Structural and functional analyses of photosystem II in the marine diatom *Phaeodactylum tricornutum*

Orly Levitan^{a,b,1}, Muyuan Chen^{c,1}, Xuyuan Kuang^{d,e,f,1}, Kuan Yu Cheong^{a,g}, Jennifer Jiang^{d,e}, Melissa Banal^{d,e}, Nikhita Nambiar^{d,e}, Maxim Y. Gorbunov^a, Steven J. Ludtke^c, Paul G. Falkowski^{a,e,h,2}, and Wei Dai^{d,e,2}

^aEnvironmental Biophysics and Molecular Ecology Program, Department of Marine and Coastal Sciences, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901; ^bDepartment of Biochemistry and Microbiology, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901; ^cVerna and Marrs Mclean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030; ^dDepartment of Cell Biology and Neuroscience, Rutgers, The State University of New Jersey, Piscataway, NJ 08854; ^eInstitute for Quantitative Biomedicine, Rutgers, The State University of New Jersey, Piscataway, NJ 08854; ^fDepartment of Hyperbaric Oxygen, Xiangya Hospital, Central South University, Changsha, 410008 Hunan Province, China; ^gDepartment of Plant Biology, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901; and ^hDepartment of Earth and Planetary Sciences, Rutgers, The State University of New Jersey, Piscataway, NJ 08854

Contributed by Paul G. Falkowski, July 10, 2019 (sent for review April 19, 2019; reviewed by Thomas S. Bibby and Kevin E. Redding)

A descendant of the red algal lineage, diatoms are unicellular eukaryotic algae characterized by thylakoid membranes that lack the spatial differentiation of stroma and grana stacks found in green algae and higher plants. While the photophysiology of diatoms has been studied extensively, very little is known about the spatial organization of the multimeric photosynthetic protein complexes within their thylakoid membranes. Here, using cryo-electron tomography, proteomics, and biophysical analyses, we elucidate the macromolecular composition, architecture, and spatial distribution of photosystem II complexes in diatom thylakoid membranes. Structural analyses reveal 2 distinct photosystem II populations: loose clusters of complexes associated with antenna proteins and compact 2D crystalline arrays of dimeric cores. Biophysical measurements reveal only 1 photosystem II functional absorption cross section, suggesting that only the former population is photosynthetically active. The tomographic data indicate that the arrays of photosystem II cores are physically separated from those associated with antenna proteins. We hypothesize that the islands of photosystem cores are repair stations, where photodamaged proteins can be replaced. Our results strongly imply convergent evolution between the red and the green photosynthetic lineages toward spatial segregation of dynamic, functional microdomains of photosystem II supercomplexes.

diatom | photosystem II | cryo-electron tomography | thylakoid membranes | functional absorption analysis

Diatoms evolved ~200 Mya as secondary symbionts (1, 2), and their primary productivity in the contemporary oceans forms the base of most marine food webs, especially along continental margins, in shallow seas, and at high latitudes (3). Despite their ecological importance, however, very little is known about the spatial organization of the proteins responsible for their photosynthetic performance. To that end, we isolated thylakoid membranes from the marine diatom *Phaeodactylum tricornutum* (*SI Appendix, Fig. S1*) and subjected them to concurrent proteomic, bioinformatic, biophysical, and structural analyses. Using phase contrast cryo-electron tomography (cryoET) (4–6), we found dimeric photosystem II (PSII) reaction centers (RCs) are composed of 2 distinct subpopulations (Fig. 1 and *Movie S1*). One subpopulation is associated with light harvesting complexes, while the other is not. This is a direct observation of heterogeneity of PSII RCs within thylakoid membranes of a photoautotroph in the red photosynthetic lineage.

Results

Proteomic Analysis of PSII in *P. tricornutum*. The PSII holocomplex in *P. tricornutum* has more than 20 known subunits, most of which are encoded in the plastid (7, 8) (*SI Appendix, Table S1*). Phylogenetic analyses of the core and oxygen evolving complex

(OEC) subunits of PSII revealed that they are monophyletic, originating from a single ancestor of the red algal lineage (9) (*SI Appendix, Fig. S2*). Proteomic analysis identified 5 OEC subunits (*Table S1*), out of which 4 are phylogenetically closely related to red algae (PsbQ', PsbO, PsbV, and PsbU) (10), and the diatom-specific Psb31. In *P. tricornutum*, Psb31 is an extrinsic subunit that functions in concert with PsbO to enhance water splitting (11). The presence and integration of the Psb31 subunit in the complex further suggests that the structure of the OEC in PSII was modified following the secondary endosymbiotic event.

Structural Determination of PSII Supercomplexes in the Thylakoid Membrane. Patches of loose PSII holocomplex clusters were identified from tomograms of isolated thylakoids (Figs. 1*E* and 2*A* and *B*). The PSII complexes displayed dominant C2 symmetry and possessed the overall morphology, size, and shape of PSII from the red algae *Cyanidium caldarium*, which shares >83% sequence identity with *P. tricornutum*'s PSII (Protein Data Bank [PDB] ID 4YUU) (10) (*SI Appendix, Fig. S3*). Within

Significance

Despite distinctions in the architecture of thylakoid membranes, the fundamental machinery responsible for photosynthetic electron transfer is highly conserved in all oxygenic organisms. Using cryo-electron tomography in conjunction with proteomic and biophysical analyses, we show the distribution of photosystem II in thylakoid membranes of a diatom is heterogeneous. There are 2 subpopulations of the PSII supercomplexes; one contains antennae complexes, and the other does not. The former is functional, while the latter appears to be photochemically inactive. We suggest, based on tomography and biophysical analysis, that photochemically damaged reaction centers are physically isolated to repair stations where the inactive proteins can be removed and replaced.

