Specific Lhc Proteins Are Bound to PSI or PSII Supercomplexes in the Diatom *Thalassiosira pseudonana*[1][OPEN]

Claudio Calvaruso,[a] Anne Rokka,[b] Eva-Mari Aro,[c] and Claudia Büchel[a,2,3]

[a]Institute for Molecular Biosciences, Goethe University of Frankfurt, 60438 Frankfurt, Germany
[b]Turku Bioscience, University of Turku and Åbo Akademi University, FI–20520 Turku, Finland
[c]Department of Biochemistry, Molecular Plant Biology, University of Turku, FI–20520 Turku, Finland

ORCID IDs: 0000-0003-1165-1258 (C.C.); 0000-0003-1482-9154 (A.R.); 0000-0002-2922-1435 (E.-M.A.); 0000-0001-6640-7610 (C.B.)

Despite the ecological relevance of diatoms, many aspects of their photosynthetic machinery remain poorly understood. Diatoms differ from the green lineage of oxygenic organisms by their photosynthetic pigments and light-harvesting complex (Lhc) proteins, the latter of which are also called fucoxanthin-chlorophyll proteins (FCP). These are composed of three groups of proteins: Lhcf as the main group, Lhcr that are PSI associated, and Lhcx that are involved in photoprotection. The FCP complexes are assembled in trimers and higher oligomers. Several studies have investigated the biochemical properties of purified FCP complexes, but limited knowledge is available about their interaction with the photosystem cores. In this study, isolation of stable supercomplexes from the centric diatom *Thalassiosira pseudonana* was achieved. To preserve in vivo structure, the separation of thylakoid complexes was performed by native PAGE and sucrose density centrifugation. Different subpopulations of PSI and PSII supercomplexes were isolated and their subunits identified. Analysis of Lhc antenna composition identified Lhc(s) specific for either PSI (Lhcr 1, 3, 4, 7, 10–14, and Lhcf10) or PSII (Lhcf 1–7, 11, and Lhcr2). Lhcx6_1 was reproducibly found in PSII supercomplexes, whereas its association with PSI was unclear. No evidence was found for the interaction between photosystems and higher oligomeric FCPs, comprising Lhcf8 as the main component. Although the subunit composition of the PSII supercomplexes in comparison with that of the trimeric FCP complexes indicated a close mutual association, the higher oligomeric pool is only weakly associated with the photosystems, albeit its abundance in the thylakoid membrane.

Photosystems have evolved in photosynthetic prokaryotes and eukaryotes, adapting pigments, reaction centers, and antenna complexes to the different environmental conditions (Blankenship, 2010). In the oxygenic organisms, photosystem reaction centers are highly similar between prokaryotes and eukaryotes (Nelson and Junge, 2015), whereas higher variability is present in the light-harvesting systems (Neilson and Durnford, 2010; Nowicka and Kruk, 2016).

Diatoms are unicellular eukaryotic microalgae that originated from a secondary endosymbiosis event between a eukaryotic cell and a red algal ancestor (Bhattacharya et al., 2007). Consequently, chloroplasts exhibit a four-membrane envelope and thylakoids are organized in a three-band structure, lacking the grana-stroma organization (Bedoshvili et al., 2009). Spatial separation of the photosystems is not as defined as in land plants, although PSI is observed at a higher abundance in the outer membranes (Pyszniak and Gibbs, 1992; Flori et al., 2017). Since the spatial separation is not as strict as in land plants, the mechanism for preventing energy spillover from PSII to PSI is still unknown. Photosynthetic pigments include chlorophyll *a*, chlorophyll *c*, fucoxanthin as the main carotenoid, and diatoxanthin/diadinoxanthin involved in the xanthophyll cycle (Kuczynska et al., 2015). The pigments are cofactors of the antenna proteins that assemble into Fucoxanthin-Chlorophyll *a*/*c* Protein (FCP) complexes. FCP subunits are encoded by different gene families: Lhcf proteins, which are mainly involved in the light-harvesting mechanism; Lhcr proteins, which resemble the only Lhcs of the red ancestor and are mainly associated with PSI; and Lhcx proteins, for some of which a photoprotective function has been proven (Büchel, 2020). From genomic sequence analysis, the so-called FCP genes have also been annotated, but their function remains unknown (Armbrust et al., 2004). In diatoms like the model organism *Thalassiosira pseudonana*, the number of expressed *Lhc* genes is higher than in organisms

[1]This work was supported by the European Union Research and Innovation Program Horizon 2020 ITN (grant no. 675006) and by the Academy of Finland (grant no. 307335).

[2]Author for contact: c.buechel@bio.uni-frankfurt.de.
[3]Senior author.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Claudia Büchel (c.buechel@bio.uni-frankfurt.de).

C.C. and C.B. designed the research plan; C.B. supervised the experiments; C.C. performed the biochemical experiments and analyzed the data; A.R. performed MS analysis; C.C., A.R., E.-M.A., and C.B. wrote the article.
[OPEN]Articles can be viewed without a subscription.
www.plantphysiol.org/cgi/doi/10.1104/pp.20.00042

of the green lineage (Teramoto et al., 2001), with 11 *Lhcf*, 14 *Lhcr*, and five *Lhcx* gene products (Armbrust et al., 2004). The influence of this higher variability on the antenna complex organization is still under debate. Trimeric and oligomeric FCP complexes were isolated from the centric *T. pseudonana* and *Cyclotella meneghiniana* (Grouneva et al., 2011; Gundermann et al., 2019), but their physical association with the photosystems is still unclear.

Isolation of native-state photosystems is required to understand the actors involved in the light-harvesting process and its regulatory mechanisms. The isolation of stable complexes also allows us to determine their structure and the interaction between subunits, which was successfully achieved for the complexes from plants and green algae (Caffarri et al., 2009; Haniewicz et al., 2015; Mazor et al., 2017; Qin et al., 2019; Shen et al., 2019). Only recently, the structure of a PSII-FCPII complex from the diatom *Chaetoceros gracilis* was reported (Nagao et al., 2019a; Pi et al., 2019), but the lack of a fully sequenced genome for this organism limited subunit identification.

By contrast, the genome sequence availability for *T. pseudonana* (Armbrust et al., 2004) allows the identification of Lhc proteins by mass spectrometric analysis. In addition, a protocol for the isolation of intact plastids was established recently (Schober et al., 2018). Previous studies employed different methods for the isolation of thylakoid complexes: native PAGE, Suc density gradient ultracentrifugation, and gel filtration (Caffarri et al., 2009; Järvi et al., 2011; Barera et al., 2012). In this study, large-pore blue native PAGE (lpBN-PAGE), clear native PAGE (CN-PAGE), and Suc density gradient centrifugation were chosen. In particular, native PAGE achieves higher complex resolution in comparison with the other techniques and allows the different photosystem assemblies to be distinguished. This facilitates conclusions about the spatial organization of the subunits, as previously observed in Arabidopsis (*Arabidopsis thaliana*; Järvi et al., 2011).

In this study, we provide insights into the organization of the multitude of Lhc proteins of diatoms around the photosystems, wherein we identify the complement of PSII-associated antenna proteins.

RESULTS

Analysis of Thylakoid Membrane Complexes

We analyzed the composition of thylakoid membrane complexes in *T. pseudonana* by using lpBN-PAGE, CN-PAGE, and Suc density gradient techniques (Fig. 1). In particular, we focused on the isolation of native photosystem supercomplexes. Thylakoid complexes isolated from Arabidopsis were used as reference material (Fig. 1, A and B, lanes 3; Supplemental Fig. S1) for the identification of bands and the estimation of the molecular mass (MW) range on gels.

As a first step, we evaluated if high-MW complex isolation was possible using thylakoid membranes as starting material or whether intact plastids were needed. The same conditions (7.5 μg of chlorophyll *a* and a 30-min solubilization with 0.75% [w/v] *n*-dodecyl-α-D-maltoside [α-DDM] at 4°C) were applied to both *T. pseudonana* samples. After incubation, complexes were separated by lpBN-PAGE and the band patterns were compared. Intact plastids revealed distinct high-MW complexes (Fig. 1A, lane 2), with four major bands visible in the region above the PSII dimer band, whereas thylakoid membrane samples provided much less of high-MW supercomplexes (Fig. 1A, lane 1). An estimation of the MW range for the four major bands from plastids was carried out by comparing *T. pseudonana* plastid and Arabidopsis thylakoid samples, solubilized with α-DDM (Supplemental Fig. S1). PSII supercomplex bands were labeled according to Järvi et al. (2011). In both cases, supercomplex bands migrated in a similar area of the gel, indicating a comparable MW range of the complexes.

Different detergent concentrations (Supplemental Fig. S2) and incubation times (Supplemental Fig. S3) did not produce any improvement. Thus, according to these results, intact plastids were chosen as starting material and used for all the successive experiments.

BN-PAGE makes use of Coomassie Brilliant Blue (Coomassie) as a negatively charged molecule that conveys the charge required for separation of the protein complexes. This bulky dye might produce artifacts, affecting the complexes with highest MW. Thus, we also used CN-PAGE as an alternative native gel system. Figure 1B shows the comparison of complexes from plastids run on CN-PAGE and BN-PAGE gels. A similar band pattern was observed in the supercomplex area, using lpBN-PAGE (Fig. 1B, lane 1) and CN-PAGE (Fig. 1B, lane 2), excluding side effects introduced by Coomassie.

