DR GAL WITTENBERG (Orcid ID : 0000-0003-1037-0040)

DR ETIENNE MEYER (Orcid ID : 0000-0003-4712-9824)

Received Date : 16-Dec-2016
Revised Date : 29-Jan-2017
Accepted Date : 31-Jan-2017
Article type : Original Article

Identification and characterization of a stable intermediate in photosystem I assembly in tobacco

Gal Wittenberg[1], Sari Järvi[2], Marta Hojka[1], Szilvia Z. Tóth[1 a], Etienne H. Meyer[1], Eva-Mari Aro[2], Mark Aurel Schöttler[1] and Ralph Bock[1 *]

[1] Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany.

[2] Department of Biochemistry, Molecular Plant Biology, University of Turku, FI-20014 Turku, Finland.

[a] present address: Institute of Plant Biology, Biological Research Centre of the Hungarian Academy of Sciences, Szeged, Temesvári krt. 62, H-6726 Szeged, Hungary.

* Corresponding author: Ralph Bock, Max-Planck-Institut für Molekulare Pflanzenphysiologie, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany. Email: rbock@mpimp-golm.mpg.de. Tel: +49 331 567-8700. Fax: +49 331 567-8408

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/tpj.13505

Running title: A PSI assembly intermediate

Key words: photosystem assembly; photosynthesis; photosystem biogenesis; thylakoid membrane

ABSTRACT

PSI is the most efficient bioenergetic nanomachine in nature and one of the largest membrane protein complexes known. It is composed of 18 protein subunits that bind more than 200 co-factors and prosthetic groups. While the structure and function of PSI have been studied in great detail, very little is known about the PSI assembly process. In this work, we have characterized a PSI assembly intermediate in tobacco plants, which we named PSI*. We found PSI* to contain only a specific subset of the core subunits of PSI. PSI* is particularly abundant in young leaves where active thylakoid biogenesis takes place. Moreover, PSI* was found to over-accumulate in PsaF-deficient mutant plants and we show that re-initiation of PsaF synthesis promotes the maturation of PSI* into PSI. The attachment of antenna proteins to PSI also requires the transition from PSI* to mature PSI. Our data could provide a biochemical entry point into the challenging investigation of PSI biogenesis and allows us to improve the model for the assembly pathway of PSI in thylakoid membranes of vascular plants.

INTRODUCTION

The light reactions of photosynthesis are performed by large membrane-bound protein complexes, embedded in the thylakoid membranes within the chloroplasts of eukaryotic algae and green plants. Photosystem I (PSI) acts as a light-driven plastocyanin (PC)-ferredoxin oxidoreductase, using light energy to excite electrons and reduce ferredoxin, thus providing

the reducing power needed for carbon fixation and many other metabolic processes in the cell. PSI is one of the most complex protein assemblies in nature. In vascular plants, it is composed of 14 subunits (PsaA-L, PsaN and PsaO) and four antenna proteins (LhcA1-4; Jensen et al., 2007; Qin et al., 2015; Mazor et al., 2015). Five PSI subunits are encoded in the chloroplast genome (PsaA-C, PsaI and PsaJ), while all other subunits are encoded in the nuclear genome and must be post-translationally imported into chloroplasts Schöttler et al., 2011). In addition to its protein subunits, PSI contains a plethora of co-factors and prosthetic groups that are required for its function. The latest crystal structures revealed that PSI of seed plants contains at least 155-156 chlorophylls, 32-35 lipids, 10-14 carotenoids, 3 iron-sulfur [Fe_4S_4] clusters and 2 phylloquinones (Qin et al., 2015; Mazor et al., 2015).

The enormous structural complexity of PSI suggests that its assembly should be tightly coordinated and highly regulated. The assembly process is thought to initiate with the co-translational insertion of its two reaction center subunits, PsaA and PsaB, into the thylakoid membrane (Göhre et al., 2006; Cai et al., 2010; Schöttler et al., 2011). This is followed by addition of the other PSI subunits to the reaction center and incorporation of the many co-factors, ultimately yielding mature PSI. However, the sequence of events leading from the reaction center dimer to the mature PSI complex remains largely unclear. Some speculations have been put forward based on subunit arrangement within the determined crystal structures, but experimental data that would resolve individual steps in the assembly pathway are largely lacking. One of the complications that have hampered the biochemical investigation of PSI biogenesis is that the assembly process appears to occur very fast. Moreover, a serious technical obstacle has been that the large reaction center heterodimer accounts for almost half of the molecular mass of PSI, and the remainder of the subunits to be added are rather small. Consequently, assembly intermediates cannot be easily resolved by molecular mass-based biochemical separation techniques (Schöttler et al., 2011). A PSI

subcomplex, suggested to be a late assembly intermediate, was identified in the unicellular alga *Chlamydomonas reinhardtii* (Ozawa et al., 2010). However, this subcomplex is nearly complete, lacks only the two small subunits PsaG and PsaK, and appears to be functional in that it participates in electron transfer (Ozawa et al., 2010).

Here, we report the characterization of PSI*, an assembly intermediate in the PSI biogenesis pathway of plants. This subcomplex was found to accumulate in very young leaf material of tobacco (*Nicotiana tabacum*) plants, where cell and organelle division and, concomitantly, thylakoid biogenesis is highly active. We find that PSI* contains only a subset of PSI subunits and lacks the LHCI antenna. Furthermore, we report that the addition of PsaF to the PSI* complex is needed for its transition into mature PSI, thus highlighting PsaF as a potential rate-limiting factor in PSI biogenesis. Our data provide a biochemical entry point into the investigation of PSI assembly and allow us to propose a revised model for the assembly pathway of PSI in thylakoid membranes of chloroplasts.

RESULTS

Identification of PSI*, a PSI subcomplex accumulating in developing tobacco leaves

To explore possible entry points into the pathways of PSI assembly, we sought to compare the thylakoid membrane complexes of young and mature tobacco leaves (Figure 1A, B). The underlying assumption was that young, developing leaves should have high synthesis rates of photosynthetic protein complexes, due to the high demand for thylakoid biogenesis during cell and organelle division. Conversely, mature, fully expanded leaves should have much lower synthesis rates of the photosynthetic complexes, because cell division has ceased, the photosynthetic apparatus is fully established and the turnover of the protein complexes in the thylakoid membrane is thought to be very low (Krech et al., 2012; Hojka et al., 2014). It, therefore, seemed reasonable to assume that

putative PSI assembly intermediates are more abundant in young, expanding leaves than in mature leaves. The ontogenetically youngest part of a developing leaf is the sector of the blade that is closest to the leaf base, where cell division and differentiation occurs (Figure 1B). These sectors (4-6 cm wide) where excised from the three youngest leaves of young tobacco plants (4-5 weeks old), pooled from a larger number of plants (typically 70-90 per isolation) and used for the extraction of thylakoid membranes. Similarly, 5-6 cm sectors from the tips of fully expanded leaves (taken from 7-8 week old plants) were used as source material for thylakoid isolation from mature leaf tissue (Figure 1A).

To compare the protein complex composition and assay for the presence of putative assembly intermediates, thylakoid membranes extracted from both types of material were solubilized with the mild non-ionic detergent n-dodecyl β-D-maltoside (DDM) and protein complexes were separated by blue-native polyacrylamide gel electrophoresis (BN-PAGE). At first glance, the overall pattern of protein complexes in mature and young leaf thylakoids was found to be similar (Figure 1C), and only minor differences could be observed. However, upon closer inspection, a faint Coomassie-stained band at a molecular weight of approximately 500 kDa was consistently observed to be enriched in the thylakoid preparations from young leaves (Figure 1C).

