isolated here contains very likely a typical, well-known PSI-LHCII with additional proteins.

Absorption and fluorescence spectra of the larger PSI-LHCII supercomplex

To further characterize this unknown and larger PSI-LHCII supercomplex, we cut the corresponding band and eluted the particles from a preparative 20-cm CN-PAGE (Figure S2c) using membranes from the *pph1* mutant, which is blocked in state 2 and is enriched in the larger PSI-LHCII supercomplex (Figure 1). The spectrum of the larger PSI-LHCII

supercomplex was then compared with the spectra of eluted PSI and PSI-LHCII supercomplexes (Figure 3a). The results indicated a high content of Chl b in the band of interest in comparison to PSI; however, the larger PSI-LHCII supercomplex showed a similar spectrum to that of the PSI-LHCII supercomplex, with possibly a very minor increase of the Chl b component. This result suggests that either the additional subunits in the larger PSI-LHCII supercomplex do not bind pigments or they bind Chls with a Chl a/b ratio lower than that of the PSI-LHCII complex (~4.9) (Galka *et al.*, 2012), but higher than that of LHCII (~1.4), which would be clearly visible in the Chl b absorption contribution at ~470 nm and 650 nm. Complexes with such an intermediate Chl a/b ratio are the Lhca complexes: both Lhca1-a4 and Lhca2-a3 dimers have a Chl a/b ratio of ~3.7 (Wientjes and Croce, 2011).

To determine if such possible additional pigment proteins are visible in fluorescence and/or able to transfer energy to the PSI core, we measured the low-temperature fluorescence emission and excitation spectra of the complexes (Figure 3b and 3c). As in the case of the absorption spectrum, the larger PSI-LHCII supercomplex showed an excitation spectrum very similar to that of the PSI-LHCII supercomplex indicating that the complex is energetically well connected. Emission spectra of PSI, PSI-LHCII and the larger PSI-LHCII had a similar emission with a maximum at ~730 nm. Some disconnected LHCII (peak at 679 nm) was visible in the PSI-LHCII sample, probably due to long purification procedure by elution from the gel; in the case of the larger PSI-LHCII, some minor contamination from PSII (peak at 684 nm) was visible, in agreement with results shown in Figure 1 and Figure S2 that reveal close migration of the PSII-C_2S complex. Spectroscopic results are in agreement with biochemical results and support the fact that the band of interest is a not yet described PSI-LHCII complex. However, considering the large similarity between the classical PSI-LHCII and the larger PSI-LHCII for absorption and fluorescence properties, these results did not allow us to conclude whether the additional subunits were Chl binding proteins or not. We therefore performed further biochemical analyses.

Immunoblot analysis and antenna quantification

The estimated size of the complex of interest on the CN-PAGE is compatible with the addition of 2 to 3 monomeric antennas (whose holoprotein MW is ~40 kDa) to the well characterized PSI-LHCII complex. To identify these proteins, we performed SDS-PAGE using different gel systems to allow the separation of individual antenna proteins as well as immunoblot analyses on all Lhc antenna proteins of Arabidopsis.

First of all, we checked if additional Lhcb proteins could be bound to the large PSI-LHCII complex (Figure 4a). Analysis of Sypro stained gels clearly indicated that there is no increase in LHCII content (Lhcb1 and Lhcb2) per PSI core in the larger PSI-LHCII supercomplex, excluding the addition of monomeric or trimeric LHCII. We also checked all the other Lhcb proteins, which are the monomeric antennas Lhcb4, Lhcb5 and Lhcb6 (also called CP29, CP26 and CP24, respectively) and the Lhcb3 subunits that is specifically located in the moderately bound LHCII-M trimer (Caffarri *et al.*, 2009). Both Sypro stained gel and immunoblots with specific antibodies indicated that none of these subunits was significantly increased in the larger PSI-LHCII supercomplex as compared to the PSI-LHCII. The small and similar increase of each Lhcb subunits visible in the immunoblots in the case of the larger PSI-LHCII supercomplex was due to a minor contamination from PSII as demonstrated by loading, as a

control, PSII membrane (BBY) to have similar stoichiometric amounts of Lhcb proteins and Lhca proteins in the PSI samples (Figure 4a).
We therefore analyzed the content of Lhca1-a5 proteins and of the PSI core subunit PsaG by Sypro stained SDS-PAGE and immunoblots (Figure 4b). Surprisingly, both on the Sypro stained gel and the immunoblots, the only signals that were increased (by about 50%, Figure 4c) corresponded to bands containing the Lhca1 and the Lhca4 proteins. We therefore concluded that the large PSI-LHCII supercomplex binds an extra Lhca1-a4 dimer, and so contains 6 Lhca in total. Hereafter we call this larger supercomplex "PSI-6LHCI-LHCII".
Note that the addition of a Lhca1-a4 dimer to the PSI-LHCII complex is practically undetectable in the absorption spectrum (see Figure S4), in agreement with our experimental data showing no significant differences between PSI-LHCII and the PSI-6LHCI-LHCII eluted fractions (Figure 3).

Single particle electron microscopy

To identify the location of the additional Lhca1-Lhca4 dimer in the PSI-6LHCI-LHCII supercomplex, we analyzed the architectural organization of this complex and of the typical PSI-LHCII by single particle negative staining electron microscopy (EM).
In two independent purifications and EM analyses, we detected a complex with the same structure as described in the past for the PSI-LHCII fraction, with a single LHCII trimer bound to the PSI core on the side of PsaK (Kouřil *et al.*, 2005; Galka *et al.*, 2012) (Figure 5a-b; Figure S3a for all 2D EM classes). For the eluted fraction corresponding to the PSI-6LHCI-LHCII complex, the two main classes were the classical PSI-LHCII complex described above and a new particle showing a larger PSI-LHCII complex with an extra density on the side of PsaB-PsaI-PsaH (Figure 5c-d; Figure S3b for all EM classes). This extra density has a size compatible with an Lhca dimer. We were indeed able to fit an Lhca1-a4 dimer in this extra density maintaining a similar topology for contacts with respect to PSI as for the normal Lhca1-a4 (Figure 5e). The extra Lhca1-a4 dimers would be close to the site where Lhca dimers have been identified in PSI complexes of red (Pi *et al.*, 2018) and green algae (Su *et al.*, 2019; Suga *et al.*, 2019) (Figure 5f). This dimer, though, seems shifted towards PsaH compared to the known structures from algae with the result that Lhca4 would be in a similar position as Lhca2 in the *Chlamydomonas* PSI complex (Figure 5e-f), while Lhca1 would be located towards the LHCII trimer with which it seems to interact.
It is important to note that in the fraction of the PSI-6LHCI-LHCII sample, we detected a significant amount of classical PSI-LHCII particles (Figure S3b) indicating a significant contamination of this fraction by the much more abundant and adjacent PSI-LHCII fraction (Figure 1 and Figure S2). This explains the immunoblot results showing an increase of Lhca1 and Lhca4 of ~1.5 times (instead of 2 as expected for a pure fraction) in the PSI-6LHCI-LHCII preparation as compared to PSI-LHCII (Figure 4c).

Searching for a PSI-6LHCI complex without LHCII

In order to determine whether PSI complexes with an additional Lhca1-a4 dimer but without LHCII exist, as is the case for algal PSI complexes, we performed additional 2D gels (CN-PAGE + SDS-PAGE) on membranes in state 1 and state 2 conditions using only digitonin (1%). This bulky detergent is considered to be particularly able to maintain weak protein

interactions compared to the other detergents used for purification of membrane complexes, such as α-DDM. If existing and stable enough *in vitro*, a PSI-6LHCI complex should migrate above the PSI band, but below the PSI-LHCII band present in state 2 membranes. In our gels, we were not able to visualize a clear band that could correspond to a PSI-6LHCI complex both in State 1 or State 2 samples (Figure 6). At the position expected, an extremely faint and smeared band migrating above PSI was visible in state 1 membrane. This suggests that such a complex would be absent or very unstable in state 1 or state 2, in agreement with the fact that the additional Lhca dimer seems in contact with LHCII (Figure 5e). However, and quite surprisingly, a band migrating at the position of the PSI-6LHCI-LHCII was clearly visible in state 1 membranes, which, as expected, lack the typical PSI-LHCII complex (Figure 6). Even if faint, the presence of a PSI-6LHCI-LHCII in state I membranes suggests that during the transition from state 2 to state 1 this supercomplex is more resistant to the dephosphorylation by the PPH1 phosphatase than the PSI-LHCII complex.

DISCUSSION

In recent years the use of cryo-EM structural methods has led to the description of high resolution PSI supercomplexes with a Lhca dimer bound adjacent to the PsaH-PsaI-PsaB core subunits, first in the red alga *Cyanidioschyzon merolae* (Pi *et al.*, 2018), and then in two different green algae, *Chlamydomonas reinhardtii* and *Bryopsis corticulans* (Ozawa *et al.*, 2018; Suga *et al.*, 2019; Su *et al.*, 2019; Qin *et al.*, 2019). The presence of a dimer on the PsaH-PsaI-PsaB side of PSI was proposed to be a specific property of algal photosystems that was lost in plants. In this article we present the isolation and characterization of a previously undescribed large plant PSI-LHCII complex that contains an Lhca dimer located in a similar though not identical position to the Lhca dimer in the algal PSI. Thus, the binding of an additional Lhca dimer on PSI was not lost in plants, although the position appears to be modified with the dimer shifted closer to PsaH compared to its position in PSI from green algae, where it is already displaced compared to red algae (Suga *et al.*, 2019). This position brings the dimer into close proximity with the LHCII trimer bound on the other side, with which it also seems to interact. Indeed, considering that we were unable to find a PSI-6LHCI complex, it seems that the LHCII trimer plays an important role in the binding of the additional Lhca dimer: it is either strictly necessary for binding or it is simply sufficient to stabilize binding and make it resistant to detergent solubilization. In the first case, it would mean that the binding of the additional Lhca1-Lhca4 dimer is likely dependent on LHCII phosphorylation and is therefore induced by state 2 conditions. It could necessitate either the LHCII binding only, or other state transition-specific modification(s) of the complex. Interestingly, acetylation of the PsaH protein, in very close proximity to this additional dimer, was found to be essential for state transitions. The loss of this acetylation, though, induced the loss of the classic, well-known PSI-LHCII supercomplex, suggesting instead a role in LHCII binding (Koskela *et al.*, 2018).