As a further control, Suc gradient separation of plastid complexes was also performed (Fig. 1C, left). After the run, three fractions were harvested and the main complexes were identified by spectroscopic analysis (Supplemental Fig. S4, A–D). The uppermost Suc density gradient band was assigned to FCP complexes, showing absorption and emission spectra similar to the isolated FCP analyzed by Gundermann et al. (2019). The middle band was identified as PSI complex because of the red shift of the Q_y absorption (Supplemental Fig. S4A) and the emission peak at 715 nm in the 77K fluorescence emission spectrum (Supplemental Fig. S4C), as described by Veith and Büchel (2007). The 715-nm peak was also observed when chlorophyll *c* was excited (Supplemental Fig. S4D), indicating energy transfer from the antenna to the reaction center of PSI. The lowest Suc density gradient fraction was attributed to PSII because the main emission occurred at 687 nm in the fluorescence emission spectra (Supplemental Fig. S4, B–D). This attribution is also in agreement with the results described by Pi et al. (2019). The complexes of those fractions were next separated by lpBN-PAGE and compared with solubilized complexes without prefractionation by Suc density gradient (Fig. 1C, right). The upper Suc density gradient band

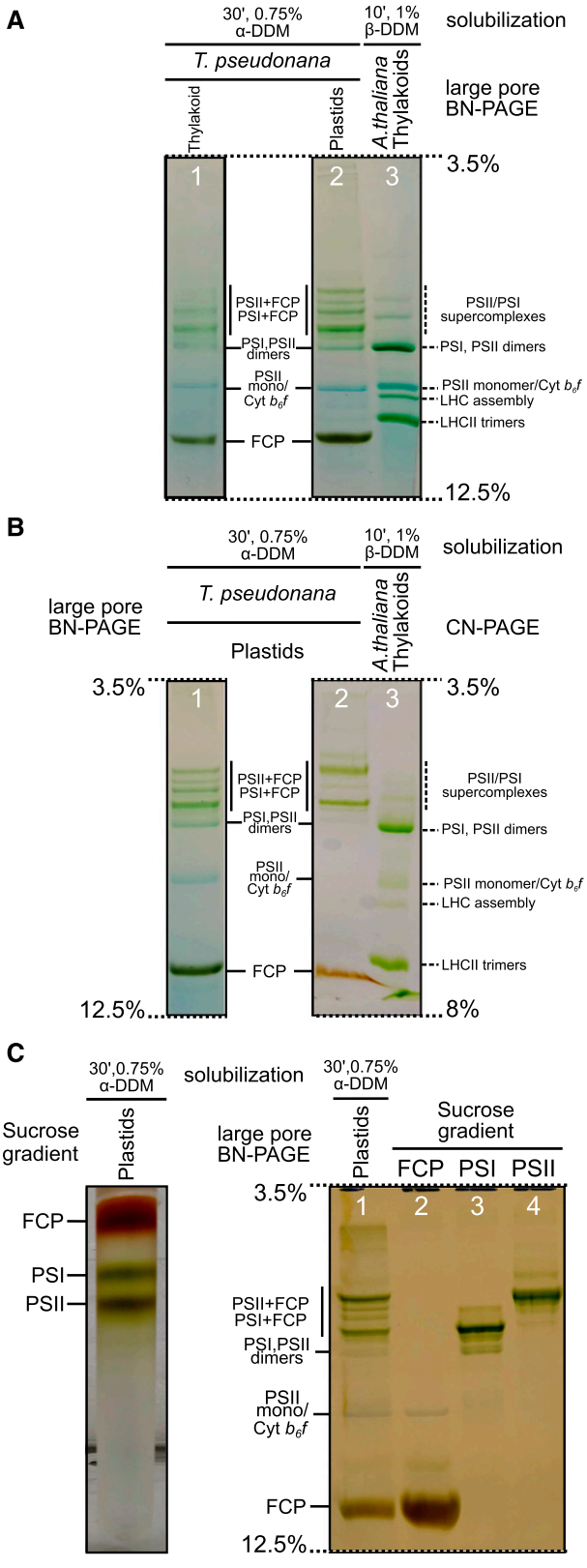

Figure 1. Separation of native-state thylakoid membrane complexes from *T. pseudonana* (with Arabidopsis complexes as a reference). A, lpBN-PAGE (3.5% to 12.5%) of solubilized photosynthetic proteins; 7.5 μg of chlorophyll *a* per lane was loaded in each case. Dotted lines are used to label complexes of Arabidopsis according to Järvi et al. (2011), and solid lines for the complexes of *T. pseudonana* according to Grouneva et al. (2011). Image contrast was enhanced to improve the visualization of the bands. Lanes 1 and 2, Comparison of the supercomplex separation using either thylakoid membranes (lane 1) or intact plastids (lane 2) of *T. pseudonana*, both solubilized using 0.75% (w/v) α-DDM at 4°C for 30 min. Lane 3, Thylakoid membranes from Arabidopsis (solubilization with 1% [w/v] *n*-dodecyl-β-D-maltoside [β-DDM] at 4°C for 10 min) were used as a reference. B, Comparison of separation patterns on CN-PAGE (3.5%–8%) and lpBN-PAGE. Lanes 1 and 2, Intact plastids of *T. pseudonana* by lpBN-PAGE (lane 1) and CN-PAGE (lane 2); solubilization and amount as in A. Lane 3, Thylakoid membranes of Arabidopsis as a reference (solubilization and amount as in A). C, Separation of *T. pseudonana* complexes by Suc density gradient centrifugation and analysis with lpBN-PAGE. Left, Thylakoids were solubilized as in A, and an amount corresponding to 100 μg of chlorophyll *a* was separated by ultracentrifugation (16 h at 132,000*g*, 4°C). The three fractions are labeled according to Pi et al. (2019). Right, Comparison between directly solubilized plastids (lane 1; solubilization and amount as in A) and Suc density gradient fractions (lanes 2–4; chlorophyll *a* not determined).

contained mainly FCP complexes, with slight contaminations of cytochrome b_6f (Fig. 1C, right, lane 2). The main complex of the PSI fraction aligned with the lowest MW band in the supercomplex area (Fig. 1C, right, lane 3). The main complex of the PSII fraction corresponded to the uppermost band of the intact plastid sample (Fig. 1C, right, lane 4). In both cases, the main complexes were accompanied by some minor bands. Most of them probably corresponded to different states of PSI and/or PSII, since they were also found when plastids were solubilized directly. The subunit composition of the main PSI and PSII supercomplexes was also comparable, whether isolated with or without prior Suc density gradient fractionation, as shown by 2D SDS-PAGE (Supplemental Fig. S4E). These results demonstrate that essentially the same complexes can be isolated by different preparation procedures.

Identification of the Complexes Separated by BN-PAGE

Mass spectrometry (MS) analysis was conducted to identify the supercomplexes found via BN-PAGE. The analysis had the scope to prove the tentative assignment to PSI and PSII made above. For sample preparation, gel bands were excised from lpBN-PAGE carried out using a 3.5% to 8% acrylamide gradient (Fig. 2A). The lower acrylamide concentration improved the resolution of the bands, reducing cross-contaminations of closely migrating complexes. Photosystem core subunits were identified and their relative abundances analyzed per band by lpBN-PAGE.

A spectral counting approach was used to determine relative abundances of identical proteins from the different supercomplex bands of BN-PAGE (BN-BANDs). This is possible because an increase in protein abundance typically results in an increase in the number of its proteolytic peptides, and vice versa. This increased number of (tryptic) digests then usually results in an increase in protein sequence coverage, the number of

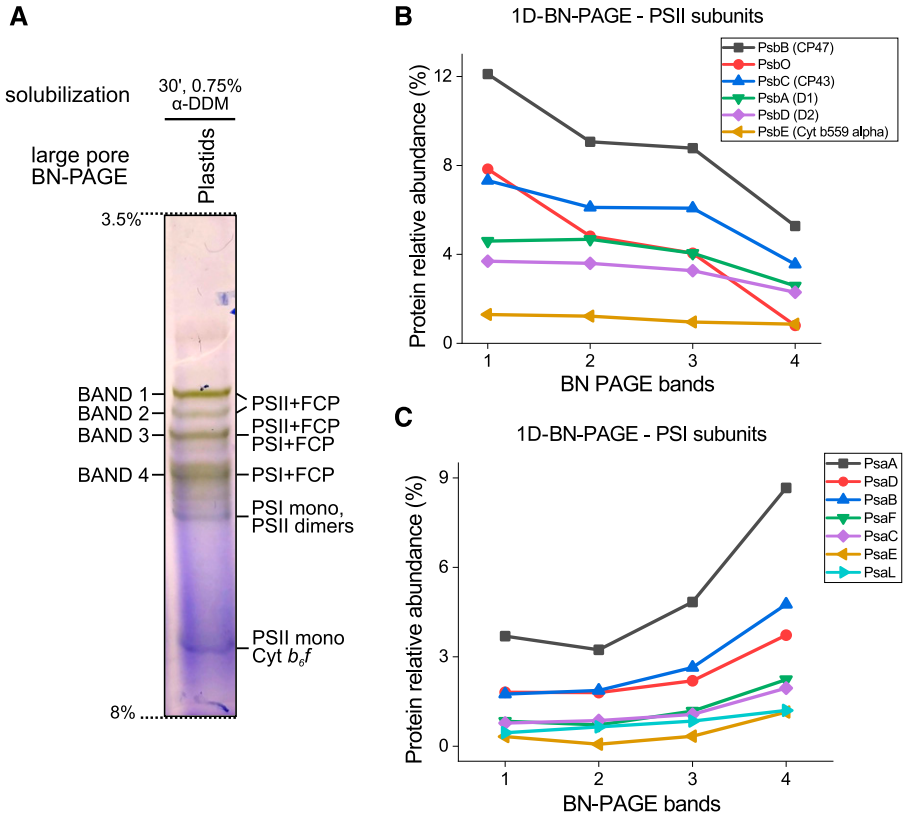

Figure 2. Analysis of the core subunits of PSI and PSII complexes in 1D lpBN-PAGE. A, Short gradient (3.5%–8%) lpBN-PAGE was performed to improve the resolution of the bands; otherwise, conditions were the same as described in Figure 1A. BN-BANDs 1 to 4 were excised, proteins digested, and the peptides analyzed with LC-electrospray ionization-MS/MS. B and C, The relative abundance of proteins in BN-BANDs 1 to 4 (for calculation, see "Materials and Methods"). B shows the analysis of PSII core subunits, and C shows the analysis of PSI core subunits. Only proteins with an abundance of minimally 1% in at least one band are depicted.

identified unique peptides, and the number of identified total MS/MS spectra (spectral count) for each protein (Washburn et al., 2001). This approach allowed us to compare the relative abundance of individual proteins along the four different supercomplex bands of the BN gel. The relative abundance (see "Materials and Methods") gives an estimate of the number of times that peptides belonging to a single protein were detected in relation to the total number of peptides in a given sample. The complete MS data from the BN-BANDs can be found in Supplemental Table S1. Note, however, that the ratios of peptides determined do not necessarily reflect the real stoichiometry in the protein pool of the sample, and comparison of the abundances of different proteins on the same gel band is impossible.