To resolve the differences between thylakoids of young and mature leaves in more detail and, at the same time, determine the subunit composition of the complexes observed by one-dimensional BN-PAGE (Figure 1C), an additional separation in the second dimension was performed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). To specifically search for PSI assembly intermediates, the two-dimensional (2D) gels were blotted onto membranes and probed using antibodies against the PSI core subunits PsaA and PsaB. As the assembly of PSI starts

with the membrane insertion of A and B subunits (Schöttler et al., 2011; Yang et al., 2015), these two proteins must be present in any assembly intermediate. Interestingly, both PsaA and PsaB were identified in a complex of approximately 500 kDa (Figure 1D). We, therefore, tentatively named this complex PSI*. PSI subcomplexs of similar size have been observed in a few previous studies and were commonly referred to as a PSI complex lacking its antenna proteins (Heinemeyer et al., 2004; Granvogl et al., 2006; Roose et al., 2014; Soursa et al., 2015). Indeed, no LhcA could be identified in PSI* (Figure 1D). Interestingly, the smaller PSI complex appeared to be relatively abundant in the young leaf samples compared to mature leaves, consistent with what one would expect from a subcomplex representing an assembly intermediate (Figure 1D,E). This prompted us to hypothesize that PSI* might actually represent an intermediate complex in PSI assembly. We, therefore, set out to characterize PSI* in more detail.

PSI* contains only a subset of PSI subunits

To determine the molecular composition of the PSI* complex, the proteins present in the gel segment corresponding to PSI* were identified. To this end, the slice containing PSI* was cut from the 2D gel, subjected to tryptic protein digestion, and its protein composition was determined by MS/MS analysis (Table 1). For comparison, the gel slice containing the mature PSI (co-migrating in the first dimension with the dimeric form of photosystem II (PSII); Figure 1C) and a gel slice adjacent to PSI* were also analyzed ("PSI" and "Background", respectively; Table 1). Several PSI subunits were identified in PSI*. These include the two large reaction center subunits, PsaA and PsaB, two of the stromal ridge subunits, PsaD and PsaE, and PsaL which has been implicated in forming the LHCII binding site during state transitions (Lunde et al., 2000; Table 1). Other subunits that were identified in mature PSI with high scores and peptide numbers were either missing from PSI* or identified only with very low scores, suggestion lower protein abundance. For example, PSI* appears to lack the PSI subunits PsaK and PsaG, and is largely devoid of PsaF. As expected, none of the LhcA proteins were

identified in PSI* (Table 1).

In order to verify the MS/MS data and confirm the absence of antenna proteins and some core subunits of PSI from PSI*, we probed the 2D membranes with antibodies against additional PSI subunits (Figure 1D). These experiments confirmed the presence of PsaD and PsaL in PSI*. Moreover, PSI* was also found to contain PsaC and PsaH (albeit at somewhat reduced amounts; Figure 1D), proteins that, due to their small size, were not reliably detectable by mass spectrometry. Presence of the C, D and E subunits in PSI* indicates that the complex contains the complete stromal ridge. Presence of both PsaL and PsaH suggests that the LHCII-binding site also assembles at this stage. In accordance with the proteomic data, PSI* was found to lack PsaF, PsaK and LhcA3, a subunit of the PSI antenna. Taken together, our mass spectrometric and immunobiochemical analyses suggest that PSI* is a PSI subcomplex that lacks a specific subset of subunits as well as the LHCI antenna proteins.

PSI* is particularly abundant in developing leaf tissue

When assessing the distribution of the PSI core subunits PsaA and PsaB in 2D gels, PSI* was identified as a protein complex in which PsaA and PsaB appeared to be more abundant in thylakoid membranes extracted from young leaf tissue than those extracted from mature leaves (Figure 1 D,E). We were able to show that the same is true when probing 2D gels with antibodies against two other PSI* components, PsaD and PsaL (Figure 1D,E). When calculating the ratio of the signal originating from PSI* to the total PSI signal, it turned out that, in mature leaf tissue, only a small fraction of the PSI subunits is present in PSI* (PsaA: 11.1% ±2.7%; PsaB: 19.4% ±2.8%; PsaD: 10.5% ±4.8%; PsaL: 10.9% ±5.5%; n=3). By contrast, in developing leaf tissue, a substantial proportion of the total PSI signal originates from PSI* (PsaA: 23.4% ±5.9%; PsaB: 30.0% ±0.9%; PsaD: 21.8% ±4.1%; PsaL: 21.3% ±4.7%; n=3), suggesting that the accumulation of PSI* might be related to the active *de*

novo synthesis of thylakoidal protein complexes. Therefore, the high abundance of PSI* in developing leaves led us to hypothesize that PSI* represents an intermediate in the assembly pathway of PSI.

PSI* is an assembly intermediate of PSI

The *de novo* biogenesis of the PSI core requires both protein subunits encoded in the plastid DNA and nucleus-encoded subunits that are synthesized in the cytosol and post-translationally imported into the chloroplast. Thus, upon inhibited translation in the chloroplast or the cytosol, some assembly intermediates can be completed to yield mature PSI, but no new subcomplexes can form. Consequently, if PSI* is indeed a PSI assembly intermediate, its amounts should decrease upon inhibition of protein biosynthesis. To test this assumption, a series of experiments with translational inhibitors was carried out. The translational machinery in the chloroplast is of the bacterial type (Tiller and Bock, 2014) and, therefore, sensitive to antibiotics that block translation on prokaryotic 70S ribosomes, such as lincomycin. Similarly, translation on eukaryotic 80S ribosomes in the cytosol can be specifically inhibited with the antibiotic cycloheximide. We used lincomycin and cycloheximide to block protein biosynthesis in either the chloroplast or the cytosol (or in both compartments) and assessed the fate of PSI*. To this end, thylakoidal protein complexes isolated from antibiotic-treated tissue were electrophoretically separated in 2D gels and the relative abundances of PSI and PSI* were assayed by western blot analysis with anti-PsaA antibodies. While PSI* was highly abundant after the incubation of leaf disks excised from young leaves in solution without antibiotics, incubation in antibiotic solution led to a strongly increased ratio of mature PSI to PSI*, in line with the assumption that PSI* is an assembly intermediate, whose presence depends upon *de novo* synthesis of PSI subunits. The decline in PSI* was seen with both antibiotics alone and also upon combined application of lincomycin and cycloheximide, consistent with the dependence of PSI accumulation on protein biosynthesis in both the chloroplast and the cytosol (Figure 2A).

To ultimately confirm that PSI* is an assembly intermediate, we sought to obtain direct *in vivo* evidence for its conversion into mature PSI. We, therefore, performed pulse-chase experiments. Due to the technical difficulty of performing such experiments in tobacco, we chose to follow the fate of PSI* in *Arabidopsis thaliana*, in which PSI* could be readily identified (Figure 2B). *De novo* synthetized proteins of *A. thaliana* leaves were pulse-labeled with [^{35}S]-methionine for two hours. Subsequently, the fate of labeled proteins was followed over a two hour chase phase in the presence of unlabeled methionine. Protein complexes were then extracted, separated by 2D gel electrophoresis, and their assembly state was monitored by autoradiography (Figure 2B). For PSI, the reaction center proteins PsaA/B could be identified and followed (filled arrowheads in Figure 2B). After the two hour pulse period, high amounts of the fully assembled PSI complex could be identified, indicating that the assembly of mature PSI occurs rather fast. Moreover, the amounts of radiolabel incorporated into PSI* (open arrowheads) was high, consistent with PSI* being an intermediate complex. We, therefore, assume that, while most of the labeled PsaA/B is found in mature PSI, the signal observed in PSI* represents a "snapshot" of the assembly process. When quantifying the signal originating from PsaA/B in PSI* in relation to that originating from PsaA/B in the fully assembled PSI, we found the PSI*/PSI ratio to be high (82% ± 9%; n=3; Figure 2C). During the chase period, the amount of radiolabeled PSI* decreased drastically. Comparison of the PsaA/B signal from PSI* to the signal from fully assembled PSI following the chase revealed a significant decrease compared with that observed after the pulse (PSI*/PSI ratio = 32% ± 1%; n=3; Figure 2C). These results indicate that PSI* is converted into mature PSI *in vivo*, and thus suggest very strongly that PSI* indeed represents a PSI assembly intermediate. Further, definitive proof of the maturation of PSI* into mature PSI is described in the following section (see below, Figure 3D).