The question remains about the origin of the additional Lhca dimer. Even if it has been proposed that the normal Lhca1-Lhca4 dimer has a weaker binding to PSI compared to Lhca2-Lhca3 (Krumova *et al.*, 2014; Nellaepalli *et al.*, 2014), this binding is considered very strong due to several interactions between Lhca1-a4 and the PSI core (Qin *et al.*, 2015; Mazor *et al.*, 2017). Indeed, a harsh detergent treatment is needed to detach the Lhca proteins from the core

(Croce *et al.*, 1998). Therefore it is unlikely that the additional Lhca1-a4 dimer originates from other PSI particles, and indeed so far no PSI complex with only Lhca2-a3 has been described in WT plants. An alternative possibility is the existence of a "reserve" of Lhca dimers, though to our knowledge, nothing of the sort has been described so far. It is also possible that after acclimation to low light or state 2 conditions, the presence of this additional dimer is due to an increased synthesis of these proteins. A variation in Lhca stoichiometry has indeed been proposed in the past for plants grown at different light intensities (Bailey *et al.*, 2001). Therefore it is possible that a pool of constitutively large PSI-LHCII supercomplexes exists *in vivo*. The fact that the PSI-6LHCI-LHCII complex was detected in state 1 conditions when the classic PSI-LHCII is undetectable (Figure 6) supports this proposition. Alternatively, this result could indicate that the PSI-6LHCI-LHCII complex is formed in state 2, and is then slowly disassembled during the transition from state 2 to state 1. Since the additional Lhca1-Lhca4 dimer seems to interact with the LHCII trimer at the level of the monomer clearly identified as the phosphorylated Lhcb2 isoform (Pan *et al.*, 2018), it could block or limit the access to the phosphorylated threonine residue by the PPH1 phosphatase.

It is also interesting to point out that mutants of individual Lhca proteins, and especially of Lhca4, have been shown to be affected for state transitions, even though the PSI-LHCII supercomplex is still visible on gel (Benson *et al.*, 2015; Bressan *et al.*, 2018). These results have been interpreted as an indication of the binding of additional LHCII trimers to the PSI complex in state 2 via interactions with Lhca proteins. Indeed, one such complex was already described by electron microscopy, with an additional LHCII trimer bound to the Lhca2-a3 dimer (Yadav *et al.*, 2017). However we were not able to isolate this complex from native gels and visualize it in EM, indicating that such a complex would be extremely delicate and disassembled during gel migration. Interestingly, another complex described in the moss *Physcomitrella patens* reveals a second LHCII trimer bound on the LHCII side of the PSI-LHCII complex by the intermediate of another antenna dimer, hypothesized to be the moss-specific Lhcb9 protein (Pinnola *et al.*, 2018). Though we have no evidence in our work, based on the results described above suggesting the existence of PSI-LHCII supercomplexes with additional LHCII (Benson *et al.*, 2015; Bressan *et al.*, 2018) and the fact that the *lhca4* mutant is the most affected for state transitions (Benson *et al.*, 2015), we can speculate that the additional Lhca1-a4 dimer that we describe here could serve as a binding site for loosely bound LHCII trimers. Note that in the case of PSII, loosely bound LHCII do exist: these LHCII can efficiently transfer energy to PSII, but their binding is too weak to allow their purification with PSII after membrane solubilization (Dekker and Boekema, 2005). A similar interaction could exist in the case of PSI.

It should also be noted that the abundance of a particular complex on a native gel is not necessarily dependent on its abundance *in vivo*, since particular weak interactions (but stable *in vivo*) can be lost after detergent solubilization. This is indeed the case of the well characterized Arabidopsis PSI-LHCII complex, which is highly abundant after digitonin solubilization, but it is totally lost with α-DDM solubilization (Galka *et al.*, 2012). Therefore, it is possible that the PSI-6LHCI-LHCII characterized here could play an important role *in vivo*, despite the apparently low abundance *in vitro*. This complex could be also part of a larger megacomplex either with only extra LHCII (as discussed above) or associated with unidentified complexes that detach during purification. We do however exclude the possibility that this larger PSI

supercomplex could be part of other recently described PSI megacomplexes, such as PSI-NDH (Kouřil *et al.*, 2014) or PSI-Cytb_6f (Yadav *et al.*, 2017). In the case of PSI-NDH, the interaction between PSI and NDH involves other parts of the PSI complex and would involve the low abundant Lhca5 and Lhca6 isoforms located at different positions to the Lhca1-a4 dimer identified here. Specifically, the Lhca5 an Lhca6 proteins are proposed to be next to the Lhca2-a3 dimer (Kouřil *et al.*, 2014), or at the place of Lhca4 and Lhca2, respectively (Otani *et al.*, 2018). In the case of the PSI-Cytb_6f complex, EM visualization indicates that the interaction between the two complexes is next to normal Lhca1 and far away from the position of the extra Lhca1-a4 extra dimer described here, and the dimer itself was also not detected (Yadav *et al.*, 2017).

In our work, using a biochemical and electron microscopy approach, we demonstrated that an additional Lhca dimer can bind to plant PSI on the PsaB-PsaI-PsaH side, a feature previously considered exclusive of algal PSI after recent structural works (Qin *et al.*, 2015; Mazor *et al.*, 2017; Pi *et al.*, 2018; Pan *et al.*, 2018; Su *et al.*, 2019; Suga *et al.*, 2019; Pan *et al.*, 2019). However, it should be noted that our results show differences between plant and algal PSI: in Arabidopsis, the extra dimer is composed of two Lhca isoforms (Lhca1 and Lhca4) that are already present in the "normal" row of Lhca proteins, and not by specific isoforms such as Lhca2-a9 in *Chlamydomonas* (Figure 5f). The extra dimer also seems shifted and closer to PsaH with only one monomer in a similar position to the algal PSI (Figure 5e-f). The extra dimer seems to be in contact with the LHCII trimer bound to the PsaO side and the presence of LHCII seems necessary for the stable binding of the Lhca dimer (Figure 6), which is not the case for algal PSI. Negative staining EM is a relatively simple structural method, but it cannot provide detailed structural features due to the limited resolution of the approach (at best ~12 Å). Future work using a cryo-EM approach, not trivial for this complex due to its low abundance and difficult purification, will allow the clarification of the exact interactions between the extra dimer and plant PSI. Future work will allow other open questions about the PSI-6LHCI-LHCII complex to be answered: What is the origin of the additional Lhca dimer? And what is the functional role of the PSI-6LHCI-LHCII complex? It is possible that a larger PSI supercomplex is useful under particular conditions of light quality or intensity. An interesting result in this direction is the recent demonstration in the moss *Physcomitrella patens* that PSI has an extended Lhc system under low light (Pinnola *et al.*, 2018; Iwai *et al.*, 2018). It is also possible that Lhca content per PSI might change under different light conditions, as previously proposed for Arabidopsis (Bailey *et al.*, 2001). In conclusion, our work adds important insights into PSI function and opens new questions about the complex structural plasticity of photosystems, a feature that is necessary for optimizing photosynthesis under variable environmental conditions.

EXPERIMENTAL PROCEDURES

Clear Native Gel Electrophoresis (CN-PAGE) and supercomplex isolation

Thylakoid membranes were prepared from *Arabidopsis thaliana* WT Col-0 in state 1 (obtained after dark incubation for several hours) and in state 2 (Galka *et al.*, 2012), or from *pph1* mutant plants (also called *tap38*), which are blocked in state 2 (Shapiguzov *et al.*, 2010; Pribil *et al.*, 2010), as described previously (Galka *et al.*, 2012). Plants were grown 5-6 weeks under short-day (8 hr light) conditions with ~120 μmol photons $m^{-2}s^{-1}$ illumination at 22°C.

Membranes were solubilized using the following protocol: the thylakoids were washed in a buffer containing 25 mM BisTris HCl pH 7 and 20% w/v glycerol (Järvi *et al.*, 2011), then centrifuged for 10 min at 12 000 g. Membranes were then resuspended in the same buffer at a concentration of 1 mg/mL, and the same volume of detergent solution was added to obtain final concentrations of 0.5 mg/mL of chlorophyll in 0.6% (w/v) digitonin and 0.1% (w/v) n-dodecyl-α-D-maltoside (α-DDM). After 20 min of solubilization on ice, sodium deoxycholate (DOC), a mild ionic detergent necessary for electrophoretic mobility, was added to a final concentration of 0.2% (w/v).

The solubilized complexes were separated on homemade Clear-Native Polyacrylamide Gels (CN-PAGE) in a similar way as described before (Järvi *et al.*, 2011) using a 3.30%-10% acrylamide gradient (%C=3), or on precast gels (Invitrogen 3-12% NativePAGE BisTris Gels; used for Figure 1). The electrode buffers contained 50 mM BisTris, 50 mM Tricine and the cathode buffer also 0.1% DOC and 0.04% α-DDM, a higher concentration than the usual protocol (Järvi *et al.*, 2011). This change allows a better separation between PSI and PSII supercomplexes in the region of interest (Figure S2a-b), probably by modification of micelle composition and size. For preparation of the complexes, the gel bands were cut from the gel, chopped and incubated overnight in a buffer containing 10 mM Hepes KOH pH 7.5 and 0.01% digitonin, in the cold with slow agitation. After this incubation, the samples were briefly centrifuged and the supernatant containing the solubilized complexes was used for further experiments.

SDS-PAGE and immunoblot analyses

Complex composition was checked on SDS-PAGE. Different systems were used in order to better separate the interesting proteins (for instance Lhca or Lhcb proteins). We numbered here the different systems to easily refer in figure legends. System 1: Laemmli system with Acrylamide/Bis-acrylamide ratio of 75:1, total concentration 15% and 6 M urea; System 2: Laemmli system with Acrylamide/Bis-acrylamide ratio of 29:1, total concentration 14% and 2 M urea (this system allow separating Lhca and Lhcb proteins, which migrate very close in the gel and are partly superimposed in other gel systems).

