Article

The cryo-EM structure of Photosystem I from *Chromera velia* with a bound superoxide dismutase heterodimer

Xukun Yuan[1], Martin Lukeš [2], Pu Qian [3], Quentin Charras[1], Tomáš Smrčka[4,5], Petr Man [4], Lucie Kovářová[2], Peter Koník [6], Emmanuel Heilmann [1], Ashraf Al-Amoudi[7], Roman Sobotka [2,6] ✉ & Andreas Naschberger [1] ✉

Photosystem I is a key component of the solar energy conversion machinery in oxygenic photosynthesis, and its core, where photochemistry occurs, is highly conserved. However, the coral-associated alga *Chromera velia* that is evolutionary linked to parasitic apicomplexans, exhibits Photosystem I with unusual features. These include the splitting of the central PsaA subunit and the binding of superoxide dismutases as regular subunits. The organization of such a unique Photosystem I was enigmatic. Here, we present the cryo-EM structure of *Chromera velia* Photosystem I at 1.84 Å resolution. Our work reveals a superoxide dismutase heterodimer bound to the stromal side of the core, stabilized by extensions of canonical subunits, a novel protein PsaV, and a reduced light-harvesting apparatus. We elucidate how the complex evolved to accommodate the superoxide dismutase, assemble the split PsaA, and integrate antenna proteins in a non-canonical orientation. Based on our data and prior physiological data, we propose that this specialized Photosystem I functions likely as an Mehler machine, redirecting electrons from Photosystem II back to water. This mechanism enables *Chromera velia* to manage redox imbalance and reduce photorespiration through localized oxygen consumption.

Oxygenic photosynthesis, the primary metabolic process on Earth, emerged around 3 billion years ago. By combining two light-driven oxidoreductase complexes (two photosystems—PSI and PSII) within a single cell[1] a cyanobacterial ancestor started to oxidize water as an unlimited electron source for carbon fixation. Eukaryotes later acquired this metabolic strategy transforming an endosymbiotic cyanobacterium into specialized organelles (plastids). Although the number and arrangement of subunits in photosystems, especially PSI, can vary in different types of phototrophs, their core structural architecture remained highly conserved, indicating strong evolutionary constraints[2].

A notable exception is represented by the photosynthetic apparatus of the Alveolate alga *Chromera velia*[3], which significantly diverged from the canonical "oxygenic" organisation. The plastid of *C. velia* derived from red algae and tracks back to the early photosynthetic origin of apicomplexan parasites, including serious pathogens such as *Plasmodium falciparum*, the causative agent of malaria. In *C. velia*, the PsaA locus underwent a gene fission and two distinct gene products, PsaA1 and PsaA2, are incorporated separately into the core[4]. Some small subunits were also lost in both photosystems[5], while PSI acquired five novel protein subunits. Of those five, three

[1]King Abdullah University of Science and Technology (KAUST), Biological and Environmental Science and Engineering Division, Thuwal, Saudi Arabia. [2]Centre Algatech, Institute of Microbiology of the Czech Academy of Sciences, Třeboň, Czech Republic. [3]School of Biosciences, University of Sheffield, Sheffield, UK. [4]Institute of Microbiology of the Czech Academy of Sciences, Prague, Czech Republic. [5]Department of Biochemistry, Faculty of Science, Charles University, Praha, Czech Republic. [6]Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic. [7]Electron Microscopy Lab, Imaging and Characterization Core Labs, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. ✉e-mail: sobotka@alga.cz; andreas.naschberger@kaust.edu.sa

(Unknown1–3) show no homology with any existing database entries, while two were homologous to superoxide dismutase (SOD). These metalloenzymes play a crucial role in oxygenic phototrophs as a first line defence against reactive oxygen species (ROS), catalyzing the dismutation of superoxide into O_2 and H_2O_2. All function as obligate oligomers (dimers, tetramers, or hexamers) and are classified according to their catalytic metal cofactor into Mn-, Fe-, Ni-, and Cu/Zn-SODs[6]. In chloroplasts of green algae, plants, and dinoflagellates[7], Fe-SODs are typically found in the chloroplast stroma, near PSI, the main site of superoxide production[6,8], and are essential for viability[9,10]. Notably, the previously isolated PSI from *C. velia* exhibited SOD activity in vitro[5] and spectroscopic data has ranked this modified PSI supercomplex among the most photochemically efficient PSI, outperforming plants and green algae[11].