PsaF integration facilitates the maturation of PSI* into PSI

Of all PSI subunits absent from PSI*, the apparent deficit in PsaF (Table 1; Figure 1D), the PC-docking subunit of PSI, was the most intriguing. PsaF deficiency has been shown to reduce overall PSI levels and destabilize the stromal ridge, highlighting the importance of the F subunit in creating stable mature PSI units (Haldrup et al., 2000). To explore the possible function of PsaF in the transition from PSI* to PSI and to provide additional genetic evidence for the role of PSI* in PSI assembly, we sought to knock down the *PSAF* gene. To avoid pleiotropic effects, we used a transgenic tobacco mutant in which the (nucleus-encoded) *PSAF* gene was knocked down by inducible RNAi. These plants were grown for six weeks prior to induction, allowing them to accumulate sufficient functional source tissue to support plant growth during RNAi induction. Interestingly, when PsaF was silenced by inducing RNAi via treatment of plants with ethanol vapor (Hojka et al., 2014), the strong drop in PsaF accumulation was not accompanied by a similar reduction in PsaA levels. While the induced *PSAF-RNAi* tissue was found to contain almost no PsaF protein, PsaA accumulation was reduced only by about 50% (Figure 3A). This could be further verified by determining redox active PSI amounts via difference absorbance spectroscopy, quantifying the light-induced difference absorbance changes of P_{700} (Figure 3B). Taken together, these results indicate that all of the PsaA present in the induced *PSAF-RNAi* tissue is assembled into a complex capable of light-induced charge separation. Induced *PSAF-RNAi* plants were found to be slightly paler compared to wild-type plants, and this phenotype was more pronounced in young leaves than in mature leaves (Figure S1A-C). In agreement with previous data (Haldrup et al., 2000), the young tissue lacking PsaF showed a decreased F_V/F_M value, an indicator of maximum photochemical efficiency (Figure S1D). This is in line with the presence of significant amounts of free, uncoupled LHCI antenna proteins in their monomeric and dimeric form in the thylakoids, resulting in increased basal chlorophyll-a fluorescence emission in darkness (see below, Figure 3G, 4A).

Remarkably, when analyzing PSI complex distribution in PsaF-deficient tissue using 2D gels, it

was found that large amounts of PSI* accumulated while only very little mature PSI assembled (Figure 3C). As the vast majority of PSI in these plants accumulates as PSI*, these results indicate that (i) PSI* is able to harvest light and perform charge separation, and (ii) the incorporation of the PsaF subunit into the assembling photosystem is an important step in the maturation of the PSI* complex into a complete PSI unit. Moreover, the overaccumulation of PSI* upon knock-down of PsaF suggests that the integration of PsaF into the complex may represent a rate-limiting step in PSI biogenesis. If this is indeed the case, then restricting the assembly process at an earlier stage should result in decreased PSI* abundance. This is because the reduced amounts of nascent PSI* would then be converted very quickly into mature PSI. Indeed, PSI* could not be detected in thylakoids of PsaA knock-down plants (Krech et al., 2012; Figure 3C), in which the limiting supply of PsaA determines the rate of PSI biogenesis.

To verify that PsaF is the limiting factor in PSI* maturation into a fully assembled PSI, we followed the fate of PSI* in the PSI*-overaccumulating *PSAF-RNAi* plants. Following one week of RNAi induction by ethanol vapor (leading to accumulation of PSI mainly as PSI*), the ethanol vapor was removed, thus allowing for the recovery of PsaF synthesis. To unambiguously confirm that PsaF provision converts PSI* into mature PSI, ethanol removal was accompanied by the addition of lincomycin, thus stopping any *de novo* synthesis of PSI cores due to blocked synthesis of the chloroplast-encoded A, B and C subunits. Consequently, any increase in the abundance of fully assembled PSI observed under these conditions must have originated from conversion of the previously accumulated PSI* into PSI. Indeed, after six hours of incubation under these conditions (absence of ethanol and presence of the chloroplast translational inhibitor lincomycin), the majority of the PSI signal (detected by the anti-PsaA antibody) had shifted from PSI* to fully assembled PSI (Figure 3D). These results clearly show that PsaF supply represents a limiting step in PSI maturation that causes accumulation of PSI* as a relatively stable intermediate. Furthermore, as no new PsaA,B

and C can be synthesized during the incubation with lincomycin, any accumulation of mature PSI must originate from the core proteins already present in the thylakoids, that can now bind the nuclear encoded subunits downstream to PsaF in the assembly process (and whose expression is not affected by lincomycin). Thus, our results clearly demonstrate that PSI* is not a dead-end product in the PSI assembly pathway, but represents a *bona fide* assembly intermediate awaiting PsaF incorporation.

PSI* contains a complete electron transport chain, but cannot efficiently receive electrons

The generation of the PSI*-overaccumulating *PSAF-RNAi* plants enabled us to investigate whether or not PSI* is an active PSI in that it is capable of performing electron transport reactions. Upon illumination, *PSAF-RNAi* plants performed P_{700} oxidation into P_{700}^{+}, as determined by a change in absorbance at near infrared. Despite the nearly complete absence of PsaF, the signal measured in thylakoids extracted from these plants was more than half of that measured for thylakoids extracted from wild-type plants (Figure 3B). This activity correlates with the PsaA contents in the *PSAF-RNAi* lines which were reduced to nearly half of those present in wild-type plants (Figure 3A). Since the vast majority of PsaA in *PSAF-RNAi* plants accumulates in PSI*, we can conclude that PSI* is capable of performing charge separation upon illumination.

To determine whether PSI* contains a fully functional electron transport chain, we measured the time needed for P_{700}^{+} charge recombination in the absence of an electron acceptor. In fully assembled PSI containing the two terminal iron-sulfur clusters F_A and F_B, charge recombination typically occurs within tens of milliseconds. However, when PsaC is not correctly assembled and these co-factors are missing, charge recombination occurs much faster, within a few milliseconds (Vassiliev et al., 2001; Wittenberg et al., 2013). Charge recombination was found to occur within the

tens of milliseconds timescale in thylakoids extracted from both wild-type tissue (t = 68.9 ms ±9.2) and PSI*-accumulating PsaF-deficient tissue (t = 48.0 ms ±0.0, Figure 3E), indicating that PSI* contains a fully assembled stromal ridge with functional F_A and F_B clusters.

PsaF has been implicated in binding PC (Fischer et al., 1999). We, therefore, wanted to determine whether PSI* contains a functional donor side. To this end, we determined the redox equilibration between P_{700} and PC in intact wild type and *PSAF-RNAi* plants. Plants were pre-illuminated for several minutes, to activate the Calvin-Benson cycle and avoid acceptor-side limitation in PSI, and to ensure rapid diffusion of PC in the thylakoid lumen (Kirchhoff et al., 2011). Then, a short saturating light pulse was applied, to fully oxidize the high-potential chain and reduce the plastoquinone pool. At the end of the saturating pulse, the actinic light was switched off, and the reduction kinetics of P_{700}^+ and PC were measured. The fully oxidized state of P_{700}^+ and PC was normalized to one, and their fully reduced state in darkness was normalized to zero. To quantitatively describe the interaction between PC and P_{700}, the normalized reduction kinetics were plotted against each other (Figure 3F). The apparent redox equilibration constant (K_{app}) was found to be lower for the *PSAF-RNAi* plants (K_{app} = 4.8 ±0.9) than that calculated for wild-type plants (K_{app} = 12.7 ±2.2), indicating that, consistent with the absence of the PsaF subunit and its proposed function in PC docking, the donor side of PSI* is slightly less efficient in oxidizing PC.

PSI* was found not to contain the PSI associated antenna proteins (Table 1; Figure 1D). To assess the functional coupling of the PSI antenna to PSI*, we measured chlorophyll-*a* fluorescence emission spectra at 77K in the wild type and the *PSAF-RNAi* plants. The maximum fluorescence emission signal from PSI-LHCI (peaking at 733 nm in wild-type tobacco) was significantly shifted to shorter wavelengths in the *PSAF-RNAi* plants (Figure 3G), and the strongly increased emission in the range of 730 and 700 nm wavelength, where uncoupled LhcA proteins mainly emit their chlorophyll-a fluorescence (Castelletti et al., 2003; Croce et al., 2004), indicates that the PSI antenna proteins are

not connected to PSI*.