Author contributions: O.L., M.C., P.G.F., and W.D. designed research; O.L., M.C., X.K., K.Y.C., J.J., M.B., N.N., M.Y.G., and W.D. performed research; M.C. and S.J.L. contributed new reagents/analytic tools; O.L., M.C., X.K., J.J., M.Y.G., S.J.L., P.G.F., and W.D. analyzed data; and O.L., M.C., M.Y.G., P.G.F., and W.D. wrote the paper.

Reviewers: T.S.B., University of Southampton; and K.E.R., Arizona State University.

The authors declare no conflict of interest.

Published under the PNAS license.

Data deposition: Electron density maps of photosystem II holocomplex and PSII cores in 2D arrays have been deposited in the EMDataResource under accession codes EMD-0539 and EMD-0540.

¹O.L., M.C., and X.K. contributed equally to this work.

²To whom correspondence may be addressed. Email: falko@marine.rutgers.edu or wei.dai@rutgers.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1906726116/-DCSupplemental.

Published online August 13, 2019.

Fig. 1. Phase contrast cryoET of *P. tricornutum* thylakoid membranes revealed spatial distribution of membrane-embedded PSII complexes. (A) Slice view of a representative tomogram. Arrows indicate features of interest on the membranes. Pink, PSII clusters; cyan, PSII 2D crystalline arrays; yellow, ATP synthases. (B) A 3D annotated view of the tomogram. Gray meshes represent the thylakoid membranes, and colored densities represent the 2 subpopulations of PSII and ATP synthase as shown in A. Cyan, 2D crystalline arrays of PSII cores; yellow, ATP synthases; pink, PSII clusters. Zoomed-in slice views of (C) 2D crystalline arrays, (D) ATP synthase, and (E) PSII clusters.

the clusters, PSII formed a semiorganized 2D distribution similar to PSII packing observed in higher plants (12–16). The nearest-neighbor distances between PSII dimers followed a narrow Gaussian distribution with a sharp peak at ~ 24 nm (Fig. 2*F, i*). Within the clusters, the orientation of PSII complexes is aligned with their neighbors; the distribution of angles between 2 neighboring complexes shows preference toward $\leq 45^\circ$ (Fig. 2*F, ii*). These clusters were primarily identified in regions where 2 or more thylakoid membranes are layered within a relatively constant distance of ~ 9 nm (Fig. 2*F, iii*, and [Movie S2](#)).

To further understand the organization of the PSII complexes, we performed subtomogram averaging. The PSII core complex, resolved at ~ 25 Å (Fig. 2*D* and [Movie S3](#)), showed 2 densities protruding toward the lumen. Lowering the isosurface threshold revealed 2 extra transmembrane densities on the sides of the PSII core (Fig. 2*D*). Based on comparison with the structure of PSII from *Arabidopsis thaliana* (PDB ID 5MDX), the overall structure of the extended densities is concordant with light harvesting complexes (LHC) ([SI Appendix, Fig. S3](#)) (17). Indeed, quantitative mass spectrometry of these thylakoid membrane preparations revealed that the most abundant proteins were 2 light-harvesting fucoxanthin chlorophyll *a/c* (FCP) proteins, Lhcf4 and Lhcf10, which were present at approximately 6-fold higher molar abundance than the PSII core ([SI Appendix, Table S1](#)). These results suggest a superstructure where 1 dimeric PSII core is associated with 12 FCP proteins, 6 per side. This stoichiometry is similar to that reported for a higher plant PSII–LHC structure (18). In our PSII subtomogram average, these FCP densities are tethered to the PSII core by interactions with the PsbC subunit. These linker densities have a less defined shape and are longer compared to their counterparts in higher plants. The lower occupancy and local resolution of these extended densities suggested that some, but not all, PSII complexes in the clusters are

associated with FCPs. Alternatively, it could also suggest that the binding between the FCPs and the PSII core is less rigid than that observed in higher plants.

Identification and Structural Determination of PSII 2D Arrays in the Thylakoid Membrane. In addition to PSII clusters, tightly packed 2D crystalline arrays of protein complexes were also observed in the tomograms (Fig. 1*C*). The complexes are arranged in a 2D lattice with constants of 11.4 and 17.9 nm at an angle of 102° between lattice vectors (Fig. 3*B*). The repeating units within the 2D array also possess C2 symmetry. Using 1,868 particles without an external initial model, we generated a 3D average map of the 2D array single unit at 10.4 Å resolution, determined by Fourier shell correlation (FSC) (19) (Fig. 3 and [SI Appendix, Fig. S4](#)). At this resolution, the overall domain organization of the averaged density map resembles the PSII core from *C. caldarium* (PDB ID 4YUU), with an overall map-model FSC of ~ 20 Å ([SI Appendix, Fig. S4](#)). Each of the 2 asymmetric units, composing the dimer, can be confidently segmented into 4 subunits, which correspond to the 4 core chains (PsbA, B, C, and D) of the PSII complex (Fig. 3*E* and [Movie S4](#)). The subtomogram average of the 2D array single unit shares morphology and domain organization with the core densities in diatom PSII holocomplexes in loose clusters ([SI Appendix, Fig. S5](#)). This observation further supports that the PSII complexes in the clusters and the single units in the 2D crystalline arrays represent 2 distinct, yet interchangeable, forms of PSII.

On average, the PSII particles associated with FCPs have a density of 111 ± 17.7 (1SD) dimers per square micrometer of thylakoid membrane, while the crystalline arrays have a density of 24.9 ± 5.7 (1SD). These densities are significantly different at a 90% confidence interval (Kruskal–Wallis, $P < 0.1$). Overall, the crystalline arrays account for $\sim 18\%$ of the total PSII superstructures. Moreover, the 2 types of PSII supercomplexes are physically isolated within the thylakoid membranes. The 2 subpopulations were not observed within the same thylakoid membrane in the tomograms used in the analysis.