Figure 2, B and C, show the relative abundances of PSII (Fig. 2B) and PSI (Fig. 2C) subunits across the four BN-BANDs. The subunits of PSII (PsbA, PsbB, PsbC, PsbD, PsbE, and PsbO) and PSI (PsaA, PsaB, PsaC, PsaD, PsaE, PsaF, and PsaL) were all detected in the different supercomplex bands. The overall result demonstrated a decreasing abundance of PSII proteins from BN-BAND 1 to BN-BAND 4, whereas the opposite trend was observed for the PSI protein group. According to these results, the two upper bands corresponded mostly to PSII complexes, BN-BAND 3 represented an overlap between PSI and PSII, and BN-BAND 4 predominantly contained PSI complexes. The analysis of the trends provided qualitative estimates for the most representative complex of each band, based on the relative abundance of the complex subunits. Nevertheless, both PSII and PSI proteins were detected on all the gel bands, so colocalization of different complexes on the same gel band cannot be excluded. During excision, smaller bands might overlap with the main bands and might contaminate the latter. Also, the possibility of smearing between different bands (cross-contamination) and the extremely high sensitivity of the mass spectrometric analysis have to be taken into account. All these factors have been considered in the interpretation of the data given below.

When analyzing the bands obtained by CN-PAGE, the photosystem core subunits showed similar trends to those above. However, the proteins of the oxygen-evolving complex (OEC) were better retained by CN-PAGE (Supplemental Table S2), whereas the same Lhc proteins were found on both gel types (Supplemental Table S3). Since we were interested mainly in the Lhc complement of the different supercomplexes, and the stability of binding of the different Lhc subunits to the cores, we further focused on the analysis of complexes separated by BN-PAGE.

Analysis of the Photosystem Core Complex Subunits with 2D SDS-PAGE

To confirm the results obtained by BN-PAGE, 2D SDS-PAGE was performed. This analysis led to further identification of the subunits of PSII and PSI

supercomplexes and, more importantly, their Lhc supplement (see next paragraph).

For the analysis, one whole BN gel lane was denatured and subjected to 2D SDS-PAGE for the separation of complex subunits. In another approach, the four most prominent bands were excised from the supercomplex area and separated individually through SDS-PAGE, yielding essentially the same results (Supplemental Fig. S5). After silver staining, the subunits were assigned according to MW by literature comparison (Nagao et al., 2010; Grouneva et al., 2011).

Figure 3 shows the 2D SDS-PAGE results from a BN gel lane in the higher MW range. For the assignment of spots, two regions were analyzed: the area between 20 and 25 kD, containing most of the antenna proteins (Lhcf, Lhcr, and Lhcx, boxed in Fig. 3), and the remaining part of the gel (120–25 kD and below 20 kD). In the latter, most of the PSI and PSII core subunits were found.

The identification of core subunits on the 2D gel confirmed the results of the MS analysis conducted using 1D BN-PAGE. The core subunits of PSII (PsbB, PsbC, PsbA, and PsbD) were found predominantly in the two bands of highest MW (BN-BANDs 1 and 2). In those bands, almost no signals belonging to PSI (i.e. PsaA/B) were detected by silver staining. In BN-BAND 3, colocalization between PSII and PSI was observed, with the occurrence of PsaA/B together with PSII subunits, while in BN-BAND 4, only PSI subunits were detected. Western blots against PSI (PsaA) and PSII (PsbC) subunits confirmed the attribution to PSI and PSII (Supplemental Fig. S5C). Low-MW proteins of PSI could also be observed in BN-BAND 4: PsaD, PsaL, and PsaE, which all have MW below 20 kD. The MW shift of the different photosystem complexes by BN-PAGE might be attributed to the partial or total loss of antenna proteins or other subunits of the complex (e.g. proteins of the OEC of PSII). For instance, the magnitude of PsbO spots decreased along the different PSII complexes (in relation to the other core subunits).

It has to be pointed out that many more BN-BANDs became visible on the 2D gel. In addition to the four already described, other minor bands were revealed by silver staining of the 2D gel. Three further complexes could be observed above the main supercomplexes by BN-PAGE, above BN-BAND 1 (Fig. 3). These complexes were composed of PSI subunits only, lacking any contamination by PSII subunits. All three uppermost complexes showed the same spot pattern, with the core proteins PsaA/B, PsaD, and PsaE (but not PsaL), and a similar antenna protein profile, suggesting a multiple aggregation state of the PSI complex. Two additional complexes of PSI could be spotted: a PSI complex, running below BN-BAND 4, and another smaller complex, proximal to the PSII dimer band (Fig. 3). Two additional complexes containing PSII subunits were also observed, apart from the three PSII complexes described so far in BN-BANDs 1 to 3. The two other complexes were located in the area around BN-BAND 4 and they contained the core subunits, but not PsbO,

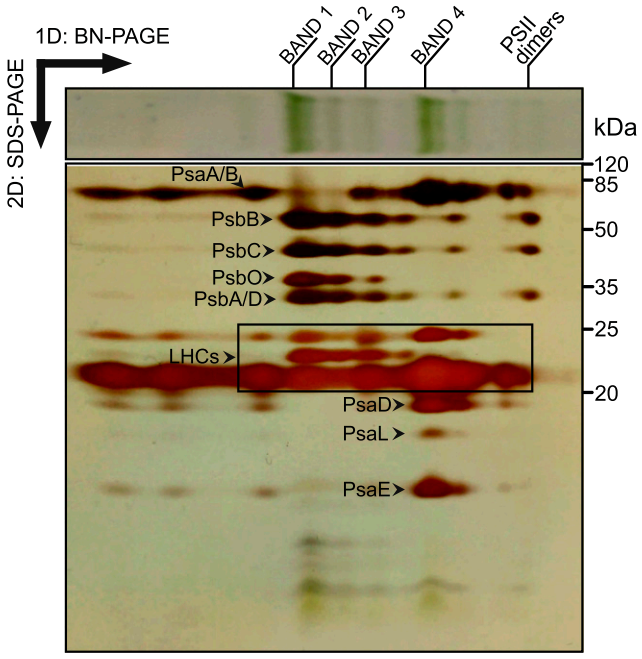

Figure 3. 2D SDS-PAGE protein identification in the PSII and PSI complexes. After separation by BN-PAGE (top), 2D SDS-PAGE was performed for the analysis of supercomplex subunits. The labeling of bands of the 1D gel (top) corresponds to that in Figure 2. The complexes in the region above the PSII dimers were resolved and subunits assigned according to Nagao et al. (2010) and Grouneva et al. (2011). Most of the photosystem core subunits were found between 120 and 25 kD and below 20 kD. The core subunits identified were as follows: PsaA/B, PsaD, PsaL, and PsaE for PSI and PsbA/D, PsaB, PsaD, and PsbO for PSII. Lhc spots were found in the area between 20 and 25 kD. The box highlights the Lhc spots that were analyzed (see Fig. 4). Contrast of the image was homogeneously enhanced in all images for the visualization of the bands.

indicating loss of the OEC. Thus, the 2D SDS-PAGE analysis gave a more detailed overview of the number of bands occurring in the area of photosystem supercomplexes.

LHC Pool of the PSI and PSII Supercomplexes

Here, we focused on the antenna proteins of the main supercomplex bands visible by BN-PAGE (Fig. 3). The analysis of the antenna proteins was conducted to identify the composition of the light-harvesting complexes bound to the reaction centers of PSII and PSI.

In Figure 4A, the gel region between 25 and 20 kD, where most of the Lhc proteins were located, is shown again. The origin of the 2D Lhc spots of this region could be assigned to the PSI or PSII supercomplexes (Fig. 3) by their relative abundances in each supercomplex (Fig. 4, B–E). The same approach that has been used for the analysis of the complexes on 1D gels was adopted for the 2D gel spots (Supplemental Table S4) as well as for the horizontal 2D gel system, with similar results (Supplemental Table S5). In protein abundance