For quantification of phosphorylation, the proteins were separated on a Laemmli system with Acrylamide/Bis-acrylamide ratio of 37.5:1, total a concentration of 12% with 6 M urea, 20 µM Phos-tag (Wako Pure Chemical Industries) and 80 µM of $MnCl_2$.

All gels were stained using Sypro Ruby.

2D gels were performed by cutting the lanes of the first dimension, incubating them for 30 min in the same buffer as the stacking gel added with 2% SDS, then blocking in the correct position on the second gel using a solution containing 0.6% agarose, the upper reservoir buffer and bromophenol blue.

Immunoblot analyses were performed using specific antibodies (Agrisera) and revelation was done by chemiluminescence using a Fusion FX7 revelation system (Vilber).

Absorption and fluorescence spectroscopy

Absorption spectra were measured on samples eluted from the CN-PAGE gel with a Cary 300 spectrophotometer (Varian). 77K fluorescence spectra were measured by using a Cary Eclipse fluorimeter (Varian) on eluted samples frozen in liquid nitrogen. Emission spectra were

measured in the range 600-800 nm using an excitation at 440 nm and excitation spectra were measured in the range 350-550 nm by monitoring emission at 730 nm.

Electron Microscopy

Eluted complexes for CN-PAGE were visualized by negative staining electron microscopy with 2% uranyl acetate on glow-discharged carbon-coated copper grids. Images were collected on a FEI Tecnai G2 20 TWIN instrument operating at 200 kV. Images of 4096 × 4096 pixels were recorded at × 67400 magnification using a Gatan OneView camera with a pixel size of 0.222 nm. Image treatment, automatic single particle picking and 2D classification were performed with Relion 3.0-beta-2 and Gctf software (Scheres, 2012; Zhang, 2016; Zivanov *et al.*, 2018).

ACKNOWLEDGMENTS

Roberta Croce (VU University Amsterdam) is thanked for the spectrum of the Lhca1-a4 dimer. Ben Field (LGBP) is thanked for ms revision. This work was partly supported by the French National Research Agency Grant ANR-2012-JCJC-0001-01 to SC. AC is now supported by the Grant Agency of the Czech Republic (project 17-02363Y). ED is supported by the "Institut National de la Santé et de la Recherche Médicale" (INSERM). ZK was supported by the Grant Agency of the Czech Republic (project 18-12178S/P501).

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

AUTHOR CONTRIBUTIONS

SC and AC conceived the experiments; AC, ZK and SC performed the biochemical experiments; SC and AC performed the EM experiments with technical help from ED; AK performed the EM data acquisitions; SC and AC analyzed the data; SC and AC wrote the article.

DATA AVAILABILITY STATEMENT

Data supporting the findings of this work are provided in the main text and the supporting information files. All data and materials used in this study will be available from the corresponding authors.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. 2D of the full lane of solubilized *pph1* membranes.

Figure S2. Optimization of the CN-PAGE and preparative CN-PAGE

Figure S3. EM projection maps obtained on PSI-LHCII and PSI-6LHCI-LHCII fractions.

Figure S4. Simulated PSI-6LHCI-LHCII and PSI-2LHCII spectra.

REFERENCES

Bailey, S., Walters, R.G., Jansson, S. and Horton, P. (2001) Acclimation of Arabidopsis thaliana to the light environment: The existence of separate low light and high light responses. *Planta*, **213**, 794–801. 10.1007/s004250100556.

Bell, A.J., Frankel, L.K. and Bricker, T.M. (2015) High yield non-detergent isolation of photosystem I-light-harvesting chlorophyll II membranes from spinach thylakoids: Implications for the organization of the PS I antennae in higher plants. *J. Biol. Chem.*, **290**, 18429–18437. 10.1074/jbc.M115.663872.

Bellaflore, S., Barneche, F., Peltler, G. and Rochalx, J.D. (2005) State transitions and light adaptation require chloroplast thylakoid protein kinase STN7. *Nature*, **433**, 892–895. 10.1038/nature03286.

Benson, S.L., Maheswaran, P., Ware, M.A., Hunter, C.N., Horton, P., Jansson, S., Ruban, A. V. and Johnson, M.P. (2015) An intact light harvesting complex i antenna system is required for complete state transitions in Arabidopsis. *Nat. Plants*, **1**. 10.1038/nplants.2015.176.

Bos, I., Bland, K.M., Tian, L., Croce, R., Frankel, L.K., Amerongen, H. van, Bricker, T.M. and Wientjes, E. (2017) Multiple LHCII antennae can transfer energy efficiently to a single Photosystem I. *Biochim. Biophys. Acta - Bioenerg.*, **1858**, 371–378. 10.1016/j.bbabio.2017.02.012.

Bressan, M., Bassi, R. and Dall'Osto, L. (2018) Loss of LHCI system affects LHCII re-distribution between thylakoid domains upon state transitions. *Photosynth. Res.*, **135**, 251–261. 10.1007/s11120-017-0444-1.

Busch, A. and Hippler, M. (2011) The structure and function of eukaryotic photosystem i. *Biochim. Biophys. Acta - Bioenerg.*, **1807**, 864–877. 10.1016/j.bbabio.2010.09.009.

Caffarri, S., Kouřil, R., Kereïche, S., Boekema, E.J. and Croce, R. (2009) Functional architecture of higher plant photosystem II supercomplexes. *EMBO J.*, **28**, 3052–3063. 10.1038/emboj.2009.232.

Crepin, A. and Caffarri, S. (2018) Functions and evolution of Lhcb isoforms composing LHCII, the major light harvesting complex of Photosystem II of green eukaryotic organisms. *Curr. Protein Pept. Sci.*, **19**. 10.2174/1389203719666180222101534.

Crepin, A. and Caffarri, S. (2015) The specific localizations of phosphorylated Lhcb1 and Lhcb2 isoforms reveal the role of Lhcb2 in the formation of the PSI-LHCII supercomplex in Arabidopsis during state transitions. *Biochim. Biophys. Acta - Bioenerg.*, **1847**, 1539–1548. 10.1016/j.bbabio.2015.09.005.

Crepin, A., Santabarbara, S. and Caffarri, S. (2016) Biochemical and spectroscopic characterization of highly stable photosystem II supercomplexes from arabidopsis. *J. Biol. Chem.*, **291**, 19157–19171. 10.1074/jbc.M116.738054.

Croce, R., Zucchelli, G., Garlaschi, F.M. and Jennings, R.C. (1998) A thermal broadening study of the antenna chlorophylls in PSI- 200, LHCI, and PSI core. *Biochemistry*, **37**, 17355–17360. 10.1021/bi9813227.

Dekker, J.P. and Boekema, E.J. (2005) Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim. Biophys. Acta - Bioenerg.*, **1706**, 12–39. 10.1016/j.bbabio.2004.09.009.

Galka, P., Santabarbara, S., Khuong, T.T.H., Degand, H., Morsomme, P., Jennings, R.C., Boekema, E.J. and Caffarria, S. (2012) Functional analyses of the plant photosystem I-light-harvesting complex II supercomplex reveal that light-harvesting complex II loosely bound to photosystem ii is a very efficient antenna for photosystem I in state II. *Plant Cell*, **24**, 2963–2978. 10.1105/tpc.112.100339.

Iwai, M., Grob, P., Iavarone, A.T., Nogales, E. and Niyogi, K.K. (2018) A unique supramolecular organization of photosystem I in the moss Physcomitrella patens. *Nat. Plants*, **4**, 904–909. 10.1038/s41477-018-0271-1.

Järvi, S., Suorsa, M., Paakkarinen, V. and Aro, E.M. (2011) Optimized native gel systems for separation of thylakoid protein complexes: Novel super- and mega-complexes. *Biochem. J.*, **439**, 207–214. 10.1042/BJ20102155.

Jensen, P.E., Bassi, R., Boekema, E.J., Dekker, J.P., Jansson, S., Leister, D., Robinson, C. and Scheller, H.V. (2007) Structure, function and regulation of plant photosystem I. *Biochim. Biophys. Acta - Bioenerg.*, **1767**, 335–352. 10.1016/j.bbabio.2007.03.004.

Koskela, M.M., Brünje, A., Ivanauskaite, A., et al. (2018) Chloroplast acetyltransferase NSI is

required for state transitions in arabidopsis Thaliana. *Plant Cell*, **30**, 1695–1709. 10.1105/tpc.18.00155.

Kouřil, R., Strouhal, O., Nosek, L., Lenobel, R., Chamrád, I., Boekema, E.J., Šebela, M. and Ilík, P. (2014) Structural characterization of a plant photosystem i and NAD(P)H dehydrogenase supercomplex. *Plant J.*, **77**, 568–576. 10.1111/tpj.12402.

Kouřil, R., Zygadlo, A., Arteni, A.A., Wit, C.D. De, Dekker, J.P., Jensen, P.E., Scheller, H.V. and Boekema, E.J. (2005) Structural characterization of a complex of photosystem I and light-harvesting complex II of Arabidopsis thaliana. *Biochemistry*, **44**, 10935–10940. 10.1021/bi051097a.

Krumova, S.B., Várkonyi, Z., Lambrev, P.H., Kovács, L., Todinova, S.J., Busheva, M.C., Taneva, S.G. and Garab, G. (2014) Heat- and light-induced detachment of the light-harvesting antenna complexes of photosystem i in isolated stroma thylakoid membranes. *J. Photochem. Photobiol. B Biol.*, **137**, 4–12. 10.1016/j.jphotobiol.2014.04.029.

Leoni, C., Pietrzykowska, M., Kiss, A.Z., Suorsa, M., Ceci, L.R., Aro, E.M. and Jansson, S. (2013) Very rapid phosphorylation kinetics suggest a unique role for Lhcb2 during state transitions in Arabidopsis. *Plant J.*, **76**, 236–246. 10.1111/tpj.12297.

Mazor, Y., Borovikova, A., Caspy, I. and Nelson, N. (2017) Structure of the plant photosystem i supercomplex at 2.6 Å resolution. *Nat. Plants*, **3**. 10.1038/nplants.2017.14.

Morosinotto, T., Breton, J., Bassi, R. and Croce, R. (2003) The nature of a chlorophyll ligand in Lhca proteins determines the far red fluorescence emission typical of photosystem I. *J. Biol. Chem.*, **278**, 49223–49229. 10.1074/jbc.M309203200.