Most parts of the averaged map of PSII cores were found to be in close agreement with the crystal structure of the red algae PSII cores. This reflects the common evolutionary origin and the high sequence identity between PSII of red algae and diatoms. However, the transmembrane density near the N terminus of PsbD was significantly shifted outward from the red algal model ([Movie S4](#)). At the current resolution, we were unable to confidently assign the mismatched density to be D2 (PsbD subunit) adopting a unique diatom conformation, or a cytochrome b559 subunit (PsbE), the fifth most abundant PSII subunit (after PsbA–D) by quantitative mass spectrometry ([SI Appendix, Table S1](#)). In the averaged map, neighboring repeating units in the arrays are connected by 2 densities extended from the D1 core subunit (PsbA) and the CP47 (PsbB) on the stromal side of the membrane (Fig. 3*C* and [SI Appendix, Fig. S4](#)). These densities were absent from PSII complexes in the loosely packed clusters. We suggest that these densities could be unique diatom polypeptides bound to the PSII core or domains of the PSII core subunits (PsbA or PsbD) that form unique conformations when disassociated from FCP proteins.

Biophysical Analysis of the PSII Complexes. Based on the kinetics of fluorescence rise under saturating, single turnover flashes, we observed a single functional absorption cross section for PSII in vivo ([Methods](#)). Depending on the light intensity under which the cells are grown, the cross section ranges from ca. 350 to 550 Å² at 455 nm, averaging ca. 450 Å². This relatively large functional cross section can only be supported by antennae pigment complexes that are energetically coupled to PSII reaction centers. Our observation implies that the 2D crystalline arrays are PSII lacking antennae and are photochemically inactive, i.e., photodamaged.

Fig. 2. Structural analysis of PSII complexes in clusters. (A) Slice view of a large PSII cluster. (B) Top view of volume rendering of PSII particles (pink) mapped back to a thylakoid membrane tomogram at their original orientation. Blue dashed line indicates the orientation of section view in C. (C) Side view of a section of the tomogram with annotated PSII particles (pink) and subtomogram (blue dashed line). Light blue regions highlight the lumen of the thylakoid membranes. (D) Bottom (lumen) and side views of the PSII complex subtomogram average at high (pink) and low (light gray) isosurface thresholds. (E) Side view of the subtomogram average filtered to 30 Å and displayed at a lower threshold (lumen side facing down), showing the additional layer of membrane and its spatial relationship with PSII. (F) (Left) diagram of PSII distribution on thylakoid membranes. (Right) Histograms of PSII packing statistics. Histogram *i* shows nearest neighbor distance of PSII particles. Histogram *ii* shows 2D angle between neighboring PSII particles (Kolmogorov–Smirnov test, $n = 1,272$, $P < 1e-6$). Histogram *iii* shows distance between center of thylakoid membranes where PSII clusters are present.

We applied a simple kinetic model to estimate the fraction of damaged reaction centers in the steady state. Assuming that the quantum yield of photodamage is $\Phi_{PI} = 10^{-7}$ and the recovery rate $k_{rec} = 10^{-4} \text{ s}^{-1}$ (20, 21), we estimate that the fraction of photo-inactivated PSII units under our experimental conditions was ca. 10% (Methods). This analysis predicts that a comparable fraction of 2D arrays lacking antenna are present in the steady state in diatoms with photochemical energy conversion efficiency of 0.55.

Discussion

Our results suggest that despite the absence of grana stacking in diatom thylakoids, there is spatial segregation of 2 subpopulations

of PSII complexes, one that is associated with FCPs (noncrystalline) and the other lacking antenna complexes (crystalline). The distribution of the 2 subpopulations is asymmetric. Our biophysical analysis suggests that these arrays represent reservoirs of photo-damaged PSII RCs. Based on the segregation of the 2 PSII subpopulations, we propose that the crystalline arrays are physically isolated repair stations where photochemically damaged proteins (e.g., D1) can be removed and replaced with functional structures. In the green photosynthetic lineage, spatial segregation of proteins on the thylakoid membrane appears to be essential to optimizing photophysiological function (12). The migration of PSII complexes to nonappressed, stromal thylakoids in higher plants seems to be

Fig. 3. Structural characterization of PSII 2D crystalline arrays. (A) Slice view of a large area of 2D array. (B) The 2D class average of the crystalline array with annotation of a unit cell. (C) Bottom (lumen side) and side view of the PSII core subtomogram average at high (cyan) and low (light gray) isosurface thresholds. (D) Pie chart showing the relative abundance of PSII subunits as detected by quantitative mass spectrum analysis of thylakoid membranes that were used for structural studies. (E) Bottom and side views of the segmented density map with corresponding PDB structures of PsbA, B, C, and D subunits (PDB ID 4YUU) shown as ribbon diagrams. The segmented map and PDB chains were color coded according to the colors used in the pie chart in D.

prerequisite for the repair of the photodamaged D1, which was inherited from anoxygenic purple bacteria and was known to undergo photodamage in all oxygenic photoautotrophs (22, 23).

Additionally, the spatial clustering of these subpopulations may function in stabilizing the local architecture of the membrane microdomains, or act to facilitate diffusion-dependent electron transport via small lipophilic molecules such as plastoquinone (24, 25). Therefore, the transition between the 2 PSII subpopulations may fine-tune the energy flow within the thylakoid membrane network of *P. tricornutum* (25) and thus represent a photophysiological strategy to sense and respond to environmental variations in spectral irradiance (12). In higher plants, it had been reported that the PsbS protein controls the microorganization of PSII complexes in grana stacks (25, 26). Since PsbS is absent in diatom genomes, it is unclear what mechanisms control the formation of those microdomains in these organisms.

The cryoET data further show that in *P. tricornutum*, there is segregation of components of the photosynthetic electron transport chain. In addition to PSII complex clusters and 2D crystalline arrays, ATP synthases were found in clusters, often distributed on membrane regions characterized by high curvature (Fig. 4). Disk-shaped densities of 18-nm diameter that may represent photosystem I (PSI) complexes were also detected in clusters on the membrane (27, 28). Yet, more particles are required to generate a reliable subtomogram average to confirm the identity of these densities.