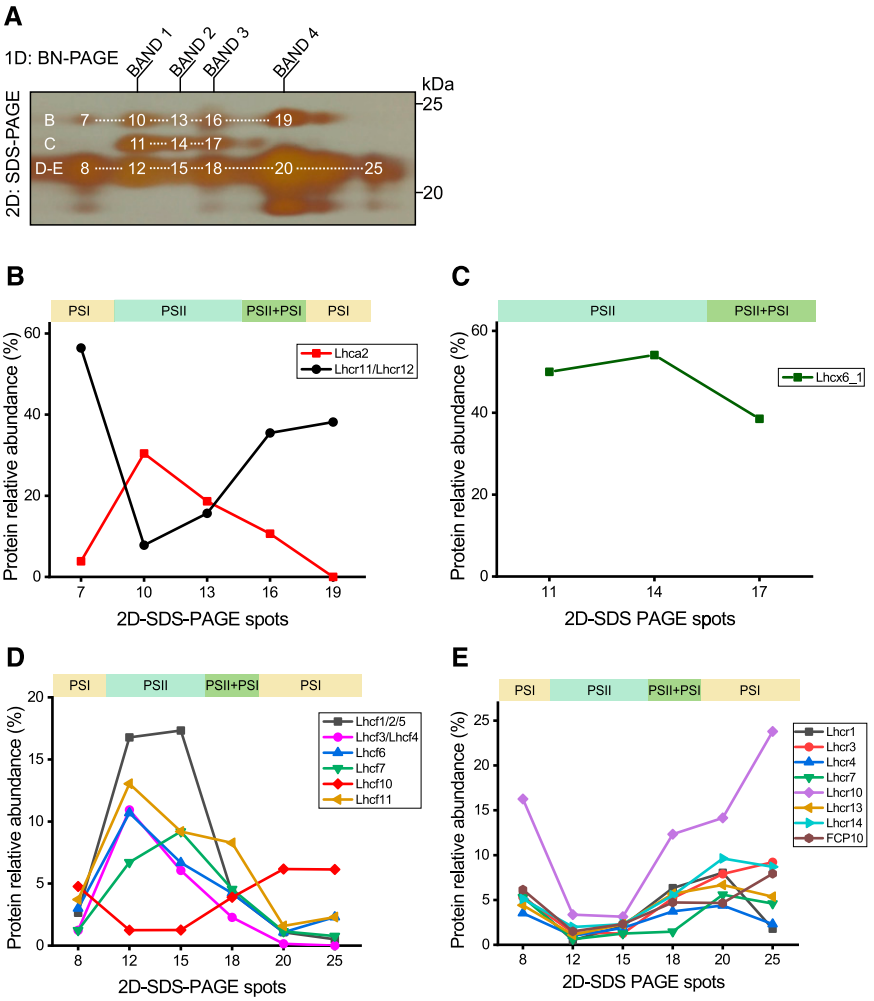

Figure 4. Analysis of the Lhc protein composition of PSI and PSII complexes by 2D SDS-PAGE. A, Magnification of the 2D SDS-PAGE results taken from Figure 3, in the MW range of 25 to 20 kD. Spots with similar MW that are compared in B to E are connected by white dotted lines. Above the box, labels indicate the positions of the 1D BN-PAGE bands as shown in Figure 3. The numbers refer to the excision order of the spots. B to E, The relative abundance of Lhcs in the spots with comparable MW. A colored bar above each graph indicates the corresponding photosystem complex on the 1D gel. Only proteins that had a relative abundance of above 1% in at least one of the spots have been used. In B, spots containing Lhc of highest MW (spots 7, 10, 13, 16, and 19) are shown, consisting mainly of Lhcr2 and the Lhcr11/12 proteins. C demonstrates spots containing Lhc of medium MW (spots 11, 14, and 17). Here, almost exclusively Lhcx6_1 was found. D and E show the analysis of spots containing Lhc of low MW (spots 8, 12, 15, 18, 20, and 25), that is, the relative abundance of Lhcf (D) and Lhcr/FCP (E) proteins.

analysis, only the most relevant proteins were included, with a consistent and clear trend for all three different gels analyzed (lpBN-PAGE, horizontal 2D, and spot 2D-PAGE). Furthermore, a relative abundance value of 1.5% at least in one of the spots/bands was used as a threshold.

The Lhc proteins were separated in three groups of spots of slightly different MWs. Spots with similar MW were analyzed together: Lhc proteins of highest MW are shown in Figure 4B, proteins of medium MW are depicted in Figure 4C, and the smallest Lhc proteins are represented in Figure 4, D and E. Assignment of spots to PSII or PSI was done as described above and is indicated by the colored bar above the graphs.

In the spots containing the Lhc proteins of highest MW (Fig. 4B), three Lhcr proteins were most abundant: Lhca2 (which is identical to Lhcr2) and Lhcr11/12, which were analyzed together because of their high sequence similarity (greater than 80%). Lhca2 showed a maximum peak of relative abundance in spots 10 and 13, with lower abundances in the other spots. Since spots 10 and 13 were derived from PSII supercomplexes, Lhca2 was assigned to the PSII antenna pool, consistent with the findings in the PSII-FCPII structure in *C. gracilis* (Nagao et al., 2019a; Pi et al., 2019). The opposite trend was observed for the distribution of Lhcr11/12, with maximum abundances in PSI-related spots.

The analysis of spots containing Lhc of medium MW is summarized in Figure 4C. These spots were mainly composed of Lhcx6_1, the sole Lhcx protein found in the study. Spots at this MW region were visible on the gel only for bands containing PSII (BN-BANDs 1–3; Fig. 4A). In BN-BAND 4, the main band of PSI, no corresponding spot was observed. Despite a lower abundance of Lhcx6_1 in the supercomplex band consisting of PSII and PSI (BN-BAND 3, spot 17), the relative abundance was high in all three spots, indicating the presence of this Lhcx protein in all the three PSII complexes identified. Most of the antenna proteins were found in the spots of lowest MW (Fig. 4, D and E). More specifically, most of the Lhcf proteins (Lhcf1/2/5, Lhcf3/4, Lhcf6, Lhcf7, Lhcf10, and Lhcf11; Fig. 4D), most of the Lhcr proteins (Lhcr1, Lhcr3, Lhcr4, Lhcr7, Lhcr10, Lhcr13, and Lhcr14; Fig. 4E), and an FCP protein (FCP10) were identified.

The Lhcf proteins are considered as the main group of light-harvesting proteins in diatoms (Büchel, 2020).

Figure 4D demonstrates that most of them exhibit a maximum relative abundance in the area of the PSII complexes. Thus, they probably belong to the antenna pool of PSII. The main peak for most of them corresponded to spot 12, a spot belonging to the PSII complex of the highest MW. Still, some exceptions were visible when analyzing the trends of relative abundance. The most evident difference concerned Lhcf10: this protein showed the opposite trend in comparison with the other Lhcf proteins. Lhcf10 relative abundance was lowest in the spots corresponding to PSII (12 and 15) and showed a higher value in the PSI-related spots (18, 20, and 25, as well as 8). The results observed here are consistent with those of Grouneva et al. (2011), who already assigned Lhcf10 to PSI.

Lhcr proteins are usually attributed to PSI, and they are evolutionarily related to the red algae PSI antenna proteins. This assumption was consistent with the results of the MS analysis (Fig. 4E). All Lhcr proteins were detected in spots assigned to PSI, with an increase in relative amount from spot 15 to spot 20, which was the maximum peak for most of the proteins analyzed. Most of the Lhcr proteins, with genes available in the databases, were found in those spots: Lhcr1, Lhcr3, Lhcr4, Lhcr7, Lhcr10, Lhcr13, and Lhcr14, and almost all showed comparable abundances in the different PSI complexes. One exception is Lhcr10, whereby higher levels were found in the smallest PSI complex (spot 25). The reason might be that all the other antenna proteins bound to the core of PSI are lost in this complex, resulting in an increased ratio between Lhcr10 and all other Lhcr proteins. Concerning the so-called FCP proteins, only FCP10 showed a clear trend, consistent in all the gel systems used and similar to the Lhcr proteins (Fig. 4E). Thus, it was attributed to the antenna pool of PSI.

In summary, by analyzing the spot composition, we were able to assign specific Lhc proteins to the supercomplexes of PSI and PSII, which were separated by 1D BN-PAGE. Lhcr proteins were found mainly in PSI complexes, with the exception of Lhca2, which was preferentially found in PSII. In total, 10 different Lhcr proteins were identified for PSI. Lhcf proteins, on the other hand, were mostly found in PSII, with the exception of Lhcf10, which seems to belong specifically to PSI. Here, nine different Lhcfs were found specifically associated with PSII. Lhcx6_1, the only Lhcx protein detected, was also enriched in PSII bands.

DISCUSSION

Factors Influencing the Stability of Multisubunit Complexes

In this study, we show how the use of isolated plastids, instead of isolated thylakoids, improves the probability of isolating nearly native-state complexes from diatoms. The isolation of diatom plastids (Schober et al., 2018) was developed with the aim to study the properties of the photosynthetic apparatus. This method allowed the isolation of high-quality plastids without strong contamination by other organelles such as mitochondria (Schober et al., 2019). In this study, we compared the outcome from solubilizing either intact plastids or isolated thylakoids with α-DDM prior to separation of the protein complexes on BN-PAGE. The result showed different band patterns using the two different starting materials. The PSI band pattern appeared to be similar between the two samples, unlike PSII, where PSII supercomplexes were almost exclusively detected in the plastid sample. Therefore, PSII complexes showed a high tendency of disassembly, which may be an intrinsic feature as a consequence of its high turnover rate (Li et al., 2016). Membrane integrity could also play a role in the stability of the photosystems. In fact, the stacking degree of membranes is the main difference between plastid and thylakoid samples. As described by Jäger and Büchel (2019), comparison between cells, isolated plastids, and thylakoids using circular dichroism spectroscopy revealed that the structural integrity of the membranes was preserved in plastids but not in thylakoids. Thus, the effectiveness of photosystem solubilization in a highly assembled state is likely influenced by membrane integrity, in particular for PSII.

Another hint for this was given by comparison of different isolation media used for the preparation of PSII complexes. The divalent cation concentration, in particular that of Mg^{2+}, influences the distance between the lamellae, regulating the stacking degree (Jäger and Büchel, 2019). Evidence for stromal interactions between PSII was found in land plants (Albanese et al., 2017), but nothing is known so far for diatoms. Using an optimal concentration of Mg^{2+} that leads to a physiological stacking state of lamellae might help the interaction of the high-MW complexes across the stromal gap. The isolation of stable PSII-FCP supercomplexes was only recently achieved in diatoms (Nagao et al., 2019a; Pi et al., 2019), and in both cases, $MgCl_2$ was present during the isolation. Conversely, for Grouneva et al. (2011), no Mg^{2+} cations were provided and no evidence for PSII supercomplexes was found.