Incorporation of the antenna proteins, and the G and K subunits into assembling PSI

PSI* identified in young leaf tissue was found to be deficient not only in PsaF, but also in the PSI subunits PsaG and PsaK and the PSI-associated antenna proteins (Table 1; Figure 1D; Figure 3G). Taking advantage of our *PSAF-RNAi* plants, we sought to determine if these proteins are able to assemble into PSI* in the absence of PsaF or if their incorporation into PSI is dependent on prior integration of PsaF into the growing complex. In the *PSAF-RNAi* plants, the LhcA antenna proteins of PSI were found to accumulate mainly as free monomers and dimers (Figure 4A). The small fraction of LhcA proteins that was attached to larger complexes was exclusively associated with the residual amounts of mature PSI (Figure 4A), suggesting that LHCI attachment requires the transition from PSI* to PSI. These results also suggest that the incorporation of PsaF into the assembling photosystem facilitates antenna binding, but might not be strictly required for it. Moreover, PsaF integration was required for the integration of PsaG into PSI, as evidenced by PsaG being undetectable in PSI* even in PSI*-overaccumulating tissue (Figure 4B). Interestingly, in these plants, PsaG was found to be associated with two smaller complexes, resembling in molecular weight the LHCI monomers and dimers. Previously, PsaG was shown to interact with LhcA1 in the fully assembled complex (Mazor et al., 2015; Qin et al., 2015). We, therefore, hypothesize that these two lower molecular weight complexes represent another PSI assembly intermediate in which PsaG associates with LHCI prior to incorporation into PSI.

By contrast, PsaK, which was found to be associated only with fully assembled PSI in wild-type plants, could be identified in PSI* in the PsaF-deficient plants (Figure 4C). Assembly of PsaK into PSI* under the conditions of PSI* overaccumulation indicates that its assembly into the photosystem should be independent of that of PsaF and the antenna proteins.

DISCUSSION

In this work, we have characterized PSI* and show it to be an intermediate complex in the PSI assembly pathway. Its identification as an assembly intermediate was facilitated by the comparison of photosynthetic protein complexes between developing leaves and mature leaves using high-resolution 2D electrophoresis techniques. When the photosynthetic apparatus is fully assembled, the protein complexes in the thylakoid membrane are highly stable and display very low turnover rates (Krech et al., 2012; Hojka et al., 2014). Therefore, assembly intermediates are unlikely to be present in readily detectable amounts in mature leaves. To search for PSI assembly intermediates, we chose to use young leaves of tobacco, a plant with a high growth rate that produces large leaves and substantial amounts of leafy biomass in a relatively short time. This facilitated the excision of the growing zone of developing leaves (where active thylakoid biogenesis takes place) and provided us with enough material for the large-scale isolation of thylakoid membranes. Together with the use of a high-resolution 2D electrophoresis method, this allowed us to identify the PSI assembly intermediate characterized here (Figure 1).

Several lines of evidence strongly support the conclusion that PSI* is a genuine intermediate in the assembly pathway of PSI in the thylakoid membrane. First, PSI* is more abundant in developing leaf tissue than in fully differentiated tissue (Figure 1). Second, PSI* accumulation is dependent on *de novo* protein synthesis and its relative abundance to mature PSI decreases upon inhibition of translation (Figure 2A). Third, pulse-labeled PSI* is rapidly converted into mature PSI *in*

vivo (Figure 2B and C). Fourth, conversion of PSI* into mature PSI is largely blocked upon knock-down of PsaF, a subunit absent from PSI* and required for the downstream assembly steps (Figure 3C). Fifth, upon resupply of PsaF in the absence of *de novo* synthesis of PsaA, PsaB and PsaC, PSI* already present in the thylakoids matures into fully assembled PSI (Figure 3D). Finally, accumulation of PSI* can be effectively prevented when an upstream step in the PSI biogenesis pathway is made rate-limiting, as demonstrated with the PsaA knock-down mutant (Figure 2C).

A putative late assembly intermediate of PSI was previously identified in exponentially growing cells of the unicellular alga *Chlamydomonas reinhardtii* (Ozawa et al., 2010). It lacked the two small PSI subunits PsaK and PsaG (Ozawa et al., 2010). Thus, the *C. reinhardtii* PSI subcomplex might be similar to the tobacco PSI*. The same (or a highly similar) PSI subcomplex in *C. reinhardtii* was suggested to interact with the PSI assembly factor Ycf4 (Ozawa et al., 2009), raising the possibility that the Ycf4 protein assists with the incorporation of the final subunits into nearly mature PSI complexes. However, the *C. reinhardtii* subcomplex was found to bind PsaF, even if loosely. The dependence of PSI* maturation into PSI on PsaF supply in tobacco suggests PSI* might represent an earlier assembly stage compared with the identified *C. reinhardtii* subcomplex.

Like in *C. reinhardtii*, PsaK also appears to be the last subunit to assemble into PSI complexes in cyanobacteria (Dühring et al., 2007). Based on the detection of a short-lived PSI complex additionally lacking PsaL, it was proposed that PsaL is the penultimate subunit to integrate into cyanobacterial PSI (Dühring et al., 2007). This is in contrast to the assembly pathway in seed plants, where PsaL is an abundant component of PSI* (Table 1; Figure 1D) and thus, integrates earlier into the plant PSI complex. Indeed, in the PSI complex of seed plants PsaL is not a peripheral subunit *sensu stricto* but, to a large part, covered from the outside by PsaI and PsaH (Qin et al., 2015; Mazor et al., 2015).

Our data presented here allow us to suggest an improved model for PSI assembly in plants (Figure 5). The assembly process begins with the co-translational insertion of the two large reaction center subunits PsaA and PsaB, mediated by the Alb3 protein insertase (Göhre et al., 2006; Cai et al., 2010; Schöttler et al., 2011). The formation of the reaction center dimer (Figure 5) is then followed by the addition of the stromal ridge subunits PsaC, PsaD and PsaE. The presence of a fully assembled stromal ridge is also supported by the largely functional electron transport chain in PSI*, as determined by charge recombination kinetics (Figure 3E). It is not yet clear whether the assembly of the LHCII-docking domain (comprised of the subunits PsaL, PsaH and PsaI) occurs in parallel or as a subsequent step, but our data presented here suggest that the addition of PsaL and, to some extent, PsaH (unfortunately, no functional antibodies could be obtained for PsaI) generates PSI*, a relatively stable assembly intermediate that accumulates to substantial amounts in developing leaf tissue. The presence of PsaL in PSI* is in agreement with the PSI crystal structure that revealed interactions of PsaL with both reaction center subunits as well as with PsaD (Amunts et al., 2010). However, as Arabidopsis plants lacking PsaL are relatively unaffected in their growth and photosynthetic performance (Lunde et al., 2000), it seems possible that the L subunit plays only a limited role in stabilizing PSI complexes during their assembly.

PSI* was found to be largely devoid of PsaF. Forming the PC-docking site, the F subunit was postulated to be required for PSI function. In constitutive Arabidopsis *PsaF*-antisense mutants suffering from a similar reduction in PsaF content as the young leaves of our inducible RNAi plants, PC oxidation was found to be slowed down. Moreover, overall PSI stability and, in particular, the stability of the stromal ridge subunits was impaired. Indeed, many constitutive PsaF knock down plants died under autotrophic growth conditions (Haldrup et al., 2000). In our PsaF knock-down mutants, the defects in PC oxidation and PSI accumulation were less severe. One major difference is that while the Arabidopsis mutants starved under autotrophic growth conditions, in our system the

source leaves, which were already fully developed prior to RNAi induction, maintained their photosynthetic activity and supported growth of the newly developing leaves. Therefore, possibly, our inducible RNAi mutants are less stressed compared with constitutive PsaF knock-down plants. However, our data show that under such conditions, loss of PsaF does, to some extent, impact PC oxidation (Figure 3F).