Nellaepalli, S., Zsiros, O., Tóth, T., Yadavalli, V., Garab, G., Subramanyam, R. and Kovács, L. (2014) Heat- and light-induced detachment of the light harvesting complex from isolated photosystem i supercomplexes. *J. Photochem. Photobiol. B Biol.*, **137**, 13–20. 10.1016/j.jphotobiol.2014.04.026.

Otani, T., Kato, Y. and Shikanai, T. (2018) Specific substitutions of light-harvesting complex I proteins associated with photosystem I are required for supercomplex formation with chloroplast NADH dehydrogenase-like complex. *Plant J.*, **94**, 122–130. 10.1111/tpj.13846.

Ozawa, S.I., Bald, T., Onishi, T., Xue, H., Matsumura, T., Kubo, R., Takahashi, H., Hippler, M. and Takahashi, Y. (2018) Configuration of Ten Light-Harvesting Chlorophyll a/b Complex I Subunits in Chlamydomonas reinhardtii Photosystem I. *Plant Physiol.*, **178**, 583–595. 10.1104/pp.18.00749.

Pan, X., Cao, P., Su, X., Liu, Z. and Li, M. (2019) Structural analysis and comparison of light-harvesting complexes I and II. *Biochim. Biophys. Acta - Bioenerg.* 10.1016/j.bbabio.2019.06.010.

Pan, X., Ma, J., Su, X., Cao, P., Chang, W., Liu, Z., Zhang, X. and Li, M. (2018) Structure of the maize photosystem I supercomplex with light-harvesting complexes I and II. *Science (80-.).*, **360**, 1109–1113. 10.1126/science.aat1156.

Pesaresi, P., Lunde, C., Jahns, P., et al. (2002) A stable LHCII-PSI aggregate and suppression of photosynthetic state transitions in the psae1-1 mutant of Arabidopsis thaliana. *Planta*, **215**, 940–948. 10.1007/s00425-002-0835-0.

Pi, X., Tian, L., Dai, H.E., Qin, X., Cheng, L., Kuang, T., Sui, S.F. and Shen, J.R. (2018) Unique organization of photosystem I–light-harvesting supercomplex revealed by cryo-EM from a red alga. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, 4423–4428. 10.1073/pnas.1722482115.

Pinnola, A., Alboresi, A., Nosek, L., et al. (2018) A LHCB9-dependent photosystem I megacomplex induced under low light in Physcomitrella patens. *Nat. Plants*, **4**, 910–919. 10.1038/s41477-018-0270-2.

Pribil, M., Pesaresi, P., Hertle, A., Barbato, R. and Leister, D. (2010) Role of plastid protein phosphatase TAP38 in LHCII dephosphorylation and thylakoid electron flow. *PLoS Biol.*, **8**. 10.1371/journal.pbio.1000288.

Qin, X., Pi, X., Wang, W., et al. (2019) Structure of a green algal photosystem I in complex with a large number of light-harvesting complex I subunits. *Nat. Plants*, **5**, 263–272. 10.1038/s41477-019-0379-y.

Qin, X., Suga, M., Kuang, T. and Shen, J.R. (2015) Structural basis for energy transfer pathways in the plant PSI-LHCI supercomplex. *Science (80-.).*, **348**, 989–995. 10.1126/science.aab0214.

Rochaix, J.-D. (2014) Regulation and Dynamics of the Light-Harvesting System. *Annu. Rev. Plant Biol.*, **65**, 287–309. 10.1146/annurev-arplant-050213-040226.

Santabarbara, S., Tibiletti, T., Remelli, W. and Caffarri, S. (2017) Kinetics and heterogeneity of

energy transfer from light harvesting complex II to photosystem i in the supercomplex isolated from Arabidopsis. *Phys. Chem. Chem. Phys.*, **19**, 9210–9222. 10.1039/c7cp00554g.

Scheres, S.H.W. (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.*, **180**, 519–530. 10.1016/j.jsb.2012.09.006.

Shapiguzov, A., Ingelsson, B., Samol, I., Andres, C., Kessler, F., Rochaix, J.D., Vener, A. V. and Goldschmidt-Clermont, M. (2010) The PPH1 phosphatase is specifically involved in LHCII dephosphorylation and state transitions in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 4782–4787. 10.1073/pnas.0913810107.

Su, X., Ma, J., Pan, X., Zhao, X., Chang, W., Liu, Z., Zhang, X. and Li, M. (2019) Antenna arrangement and energy transfer pathways of a green algal photosystem-I–LHCI supercomplex. *Nat. Plants*, **5**, 273–281. 10.1038/s41477-019-0380-5.

Suga, M., Ozawa, S.I., Yoshida-Motomura, K., Akita, F., Miyazaki, N. and Takahashi, Y. (2019) Structure of the green algal photosystem I supercomplex with a decameric light-harvesting complex I. *Nat. Plants*, **5**, 626–636. 10.1038/s41477-019-0438-4.

Wientjes, E. and Croce, R. (2011) The light-harvesting complexes of higher-plant Photosystem I: Lhca1/4 and Lhca2/3 form two red-emitting heterodimers. *Biochem. J.*, **433**, 477–485. 10.1042/BJ20101538.

Wientjes, E., Stokkum, I.H.M. Van, Amerongen, H. Van and Croce, R. (2011) Excitation-energy transfer dynamics of higher plant photosystem i light-harvesting complexes. *Biophys. J.*, **100**, 1372–1380. 10.1016/j.bpj.2011.01.030.

Yadav, K.N.S., Semchonok, D.A., Nosek, L., Kouřil, R., Fucile, G., Boekema, E.J. and Eichacker, L.A. (2017) Supercomplexes of plant photosystem I with cytochrome b6f, light-harvesting complex II and NDH. *Biochim. Biophys. Acta - Bioenerg.*, **1858**, 12–20. 10.1016/j.bbabio.2016.10.006.

Zhang, K. (2016) Gctf: Real-time CTF determination and correction. *J. Struct. Biol.*, **193**, 1–12. 10.1016/j.jsb.2015.11.003.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J.H., Lindahl, E. and Scheres, S.H.W. (2018) New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife*, **7**. 10.7554/eLife.42166.

FIGURES LEGENDS

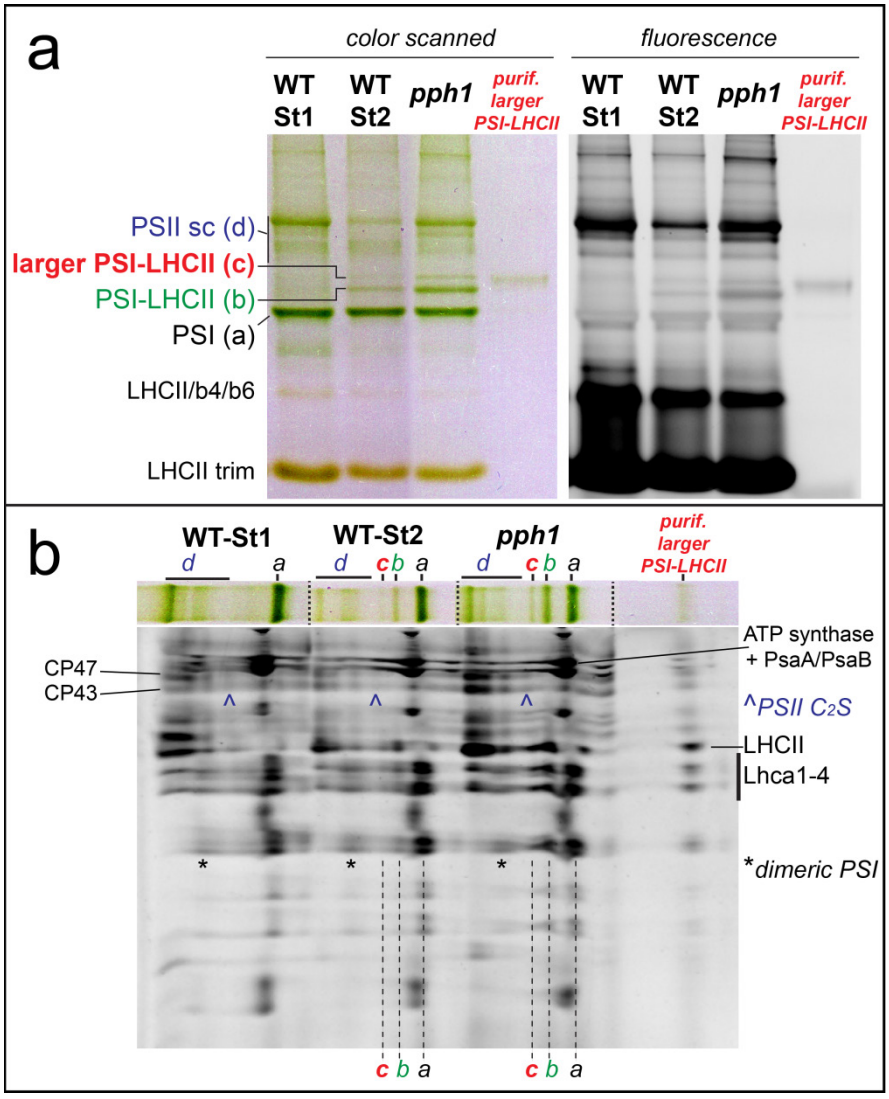

Figure 1. Identification of an undescribed larger PSI-LHCII supercomplex.
a) CN-PAGE of solubilized membranes from state 1 and state 2 WT plants and from *pph1* plants blocked in state 2. In state 2 samples an uncharacterized green band (band *c*) is visible above the PSI-LHCII complex (Galka *et al.*, 2012). This band has low fluorescence compared to PSII supercomplexes (PSII sc). The purified supercomplex was also loaded on the gel (right lane). b) Second dimension by SDS-PAGE (system 1, see Methods) of the lanes of panel A. For simplicity, each green band is indicated with a letter (see panel A). The band of the supercomplex of interest (c) has a similar profile as the PSI-LHCII supercomplex and contains both LHCII and PSI subunits (the region of Lhca is indicated), suggesting that the complex is a larger PSI-LHCII supercomplex. CP47 and CP43 are also indicated to localize PSII supercomplexes: note in particular that the PSII-C_2S supercomplex migrates very close to the newly identified larger PSI-LHCII.