Another interesting feature of *C. velia* is the presence of three types of light-harvesting proteins, two of which are found in distant algal groups. In particular, the fucoxanthin chlorophyll protein (FCP)-like-antennae have homologs in dinoflagellates and diatoms[12,13], while the Rhodophyta-type Light-Harvesting Complex (LHCr) antennae, which are associated with PSI, have homologs in red algae[13]. Although both LHCr and FCP proteins have been co-isolated with *C. velia* PSI[13,14], the main antenna bound to PSI consists of FCPs; and LHCr play only a minor role in the collection of light for PSI[12]. *C. velia* can also express far-red absorbing antennae specifically associated with PSII with an absorption maximum at 705 nm[15,16].

The pigment composition of *C. velia* is however, less diverse, with fewer cofactor types compared to related dinoflagellates or e.g., green microalgae. Together with chlorophyll *a* (CLA), β-carotene (BCR), and violaxanthin (XAT), *C. velia* exhibits a novel xanthophyll pigment (named chromeraxanthin; LRX). NMR data classified LRX as a new class of carotenoids, combining an acetylated cyclohexane end-ring from fucoxanthin with a methylated cyclopentane ring typical of capsanthin[17].

C. velia is preferentially associated with macroalgae assemblages of coral[18] but can also colonise coral larvae[19], similar to *Symbiodinium*, a related dinoflagellate whose common ancestor with *C. velia* lived 0.75–1.1 billion years ago[20]. The exact nature of this interaction is unclear[21], but *C. velia* is more likely a facultative (photo)parasite[22]. Indeed, this alga can be photoautotrophically cultivated and displays exceptionally robust and efficient photosynthesis. *C. velia* cells achieve similar maximum photosynthetic rates under both light-limited and light-saturated conditions[23,24], maintaining high efficiency across a wide wavelength range. Under sinusoidal light cycles peaking at 500 μmol photons m^{-2} s^{-1} that closely resemble natural conditions, *C. velia* exhibited very high rates of O_2 evolution and ^{14}C fixation during the light phase. These rates were 4–5 times higher than those measured in cells under continuous light.

Here, we report the structure of PSI from *C. velia* at a near-atomic resolution of 1.84 Å, revealing unprecedented structural details in the PSI supercomplex. We found an SOD heterodimer bound to the PSI core, stabilized by protein extensions of the canonical PSI subunits, and a previously unknown protein PsaV, which likely stabilizes the core PsaA subunit that is split into two separate polypeptides. Further, *C. velia* exhibits a minimized light-harvesting apparatus, as many of its FCPs have been lost due to structural rearrangements that would otherwise lead to steric clashes in a canonical PSI. The role of this unique PSI-SOD assembly in the physiology of *C. velia* is discussed.

Results

The overall structure of PSI from *C. velia*

To prepare PSI supercomplexes for cryo-EM, photosynthetic membranes were solubilized with β-dodecyl-n-maltoside (β-DDM) and fractionated using sucrose density gradient ultracentrifugation in lauryl maltose neopentyl glycol (LMNG) (Supplementary Fig. 1). The main PSI-containing fraction was collected and further purified either by (i) size exclusion chromatography (SEC) or (ii) native-PAGE