In the absence of PsaF, LhcA proteins were found to accumulate mainly as monomers and dimers (Figure 4A). Together with the 77K emission spectra determined for these plants (Figure 3G), these data lead us to hypothesize that the association of PsaF with PSI* is needed to facilitate binding of LhcA subunits to the complex. PSI* also lacks the two small PSI subunits PsaK and PsaG. The binding of the antenna to PSI requires interactions with the subunits PsaG, PsaF, PsaJ and PsaA (Mazor et al., 2015). Our results suggest that PsaG might be bound to the antenna proteins prior to its integration into PSI, and that the incorporation of both the antenna and the G subunit into PSI is facilitated by the prior insertion of PsaF (Figure 4B). However, the presence of the antenna proteins was reported not to be required for PsaG and PsaK accumulation (Wientjes et al., 2009), and the presence of PsaG was found not to be required for the binding of the antenna to PSI (Jensen et al., 2002). Taken together, these results imply that PsaG does not function in targeting the antenna to the assembling photosystem. Unlike PsaG, it appears that the K subunit readily assembles into PSI lacking PsaF. As in wild-type plants no PsaK can be detected in PSI*, it seems possible that the addition of PsaF facilitates PsaK incorporation, but to a limited extent (Figure 5).

Our analysis of PsaF knock-down plants in tobacco revealed that the addition of PsaF to the nascent complex assists PSI* to mature into PSI (Figure 3D). Thus, the addition of PsaF to PSI* could be a rate-limiting step in PSI assembly, and hence, the cause of PSI* accumulation. In contrast to all other PSI subunits, PsaF requires the chloroplast Sec pathway for its integration into the thylakoid membrane (Voelker and Barkan, 1995). Other PSI subunits utilize either the signal recognition particle pathway (e.g., PsaA, PsaB and the LhcA proteins; Göhre et al., 2006), may insert

spontaneously (PsaG and PsaK; Mant et al., 2001; Zygadlo et al., 2006) or are directed to the lumen via the Tat pathway (PsaN; Nielsen et al., 1994). Considering that provision of PsaF to PSI* likely limits the transition to subsequent steps in PSI assembly, PSI* accumulation could be the result of overloading of the Sec pathway with client proteins under conditions of highly active thylakoid biogenesis. Besides PsaF, a number of other abundant thylakoid proteins are client proteins of the Sec pathway, including several PSII subunits and PC Voelker et al., 1997). It will be interesting to determine whether or not manipulation of the Sec pathway and/or of the levels of its other client proteins affect PSI* accumulation and the fate of the peripheral, non-essential PSI subunits.

In sum, our work reported here describes PSI*, a PSI subcomplex, as an intermediate in the PSI assembly pathway of vascular plants. PSI* could serve as a fresh entry point into the study of PSI biogenesis, which heretofore was not amenable to biochemical investigation. Future work will be directed towards the detailed molecular analysis of PSI*, the determination of its content of co-factors and prosthetic groups and the elucidation of its interaction with PSI assembly chaperones (e.g., Ycf3, Y3IP1, Ycf4, PYG7, PPD1, PSA2 and FTSH; Ruf et al., 1997; Boudreau et al., 1997; Stöckel et al., 2006; Albus et al., 2010; Krech et al., 2012; Liu et al., 2012; Fristedt et al., 2014) acting upstream or downstream of PSI* formation. The experimental approaches that led to the identification of PSI* will likely also be applicable to the analysis of the assembly pathways of other multiprotein complexes in the thylakoid membrane (e.g., the cytochrome b_6f complex and the chloroplast ATP synthase) whose biogenesis is currently poorly understood.

EXPERIMENTAL PROCEDURES

Plant material and growth conditions

Tobacco (*Nicotiana tabacum* cv. Petit Havana) plants were grown in the greenhouse at an average growth light intensity of 300 µE m^{-2} s^{-1}. PsaA-deficient mutant plants (KD-*psaA*) were generated by stable transformation of the tobacco chloroplast genome as reported previously (Krech et al., 2012). Transgenic PsaF-deficient mutant plants (*PSAF-RNAi*) were produced by nuclear transformation of tobacco plants with a standard hairpin-type RNAi construct (a full characterization of the mutant plants will be reported elsewhere). The RNAi construct was placed under the control of an ethanol-inducible promoter, as described previously (Hojka et al., 2014). *Arabidopsis thaliana* ecotype Col-0 plants were grown in growth chambers at a light intensity of 120 µE m^{-2} s^{-1} in a 8 h light/16 h dark regime at 23°C.

Thylakoid extraction and isolation of thylakoid-rich membrane fractions

Thylakoid membranes were isolated according to previously published procedures (Schöttler et al., 2004). Following isolation, thylakoids were washed in 25BTH20G buffer (25 mM Bis-Tris pH 7.0, 20% glycerol v/v) supplemented with 1M NaCl, and then collected again by centrifugation at 20,000 g. For isolation of thylakoid-rich membrane fractions, leaf material was ground with a mortar and pestle in isolation buffer (50 mM MES pH 6.1, 5 mM $MgCl_2$, 30 mM KCl, 350 mM D-Sorbitol). The homogenate was filtered through gauze and Miracloth (Millipore, www.merckmillipore.com) and the membranes were collected by centrifugation for 20 min at 20,000 *g*. All steps were carried out at 4°C. The isolated membranes were stored in 25BTH20G buffer at -80°C.

Gel electrophoresis and immunoblotting

Sample preparation and gel electrophoretic separation of membrane protein complexes by blue-native polyacrylamide gel electrophoresis (BN-PAGE) were performed according to published procedures (Järvi et al., 2011). Membranes were solubilized with 1% n-dodecyl β-D-maltoside (w/v). For a second dimension SDS-PAGE, BN-PAGE gels were incubated for 1 h in 100 mM Tris pH 6.8, 1% SDS (w/v) and 1% β-mercaptoethanol (v/v), followed by washes with 100 mM Tris pH 6.8. Denatured proteins were separated in 8% or 10% SDS-polyacrylamide (PAA) gels. For immunoblotting, proteins were transferred to Hybond-P polyvinylidene difluoride membranes (GE Healthcare, www.gelifesciences.com) using standard protocols. Immunochemical detection was performed with specific antibodies using the ECL Plus system (GE Healthcare, www.gelifesciences.com). Polyclonal antibodies against PsaA, PsaB, PsaC, PsaD, PsaF, PsaK, PsaL and LhcA1 (produced in rabbits) were purchased from Agrisera (www.agrisera.com). Goat anti-rabbit IgG (H+L)-horseradish peroxidase conjugate (Bio-Rad, www.bio-rad.com) was used as a secondary antibody.

Chemiluminescence was imaged by a Syngene G:BOX (www.syngene.com). Measurements were stopped prior to signal saturation, thus ensuring that the signals collected were in the linear range. Quantification was carried out using the accompanying GeneTools software.

Protein identification by mass spectrometry

Gel slices excised from PAA gels were cut into 10 pieces each and subjected to tryptic protein digestion according to a standard protocol (Shevchenko et al., 2006). The resulting peptides were desalted using a C18 Ziptip (Millipore, www.merckmillipore.com) and resuspended in 5% (v/v) acetonitrile and 0.1% (v/v) trifluoroacetic acid. Peptides were then analyzed by liquid chromatography (Proxeon EASY nLC, Thermo Fisher Scientific, San Jose, CA) using a Chromolith CapRod RP-18e 150-0.2 (Merck, Darmstadt, Germany) column. To this end, the peptides were

separated by a 30 min linear acetonitrile gradient (0-32% v/v) followed by a final elution step with 64% (v/v) acetonitrile for 5 min. The HPLC was coupled via a nano ESI ion source to a high-resolution Orbitrap hybrid mass spectrometer (LTQ-Orbitrap, Thermo Fisher Scientific). The spectral acquisition for the full scan MS spectra was performed at a full-width half-maximum resolution of 60,000 in the Orbitrap section of the MS, while the data-dependent MS/MS, with up to five spectra per preceding full scan, were obtained in the linear ion trap of the LTQ. Database searches were performed by Mascot 2.2 (Matrix Science Ltd., London, UK; www.matrixscience.com) against an in-house database of plants proteins with the following parameters: proteolytic enzyme: trypsin allowing up to one missed cleavage; maximal precursor and fragment ion errors: 10 ppm and 0.8 Da; peptide charge: +2 or +3; variable modification: oxidation of methionine. Standard scoring was used.