Overall, our results reveal distinctly different structures of 2 subpopulations of PSII supercomplexes within thylakoid membranes. The observed segregation suggests that despite the evolutionary differences between the red and the green eukaryotic algal lineages, there is convergence in the spatial segregation in distribution of the PSII supercomplexes. The spatial segregation implies a common functional process that was inherited before the evolutionary split of eukaryotic algae (1).

Methods

Cell Culture and Isolation of Intact Plastids. Cultures of *P. tricornutum* Bohlin (CCAP 1055/1), strain Pt1 8.6, were grown in artificial seawater enriched with nutrients according to F/2 medium preparation protocol (29, 30). The cells

Fig. 4. Spatial distributions of ATP synthase and putative photosystem I complexes in thylakoid membrane tomograms supported segregation of photosynthetic machinery in diatoms. (A) Slice view of thylakoid membrane tomogram showing clusters of PSII holocomplexes (pink arrows) and ATP synthase (yellow arrows). (B) Subtomogram average of ATP synthase at ~25 Å resolution, fitted with PDB model 6FKF. (C) A cluster of putative photosystem I density (green arrows) on thylakoid membrane.

were grown in continuous white LED light at 100–120 $\mu\text{mol photon m}^{-2} \text{s}^{-1}$. Optically thin cultures were harvested at a density of $\sim 3\text{--}5 \times 10^5$ cells per mL by continuous centrifugation at 10,000 g and at 4 °C. The cell pellet was kept and resuspended in 10 mL of Breaking Buffer A (2 mM Na_2EDTA , 1 mM $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$, 1 mM $\text{MnCl}_3 \cdot 4\text{H}_2\text{O}$, 50 mM Hepes [adjusted to pH 7.5], and 0.66 M Sorbitol) in a 50-mL Falcon tube. A variant of the breaking buffer with higher magnesium concentration, Breaking Buffer B (2 mM Na_2EDTA , 40 mM $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$, 1 mM $\text{MnCl}_3 \cdot 4\text{H}_2\text{O}$, 50 mM Hepes [adjusted to pH 7.5], and 0.66 M Sorbitol), was used in some sample preparations to enrich the 2D crystalline arrays in sample. Intact plastids were isolated using a French Press set to 6,000 psi. Large cell debris and unbroken cells were removed by discarding the cell pellet formed after centrifugation for 5 min at 5,000 g and at 4 °C. The supernatant was gently layered onto a 2 mL 2 M sucrose cushion, followed by centrifugation in a swinging bucket rotor (HB-6; DuPont) for 15 min at 16,000 g , 4 °C. At the end of the centrifugation, a brown band consisting of intact plastids, formed at the interface between the sucrose cushion and breaking buffer, was transferred to a 1.7-mL Eppendorf tube on ice. The centrifugation step at 16,000 g was repeated for a total of 3 times, every time with a fresh supernatant after the previous was collected into a separate fresh tube, to collect the maximal number of intact plastids in the brown band. Next, the isolated brown band was centrifuged for 10 min at 10,000 g and 4 °C. The supernatant was removed and 1 mL of the Washing Buffer (6 mM Na_2EDTA , 5 mM $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$, 1 mM $\text{MnCl}_3 \cdot 4\text{H}_2\text{O}$, 50 mM Hepes [adjusted to pH 7.5], 10 mM KCl, 0.66 M Sorbitol) was added to the plastid pellet. The washing step was repeated for a total of 3 times.

Osmotic Shock to Isolate Thylakoid Membranes. Isolated plastids, resuspended in the Washing Buffer, were centrifuged for 10 min at 10,000 g and 4 °C. Next, 1.5 mL of High Salt Buffer (5 mM Na_2EDTA , 0.2 M KCl, 10 mM Hepes [adjusted to pH 7.5]) was added to the plastid pellet, followed by centrifugation for 15 min at 48,000 g in 4 °C using Thermo Fisher SS-34 Fixed Angle Rotor. Then, 1.5 mL of Low Salt Buffer (5 mM $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$, 50 mM Hepes [adjusted to pH 7.5]) was added to the pellet, followed by centrifugation for 15 min, at 48,000 g and 4 °C. The pelleted thylakoid membranes were resuspended in Low Salt Buffer for subsequent image analysis. Chlorophyll a per μL was determined spectrophotometrically by diluting a known volume of sample in 90% acetone (31). The concentration of thylakoid membranes was adjusted to chlorophyll a concentration of $\sim 0.05 \mu\text{g/mL}$ and used for image processing.

Tomography Data Collection. EM grids for structural studies were prepared from 7 independent batches of extracted thylakoid membrane preparations. An aliquot of 3.5 μL extracted thylakoid membrane sample was mixed with 10 nm gold fiducial markers (EMS) and applied to Quantifoil holey grids (R 2.0/1.0, Cu, 200 mesh; Quantifoil) before they were plunge frozen using a Leica EM GP plunger. Images and tilt series of the samples were imaged in a Talos Arctica cryo-electron microscope (Thermo Fisher Scientific). This electron microscope has a post-column BioQuantum energy filter (the slit was set to 20 eV), K2 direct electron detector, and Volta phase plates at the back focal plane. Tilt series of extracted thylakoid membrane were collected under the following conditions: 39,000 $\times$ microscope magnification, spot size 8, 100- μm condenser aperture, and defocus close to $-0.5 \mu\text{m}$ with phase plate or $-5 \mu\text{m}$ without phase plates. The image pixel size was 3.49 $\text{\AA}/\text{pixel}$. Typically, a tilt series ranges from -69° to 69° at 3° step increments. Tilt series used for high-resolution subtomogram averaging of 2D arrays were collected on a Titan Krios at Purdue University on a post-GIF K2 Summit camera at 51,000 $\times$ microscope magnification and defocus close to $-0.5 \mu\text{m}$ with Volta phase plates under counting mode. The illumination settings used were spot size 10 and condenser aperture 100 μm . The image pixel size was 2.65 $\text{\AA}/\text{pixel}$. The accumulated dose for each tilt series was 50–60 $\text{e}/\text{\AA}^2$. Automated data collection was carried out using SerialEM (32) software.