separation followed by gel elution (Supplementary Fig. 1). The subunit composition analysis using SDS-PAGE in the second dimension, followed by protein mass spectrometry (MS), confirmed a protein spectrum consistent with previous findings by Sobotka et al.[5] (Fig. 1a, b). High-performance liquid chromatography (HPLC)-based pigment analysis of the native gel-purified PSI identified the expected pigments: CLA, BCR, XAT, and the *C. velia*-specific LRX (Fig. 1c)[3]. For both purification approaches, we prepared EM grids by plunge-freezing and collected two independent datasets: 20,980 movies for the SEC purification (Dataset 1), and 11,917 movies for native-PAGE elution purification (Dataset 2). Dataset 1 refined to an overall resolution of 2.36 Å, while the gel-eluted PSI refined to 2.48 Å (Supplementary Fig. 2). Superposition of both maps showed very similar densities except for a clear density for a lumenal protein bound to PsaB, which was poorly ordered in Dataset 1, indicating that the separation on the native PAGE preserves loosely bound proteins better than SEC (Supplementary Fig. 2). No significant differences between maps were found in the core. Therefore, both datasets were merged, refining the consensus map (the full map without focused refinement) to a final global resolution of 1.84 Å (Supplementary Figs. 2 and 3 and Supplementary Table 1). The peripheral regions, particularly where the antennae are located, exhibited flexibility relative to the PSI core. We improved these areas by applying masked refinement using different masks (Supplementary Figs. 2 and 3). All maps were utilized to generate a merged composite map, into which an atomic model of PSI was built.

The structure of PSI from *C. velia* consists of 14 core subunits (PsaA1, PsaA2, PsaB, PsaC, PsaD, PsaE, PsaF, PsaI, PsaL, PsaM, PsaR, PsaSOD1, PsaSOD2, and PsaV) and five FCP antennae (FCPa, FCPb, FCPc, FCPd, and FCPe) (Fig. 2a). We found a total of 245 cofactors, specifically 114 CLA, 1 chlorophyll a' (CL0), 8 BCR, 22 XAT, 4 LRX, 11 monogalactosyl-diacylglycerol (LMG), 6 digalactosyl-diacylglycerol (DGD), 71 β-DDM (LMT), 1 lauryl maltose neopentyl glycol (AV0), 1 sulfoquinovosyl-diacylglycerol (SQD), 2 phylloquinone (PQN), 3 iron sulfur cluster (SF4) and 1 acylated monogalactosyl-diacylglycerol (RRL) (Fig. 1d, Supplementary Fig. 4a–d, and Supplementary Tables 2 and 3)[25,26].

The high-resolution, especially in the core, allowed us to unambiguously identify most rotameric conformations of the protein side chains and the identities of most of the ligands, including their conformations, cis-trans isomerization, and puckering of the head groups (Supplementary Fig. 5a). We also built 1558 water molecules, which are observable across all protein subunits, including the transmembrane part, with an accumulation around the iron-sulfur clusters and the reaction center (Supplementary Fig. 5b).

The previously reported unknown proteins 1, 2, and 3[5] could be identified by building side chains according to their appearance in the map, followed by BLAST searches against the available genomic and transcriptomic databases[27,28]. Despite low conservation with *Symbiodinium*[29-31] unknown proteins 2 and 3 could be assigned as PsaR and PsaI, respectively (Supplementary Fig. 6a, b). The protein previously named Unknown 1[5] was identified in the density but had no counterpart in other known PSI structures, therefore, this new subunit was named PsaV. Additionally, we identified density attributed to one unreported and unidentified protein[5]. We could assign it as PsaM (Fig. 2 and Supplementary Fig. 6a, b).