Inhibition of protein biosynthesis in the cytosol and in chloroplasts

Leaf disks (1.4 cm diameter) were cut from 5-6 week old plants and incubated in tap water with or without inhibitors of protein biosynthesis. Chloroplast translation was inhibited with 2.5 mg/mL lincomycin (Sigma, www.sigmaaldrich.com), cytosolic translation was blocked with 20 µg/mL cycloheximide (Sigma). To inhibit protein biosynthesis in both compartments, the two antibiotics were combined. The leaf disks were incubated in the antibiotic solutions for 3 h in the dark, followed by 1h incubation at growth light intensity and extraction of thylakoid membranes.

***In vivo* labeling of thylakoid proteins**

Radioactively labeled methionine was incorporated into detached *Arabidopsis* leaves (via their petioles) overnight by incubation in a solution containing 20 µCi mL^{-1} [^{35}S]-methionine and 0.4% (v/v) Tween 20, followed by labeling of *de novo* synthesized proteins at 120 µE m^{-2} s^{-1} light intensity and 23°C for 2 h. Subsequently, the leaves were washed with unlabeled methionine (10 mM methionine,

0.4% (v/v) Tween 20) followed by incorporation of unlabeled methionine into the leaves by incubation in the same solution for 1 h in the dark. The chase experiment was performed by subsequent incubation for 2 h in the light (120 µE m^{-2} s^{-1}) at 23°C. Afterwards, samples were collected, thylakoidal protein complexes were separated on 2D gels, and the synthesis and degradation of the PSI reaction center proteins PsaA and PsaB was assessed by autoradiography using X-ray films. Quantification of differences in accumulation of labeled proteins was carried out by densitometric analysis of the films with a FluorChem™ 8000 image analyzer (Alpha Innotech Corp., San Leandro, CA). The standard error was calculated for the PSI*/PSI ratios.

Induction and release of RNAi suppression of *PSAF*

Leaf disks were cut from *PSAF-RNAi* plants after one week of RNAi induction with ethanol vapor. The leaf disks were washed twice in tap water to remove traces of ethanol, and then incubated in tap water containing 2.5 mg/mL lincomycin for 2 h in the dark, followed by a 6 h incubation at 25 µE m^{-2} s^{-1}. Thylakoid membranes were then extracted and analyzed.

Biophysical analyses

PSI was quantified from light-induced difference absorbance changes of the chlorophyll-*a* dimer special pair, P_{700}. Thylakoids equivalent to 50 µg chlorophyll mL^{-1} were solubilized in the presence of 0.2 % (w/v) β-dodecylmaltoside, 10 mM sodium ascorbate as electron donor and 100 µM methylviologen as electron acceptor. Photooxidation was achieved by the application of a light pulse of 250 ms length (2000 µE m^{-2} s^{-1}). Measurements were performed with the PC-P700 version of the Dual-PAM instrument (Heinz Walz GmbH, www.walz.com; Schöttler et al., 2007). PSI charge recombination kinetics were measured as reported previously (Wittenberg et al., 2013). For intact leaf redox equilibration measurements, leaves were pre-illuminated with saturating light. This

ensures that the Calvin-Benson cycle is fully activated, thus avoiding an acceptor-side limitation of PSI. Then, a short saturating light pulse (5000 µE m^{-2} s^{-1}, 250 ms length) was applied, to achieve full oxidation of both PC and P_{700}. At the end of that light pulse, the actinic illumination was completely switched off, and the reduction kinetics of PC and P_{700} were measured with a PC-P_{700} version of the Dual-PAM instrument.

77K chlorophyll-*a* fluorescence emission spectra of freshly isolated thylakoids equivalent to 10 µg chlorophyll mL^{-1} were measured using an F-6500 fluorometer (Jasco GmbH, www.jascoinc.com). The sample was excited at 430 nm wavelengths with a 10 nm bandwidth, and the emission spectrum between 660 and 800 nm wavelengths was recorded with a bandwidth of 1 nm. The spectra were corrected for the instrument's response and normalized to the PSII-LHCII maximum emission at 686 nm wavelength.

ACKNOWLEDGEMENTS

We thank Monika Suchoszek (MPI-MP) for help with the cultivation and induction of the *PSAF-RNAi* plants, Caroline Runge, Lisa Heinke and Martina Schneemann (MPI-MP) for technical assistance, Dr. Marjaana Suorsa (University of Turku) for discussion and the MPI-MP Green Team for plant cultivation. This work was supported by a grant from the Deutsche Forschungsgemeinschaft to R.B. (FOR 2092; BO 1482/17-1), the Academy of Finland projects 271832 and 273870 (E.-M.A.) and the Alexander von Humboldt Foundation (research grant to S.Z.T.). The authors declare no conflicts of interest.

Author contributions: G.W., S.J., M.A.S., S.Z.T and M.H. designed and performed research and analyzed data. E.H.M. performed research and analyzed data. E.-M. A. designed research and analyzed data. R.B. designed research, analyzed data and wrote the paper together with G.W.

Short Supporting Information Legend

Supplemental Figure 1. Phenotype of *PSAF-RNAi* plants following one week of RNAi induction by ethanol vapor.

REFERENCES

Albus, C., Ruf, S., Schöttler, M.A., Lein, W., Kehr, J. and Bock, R. (2010) Y3IP1, a nucleus-encoded thylakoid protein, co-operates with the plastid-encoded Ycf3 protein in photosystem I assembly. *Plant Cell* **22**, 2838-2855.

Amunts, A., Toporik, H., Borovikova, A. and Nelson, N. (2010) Structure determination and improved model of plant photosystem I. *J. Biol. Chem* **285**, 3478-3486.

Boudreau, E., Takahashi, Y., Lemieux, C., Turmel, M. and Rochaix, J.-D. (1997) The chloroplast ycf3 and ycf4 open reading frames of Chlamydomonas reinhardtii are required for the accumulation of the photosystem I complex. *EMBO J.* **16**, 6095-6104.

Cai, W., Ma, J., Chi, W., Zhou, M., Guo, J., Lu, C. and Zhang, L. (2010) Cooperation of LPA3 and LPA2 is essential for photosystem II assembly in Arabidopsis. *Plant Physiol.* **154**, 109-120.

Castelletti, S., Morosinotto, T., Robert, B., Caffarri, S., Bassi, R. and Croce, R. (2003) Recombinant Lhca2 and Lhca3 subunits of the photosystem I antenna system. *Biochemistry* **42**, 4226-4234.

Croce, R., Morosinotto, T., Ihalainen, J.A., Chojnicka, A., Breton, J., Dekker, J.P., van Grondelle, R. and Bassi, R. (2004) Origin of the 701-nm fluorescence emission of the Lhca2 subunit of higher plant photosystem I. *J. Biol. Chem.* **279**, 48543-48549.

Dühring, U., Ossenbühl, F. and Wilde, A. (2007) Late assembly steps and dynamics of the cyanobacterial photosystem I. *J. Biol. Chem.* **282**, 10915-10921.

Fischer, N., Boudreau, E., Hippler, M., Drepper, F., Haehnel, W. and Rochaix, J.-D. (1999) A large fraction of PsaF is nonfunctional in photosystem I complexes lacking the PsaJ subunit. *Biochemistry* **38**, 5546-5552.

Fristedt, R., Williams-Carrier, R., Merchant, S.S. and Barkan, A. (2014) A thylakoid membrane protein harboring a DnaJ-type zinc finger domain is required for photosystem I accumulation in plants. *J. Biol. Chem.* **289**, 30657-30667.

Göhre, V., Ossenbühl, F., Crèvecoeur, M., Eichacker, L.A. and Rochaix, J.-D. (2006) One of two Alb3 proteins is essential for the assembly of the photosystems and for cell survival in Chlamydomonas. *Plant Cell* **18**, 1454-1466.

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

Haldrup, A., Simpson, D.J. and Scheller, H.V. (2000) Down-regulation of the PSI-F subunit of photosystem I (PSI) in Arabidopsis thaliana. *J. Biol. Chem.* **275**, 31211-31218.

Heinemeyer, J., Eubel, H., Wehmhöner, D., Jänsch, L. and Braun, H.-P. (2004) Proteomic approach to characterize the supramolecular organization of photosystems in higher plants. *Phytochemistry* **65**, 1683-1692.