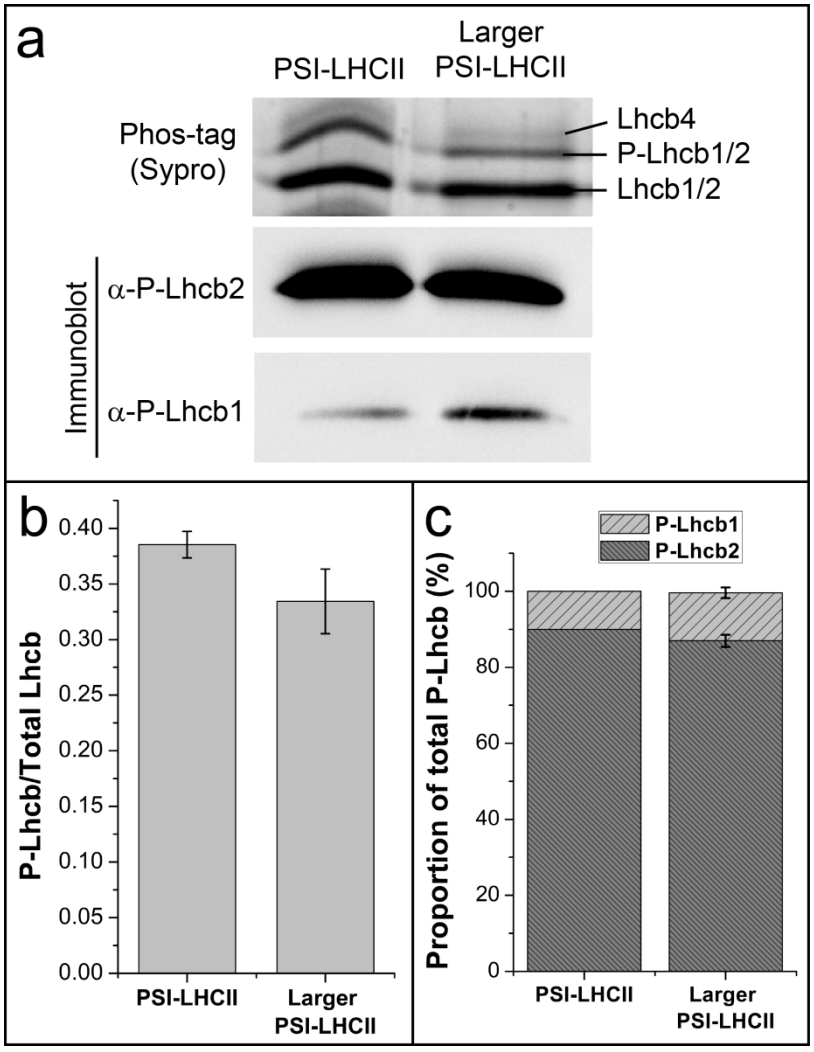

Figure 2. LHCII phosphorylation in the typical and in the larger PSI-LHCII complexes
a) Phosphorylation of the larger PSI-LHCII supercomplex compared to the PSI-LHCII. Purified supercomplexes were loaded in a similar amount for LHCII content (0.03 μg in Chls) on SDS-PAGE containing Phos-tag, allowing separation of phosphorylated LHCII from its non-phosphorylated form. Immunoblots using specific antibodies against P-Lhcb2 and P-Lhcb1 isoforms were also performed, revealing a similar phosphorylation in both complexes. b) Quantification of the total phosphorylation, as estimated from Phos-tag gels. In both complexes, the quantification indicates the phosphorylation of around one Lhcb monomer per LHCII trimer. Error bars: SE, n=4. c) Quantification of Lhcb2 and Lhcb1 phosphorylation, based on immunoblot results. The values for the PSI-LHCII supercomplex were established to be 90% P-Lhcb2 and 10% P-Lhcb1, based on previous results (Crepin and Caffarri, 2015). The values for the larger supercomplex are calculated from the relative amount of each isoform normalized to PSI-LHCII. Error bars: SE, n=3.

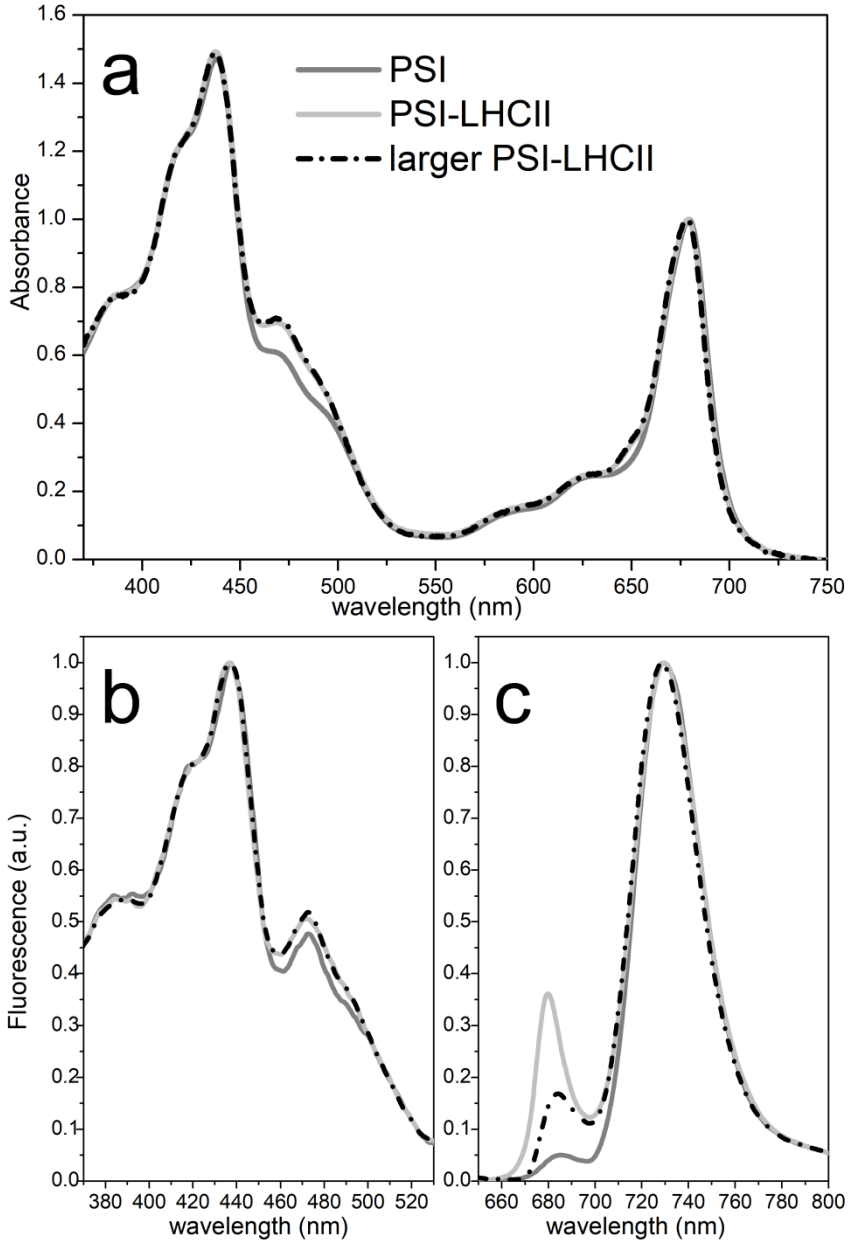

Figure 3. Absorption and fluorescence spectra.
a) Absorption spectrum of the purified larger PSI-LHCII supercomplex compared with spectra of PSI and PSI-LHCII. Note the similarity between the larger and the normal PSI-LHCII supercomplexes. b) Excitation spectra (detection at 730 nm) of the same samples as in panel A. Note that excitation spectra resemble to absorption spectra, indicating the LHCII in the larger complex is energetically well connected to PSI. c) Fluorescence emission spectra (excitation at 440 nm) of the three samples. The emission at 679 nm for the PSI-LHCII sample is due to some disconnected LHCII (probably during elution from CN-PAGE); in the case of the larger PSI-LHCII sample, the left peak (684 nm) is mainly due to a minor contamination from PSII supercomplexes closely migrating on the CN-PAGE. Spectra are normalized at the maximum in the red region (absorption and emission) and in the Soret (excitation).