The influence of the stacking degree on the stabilization of PSII might also reflect the different behavior of PSI supercomplexes. PSII was observed to be localized predominantly in the appressed region of the thylakoid membranes, unlike the PSI complex, which is localized preferentially in the outer lamellae, even though the separation between PSII and PSI is not as defined as in land plants (Flori et al., 2017). Because of the different environment, the stability of PSI complexes might not be influenced by stromal gap interactions, and this could explain why it is not affected by the loss of the stacking in the thylakoid sample during solubilization.

Therefore, despite the limited knowledge about the stability of photosystem supercomplexes in diatoms, the use of isolated plastids severely improved the biochemical characterization of the thylakoid complexes.

LHC Composition of PSI and PSII Supercomplexes

In diatoms, previous studies about the Lhc protein composition were mostly conducted on the free pool of FCP complexes (Lepetit et al., 2007; Röding et al., 2018; Gundermann et al., 2019), and only limited information is available about the Lhcs bound to the photosystem cores. In our study, we underlined the differences between the Lhcs bound in PSI and PSII supercomplexes, focusing on their specificity. To assign the antenna protein to the corresponding photosystem, the relative abundance of each Lhc protein was checked against those of photosystem core proteins. Thus, proteins that were coherently found in the respective bands were assigned to the same complex.

The antenna pool of PSII was composed of most of the Lhcf proteins (Lhcf 1/2, 3/4, 5, 6, 7, and 11), one Lhcr protein (Lhcr2), and the Lhcx6_1 protein. Members of the Lhcf family were detected in all the PSII complexes, indicating stable binding and thus a localization closer to the reaction center. In the PSII-FCPII structure of *C. gracilis*, the antennae are symmetrically organized, with two tetramers and three monomers bound to each side of the PSII core dimer. The subunits of the tetramers were either identified as Lhcf1 (Nagao et al., 2019a) or as Lhcf8 (Pi et al., 2019), which were both found in PSII in our study as well. Lhcf1 was very prominent in our PSII complexes, whereas Lhcf8 had a very low relative abundance (below the threshold) compared with the other members of the Lhcf family (Supplemental Tables S1, S4, and S5). According to Grouneva et al. (2011), Lhcf8 is the main subunit of the oligomeric FCP pool in *T. pseudonana*, and the same holds for the oligomeric FCPb complex in the closely related centric diatom *C. meneghiniana* (named Fcp5 or Lhcf3 in this species; Gundermann et al., 2019). So apparently in *T. pseudonana*, the FCP oligomers are not tightly connected to PSII supercomplexes and detach easily during solubilization, although energy transfer from FCPb to PSII was demonstrated at least for *C. meneghiniana* (Chukhutsina et al., 2013). Another interesting fact is the presence of Lhcr2 in the antenna pool of PSII. Lhcr proteins were considered to be part of the PSI antenna pool, so the occurrence of Lhcr2 in the PSII supercomplex contradicts this general consensus. Furthermore, Lhcr2 is probably bound to PSII as a monomer, since it has neither been found in the pool of FCP trimers nor oligomers (Grouneva et al., 2011). This hypothesis is also confirmed by the PSII-FCPII structure of *C. gracilis*, where one monomer was identified as Lhca2 (Pi et al., 2019).

Concerning the Lhcx group, only Lhcx6_1 was detected in PSII supercomplexes, so it seems to be constitutively expressed even under the very-low-light conditions used here. Another Lhcx protein known to be expressed under low light is Lhcx1 (Zhu and Green, 2010), but no evidence for its presence in PSII supercomplexes was found. Its absence might be explained by the much lower growth light intensities used here when, due to high cell densities, self-shading reduced the amount of light experienced by the cells even further, inducing a down-regulation of the photoprotective proteins. Another possibility is that Lhcx1 is solely found in the peripheral antenna complexes (i.e. the trimeric FCPs; Grouneva et al., 2011), detached from the supercomplex.

The antenna pool of PSI supercomplexes consisted of Lhcr proteins (except Lhcr2 as described above), Lhcf10, and FCP10. Although no Lhcx protein spots related to the PSI supercomplex were detected on the silver-stained 2D gel, the mass spectrometric analysis of the 1D gel (Supplemental Table S1) showed the presence of Lhcx6_1 in BN-BAND 4, where PSI is the main complex. Since Lhcx6_1 was also detected in PSI by Grouneva et al. (2011), we cannot exclude its presence in the antenna pool of PSI supercomplexes.

Populations of Isolated Supercomplexes

In this study, we focused on the different organization states of PSII and PSI supercomplexes. Photosystems are highly dynamic multisubunit complexes, and their subunits change according to the environmental conditions, on short- and long-term time scales (Rochaix, 2014). So structural reorganization is an intrinsic feature of those complexes.

During solubilization, the detergent disrupts the interactions between proteins, generating smaller forms of the same complex. Using Suc density centrifugation, the biggest PSII supercomplex was mainly isolated, but bands of lower MW were seen on both native gel systems. Thus, those bands are either due to the action of the detergent or they represent assembly states. In both cases, the presence or absence of subunits is indicative of the binding strength and also contains information about the subunit localization with respect to the core complex. This is particularly interesting for the antenna pool, because it helps to distinguish between outer and inner LHC proteins.

PSII

On the 2D gel, five different PSII-FCP states were identified, with different MWs. On the 1D gel, the three highest MW PSII complexes correspond to BN-BANDs 1 to 3, whereas the two lowest are located above and below BN-BAND 4. Differences between the bands could be assigned to the detachment of pigmented proteins (like Lhc) or other unpigmented subunits (like PsbO). Another explanation might be the presence of dimers of supercomplexes, but this hypothesis appears improbable because the bands are all found in the same gel region as Arabidopsis supercomplexes, but not where megacomplexes are found (Supplemental Fig. S1). Although complexes from different species probably show slightly different running behavior, such strong differences seem unlikely.

PsbO is the main protein of the OEC of PSII. The corresponding spot runs at ~35 kD, above the D1/D2

spot. Comparing the magnitude of the spots, PsbO shows a decreasing trend of its relative amount in the three PSII BN-BANDs, compared with PSII core proteins. This trend is confirmed by the results of the MS analysis. This sensitivity is probably due to the lumenal location but was less pronounced when using CN-PAGE instead of BN-PAGE (Supplemental Table S2). Therefore, the detachment could be favored by charged molecules (e.g. Coomassie) absent in the CN-PAGE.

The composition of the antenna subunits is more complex, but some hypotheses can be postulated. Lhca2 (annotated as Lhcr2 in the Joint Genome Institute database) signals were detected in all the PSII complexes, indicating a strong interaction with the core. This observation is fortified by the PSII-FCPII structure of *C. gracilis* (Pi et al., 2019), where Lhca2 is located in the inner part of the antenna pool, in direct contact with the reaction center. A different behavior was observed for Lhcx6_1. The PSII complex with lowest MW was detected with 2D SDS-PAGE, running below BN-BAND 4. In this complex, the Lhca2 spot is still present, whereas the spot of Lhcx6_1 is missing. Its position should thus be more peripheral than that of Lhca2. No Lhcx proteins were detected in the structure of *C. gracilis*, but two of the three monomers (FCP-E and FCP-F) could not be assigned to specific proteins, so we cannot rule out that one of them corresponds to Lhcx6_1. As reported by Grouneva et al. (2011), Lhcx6_1 was detected as part of the peripheral FCP trimers, and in *C. meneghiniana* (Gundermann et al., 2019), also a minor population of the trimeric FCPa contained this protein. So it appears more probable that Lhcx6_1 has a peripheral localization, and, depending on isolation method, is either found in PSII supercomplexes or as part of the free pool of trimeric FCP.

The analysis of Lhcf proteins is more difficult because of the presence of several Lhcf(s) in the same spot on the gels and the high sequence similarity of some of them (Lhcf1/2, Lhcf3/4, Lhcf5, Lhcf6, Lhcf7, and Lhcf11). Lhcf proteins were observed here in all the forms of PSII supercomplexes and in the pool of free FCPs (trimeric as well as higher oligomeric; Grouneva et al., 2011), so we assume their occurrence in both the peripheral and inner districts of the antenna pool. For those proteins, similar trends were observed and the small variations (like Lhcf7) could be due to the MS analysis. Thus, the smaller PSII complexes had gradually lost Lhcx6_1, certainly some of the Lhcf proteins, and part of the OEC.

In the largest PSII complex, we identified six Lhcf proteins of high abundance, one Lhcx protein, one Lhcr protein, and some other Lhc proteins of low abundance. This is in contrast to the studies by Nagao et al. (2019a) and Pi et al. (2019), where also one Lhcr but no Lhcx and only three different Lhcf proteins were assigned (two monomers and the Lhcf that builds both the M- and S-tetramers). However, *C. gracilis* is not sequenced, which made attributions difficult. Lhcf proteins are all highly similar in sequence, and the authors give some indications for a nonidentical build of the M- and S-tetramers. If one considers the high similarity of Lhcf proteins, and follows this line of argument, the PSII-FCP supercomplexes that were structurally analyzed could accommodate maximally 10 different Lhcf proteins, whereas our study revealed at least six Lhcf proteins and one Lhcx. For the latter, the argument of sequence homology does not hold, but on the level of resolution of the structure, its presence cannot be ruled out. So the question arises whether the largest PSII supercomplexes analyzed here are of the same size as those of *C. gracilis*. The fact that complexes isolated via Suc density centrifugation, as done for the structural analysis, were the largest found on BN gels might argue for a comparable size, despite small differences in the isolation protocols and the species difference. However, we cannot rule out that exposing the PSII samples from Suc gradients to an additional electrophoretic separation led to a loss of subunits, preventing direct comparison with the published structure. On the other hand, the presence of Lhcx6_1 in the largest PSII complexes might indicate slightly bigger complexes. Our methods are thus not sufficient to estimate the precise size, and further studies are necessary to investigate the exact location of each single Lhcf protein in the antenna pool of PSII as well as the actual sizes of the complexes.