Hojka, M., Thiele, W., Tóth, S.Z., Lein, W., Bock, R. and Schöttler, M.A. (2014) Inducible repression of nuclear-encoded subunits of the cytochrome b_6f complex in tobacco reveals an extraordinarily long lifetime of the complex. *Plant Physiol.* **165**

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.

Jensen, P.E., Rosgaard, L., Knoetzel, J. and Scheller, H.V. (2002) Photosystem I activity is increased in the absence of the PSI-G subunit. *J. Biol. Chem.* **277**, 2798-2803.

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* **1767**, 335-352.

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

Krech, K., Ruf, S., Masduki, F.F., Thiele, W., Bednarczyk, D., Albus, C.A., Tiller, N., Hasse, C., Schöttler, M.A. and Bock, R. (2012) The plastid genome-encoded Ycf4 protein functions as a nonessential assembly factor for photosystem I in higher plants. *Plant Physiol.* **159**, 579-591.

Liu, J., Yang, H., Lu, Q., Wen, X., Chen, F., Peng, L., Zhang, L. and Lu, C. (2012) PSBP-domain protein1, a nuclear-encoded thylakoid lumenal protein, is essential for photosystem I assembly in Arabidopsis. *Plant Cell* **24**, 4992-5006.

Lunde, C., Jensen, P.E:, Haldrup, A., Knoetzel, J. and Scheller, H.V. (2000) The PSI-H subunit of photosystem I is essential for state transitions in plant photosynthesis. *Nature* **408**, 613-615.

Mant, A., Woolhead, C.A., Moore, M., Henry, R. and Robinson, C. (2001) Insertion of PsaK into the thylakoid membrane in a "horseshoe" conformation occurs in the absence of signal recognition particle, nucleoside triphosphates, or functional Albino3. *J. Biol. Chem.* **276**, 36200-36206.

Mazor, Y., Borovikova, A. and Nelson, N. (2015) The structure of plant photosystem I super-complex at 2.8 A resolution. *eLife* **4**, e07433.

Nielsen, V.S., Mant, A., Knoetzel, J., Møller, B.L. and Robinson C. (1994) Import of barley photosystem I subunit N into the thylakoid lumen is mediated by a bipartite presequence lacking

an intermediate processing site. *J. Biol. Chem.* **269**, 3762-3766.

Ozawa, S.-I., Nield, J., Terao, A., Stauber, E.J., Hippler, M., Koike, H., Rochaix, J.-D. and Takahashi, Y. (2009) Biochemical and structural studies of the large Ycf4-photosystem I assembly complex of the green alga Chlamydomonas reinhardtii. *Plant Cell* **21**, 2424-2442.

Ozawa, S.-I., Onishi, T. and Takahashi, Y. (2010) Identification and characterization of an assembly intermediate subcomplex of photosystem I in the green alga Chlamydomonas reinhardtii. *J. Biol. Chem.* **285**, 20072-20079.

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

Roose, J.L., Frankel, L.K. and Bricker, T.M. (2014) The PsbP domain protein 1 functions in the assembly of lumenal domains in photosystem I. *J. Biol. Chem.* **289**, 23776-23785.

Ruf, S., Kössel, H. and Bock, R. (1997) Targeted inactivation of a tobacco intron-containing open reading frame reveals a novel chloroplast-encoded photosystem I-related gene. *J. Cell Biol.* **139**, 95-102.

Schöttler, M.A., Kirchhoff, H. and Weis, E. (2004) The role of plastocyanin in the adjustment of the photosynthetic electron transport to the carbon metabolism in tobacco. *Plant Physiol.* **136**, 4265-4274.

Schöttler, M.A., Flügel, C., Thiele, W., Stegemann, S. and Bock, R. (2007) The plastome-encoded PsaJ subunit is required for efficient photosystem I excitation, but not for plastocyanin oxidation in tobacco. *Biochem. J.* **403**, 251-260.

Schöttler, M.A., Albus, C.A. and Bock, R. (2011) Photosystem I: Its biogenesis and function in higher plants. *J. Plant Physiol.* **168**, 1452-1461.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. and Mann, M. (2006) In-gel digestion for mass

spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856-2860.

Soursa, M., Rantala, M., Mamedov, F., Lespinasse, M., Trotta, A., Grieco, M., Vuorio, E., Tikkanen, M., Järvi, S. and Aro, E.-M. (2015) Light acclimation involves dynamic re-organization of the pigment-protein megacomplexes in non-appressed thylakoid domains. *Plant J.* **84**, 360-373.

Stöckel, J., Bennewitz, S. and Oelmüller, R. (2006) The evolutionarily conserved tetratrico peptide repeat protein pale yellow green7 is required for photosystem I accumulation in Arabidopsis and copurifies with the complex. *Plant Physiol.* **141**, 870-878.

Tiller, N. and Bock, R. (2014) The translational apparatus of plastids and its role in plant development. *Mol. Plant* **7**, 1105-1120.

Vassiliev, I.R., Antonkine, M.L. and Golbeck, J.H. (2001) Iron-sulfur clusters in type I reaction centers. *Biochim. Biophys. Acta* **1507**, 139-160.

Voelker, R. and Barkan, A. (1995) Two nuclear mutations disrupt distinct pathways for targeting proteins to the chloroplast thylakoid. *EMBO J.* **14**, 3905-3914.

Voelker, R., Mendel-Hartvig, J. and Barkan, A. (1997) Transposon-disruption of a maize nuclear gene, tha1, encoding a chloroplast SecA homologue: in vivo role of cp-SecA in thylakoid protein targeting. *Genetics* **145**, 467-478.

Wientjes, E., Oostergetel, G.T., Jansson, S., Boekema, E.J. and Croce, R. (2009) The role of Lhca complexes in the supramolecular organization of higher plant photosystem I. *J. Biol. Chem.* **284**, 7803-7810.

Wittenberg, G., Sheffler, W., Darchi, D., Baker, D. and Noy, D. (2013) Accelerated electron transport from photosystem I to redox partners by covalently linked ferredoxin. *Phys. Chem. Chem. Phys.* **15**, 19608-19614.

Yang, H., Liu, J., Wen, X. and Lu, C. (2015) Molecular mechanism of photosystem I assembly in

oxygenic organisms. *Biochim. Biophys. Acta* **1847**, 838-848.

Zygadlo, A., Robinson, C., Scheller, H.V., Mant, A. and Jensen, P.E. (2006) The properties of the positively charged loop region in PSI-G are essential for its "spontaneous" insertion into thylakoids and rapid assembly into the photosystem I complex. *J. Biol. Chem.* **281**, 10548-10554.

TABLES

Table 1. Identification of PSI subunits in mature PSI (PSI), PSI* and in a control segment of the same gel (Background) by MS/MS analysis.

Protein	PSI		PSI*		Background	
	MS	P	MS	P	MS	P
PsaA	553	31	223	16	176	9
PsaB	491	52	138	6	108	3
PsaD	177	3	104	3	n.d.	n.d.
PsaE	214	11	117	4	n.d.	n.d.
PsaF	300	19	38	3	n.d.	n.d.
PsaG	96	3	n.d.	n.d.	n.d.	n.d.
PsaK	45	7	n.d.	n.d.	n.d.	n.d.
PsaL	127	4	60	2	n.d.	n.d.
Lhca1	105	12	n.d.	n.d.	43	2

Lhca3	180	9	n.d.	n.d.	90	2
Lhca4	206	11	n.d.	n.d.	n.d.	n.d.
Trypsin[1]	119	114	119	133	118	128

[1]As unrelated control protein, trypsin (the protease used for in-gel protein digestion) is also included. Proteins were identified using Mascot 2.2. MS: Mascot score; P: number of peptides; n.d.: not detected.

FIGURE LEGENDS

Figure 1. Characterization of PSI*, a PSI sub-complex that contains only a subset of PSI subunits, and is enriched in young tissue in tobacco (*Nicotiana tabacum*) plants.

(A,B) Plant material used to extract thylakoids from mature (A) and young (B) leaf tissue, respectively. White arrows indicate the parts of the leaves that were harvested. The tip is the ontogenetically oldest part of the leaf and, therefore, the distal parts were taken from mature leaves (A). By contrast, the leaf base is the ontogenetically youngest part and was excised from the young leaves (B). Scale bars: 5 cm.