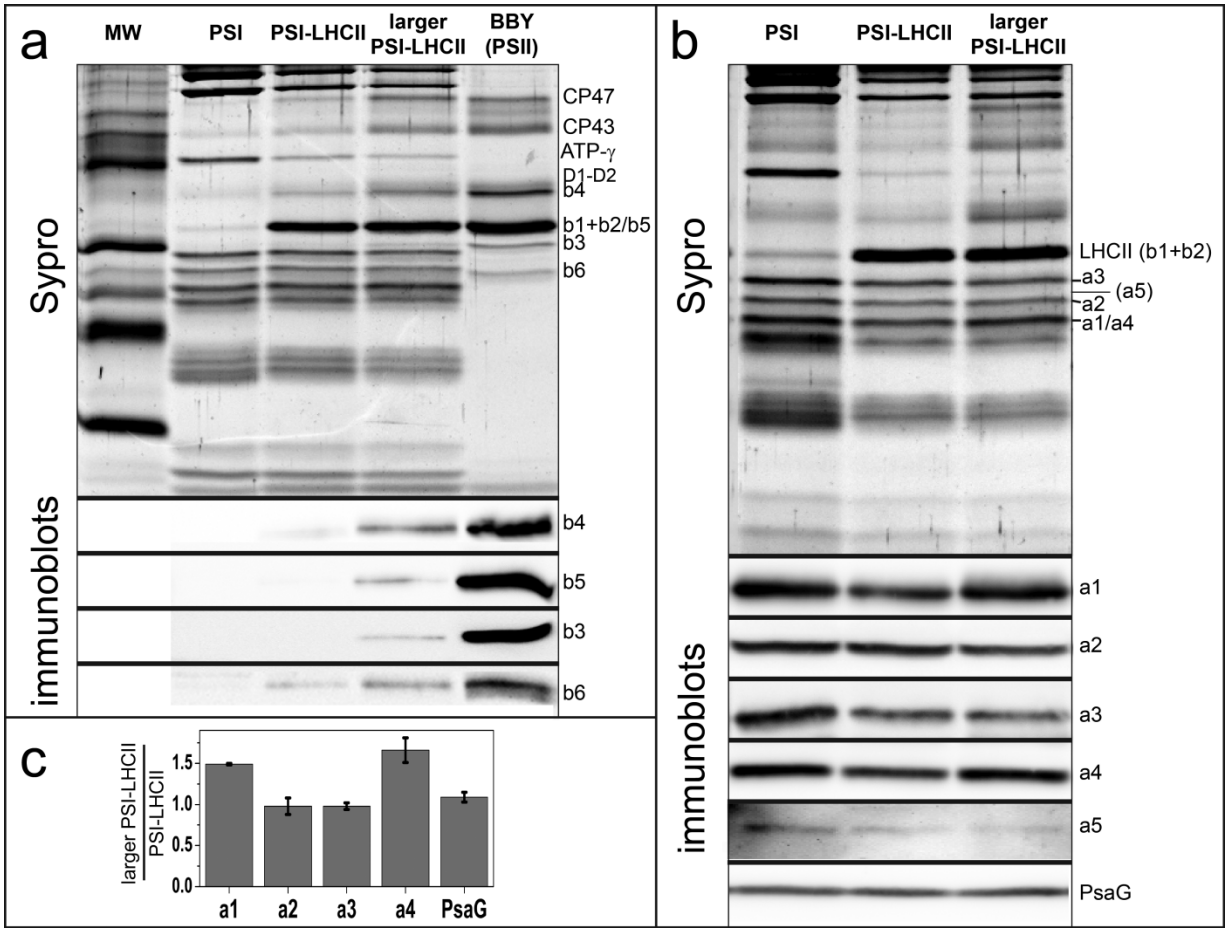

Figure 4. Identification of the extra subunits in the larger PSI-LHCII supercomplex.
a) Analysis of Lhcb proteins. Similar amounts of PSI supercomplexes (0.2 μg Chl) were separated by SDS-PAGE (system 2, see Methods) together with PSII membranes (0.1 μg Chl) to obtain similar amounts of Lhca and Lhcb monomeric proteins. Both Sypro staining and immunoblotting with specific antibodies revealed that no Lhcb component is present in a significantly higher amount in the larger PSI-LHCII supercomplex as compared to the typical PSI-LHCII. A similar and small increase of all Lhcb proteins in the larger PSI-LHCII complex is due to a minor contamination by closely migrating PSII in the CN-PAGE (see also Figure 1); b) Analysis for Lhca and PsaG proteins. Lhca1-5 proteins and the PsaG subunit were tested by Sypro staining and immunoblotting with specific antibodies. The Sypro stained gel revealed that only one band had an increased relative intensity (~1.5 times) in the larger PSI-LHCII complex as compared with PSI and PSI-LHCII complexes. This band corresponds to Lhca1 and Lhca4 (co-migrating in this gel system) and the result was confirmed by immunoblot analyses. The position of each Lhca protein on the Sypro stained gel was obtained by successive incubation of membranes with different antibodies and careful alignment of gels and immunoblots. c) Relative increase (±SD) of Lhca1-4 and PsaG proteins in the larger PSI-LHCII supercomplex as compared with the classic PSI-LHCII. Gels and immunoblots were repeated at least in triplicate on two different preparations (excluding Lhca5, not further analyzed as it was almost undetectable by immunoblot and by Sypro staining, in agreement with its very low abundance).

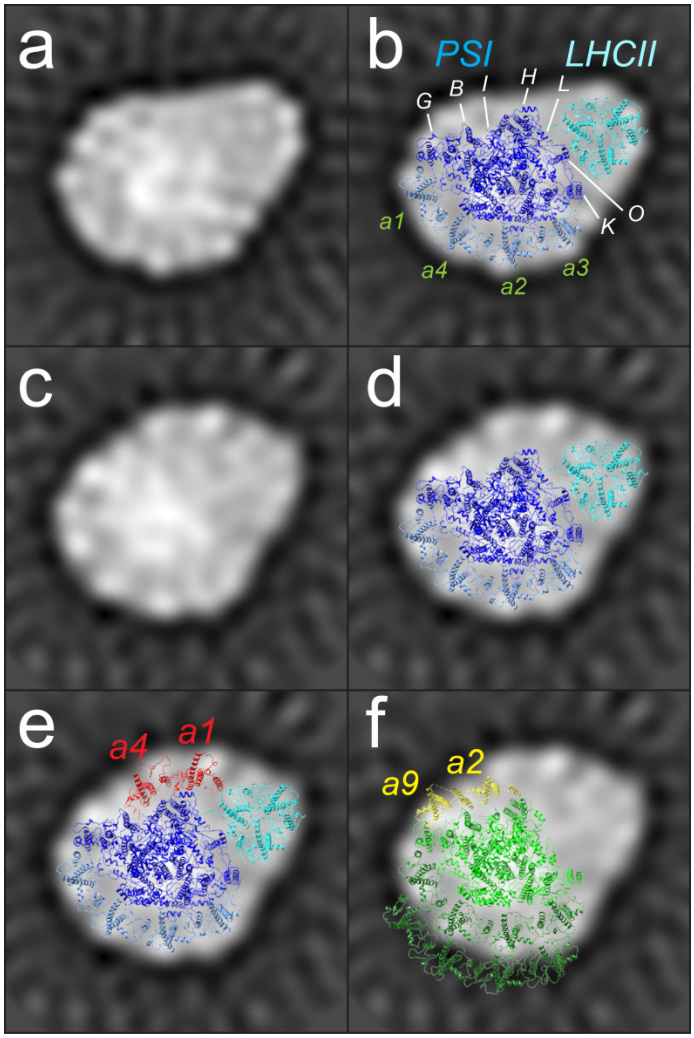

Figure 5. Single particle EM analysis

a) Negative staining single particle electron microscopy (EM) analysis of the eluted PSI-LHCII fraction showed that the main component was the well characterized particle composed of PSI and a single LHCII trimer (Kouřil *et al.*, 2005; Galka *et al.*, 2012). b) The atomic model of PSI-LHCII (5ZJI) (Pan *et al.*, 2018) is superimposed on the EM density map to show the position of LHCII and of the Lhca1-4 subunits (view from the stromal side). c) EM analysis of the PSI-6LHCI-LHCII fraction showed the presence of a particle never described previously: a particle with a large extra density on the PsaB-PsaI-PsaH side. d) The atomic model of PSI-LHCII is superimposed to highlight the position of the extra density in the PSI-6LHCI-LHCII complex. e) According to biochemical data, the extra density is fitted with an Lhca1-4 dimer (in red): its position suggests an interaction of this extra dimer with the PSI core and the LHCII trimer. f) The atomic model of the PSI of *Chlamydomonas reinhardtii* (6JO5) (Suga *et al.*, 2019) is superimposed maintaining the same positions for core and internal Lhca as plant PSI in order to show the position of the additional algal Lhca dimer (Lhca2-9, in yellow). In the alga, the extra dimer is in a different position compared to the extra density in the Arabidopsis PSI-6LHCI-LHCII complex. Note that the outer row of algal Lhca proteins is absent in plant PSI. See Figure S3 for further information about EM analysis.