Concerning the physiological relevance of the complexes, Levitan et al. (2019) already observed two subpopulations of PSII in *Phaeodactylum tricornutum*: PSII-FCP complexes and PSII core clusters, the latter assumed as a repair station for the PSII reaction centers. The PSII-FCP complexes resembled those of *C. gracilis* in size; however, due to the resolution of the method, bigger supercomplexes cannot be ruled out. Here, especially by CN-PAGE, the amounts of PSII core dimers were minor. Despite the evolutionary distance between the two species and differences in growth conditions that might change the amount of certain subpopulations, most probably PSII cores and one or two PSII supercomplexes coexist in vivo, resembling the situation in land plants, with core dimers found as assembly states as well as $C_2S_2M_2$ supercomplexes and those containing L-trimers as well (Boekema et al., 1999).

PSI

Concerning PSI supercomplexes, different studies analyzed isolated PSI-FCPI complexes in diatoms, but usually only one type was isolated and PSI was found to be monomeric, as in land plants (Veith and Büchel, 2007; Ikeda et al., 2013; Nagao et al., 2019b). In our study, four different states could be observed by lpBN-PAGE in the area between the most abundant PSII supercomplexes and the PSII core dimer. Above this area, another three PSI complexes were present, but these are probably the result of aggregation of several PSI complexes, since the protein composition did not change. The MW differences of the smaller complexes are most probably due to a detachment of Lhcr proteins.

Most Lhcr proteins showed similar abundance in the different PSI complexes. The only exception is Lhcr10,

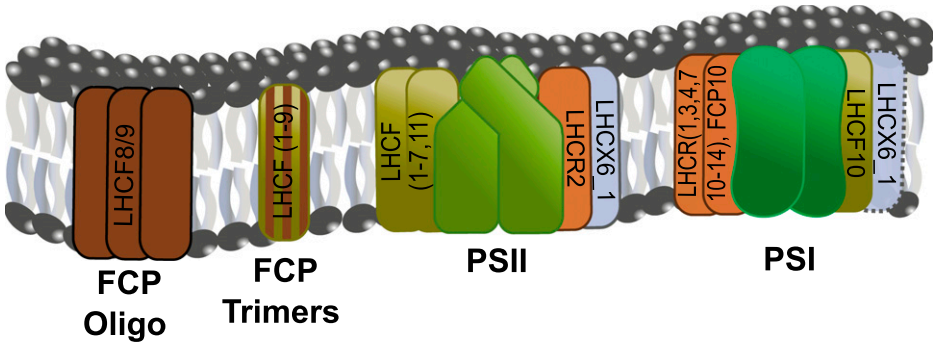

Figure 5. Model for photosystem organization under low-light conditions. The model summarizes the main results of the study. Four complexes are represented: FCP higher oligomers (FCP Oligo), FCP trimers, PSII supercomplexes, and PSI supercomplexes. Lhcf1 to -9 constitute the trimers, and FCP higher oligomers have Lhcf8/9 as main components according to Grouneva et al. (2011). They are not strongly associated with supercomplexes according to our findings. The composition of the antenna pool of both supercomplexes is represented. The dotted line around Lhcx6_1 in PSI supercomplexes indicates that its location is still uncertain, since it might be present in PSI only under high-light conditions (Grouneva et al., 2011). The model represents the antenna composition of the complexes of the photosystems but contains no information about the oligomeric state, the stoichiometry, and the precise localizations of subunits within the complexes.

which showed a higher relative abundance in comparison with all the other Lhcr proteins in the smallest complex. This higher value can be interpreted as a change in ratio between Lhcr10 and all the other Lhcr(s), due to the detachment of the other Lhcr proteins. According to this hypothesis, Lhcr10 is the antenna protein most strongly bound to the PSI reaction center.

The lack of a PSI-FCPI structure limits our knowledge about the oligomeric state of Lhcr proteins. The structure of a PSI-LHCR supercomplex in red algae (Pi et al., 2018), however, gives some hints about the structural organization. The antenna proteins are arranged as monomers around the reaction center, as observed also in the structures of other photosynthetic eukaryotes (Mazor et al., 2017; Qin et al., 2019). So most probably Lhcrs in diatoms are also organized as monomers. Since red algae contain only Lhcr proteins, no Lhcf proteins could be found in the structure of PSI of red algae. According to Grouneva et al. (2011), Lhcf10 was also detected in PSI complexes, as demonstrated here. Considering the total amount of Lhc proteins found, diatom PSI complexes have an antenna size more comparable to that of green algae (Qin et al., 2019) than that of land plants (Mazor et al., 2017).

CONCLUSION

Analysis of the two photosystems of *T. pseudonana* revealed distinct and specific antenna pools. Lhcf proteins mostly serve PSII and Lhcr proteins mostly serve PSI, with one exception for each: Lhcr2 is an antenna protein of PSII and Lhcf10 is bound to PSI. These divergences could be a constitutive feature of the complexes or due to a movement of the antenna from PSII to PSI and vice versa. This last hypothesis appears improbable for Lhca2, because of its position close to the core in the supercomplex structure of *C. gracilis*, although we cannot rule out a different arrangement for *T. pseudonana*. However, the fact that no state transitions including antenna movements have been proven for diatoms so far also argues against this hypothesis. Concerning Lhcf8, the main constituent of the free pool of FCP high oligomers, our results argue strongly in favor of the protein identification by Nagao et al. (2019a) but against Pi et al. (2019), since this subunit and thus the oligomers are apparently not strongly bound to PSII, although transferring energy to PSII (Chukhutsina et al., 2013). An open question is the localization of the Lhcx6_1 protein. Our results indicate a higher probability to find Lhcx6_1 in PSII complexes, but other studies have reported it for PSI complexes too (Grouneva et al., 2011). The reason could be related to the amount of light experienced by the cell during growth that fine-tunes the expression of those proteins.

In conclusion, we present a picture of the supercomplex organization in cells grown under very low light intensities (Fig. 5). The photosynthetic apparatus flexibly acclimates to environmental changes, and thus the organization described here should be considered as one of the multiple scenarios occurring in the thylakoid membranes. How complexes rearrange their structure under biotic and abiotic stimuli is still unknown for diatoms, but this work can serve as a basis to better understand the dynamic of the photosynthetic complexes.

MATERIALS AND METHODS

Growth Conditions

Maintaining cultures of *Thalassiosira pseudonana* (strain CCMP1335, Hustedt) cells were grown at 15°C in f/2 medium (Guillard, 1975), with a light photoperiod of 16 h of light/8 h of dark at approximately 45 μmol photons s^{-1} m^{-2}, with constant shaking at 110 rpm. For experiments, a 2-d-old cell culture was used to inoculate a 4-L culture at a starting cell concentration of 100,000 cells mL^{-1}.

The 4-L culture was grown under a light intensity of ~50 μmol photons s^{-1} m^{-2} and bubbled with air. After 6 d, cells had reached a concentration of 5 to 6 × 10^6 cells mL^{-1}.

Plastid and Thylakoid Purification from *T. pseudonana*

For plastid isolation, cells were harvested by centrifugation (5,000*g*, 10 min, 4°C) 1 h after the onset of light. The isolation was performed according to the protocol published by Jäger and Büchel (2019), with minor modifications. During all the steps, the samples were kept on ice at 4°C, unless otherwise specified. After cell harvest, the pellet was resuspended in a final volume of 20 mL of isolation medium (0.5 M sorbitol, 50 mM HEPES-KOH, 6 mM Na-EDTA, 5 mM $MgCl_2$, 10 mM KCl, 1 mM $MnCl_2$, and 1% [w/v] polyvinylpyrrolidone 40 [K30], pH 7.4), and 0.5% (w/v) fatty acid-free BSA as well as 0.1% (w/v) Cys were added. The osmolality was adjusted to the optimal value of 750 mOsmol kg^{-1} using 2 M sorbitol or water, helping to maintain membrane stability after cell disruption. For cell disruption, a French press was used at a pressure of 14.5 MPa (2,100 p.s.i. with a 1-inch piston). After disruption, cell debris was removed by centrifugation (300*g*, 9 min) and crude plastids were pelleted at 6,000*g* (10 min). The pellet was gently resuspended, adding 2 to 3 mL of isolation medium, and layered on a Percoll step gradient (10%, 20%, and 30% [v/v] Percoll in isolation medium; osmolality of ~960 mOsmol kg^{-1}). The gradients were centrifuged at 14,400*g* (30 min) using a Sorvall Discovery 90SE ultracentrifuge. After the run, plastids were harvested from the interphase between the 20% and 30% layers, washed with isolation medium, and centrifuged at 4,000*g* for 10 min. After pellet resuspension, the chlorophyll *a* concentration was measured in 90% (v/v) acetone according to Jeffrey and Humphrey (1975) and adjusted to 1.5 μg μL^{-1} chlorophyll *a*, as optimal for the further experiments. Sample aliquots were flash frozen using liquid nitrogen and then stored at −80°C.

Thylakoid samples were prepared from isolated plastids by treating them with an additional French press cycle at 18 MPa (120 p.s.i. with a 3/8-inch piston). Samples were then centrifuged for 10 min at 1,000*g* (4°C) to remove unbroken plastids. The chlorophyll *a* concentration of supernatant was determined and used for complex solubilization.