PsaA and PsaB are the major subunits of PSI and form a heterodimer hosting the reaction center. Their sequence and structure are highly conserved in all oxygenic photosynthetic organisms tracking back to cyanobacteria. In *C. velia* the reaction center subunit PsaA is split into two subunits, PsaA1 and PsaA2, which form a heterodimer within the PSI core (Fig. 3 and Supplementary Fig. 7a). Superposition with PsaA from dinoflagellates reveals that the PsaA1-PsaA2 heterodimer closely resembles the unsplit PsaA counterpart in *Symbiodinium*. However, several local structural differences were observed in PsaA1. The N-terminus is partially disordered (Fig. 3a), transmembrane helix 2
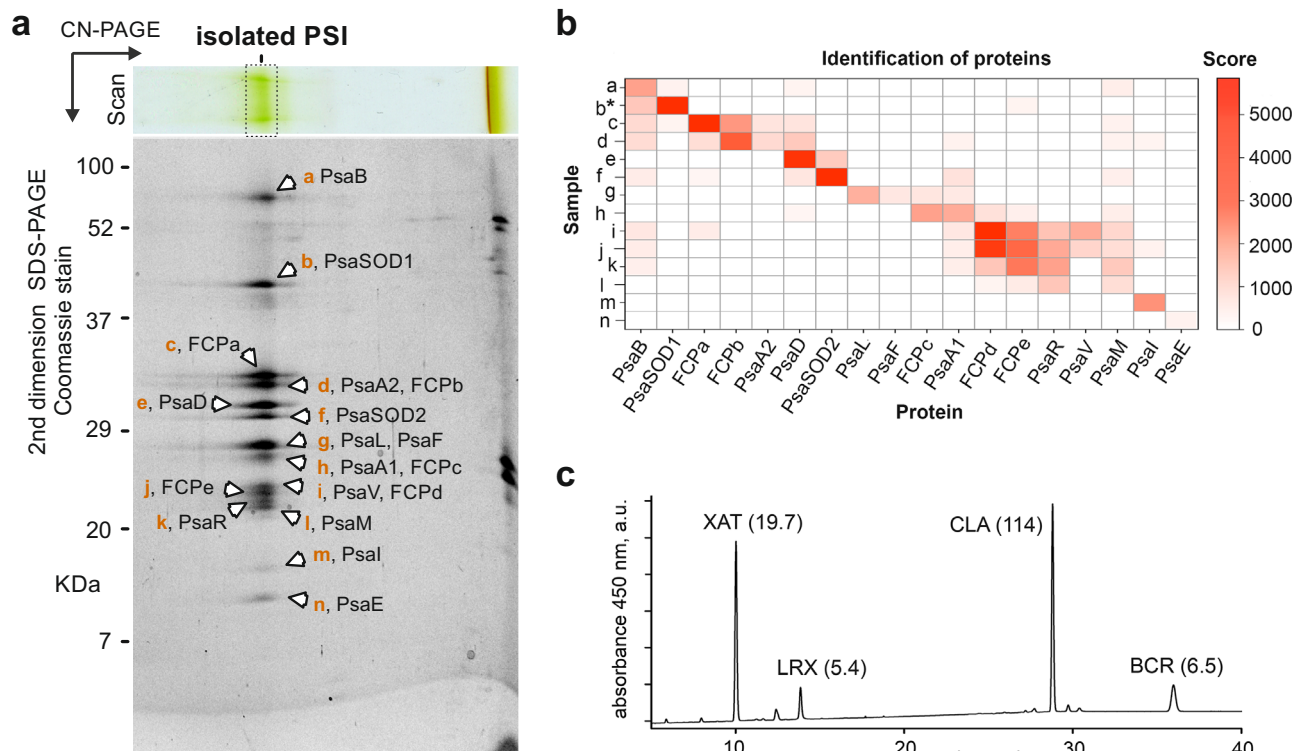
Fig. 1 | Analysis of protein and pigment composition of the purified PSI from *C. velia*. a The PSI was purified twice and in both cases, the PSI-enriched fraction was first isolated using a sucrose gradient. Subsequent purification steps involved either clear-native (CN) gel or size-exclusion chromatography (Supplementary Fig. 1). A single strip from the CN gel was further separated in the second dimension by SDS-PAGE and stained with Coomassie Blue. All visible protein spots were analyzed by MS (**b**). The proteins identified in selected spots are shown in a form of heat map, where each protein is colored based on its identification score (see Supplementary Data File for all identified proteins). * indicates that the tryptic digestion of the sample was combined with deglycosylation by PNGase F. **c** HPLC pigment analysis of the PSI complex eluted from the CN gel; the cut band is indicated by a dashed rectangle in (**a**). The ratio of individual carotenoids, showed in parentheses, is calculated to 114 CLA molecules that were identified in the cryo-EM structure of *C. velia* PSI from this work. The data for this figure are provided in the Source data file.