Tomography Data Processing. All tomographic data processing was done using the latest EMAN2 (Electron Micrograph Analysis 2) tomography workflow unless stated otherwise. A detailed step-by-step tutorial can be found at <http://eman2.org/e2tomo>.

For subtomogram averaging of PSII holocomplexes and PSII RCs, only data collected with Volta phase plates were used. A total of 315 tomograms of thylakoid membrane were successfully reconstructed, 173 from the Talos Arctica microscope and 142 from the Krios microscope at Purdue University. If we assume that each tomogram contains a single layer of flat thylakoid membrane, the total area of thylakoid reconstructed from our tomograms is ~ 500 square μm .

PSII holocomplex subpopulation tends to appear in thick stacks of membranes with lower contrast. From the dataset collected from the Talos Arctica

microscope, we selected 11 tomograms with the best image contrast, and extracted 1152 subtomograms using a box size of 128 pixels. A reference free initial model was generated from the particles with C1 symmetry using a stochastic gradient descent-based algorithm. Subtomogram averaging was performed on volumes of PSII holocomplexes with a soft spherical mask and C2 symmetry. Even and odd half-sets of particles were separated during the refinement process. For the final subtomogram average of PSII holocomplexes, 382 particles were included in the average based on tomogram quality and confidence in orientation determination. The final map was filtered by the local “gold-standard” FSC between the maps from the 2 subsets. Resolution of the PSII holocomplex average was estimated to be 26 $\text{\AA}$ (cutoff 0.143) under a soft mask that excludes LHC densities. The map agreed with the homologous crystal structure (4YUU) at 49 $\text{\AA}$ (at cutoff 0.5).

A subtomogram average of PSII cores in 2D crystalline array was carried out on tomograms reconstructed from tilt series collected from a Titan Krios microscope with phase plate to achieve high-resolution subtomogram average. Out of a total of 142 tomograms reconstructed from tilt series in this dataset, we extracted 1,868 particles from 28 tomograms from which we can confidently identify PSII arrays. The 2D class average of the crystalline array was generated using bispectrum-based reference free 2D refinement, using projections of 205 3D subtomograms of $256 \times 256 \times 64$ pixels. A 3D subtomogram average was generated from the 1,868 particles. To improve subtomogram average resolution, defocus and phase shift were estimated for each tilt image. Per-particle-per-tilt CTF correction was performed before reconstructing 2D subtilt images into 3D particles. Subtomogram alignment was performed with C2 symmetry with a soft spherical mask starting from a reference free initial model that was directly generated from the particles. The subtomogram alignment was followed by 1 round of subtilt refinement to compensate for local misalignment of the tilt series. Even and odd subsets of particles were aligned separately, and the map was filtered by local gold-standard FSC. Resolution of the averaged map was 10.4 $\text{\AA}$ at cutoff value of 0.143, under a soft mask that only covered 1 unit of the array. Simulated maps were generated from PDB structures as models to compare with averaged density map and to be used to calculate map-model FSC.

Tomogram annotation was performed by mapping subtomogram average maps back to the original tomograms based on original particle coordinates. Corresponding densities were masked out after the subtomogram average was fitted back. Contour of membranes was annotated semiautomatically using the “draw contour” tool in EMAN2 (33). Membrane annotation in Fig. 2C was drawn manually in 2D.

Chimera (University of California, San Francisco) (34) was used for data visualization. PDB structures were fitted into the density maps by the “Fit in Map” tool in Chimera. Subunit segmentation of the 2D array averaged map was performed with Segger in Chimera.

Analysis of PSII Spatial Distribution. Aligned particles were mapped back to the original coordinates in tomograms to compute the PSII distribution and packing in the membranes. The statistics were calculated using a customized Python script. For nearest neighbor distance, only 2D distances of particles that are approximately at the same plane were used in the calculation. Particle pairs with more than 10 nm distance along the z axis were excluded. For the neighbor angle distribution, 2D angles between particle pairs within 35 nm (mean + SD of nearest neighbor distance) distance were considered. The angles were constrained to 0° to 90° given the C2 symmetry of PSII. Single-sided Kolmogorov–Smirnov test was performed with the statistical module against a uniform distribution to show the preferred orientation of PSII packing. For distance between membranes, 128 tomographic slices of 72×72 pixels containing 2 layers of vertical thylakoid membranes where PSII clusters were identified were manually selected from tomograms. These slices were binned by 2, contrast-inverted, and then aligned to a common coordinate (a vertical line) and projected to the x axis. Wavelet-based peak search (find_peaks_cwt) from Scipy was used to determine peaks in the 1D projections. Distances between membranes were calculated from the peak locations.

Protein Analyses: Sample Collection, Total Protein Quantification, and Western Blots. For total protein from isolated thylakoid membranes, thylakoid membranes were resuspended in 1 $\times$ denaturing lithium dodecyl sulfate (LDS) extraction buffer containing a 1/200 protease inhibitor mixture (Sigma–Aldrich). The resuspended thylakoid membranes were lysed using FastPrep-24 5G Homogenizer (MP Biomedicals) with lysing matrix Y. The lysed solution was spun down at 4 °C, 14,000 g for 10 min. The supernatant was collected. Total protein concentration in the supernatant was quantified using Lowry assay (DC 500-0111; Bio-Rad) and a SpectraMax M3

microplate reader (Molecular Devices) at 750 nm. BSA was used as the reference protein standard.