Isolation Arabidopsis Thylakoid Protein Complexes

Isolation of thylakoid membranes and complexes from Arabidopsis (*Arabidopsis thaliana*) was performed according to Järvi et al. (2011), with minor modifications. All the steps were carried out at 4°C and under dim light. One to 2 g of fresh leaves was ground in precold grinding buffer (50 mM HEPES/KOH [pH 7.5], 330 mM sorbitol, 2 mM $EDTA·2H_2O$, 1 mM $MgCl_2$, 5 mM sodium ascorbate, and 0.05% [w/v] BSA). The mixture was filtered over two layers of Miracloth, centrifuged for 4 min at 5,000*g*, and the supernatant was discarded. The pellet was suspended in shock buffer (50 mM HEPES/KOH [pH 7.5], 5 mM sorbitol, and 5 mM $MgCl_2$) and incubated 5 min in total darkness. After incubation, the sample was centrifuged, and the remains of shock buffer were removed. The pellet was washed in storage buffer (50 mM HEPES/KOH [pH 7.5], 100 mM sorbitol, and 10 mM $MgCl_2$), centrifuged again, suspended in a minimal volume of storage buffer, and chlorophyll *a* concentration was estimated. The chlorophyll estimation was done according to Porra et al. (1989).

For solubilization of the complexes, 7.5 μg of chlorophyll *a* of isolated thylakoids was incubated 10 min in solubilization buffer (25 mM Bis-Tris [pH 7], 20% [v/v] glycerol, 1% [w/v] *n*-dodecyl-β-D-maltoside, and protease inhibitor cocktail). Identical conditions were applied when using α-DDM for solubilization. After incubation, insolubilized material was removed by centrifugation (17,383*g*, 1 min), and the supernatant was stored at −80°C or directly used for native PAGE.

lpBN-PAGE and 2D SDS-PAGE

For plastids and thylakoid solubilization, the mild detergent α-DDM was used at a final concentration of 15 mM (0.75%, w/v). The samples were incubated 30 min on ice and then centrifuged for 1 min at 17,383*g*. The supernatant was collected and applied to lpBN-PAGE. The lpBN-PAGE device was prepared according to Järvi et al. (2011), with minor modifications. Acrylamide gradients of 3.5% to 8% or 3.5% to 12.5% (v/v) were chosen. The samples were run 90 min at 6 mA (150 maximum voltage) with cathode buffer (15 mM Bis-Tris and 50 mM Tricine, pH 7) including 0.01% (w/v) Serva Blue G stain (Coomassie) and then overnight at 50 V in the same buffer without Coomassie. For the CN-PAGE, Coomassie was substituted by deoxycholic acid in the cathode buffer, at a concentration of 0.57 mg mL^{-1}, and α-DDM was also supplemented at a concentration of 0.258 mg mL^{-1}. The same anode buffer (50 mM Bis-Tris, pH 7) was used for all native PAGE.

After the run, BN gel strips were excised and incubated in equilibration buffer for the 2D gel (3% [v/v] β-mercaptoethanol, 2% [v/v] glycerol, 2.5% [w/v] SDS, and 0.45 M Tris, pH 8.4) for 4 min on ice. After a short incubation in Tricine cathode buffer (0.1 M Tris, 0.1 M Tricine, and 0.1% [w/v] SDS, pH 8.25), the gel strip was mounted on top of the 2D gel, a Tricine-SDS-PAGE (Schägger, 2006), and run 3 to 4 h at 100 V (15 mA maximum). The same procedure was followed for the horizontal configuration of the 2D gel, where gel bands from the BN-PAGE were first excised and then run separately on a single lane. The 2D gel was silver stained according to Blum et al. (1987). Gel bands and spots from the 1D and 2D gels were excised and used for the identification of proteins by MS analysis.

Western blotting was performed according to Beer et al. (2006) via 2D SDS-PAGE. Two antibodies were used for detection of PSII (α-PsbC [CP43]; Agrisera no. AS111787, diluted 1:5,000) and PSI (α-PsaA; Agrisera no. AS06172, diluted 1:2,000). The secondary antibody used was goat anti-rabbit specific peroxidase conjugate (Calbiochem catalog no. 401315, diluted 1:10,000) for both the primary antibodies. The development was performed with the enhanced chemiluminescence method (Alegria-Schaffer et al., 2009), exposing the blot to the x-ray film with variable exposure times.

Suc Density Centrifugation

Suc density gradient centrifugation was performed according to Pi et al. (2019), with minor modification. Isolated plastids were centrifuged for 10 min at 17,383*g*. The pellet was suspended in MMKB buffer containing 30 mM MES, pH 6.5, 5 mM $MgCl_2$, 10 mM KCl, and 1 M betaine. Solubilization was performed for 30 min at 4°C, with 0.75% (w/v) α-DDM, using 100 μg of chlorophyll *a* per sample at a concentration of 0.35 μg μL^{-1}. After solubilization, the sample was centrifuged for 1 min at 17,383*g*, and the supernatant was layered on top of the Suc gradient. The gradient was prepared using 0.55 M Suc in MMKB buffer by three freeze-thaw cycles at −80°C and 4°C. After centrifugation at 132,000*g* for 16 h at 4°C, fractions were collected with a syringe and concentrated using Amicon-ultra 100-kD cutoff devices (Merck-Millipore). The concentrated samples were applied to lpBN-PAGE for the analysis of the native complexes or spectroscopic analysis.

Spectroscopy

Absorption spectra were measured using a Jasco V-650 spectrophotometer. MMKB buffer was used for the dilution of samples, and spectra were recorded between 370 and 750 nm, with a 1-nm bandwidth.

The fluorescence spectra were recorded at room temperature or at 77K in a Jasco FP6500 fluorospectrometer, using MMKB buffer with 60% (v/v) glycerol for the latter. For the measurements, samples were diluted to an absorbance of about 0.03 in the Q_y maximum. Two excitation wavelengths were used to preferentially excite chlorophyll *a* (440 nm) or chlorophyll *c* (465 nm). Emission spectra were measured between 600 and 800 nm. Both emission and excitation bandwidths were set to 3 nm, and spectra were corrected using a calibrated lamp spectrum.

MS Analysis

Protein spots or bands excised from gels were subjected to in-gel digestion with trypsin (Promega) according to Shevchenko et al. (1996). Tryptic peptides were dried in a vacuum centrifuge and stored at −20°C. Directly prior to analyses, the peptides we dissolve in 0.1% (v/v) formic acid, and 5 μL was injected for LC-MS/MS analysis performed by a Q Exactive HF mass spectrometer (Thermo Fisher Scientific) connected to an Easy NanoLC 1200 system (Thermo Fisher Scientific). Peptides were first loaded on a trapping column and subsequently separated inline on a 15-cm C18 column (75 μm × 15 cm, ReproSil-Pur 5 μm 200 Å C18-AQ, Dr. Maisch HPLC). The mobile phase consisted of water with 0.1% (v/v) formic acid (solvent A) and acetonitrile:water (80:20, v/v) with 0.1% (v/v) formic acid (solvent B). A linear 30-min gradient from 6% to 35% B was used to elute peptides from samples cut from BN- or CN-PAGE gels, and a 15-min gradient was used for other peptide samples. MS data were acquired automatically by using Thermo Xcalibur 4.1 software (Thermo Fisher Scientific). An information-dependent acquisition method consisted of

an Orbitrap MS survey scan of mass range 300 to 2,000 mass-to-charge ratio followed by higher-energy collisional dissociation fragmentation for the 10 most intense peptide ions.

Bioinformatics Analysis

MS data were searched for protein identifications using the Proteome Discoverer 2.3 software (Thermo Fisher Scientific) connected to an in-house server running the Mascot 2.6.1 software (Matrix Science). The database consisted of *T. pseudonana* sequences (12,014 entries) downloaded from UniProt (https://www.uniprot.org/). Two missed cleavages were allowed. Peptide mass tolerance of ±10 ppm and fragment mass tolerance of ±0.02 D were used. Carbamidomethyl was set as a fixed modification. Met oxidation and acetylation of the proteins' N termini were included as variable modifications. For 2D gel spots or bands, fixed value PSM validator was used, and for BN-BAND samples, Percolator was used. For protein identification, a minimum of two peptides including at least one high-confidence peptide were required. Annotations of some proteins were cross-checked by using the Joint Genome Institute (https://mycocosm.jgi.doe.gov/Thaps3/Thaps3.home.html) database.

Relative abundances of proteins detected in 1D or 2D gel bands or spots were calculated in the following way. First, proteins were quantified by spectral counting. The number of spectra identified for a given protein in 1D or 2D gel bands/spots was then divided by the total number of peptide spectrum matches detected from the same band or spot. The variation of this value along the different bands/spots of the same gel was used to assign the protein occurrence in that band/spot. Only the unique peptides that matched with a single protein were taken into account. Since Lhc proteins have high sequence similarity, many peptides were present in more than one Lhc. Thus, proteins showing high similarity (greater than 80%) were analyzed as a group, and the sum of peptide matches was considered. However, peptides matching with more than two Lhc proteins were excluded from the analysis.

Accession Numbers

Sequence data from this article can be found in the UniProt (https://www.uniprot.org/) data libraries under the accession numbers reported in Supplemental Tables S1, S4, and S5.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Comparison of *T. pseudonana* and Arabidopsis complexes solubilized with α-DDM and MW range estimation.

Supplemental Figure S2. Effect of the detergent concentration on the thylakoid complex solubilization from isolated plastids of *T. pseudonana* (with Arabidopsis complexes as a reference).

Supplemental Figure S3. Effect of the solubilization time on the thylakoid complex solubilization of *T. pseudonana* (with Arabidopsis as a reference).

Supplemental Figure S4. Spectroscopic and biochemical characterization of complex bands isolated by Suc density gradient centrifugation.

Supplemental Figure S5. Horizontal 2D SDS-PAGE: protein identification and western-blot analysis of complex subunits.

Supplemental Table S1. BN- and CN-PAGE bands MS data.

Supplemental Table S2. OEC comparison of BN- and CN-PAGE MS data.

Supplemental Table S3. LHC comparison of BN- and CN-PAGE MS data.

Supplemental Table S4. Standard 2D PAGE spots MS data.

Supplemental Table S5. Horizontal 2D PAGE bands MS data.