(TMH2) is shifted by around 8 Å due to a protein truncation and appears to be flexible (Fig. 3b), and the C-terminus, which comprises a linker between TMH4 and TMH5 in the unsplit PsaA, changes direction, forming a β-strand that interacts via a β-sheet with the FCPa antenna (Fig. 3c). Additional changes were observed in PsaA2, which include a truncation and a helical alteration of the polypeptide chain (Fig. 3c). Overall, the splitting of PsaA leads to greater conformational flexibility of PsaA1, and appears to be an evolutionarily adaption that connects the FCPa antenna to the PSI core via the C-terminus of PsaA1.

Surprisingly, we identified a polypeptide chain in the position of PsaJ, which was found to be absent in the genome of *C. velia*[5]. Instead of PsaJ, the position is occupied by the C-terminus of PsaF, which in *C. velia* was modeled with a 40 amino acid (AA) long extension. A non-conserved stretch of 20 amino acids links the canonical PsaF domain to the PsaJ-like domain. Sequence alignment shows a similarity of approximately 30% between the new PsaJ-like C-terminus of PsaF and PsaJ of *Symbiodinium*, implying a gene fusion event between PsaF and PsaJ rather than an extension of the PsaF gene (Supplementary Fig. 7b, c). Interestingly, in cyanobacteria, psaF and psaJ form a bicistronic operon[32], and in the red-lineage alga *Cyanidioschyzon merolae*, they are directly adjacent in the plastid genome[33], supporting this hypothesis.

Finally, we also discovered superoxide dismutase dimers linked to the PSI core, forming enzymatic subunits never previously resolved in PSI structures (Fig. 2). Hence, we named these SOD1 and SOD2 subunits as PsaSOD1 and PsaSOD2, respectively. In summary, our structure reveals an unusual composition of PSI, with several novel structural scaffolds and enzymatic subunits acquired in the core.

The PsaSOD1/2 heterodimer is stably attached to the stromal PSI PsaCDE triad

The well-ordered PsaSOD1/2 heterodimer is attached to the stromal subunits PsaCDE (Fig. 4a) via the N-terminal extension of PsaSOD1 (linker), which is absent in PsaSOD2. This linker extends about 100 AA and is positioned on the surface of the stromal side of PSI. It forms interactions with an extensive number of subunits, including PsaD, PsaL, PsaB, and FCPb (Fig. 4b). The N-terminus folds back to the attachment site, where it interacts with PsaM at its C-terminus. Several other interactions, although less extensive, take place with the canonical subunits PsaB, PsaC, PsaD, PsaE, PsaI, and PsaM (Fig. 4b).

Multiple sequence analyses of SODs from red and green lineages show the PsaSOD1 linker is conserved (34% identity) in *C. velia* and another chromerid alga *Vitrella brassicaformis* (Supplementary Fig. 8a). Some diatoms have a shorter, less conserved N-terminus (around 40–50 residues, 13% and 20% identity), which might serve as a weaker SOD-PSI linker. However, the Psa29 (also known as PsaS) subunit, which is found only in diatoms[34], would likely block the attachment of SOD (Supplementary Fig. 8b). Hence, the stably bound SODs may be a chromerid-specific adaptation, possibly present in ancestral diatoms but lost during evolution, as suggested by the weakly conserved extension in modern species.