One to 7 μg of total protein extracted from thylakoid membranes was separated by gel electrophoresis on precast 4 to 20% Tris-glycine extended (TGX) gels (Bio-Rad). Separated proteins were transferred to PVDF membranes using a Trans-Blot Turbo Dry Transfer System and Transblot Turbo PVDF Transfer Packs (Bio-Rad). The presence of the PSII core subunit D1 (PsbA), PSI core subunit (PsaC), Cyt b6 (PetB), and ATP synthetase (AtpB) were detected using Western blot (*SI Appendix, Fig. S1*). The rabbit polyclonal antibodies (Agrisera) used were anti-PsbA (AS05 084, 1:10,000 dilution), anti-PsaC (AS10 939, 1:3,000 dilution), and anti-Cyt b6 (AS03 034, 1:10,000 dilution). An HRP-conjugated chicken anti-rabbit IgG (AS10 833, 1:20,000 dilution) was used as the secondary antibody for these rabbit polyclonal antibodies. For anti-AtpB, a chicken polyclonal from Agrisera (AS03 030, 1:6,500 dilution) was used, and HRP-conjugated goat anti-chicken IgY (AS09 603, 1:15,000 dilution) was used as the secondary antibody. Positive control protein standards (Agrisera) used include PsbA (AS01 016S), PsaC (AS04 042S), and AtpB (AS03 030S). The blots were developed with Amersham ECL Select Detection Reagent (RPN2235; GE Healthcare) and quantified using ChemiDoc MP Imaging System (Bio-Rad) with Image Lab software (Bio-Rad).

Quantitative Proteomic Analysis. Quantitative proteomic analysis was done using a label-free quantification method based on a peak intensity-based absolute quantification method called iBAQ. For calculations, we have used a Universal Proteomics Standard (UPS2; Sigma-Aldrich), which is a mixture of 48 precisely quantified human proteins with dynamic concentrations range spanning 5 orders of magnitude. UPS2 were quantitatively mixed with samples based on manufacturer's instructions. The mixed samples were loaded into SDS/PAGE as gel plugs and digested using a standard procedure for Liquid chromatography-tandem mass spectrometry (LC-MS/MS). Samples were analyzed using a Q Exactive HF tandem mass spectrometer coupled to a Dionex Ultimate 3000 RSLCnano System (Thermo Scientific). Samples were loaded onto a fused silica trap column Acclaim PepMap 100, 75 $\mu\text{m} \times 2 \text{ cm}$ (Thermo Fisher Scientific). After washing for 5 min at 5 $\mu\text{L}/\text{min}$ with 0.1% TFA, the trap column was brought in line with an analytical column (Nanoease MZ peptide BEH C18, 130A, 1.7 μm , 75 $\mu\text{m} \times 250 \text{ mm}$; Waters) for LC-MS/MS. Peptides were eluted using a segmented linear gradient from 4 to 90% B (A, 0.2% formic acid; B, 0.08% formic acid; 80% ACN): 4 to 15% B in 5 min, 15 to 50% B in 50 min, and 50 to 90% B in 15 min. Mass spectrometry data were acquired using a data-dependent acquisition procedure with a cyclic series of a full scan with resolution of 120,000 followed by MS/MS (HCD, relative collision energy 27%) of the 20 most intense ions and a dynamic exclusion duration of 20 s. The data were searched against both National Center for Biotechnology Information and Ensembl databases. The addition and quantification of UPS2 protein sequences was done using Maxquant 1.16.1.1. An iBAQ value was used to quantify protein components based on the calibration curve of the UPS2 standard (35).

Estimation of the Amount of Dynamically Photoinactivated PSII Centers. We estimated the amount of dynamically photoinactivated PSII centers using a first-order kinetic model for photoinactivation and recovery processes (20):

where S_0 and S_1 represent fractions of active and photodamaged PSII units, respectively; k_{PI} is the rate of photoinactivation; and k_{rec} is the rate of recovery. k_{PI} is proportional to the photon flux density of growth light (E), the functional absorption cross section of PSII (σ_{PSII}), and the quantum yield of photoinactivation of PSII (Φ_{PI}):

$$k_{\text{PI}} = \Phi_{\text{PI}} E \sigma_{\text{PSII}}. \quad [2]$$

This kinetic model is described by following equations:

$$d[S_0]/dt = -k_{\text{PI}}[S_0] + k_{\text{rec}}[S_1], \quad [3]$$

$$[S_0] + [S_1] = 1, \quad [4]$$

where $[S_0]$ and $[S_1]$ are the relative concentrations of active and photodamaged PSII units, respectively.

This model predicts that under steady-state conditions the proportion of photodamaged PSII units is proportional to the rate constant for photoinactivation:

$$S_1 = k_{\text{PI}} / (k_{\text{PI}} + k_{\text{rec}}). \quad [5]$$

Assuming that $\Phi_{\text{PI}} = 10^{-7}$, $k_{\text{rec}} = 10^{-4} \text{ s}^{-1}$ (20, 21), and the measured $\sigma_{\text{PSII}} = 200 \text{ \AA}^2$ for white light with $E = 120 \text{ } \mu\text{mole quanta m}^{-2} \text{ s}^{-1}$ used in our experiments, we estimate that the fraction of photoinactivated PSII units under our experimental conditions was ca. 10%.

Estimation of the Contributions of 2 Populations of PSII Complexes to Variable Fluorescence Kinetics. Structural analysis revealed 2 populations of PSII supercomplexes with strikingly different sizes. We therefore hypothesized that if PSII centers in both populations are active, this should manifest in 2 components with distinct cross sections in the shape of fluorescence induction kinetics recorded in response to a saturating single-turnover flash.