(C) BN-PAGE analysis of thylakoid membrane protein complexes solubilized with the non-ionic detergent DDM. PSI*, a faint band representing a protein complex at 500 kDa can be seen to be enriched in the young leaf tissue. The sizes of the marker bands of the molecular weight marker (MW) are given in kDa. Protein samples equivalent to 20 µg chlorophyll were loaded. SC: supercomplexes; I: PSI; II(2): PSII dimers; II(1): PSII monomers; Cyt b_6f: cytochrome b_6f complex; LII(3): LHCII trimers; LII(1): LHCII monomers.

(D) Determination of the protein subunit composition of PSI* by immunoblot analysis of thylakoid protein complexes isolated from young leaf material. Protein complexes were separated by 2D BN-PAGE/SDS-PAGE analysis and PSI subunits were detected with specific antibodies. Note that PSI* does not contain detectable amounts of the PSI subunits PsaF, PsaK, and of the light-harvesting antenna protein LhcA3.

(E) Decreased PSI* abundance in mature leaf tissue. The presence of PSI* in the thylakoids of mature leaves was determined by immunoblot analysis using antibodies against subunits identified as PSI* components. While PSI* is highly abundant in young leaves (cf. panel A), it is barely detectable in mature leaves. Samples containing 20 µg chlorophyll were analyzed.

Figure 2. PSI* is a PSI assembly intermediate.

(A) Dependence of PSI* accumulation on *de novo* protein biosynthesis. Immunoblot analysis using specific antibodies against the PsaA subunit reveals that PSI* amounts decrease in the presence of protein synthesis inhibitors that block translation in the chloroplast (lincomycin: lin) and/or the cytoplasm (cycloheximide: cyc).

(B) Confirmation of PSI* as an assembly intermediate by *in vivo* pulse labeling in Arabidopsis leaves. Fresh leaves were pulse labeled for two hours with [^{35}S]-methionine followed by a two-hour chase in the presence of unlabeled methionine. Autoradiography after separation of protein complexes by 2D BN-PAGE/SDS-PAGE analysis shows that PSI* is detected during the pulse, but largely disappears after the chase. Migration of the different complexes in the first dimension is indicated as in Figure 1. The identity of the most prominently labeled proteins is indicated in the second dimension. Filled arrows denote the identifiable PSI subunits PsaA/B from mature PSI complexes. Open arrows show the location of these subunits in PSI*. A representative 2D gel from three independent biological replications is shown.

(C) Quantitation of the signals originating from PsaA/B in PSI* relative to PSI. The PsaA/B signal intensities from PSI* and fully assembled PSI were measured for both the pulse and the chase periods, and the PSI*/PSI ratio of the signals was calculated. Data are presented with SEM, and significant differences according to Student's *t*-test are denoted by an asterisk ($p < 0.05$). N=3.

Figure 3. The PsaF subunit is required for the transition from PSI* to PSI.

(A) Immunoblot analysis of wild-type tobacco plants and transgenic PsaF-deficient mutant plants (*PSAF-RNAi*) following ethanol induction of *PSAF* gene expression. The blots were loaded based on chlorophyll (Chl) amounts, and probed with anti-PsaF and anti-PsaA antibodies. While PsaF is absent from the *PSAF-RNAi* plants, PsaA still accumulates to substantial amounts.

(B) PSI contents of wild type and *PSAF-RNAi* plants as determined by differential absorbance spectroscopy. The measured contents represent photoactive PSI capable of charge separation. Data are presented with SEM, and significant differences according to Student's *t*-test are denoted by two asterisks ($p < 0.01$). $N_{\text{wild type}}=10$, $N_{PSAF\text{-}RNAi}=5$.

(C) Immunoblot analysis of 2D BN-PAGE/SDS-PAGE-separated thylakoid samples from wild-type tobacco plants, transgenic PsaF-deficient mutant plants (*PSAF-RNAi*) and transplastomic PsaA-deficient mutant plants (KD-*psaA*; Krech et al., 2012). The blots were probed with anti-PsaA and anti-PsaD antibodies. Note that PSI* strongly overaccumulates in thylakoids of the PsaF-deficient mutant. By contrast, when PsaA supply limits PSI biogenesis, no PSI* is detected, indicating its rapid conversion into mature, fully assembled PSI. Samples containing 20 µg chlorophyll were analyzed.

(D) PsaF is the limiting factor in the maturation of PSI* into PSI. Induction of RNAi suppression of *PSAF* expression was abrogated by removing the ethanol vapor in the presence of 2.5 mg/mL lincomycin. Immunoblot analysis of 2D BN-PAGE/SDS-PAGE-separated thylakoid samples probed with anti-PsaA antibodies revealed that, after 6 hours without ethanol, a large fraction of PsaA

previously found in PSI* had matured and is now found in fully assembled PSI.

(E) Charge recombination kinetics of PSI from wild type and *PSAF-RNAi* plants following a light pulse reveals similar recombination times, within the tens of millisecond range, indicating that the PSI in both plants contains a the terminal iron sulfur clusters F_A and F_B, and thus a fully assembled electron transport chain.

(F) Redox equilibration between PC and PSI. The reduced state of PC and P_{700} in darkness was normalized to zero, and their fully oxidized state during a saturating light pulse was normalized to one. When the reduction kinetics of PC and P_{700}^+ are plotted against each other, PSI from *PSAF-RNAi* plants, largely composed of PSI*, has a lower apparent redox equilibration constant (K_{app}) than the wild type, indicating a slightly less efficient reduction of the PSI donor side by PC.

(G) 77 K chlorophyll-*a* fluorescence emission spectra of wild type and *PSAF-RNAi* plants. *PSAF-RNAi* plants show a clear blue shift in their PSI emission peak indicating PSI antenna proteins that are disconnected from PSI*.

Figure 4. Assembly of peripheral subunits into PSI*.

(A) Immunoblot analysis of 2D BN-PAGE/SDS-PAGE-separated thylakoid protein complexes from the wild type and *PSAF-RNAi* plants with anti-LhcA1 antibodies. Absence of LhcA1 from PSI* and presence of large amounts of free PSI antenna proteins (accumulating mainly as monomers and dimers) when PsaF amounts are limiting demonstrate that antenna proteins attach only to mature PSI. LI(1): LhcA monomers; LI(2): LhcA dimers. Samples containing 20 µg chlorophyll were analyzed.

(B) Immunoblot analysis of the same samples with anti-PsaG antibodies. PsaG does not associate with PSI*. However, in addition to its migration with mature PSI, PsaG is found to co-migrate with protein complexes of similar size as the LhcA1 monomers and dimers.

(C) Immunoblot analysis of the same samples with anti-PsaK antibodies. While in wild-type tissue, PsaK is identified only in mature PSI, the majority of PsaK is found in PSI* in *PSAF-RNAi* plants.

Figure 5. Updated model of PSI assembly.

PSI assembly is thought to begin with the co-translational insertion of the two large core subunits PsaA and PsaB into the thylakoid membrane. Formation of the reaction center protein dimer is followed by the integration of the three "stromal ridge" subunits (PsaC, PsaD and PsaE) and the three subunits forming the LHCII-docking site (PsaH, PsaI and PsaL – note that, due to its small size, PsaI could not be detected in this study and currently is only assumed to co-assemble with PsaH and PsaL). The resulting complex, PSI*, is reasonably stable and can be detected as an assembly intermediate under conditions of high *de novo* synthesis of PSI and/or when a component required for the subsequent assembly steps is present in limiting amounts. At this point, PsaK can already attach to the growing PSI complex, although, in wild-type plants, the majority of it appears to be incorporated later. The next subunit to assemble into PSI* is PsaF, which is required for the transition from PSI* to PSI. Incorporation of PsaF is accompanied or followed by the addition of the remaining PSI subunits and attachment of the LHCI antenna proteins. The latter preassemble with the PsaG subunit into another intermediate subcomplex. Fully assembled, mature PSI is distinguished from PSI* by the presence of the G, F, J, K and N subunits and the four LhcA proteins forming the LHCI antenna.