Superposition with a ferredoxin-bound PSI structure of plants (PDB_ID: 6YAC) revealed the attachment of the SOD heterodimer on the side opposite to the expected ferredoxin docking site (Supplementary Fig. 8c). It further showed the absence of steric clashes between the position of ferredoxin and PSI, excluding a potential blockage of the canonical electron transfer by the SOD
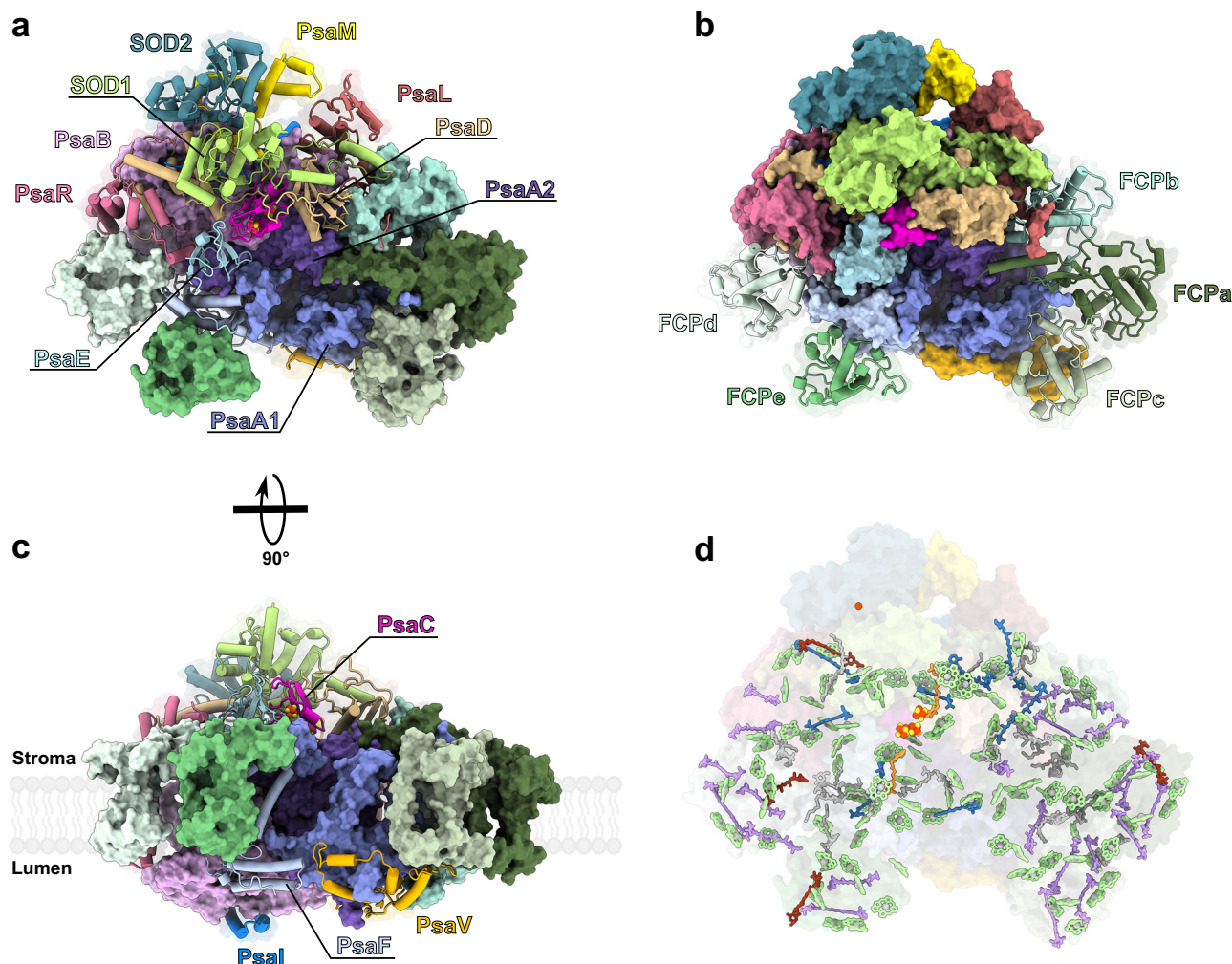
Fig. 2 | Overall structure of PSI from *C. velia*. a Top view from the stromal side showing the core subunits in ribbon style and PsaA1/A2, PsaB, and antenna proteins in surface representation. **b** Top view highlighting the antenna proteins in ribbon style and core subunits as surface. **c** Side view showing the transmembrane part. **d** Network of pigments and cofactors shown in stick and ball representations. CLAs are shown in green, XATs in purple, BCRs in blue, LRX in red, PQN in orange, lipids in light gray, iron-sulfur clusters in yellow-orange, and the iron cofactor in orange.