To test this hypothesis, we analyzed fluorescence induction kinetics in intact cells in vivo using a 2-component model. The induction of variable fluorescence was recorded in response to a saturating flash of 300 μs duration and the peak optical power density of 0.7 W/cm^2 , using a highly sensitive Fluorescence Induction and Relaxation instrument (a mini-FIRE) as described (36).

During the single turnover saturating flash of light, fluorescence of the yield rises from its minimum level (F_0) to maximum ($F_0 + F_v$) as follows:

$$F(t) = F_0 + a_1 F_v (C_1(t)(1-p)/(1-C_1(t)p)) + a_2 F_v (C_2(t)(1-p)/(1-C_2(t)p)),$$

where a_1 and a_2 are relative magnitudes of the 2 components, $C_1(t)$ and $C_2(t)$ are fractions of dynamically closed reaction centers in each cluster of PSII supercomplexes, and p is the probability of energy transfer between PSII units.

The temporal evolution of each $C_i(t)$ is described by an exponential rise function with distinct optical cross sections (σ_i) of respective PSII supercomplexes:

$$C_i(t) = 1 - \exp(-\sigma_i I t), i=1,2,$$

where I is the peak photon flux density of the excitation light.

Our analysis has revealed that the measured fluorescence kinetics dominated with a single component ($a_1 = 0.99 \pm 0.01$) with a cross section of ca. $450 \text{ \AA}^2$ typical for PSII complexes with large antennae. Thereby, the second component with a small cross section (60 to $70 \text{ \AA}^2$) was either negligible ($a_2 = 0.01 \pm 0.008$) or undetectable ($a_2 = 0$) at all.

Because the population of photoinhibitory-damaged PSII centers under our growth conditions was small (ca. 10%) and the second kinetic component could be difficult to resolve, we further examined the shape of fluorescence induction in the cells with a large population of such centers. For this purpose, we exposed cells to high light for prolonged period (20 to 30 min) to cause more severe photoinhibition and to increase the population of damaged PSII centers (up to 50 to 60%). These experiments also revealed only 1 component with a large cross section, clearly suggesting that PSII complexes in crystalline arrays are inactive, i.e., do not exhibit a variable fluorescence component.

Statistical Analysis of PSII Holocomplex and PSII RC Subpopulation Density.

From a subset of high-contrast tomograms collected on thylakoid membranes prepared under physiological conditions, we identified 924 PSII particles, 755 associated with LHC and 169 not, yielding an approximate census of 82% that appear to contain both the PSII core dimer and antenna. We applied Kruskal-Wallis ANOVA analysis on particle densities within the membrane (particle number per unit membrane area) which showed that the density of PSII associated with LHC is significantly higher than the density of PSII RCs in crystalline array at $P < 0.1$ level.

Data Availability. Programs used for tomographic data analysis are available from EMAN2.org. Electron density maps of photosystem II holocomplex and PSII cores in 2D arrays have been deposited in the EMDDataBank under accession codes EMD-0539 (37) and EMD-0540 (38).

ACKNOWLEDGMENTS. Tilt series of thylakoid membrane were collected at Rutgers Cryo-Electron Microscopy and Cryo-Electron Tomography Core Facility and the Cryo-Electron Microscopy Facility at Purdue University. We thank Jason Kaelber at Rutgers Cryo-Electron Microscopy and Cryo-Electron Tomography Core Facility for technical support in data collection. Focused ion beam milling of plunge frozen diatom cells was carried out at Thermo Fisher Scientific Factory at Brno, Czechia. We thank Jakub Kuba and Radovan Spurný for their support in milling and imaging of frozen lamellae. We would like to acknowledge Haiyan Zheng and Caifeng Zhao at the Biological Mass Spectrometry Facility at Robert Wood Johnson Medical School for conducting the quantitative mass spectrometry analysis. We also thank

Ehud Zelzion and Dana Price of the Rutgers Genome Cooperative for support with bioinformatic and phylogenetic analyses. Tomographic data acquisition at Purdue University was supported by the National Institute of Health Midwest Consortium for High Resolution Cryo-electron Microscopy (U24 GM116789-01A1). This research was supported by grants from National

Science Foundation EAGER award 1558128 to O.L. and P.G.F. and from National Institute of Health (R01GM080139 and P01GM121203 to S.J.L.). Partial funding for this study came from the Rutgers Busch Biomedical Research Grant to W.D. and Bennett L. Smith Endowment to P.G.F. X.K. was partially supported by the China Scholarship Council (File 201606375128).