ACKNOWLEDGMENTS

We thank Stefanie Jäger for the introduction to diatom plastid isolation. MS analysis was performed at the Turku Proteomics Facility, University of Turku, and at Åbo Akademi University. The facility is supported by Biocenter Finland.

Received January 17, 2020; accepted March 11, 2020; published March 20, 2020.

LITERATURE CITED

Albanese P, Melero R, Engel BD, Grinzato A, Berto P, Manfredi M, Chiodoni A, Vargas J, Sorzano C, Marengo E, et al (2017) Pea PSII-LHCII supercomplexes form pairs by making connections across the stromal gap. Sci Rep **7:** 10067

Alegria-Schaffer A, Lodge A, Vattem K (2009) Performing and optimizing Western blots with an emphasis on chemiluminescent detection. Methods Enzymol **463:** 573–599

Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, et al (2004) The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism. Science **306:** 79–86

Barera S, Pagliano C, Pape T, Saracco G, Barber J (2012) Characterization of PSII-LHCII supercomplexes isolated from pea thylakoid membrane by one-step treatment with α- and β-dodecyl-D-maltoside. Philos Trans R Soc Lond B Biol Sci **367:** 3389–3399

Bedoshvili YD, Popkova TP, Likhoshway YV (2009) Chloroplast structure of diatoms of different classes. Cell Tissue Biol **3:** 297–310

Beer A, Gundermann K, Beckmann J, Büchel C (2006) Subunit composition and pigmentation of fucoxanthin-chlorophyll proteins in diatoms: Evidence for a subunit involved in diadinoxanthin and diatoxanthin binding. Biochemistry **45:** 13046–13053

Bhattacharya D, Archibald JM, Weber APM, Reyes-Prieto A (2007) How do endosymbionts become organelles? Understanding early events in plastid evolution. BioEssays **29:** 1239–1246

Blankenship RE (2010) Early evolution of photosynthesis. Plant Physiol **154:** 434–438

Blum H, Beier H, Gross HJ (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis **8:** 93–99

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

Büchel C (2020) Light harvesting complexes in chlorophyll c-containing algae. Biochim Biophys Acta Bioenerg **1861:** 148027

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

Chukhutsina VU, Büchel C, van Amerongen H (2013) Variations in the first steps of photosynthesis for the diatom *Cyclotella meneghiniana* grown under different light conditions. Biochim Biophys Acta **1827:** 10–18

Flori S, Jouneau PH, Bailleul B, Gallet B, Estrozi LF, Moriscot C, Bastien O, Eicke S, Schober A, Bártulos CR, et al (2017) Plastid thylakoid architecture optimizes photosynthesis in diatoms. Nat Commun **8:** 15885

Grouneva I, Rokka A, Aro EM (2011) The thylakoid membrane proteome of two marine diatoms outlines both diatom-specific and species-specific features of the photosynthetic machinery. J Proteome Res **10:** 5338–5353

Guillard RRL (1975) Culture of phytoplankton for feeding marine invertebrates. In WL Smith, and MH Chanley, eds, Culture of Marine Invertebrate Animals: Proceedings—1st Conference on Culture of Marine Invertebrate Animals Greenport. Springer, Boston, MA, pp 29–60

Gundermann K, Wagner V, Mittag M, Büchel C (2019) Fucoxanthin-chlorophyll protein complexes of the centric diatom *Cyclotella meneghiniana* differ in Lhcx1 and Lhcx6_1 content. Plant Physiol **179:** 1779–1795

Haniewicz P, Floris D, Farci D, Kirkpatrick J, Loi MC, Büchel C, Bochtler M, Piano D (2015) Isolation of plant photosystem II complexes by fractional solubilization. Front Plant Sci **6:** 1100

Ikeda Y, Yamagishi A, Komura M, Suzuki T, Dohmae N, Shibata Y, Itoh S, Koike H, Satoh K (2013) Two types of fucoxanthin-chlorophyll-binding proteins I tightly bound to the photosystem I core complex in marine centric diatoms. Biochim Biophys Acta **1827:** 529–539

Jäger S, Büchel C (2019) Cation-dependent changes in the thylakoid membrane appression of the diatom Thalassiosira pseudonana. Biochim Biophys Acta Bioenerg **1860:** 41–51

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

Jeffrey SW, Humphrey GF (1975) New spectrophotometric equations for determining chlorophylls a, b, c1 and c2 in higher plants, algae and natural phytoplankton. Biochem Physiol Pflanz **167:** 191–194

Kuczynska P, Jemiola-Rzeminska M, Strzalka K (2015) Photosynthetic pigments in diatoms. Mar Drugs **13:** 5847–5881

Lepetit B, Volke D, Szabó M, Hoffmann R, Garab G, Wilhelm C, Goss R (2007) Spectroscopic and molecular characterization of the oligomeric antenna of the diatom *Phaeodactylum tricornutum*. Biochemistry **46:** 9813–9822

Levitan O, Chen M, Kuang X, Cheong KY, Jiang J, Banal M, Nambiar N, Gorbunov MY, Ludtke SJ, Falkowski PG, et al (2019) Structural and functional analyses of photosystem II in the marine diatom *Phaeodactylum tricornutum*. Proc Natl Acad Sci USA **116:** 17316–17322

Li G, Woroch AD, Donaher NA, Cockshutt AM, Campbell DA (2016) A hard day's night: Diatoms continue recycling photosystem II in the dark. Front Mar Sci **3:** 203

Mazor Y, Borovikova A, Caspy I, Nelson N (2017) Structure of the plant photosystem I supercomplex at 2.6 Å resolution. Nat Plants **3:** 17014

Nagao R, Kato K, Suzuki T, Ifuku K, Uchiyama I, Kashino Y, Dohmae N, Akimoto S, Shen JR, Miyazaki N, et al (2019a) Structural basis for energy harvesting and dissipation in a diatom PSII-FCPII supercomplex. Nat Plants **5:** 890–901

Nagao R, Tomo T, Noguchi E, Nakajima S, Suzuki T, Okumura A, Kashino Y, Mimuro M, Ikeuchi M, Enami I (2010) Purification and characterization of a stable oxygen-evolving photosystem II complex from a marine centric diatom, *Chaetoceros gracilis*. Biochim Biophys Acta **1797:** 160–166

Nagao R, Ueno Y, Akita F, Suzuki T, Dohmae N, Akimoto S, Shen JR (2019b) Biochemical characterization of photosystem I complexes having different subunit compositions of fucoxanthin chlorophyll a/c-binding proteins in the diatom *Chaetoceros gracilis*. Photosynth Res **140:** 141–149

Neilson JAD, Durnford DG (2010) Structural and functional diversification of the light-harvesting complexes in photosynthetic eukaryotes. Photosynth Res **106:** 57–71

Nelson N, Junge W (2015) Structure and energy transfer in photosystems of oxygenic photosynthesis. Annu Rev Biochem **84:** 659–683

Nowicka B, Kruk J (2016) Powered by light: Phototrophy and photosynthesis in prokaryotes and its evolution. Microbiol Res **186–187:** 99–118

Pi X, Tian L, Dai HE, Qin X, Cheng L, Kuang T, Sui SF, Shen JR (2018) Unique organization of photosystem I-light-harvesting supercomplex revealed by cryo-EM from a red alga. Proc Natl Acad Sci USA **115:** 4423–4428

Pi X, Zhao S, Wang W, Liu D, Xu C, Han G, Kuang T, Sui SF, Shen JR (2019) The pigment-protein network of a diatom photosystem II-light-harvesting antenna supercomplex. Science **365:** eaax4406

Porra RJ, Thompson WA, Kriedemann PE (1989) Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: Verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. Biochim Biophys Acta Bioenerg **975:** 384–394

Pyszniak AM, Gibbs SP (1992) Immunocytochemical localization of photosystem I and the fucoxanthin-chlorophyll a/c light-harvesting complex in the diatom *Phaeodactylum tricornutum*. Protoplasma **166:** 208–217

Qin X, Pi X, Wang W, Han G, Zhu L, Liu M, Cheng L, Shen JR, Kuang T, Sui SF (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

Rochaix JD (2014) Regulation and dynamics of the light-harvesting system. Annu Rev Plant Biol **65:** 287–309

Röding A, Boekema E, Büchel C (2018) The structure of FCPb, a light-harvesting complex in the diatom *Cyclotella meneghiniana*. Photosynth Res **135:** 203–211

Schägger H (2006) Tricine-SDS-PAGE. Nat Protoc **1:** 16–22

Schober AF, Flori S, Finazzi G, Kroth PG, Bártulos CR (2018) Isolation of plastid fractions from the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*. Methods Mol Biol **1829:** 189–203

Schober AF, Río Bártulos C, Bischoff A, Lepetit B, Gruber A, Kroth PG (2019) Organelle studies and proteome analyses of mitochondria and plastids fractions from the diatom *Thalassiosira pseudonana*. Plant Cell Physiol **60:** 1811–1828

Shen L, Huang Z, Chang S, Wang W, Wang J, Kuang T, Han G, Shen JR, Zhang X (2019) Structure of a $C_2S_2M_2N_2$-type PSII-LHCII supercomplex from the green alga *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA **116:** 21246–21255

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

Teramoto H, Ono T, Minagawa J (2001) Identification of Lhcb gene family encoding the light-harvesting chlorophyll-a/b proteins of photosystem II in *Chlamydomonas reinhardtii*. Plant Cell Physiol **42:** 849–856

Veith T, Büchel C (2007) The monomeric photosystem I-complex of the diatom *Phaeodactylum tricornutum* binds specific fucoxanthin chlorophyll proteins (FCPs) as light-harvesting complexes. Biochim Biophys Acta **1767:** 1428–1435

Washburn MP, Wolters D, Yates JR III (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol **19:** 242–247

Zhu SH, Green BR (2010) Photoprotection in the diatom *Thalassiosira pseudonana*: Role of LI818-like proteins in response to high light stress. Biochim Biophys Acta **1797:** 1449–1457