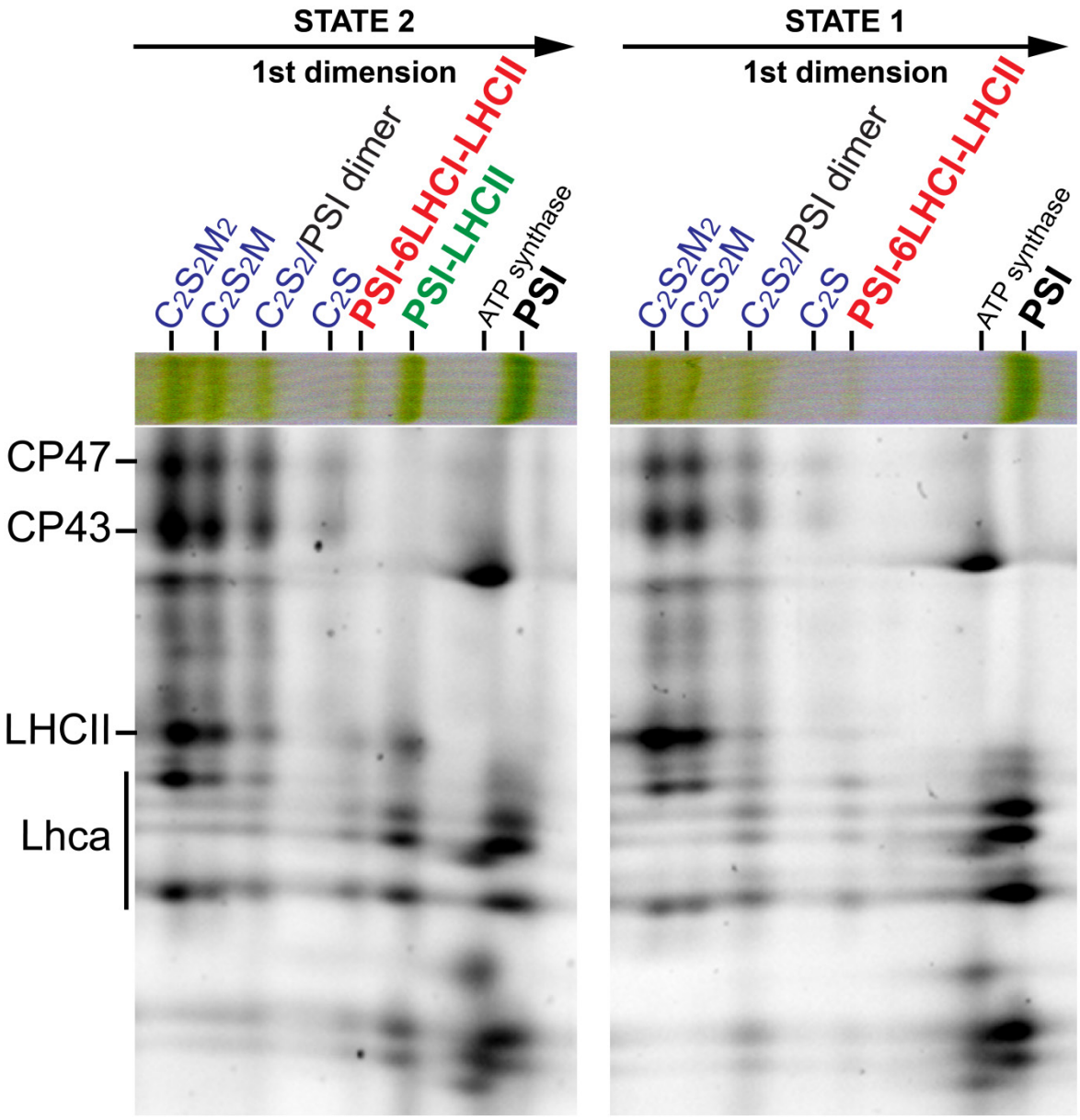

Figure 6. 2D gel of very mildly solubilized thylakoid membranes.
State 2 and State 1 membranes were very mildly solubilized using only digitonin as detergent and then separated in 2D on a CN-PAGE and on a SDS-PAGE (system 1, see Methods). State 2 membranes showed a similar profile as in Figure 1. From the bottom (right) of the first dimension: PSI, ATP synthase, PSI-LHCII, PSI-6LHCI-LHCII and the four PSII supercomplexes previously characterized (Caffarri *et al.*, 2009). State 1 membranes lacked the PSI-LHCII complex, but had a faint band corresponding to the PSI-6LHCI-LHCII complex. On the gel, CP43, CP47, LHCII and the Lhca region are indicated for supercomplex identification.

SUPPORTING INFORMATION

Isolation and characterization of a large Photosystem I-Light Harvesting complex II supercomplex with an additional Lhca1-a4 dimer in Arabidopsis

Aurélie Crepin, Zuzana Kučerová, Artemis Kosta, Eric Durand, Stefano Caffarri

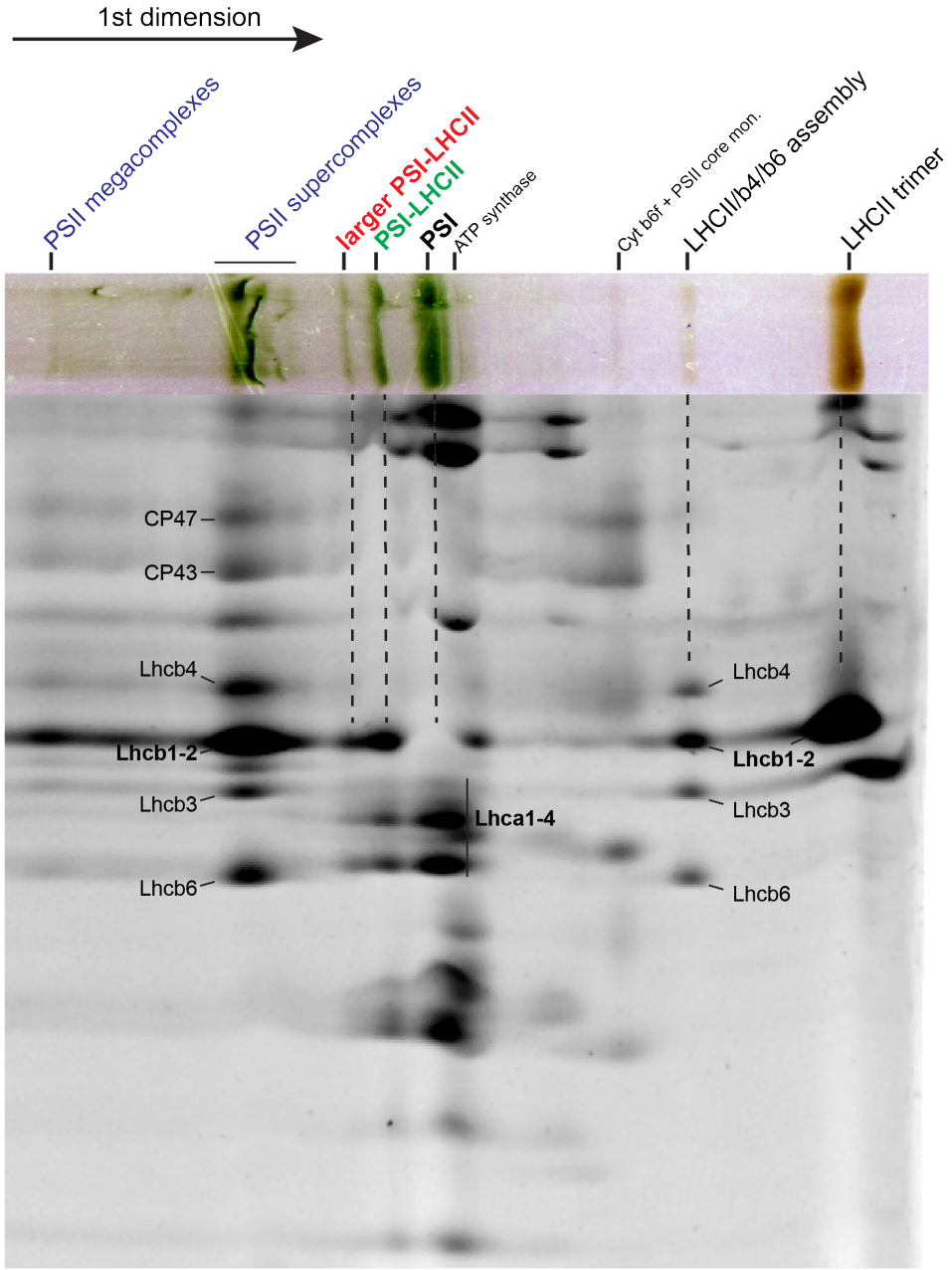

Figure S1. 2D of the full lane of solubilized *pph1* membranes.
2D gel (CN-PAGE + SDS-PAGE) using the same conditions as in Figure 1 showing the full lane for solubilized *pph1* membranes. The Lhca region is indicated, as well the position of Lhcb proteins, CP43 and CP47 for supercomplex identification.

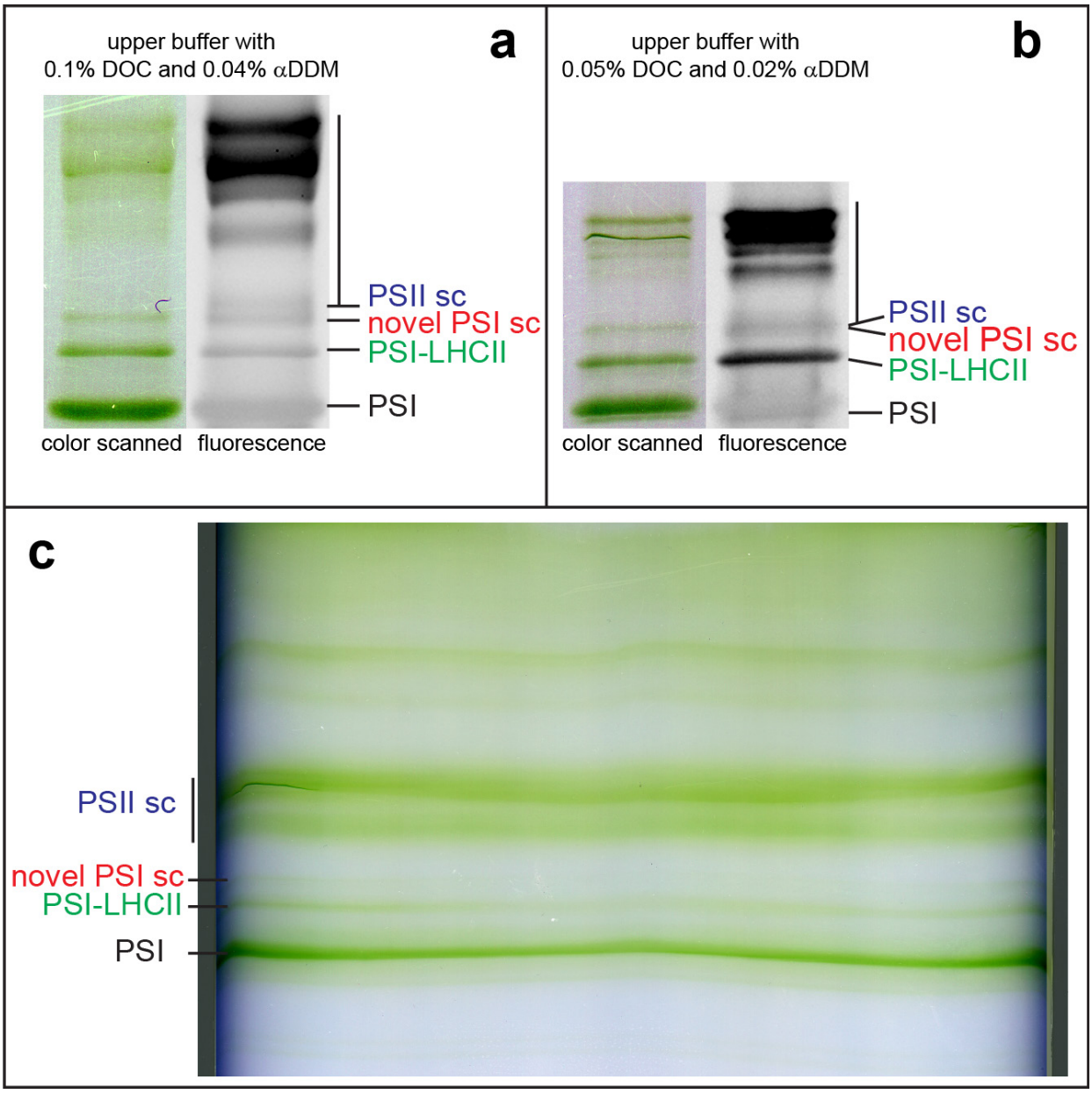

Figure S2. Optimization of the CN-PAGE and preparative CN-PAGE

a) Separation of PSI and PSII supercomplexes (sc) by native PAGE with an increased detergent concentration in the upper buffer allows a better separation between the larger PSI-LHCII and lowest PSII-C_2S supercomplex (the faint green band just above the uncharacterized PSI supercomplex better visible in fluorescence). b) Separation on the classical buffer system is shown for comparison: in this case the region of interest is less resolved. c) Preparative 20-cm gel: 1 mg of *pph1* thylakoids (expressed in total Chls) have been solubilized as described in Methods section and separated on the native gel run as in panel A. The regions corresponding to PSI, PSI-LHCII and the larger PSI-LHCII supercomplex have been cut and particles eluted by spontaneous diffusion in solution.