1. P. G. Falkowski *et al.*, The evolution of modern eukaryotic phytoplankton. *Science* **305**, 354–360 (2004).
2. W. H. C. F. Kooistra, R. Gersonde, L. K. Medlin, D. G. Mann, “The origin and evolution of the diatoms: Their adaptation to a planktonic existence” in *Evolution of Primary Producers in the Sea*, P. G. Falkowski, A. H. Knoll, Eds. (Academic Press, 2007), pp. 207–249.
3. V. Smetacek, Diatoms and the ocean carbon cycle. *Protist* **150**, 25–32 (1999).
4. W. Baumeister, Electron tomography: Towards visualizing the molecular organization of the cytoplasm. *Curr. Opin. Struct. Biol.* **12**, 679–684 (2002).
5. M. Marko, C. E. Hsieh, Three-dimensional cryotransmission electron microscopy of cells and organelles. *Methods Mol. Biol.* **369**, 407–429 (2007).
6. W. Dai *et al.*, Visualizing virus assembly intermediates inside marine cyanobacteria. *Nature* **502**, 707–710 (2013).
7. C. Bowler *et al.*, The Phaeodactylum genome reveals the evolutionary history of diatom genomes. *Nature* **456**, 239–244 (2008).
8. M. P. Oudot-Le Secq *et al.*, Chloroplast genomes of the diatoms Phaeodactylum tricornutum and Thalassiosira pseudonana: Comparison with other plastid genomes of the red lineage. *Mol. Genet. Genomics* **277**, 427–439 (2007).
9. H. S. Yoon, J. D. Hackett, G. Pinto, D. Bhattacharya, The single, ancient origin of chromist plastids. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 15507–15512 (2002).
10. H. Ago *et al.*, Novel features of eukaryotic photosystem II revealed by its crystal structure analysis from a red alga. *J. Biol. Chem.* **291**, 5676–5687 (2016).
11. S. Okumura *et al.*, Transformation of poplar (*Populus alba*) plastids and expression of foreign proteins in tree chloroplasts. *Transgenic Res.* **15**, 637–646 (2006).
12. R. Kouril, E. Wientjes, J. B. Bultema, R. Croce, E. J. Boekema, High-light vs. low-light: Effect of light acclimation on photosystem II composition and organization in *Arabidopsis thaliana*. *Biochim. Biophys. Acta* **1827**, 411–419 (2013).
13. W. Kuhlbrandt, *Photosynthetic Membranes and Membrane Proteins* (Academic Press, London, 1987).
14. P. Albanese *et al.*, Pea PSII-LHCII supercomplexes form pairs by making connections across the stromal gap. *Sci. Rep.* **7**, 10067 (2017).
15. D. J. Simpson, Freeze-fracture studies on barley plastid membranes II. Wild-type chloroplast. *Carlsberg Res. Commun.* **43**, 365–389 (1978).
16. L. A. Staehelin, Chloroplast membrane structure. Intramembranous particles of different sizes make contact in stacked membrane regions. *Biochim. Biophys. Acta* **408**, 1–11 (1975).
17. L. S. van Bezouwen *et al.*, Subunit and chlorophyll organization of the plant photosystem II supercomplex. *Nat. Plants* **3**, 17080 (2017).
18. S. Caffarri, T. Tibiletti, R. C. Jennings, S. Santabarbara, A comparison between plant photosystem I and photosystem II architecture and functioning. *Curr. Protein Pept. Sci.* **15**, 296–331 (2014).
19. P. B. Rosenthal, R. Henderson, Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
20. D. A. Campbell, E. Tyystjärvi, Parameterization of photosystem II photoinactivation and repair. *Biochim. Biophys. Acta* **1817**, 258–265 (2012).
21. T. Key *et al.*, Cell size trade-offs govern light exploitation strategies in marine phytoplankton. *Environ. Microbiol.* **12**, 95–104 (2010).
22. W. S. Chow, E. M. Aro, “Photoinactivation and mechanisms of recovery” in *Photosystem II. Advances in Photosynthesis and Respiration*, T. J. Wydrzynski, K. Satoh, J. A. Freeman, Eds. (Springer, Boston, MA, 2005), vol. **22**, pp. 627–648.
23. K. Yokthongwattana, A. Melis, “Photoinhibition and recovery in oxygenic photosynthesis: Mechanism of a photosystem II damage and repair cycle” in *Photoprotection, Photoinhibition, Gene Regulation and Environment*, B. Demmig-Adams, Ed. (Springer, The Netherlands, 2005), pp. 175–191.
24. J. P. Dekker, E. J. Boekema, Supramolecular organization of thylakoid membrane proteins in green plants. *Biochim. Biophys. Acta* **1706**, 12–39 (2005).
25. S. Tietz *et al.*, Functional implications of photosystem II crystal formation in photosynthetic membranes. *J. Biol. Chem.* **290**, 14091–14106 (2015).
26. A. Z. Kiss, A. V. Ruban, P. Horton, The PsbS protein controls the organization of the photosystem II antenna in higher plant thylakoid membranes. *J. Biol. Chem.* **283**, 3972–3978 (2008).
27. X. Pi *et al.*, Unique organization of photosystem I-light-harvesting supercomplex revealed by cryo-EM from a red alga. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4423–4428 (2018).
28. T. Veith, C. Büchel, 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 (2007).
29. R. R. L. Guillard, “Culture of phytoplankton for feeding marine invertebrates” in *Culture of Marine Invertebrate Animals*, W. L. Smith, M. H. Chanley, Eds. (Springer, Boston, MA, 1975).
30. R. R. Guillard, J. H. Ryther, Studies of marine planktonic diatoms. I. *Cyclotella nana* Hustedt, and *Detonula confervacea* (Cleve) Gran. *Can. J. Microbiol.* **8**, 229–239 (1962).
31. S. W. Jeffrey, G. F. Humphrey, New spectrophotometric equations for determining chlorophylls a, b, c1 and c2 in higher plants, algae and natural phytoplankton. *Biochem. Physiol. Pflanz.* **167**, 191–194 (1975).
32. D. N. Mastrorade, Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
33. G. Tang *et al.*, EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
34. E. F. Pettersen *et al.*, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
35. L. Arike *et al.*, Comparison and applications of label-free absolute proteome quantification methods on *Escherichia coli*. *J. Proteomics* **75**, 5437–5448 (2012).
36. M. Y. Gorbunov, P. G. Falkowski, “Fluorescence induction and relaxation (FIR) technique and instrumentation for monitoring photosynthetic processes and primary production in aquatic ecosystems” in *Photosynthesis: Fundamental Aspects to Global Perspectives*, A. van der Est, D. Bruce, Eds. (Allen Press, 2005), vol. **2**, pp. 1029–1031.
37. M. Chen, X. Kuang, O. Levitan, P. Falkowski, W. Dai, Averaged structure of Photosystem II holocomplex on thylakoid membrane from diatom. EMDDataResource. <https://www.emdataresource.org/EMD-0539>. Deposited 2 August 2019.
38. M. Chen, X. Kuang, O. Levitan, P. Falkowski, W. Dai, Averaged structure of Photosystem II 2D crystalline array on thylakoid membrane from diatom. EMDDataResource. <https://www.emdataresource.org/EMD-0540>. Deposited 2 August 2019.