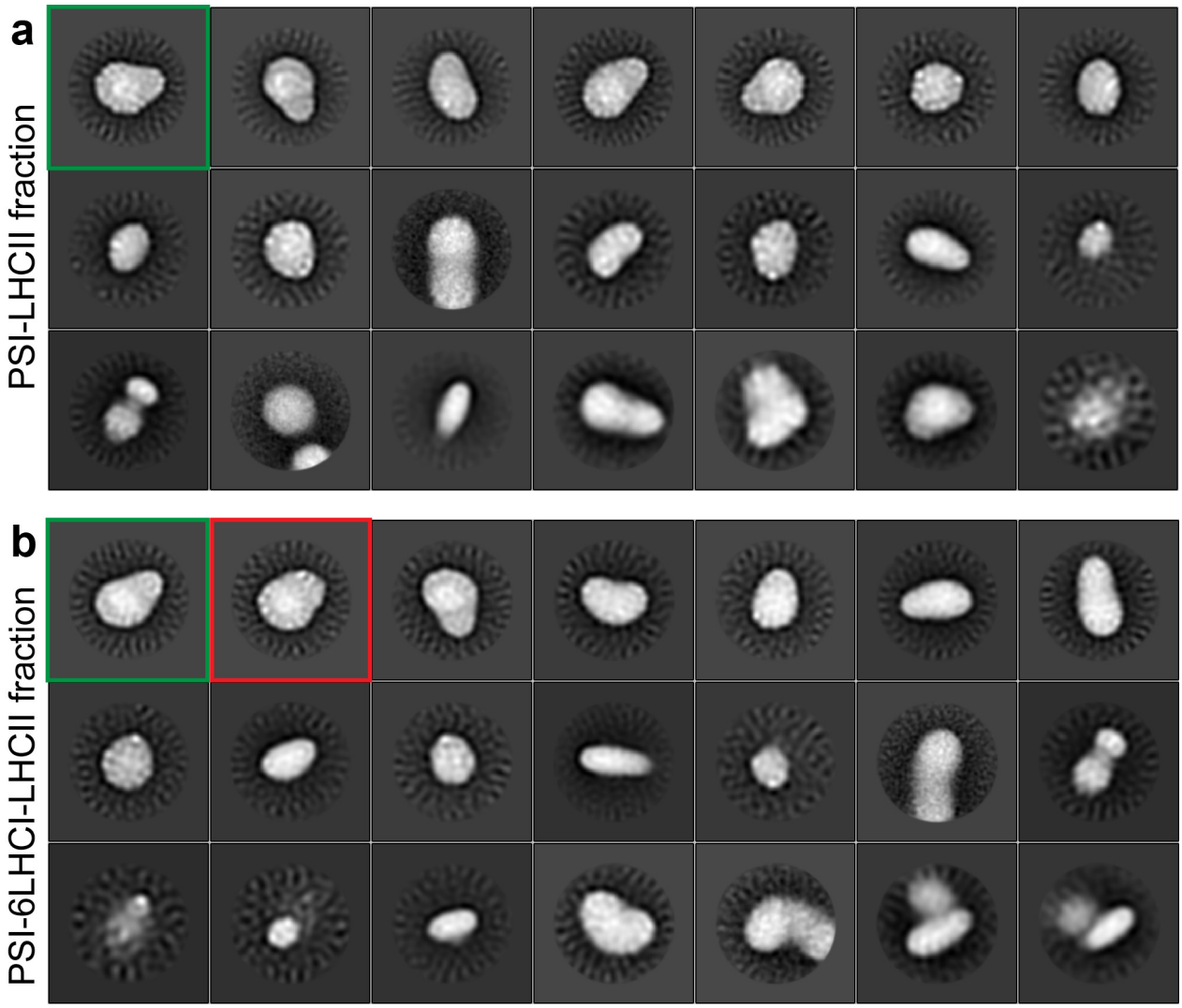

Figure S3. EM projection maps obtained on PSI-LHCII and PSI-6LHCI-LHCII fractions.
a) 2D classes (ordered by class abundance) obtained from ~25k particles auto-picked from 380 EM images on the PSI-LHCII fraction. The first 2D class (green square; average of 2022 particles) is the stromal view of the PSI-LHCII complex used in Figure 5. Other PSI-LHCII projections (lumenal view, as the second class, and tilted projections), probably some PSI core, ATP synthase and several junk particles (selected by the autopicking process and often representing aggregated particles or disassembled complexes) are also visible. b) 2D classes obtained from ~30k particles auto-picked from 530 EM images on the PSI-6LHCI-LHCII fraction. The second most abundant class (red square, average of 2203 particles) is represented by a larger PSI-LHCII complex not yet described in the literature, which is absent in the PSI-LHCII fraction and which is further analyzed in Figure 5. The presence of PSI-LHCII particles (class 1, green square, stromal view obtained from 2463 particles) indicates that the PSI-6LHCI-LHCII fraction is contaminated by the close migrating and more abundant PSI-LHCII complex. Tilted projections cannot be univocally assigned to PSI-LHCII or to PSI-6LHCI-LHCII complexes. Again, several junk particles are also visible.

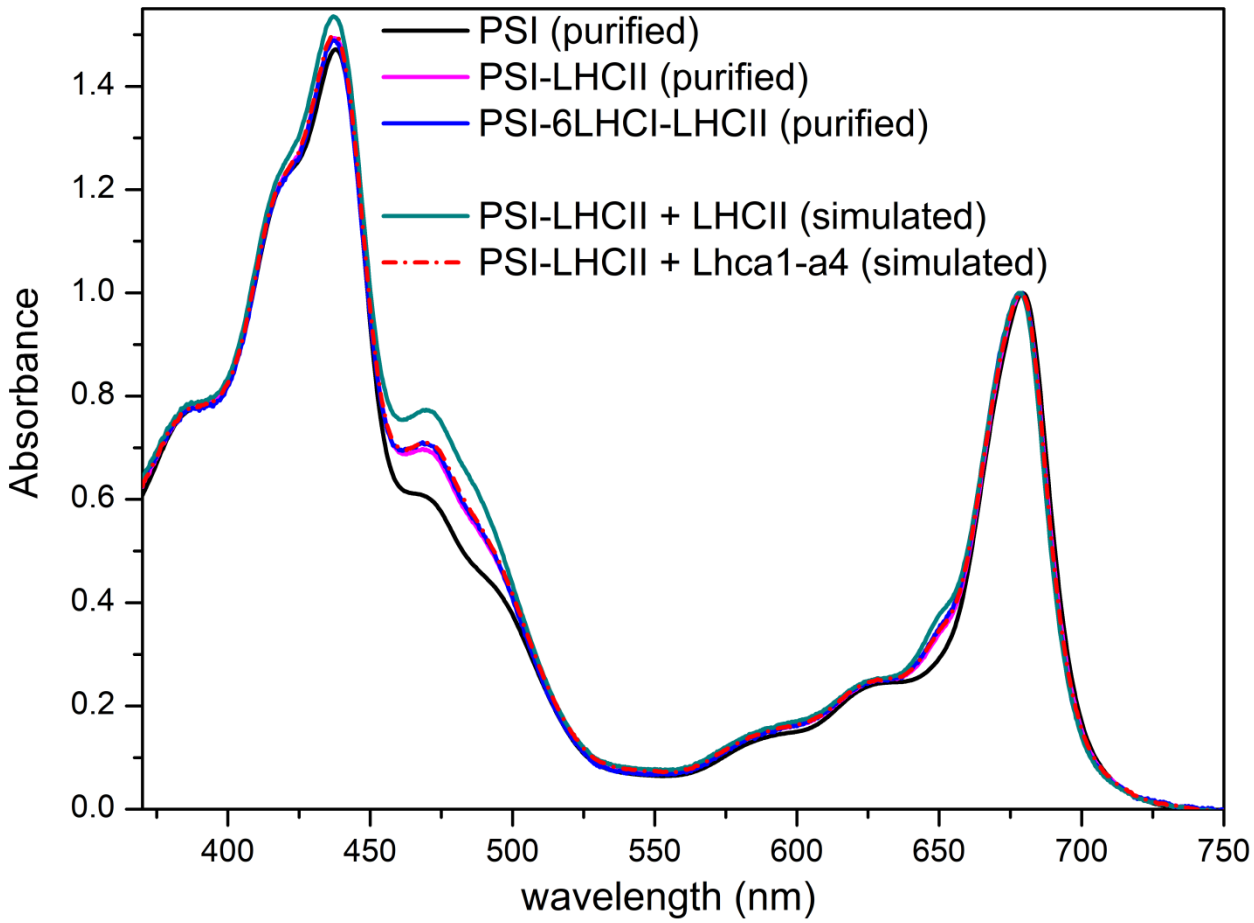

Figure S4. Simulated PSI-6LHCI-LHCII and PSI-2LHCII spectra.
In order to understand if the addition of a Lhca1-a4 dimer to the PSI-LHCII could be visible in the absorption spectrum, the spectra of the purified PSI, PSI-LHCII and PSI-6LHCI-LHCII supercomplexes are compared with the simulated spectra of the PSI-LHCII complex added with the spectrum of the Lhca1-a4 dimer (Wientjes *et al.*, 2011) or of the LHCII trimer (normalized to have the same stoichiometric amount). All the spectra are normalized to the maximum in the red region. While the addition of LHCII is clearly visible due to the high abundance of Chl b in the trimer, the simulated PSI-LHCII+Lhca1-a4 complex is almost identical to the purified PSI-LHCII and PSI-6LHCI-LHCII complexes, in agreement with experimental data.