(Supplementary Fig. 8d, f). Interestingly, the binding affinity of ferredoxin to PSI in *C. velia* could be weaker than in other groups of algae due to less negatively charged surface of the *C. velia* ferredoxin and conformational changes at the N-terminus of PsaA1 (Supplementary Fig. 8d, f).

Functional and structural differentiation of the PsaSOD1/2 paralogues

A closer inspection of the catalytic sites of PsaSOD1 and PsaSOD2 revealed that only PsaSOD2 contains the prosthetic Fe coordinated by a typical iron coordination site[35], composed of three histidine and one aspartate (Fig. 4c). In PsaSOD2 we observed that the entry of the catalytic pocket points toward the stroma and is ~70 Å apart from phylloquinone, the major site of O_2^- generation in PSI[36] (Supplementary Fig. 9a). This finding suggests that diffusing O_2^- is scavenged by the catalytic site from the stroma rather than directly from the PSI acceptor site.

In PsaSOD1, one of the histidine of the expected active site was substituted by a threonine, making iron coordination here unlikely (Fig. 4c); instead, several ordered water molecules were found in the pocket. Each protomer thus apparently fulfills a distinct functional role. The PsaSOD2 is most likely a catalytically active subunit as previously reported[5], whereas PsaSOD1, which lacks a functional catalytic site, evolved into a structural support bridging PsaSOD2 to PSI.

should be noted that the formation of heterodimers is important also for the function of plastid Fe-SODs[10] and the superposition of the plant Fe-SOD homodimer[37] and *C. velia* PsaSOD1/2 reveals an overall structural similarity (RMSD = 2.2 Å).

Remarkably, in PsaSOD1, ASN76 and ASN202 each harbor a N-linked glycosylation (Fig. 4d and Supplementary Fig. 9b). Although the density beyond the core sugar moieties was too diffuse to model, MS revealed heterogeneous high-mannose N-glycans at Asn76 and Asn202, with $HexNAc_2Hex_7$ (i.e., $Man_7GlcNAc_2$) as the predominant form (Supplementary Fig. 10). The glycans project toward the stromal side, where they could modulate PsaSOD1 interface by local hydration, thereby stabilizing its association with the PSI core.

Stabilization of PsaA1/A2 and PsaSOD1/2 heterodimers by PsaV protein and by extended canonical subunits

The new protein PsaV is attached to the lumenal side of PSI, centered on the PsaA1/A2 heterodimer, where it stabilizes and clamps the two subunits together (Fig. 5a). PsaV forms a six-stranded beta sheet involving two times two antiparallel beta strands of PsaA2 (Fig. 5a). This inter-subunit β-sheet forms a closed, β-barrel-like structure, with antiparallel strand orientation within PsaA2 and within PsaV, whereas the interface between the two subunits exhibits a parallel strand arrangement (Supplementary Fig. 11). In addition, TMH3 and its
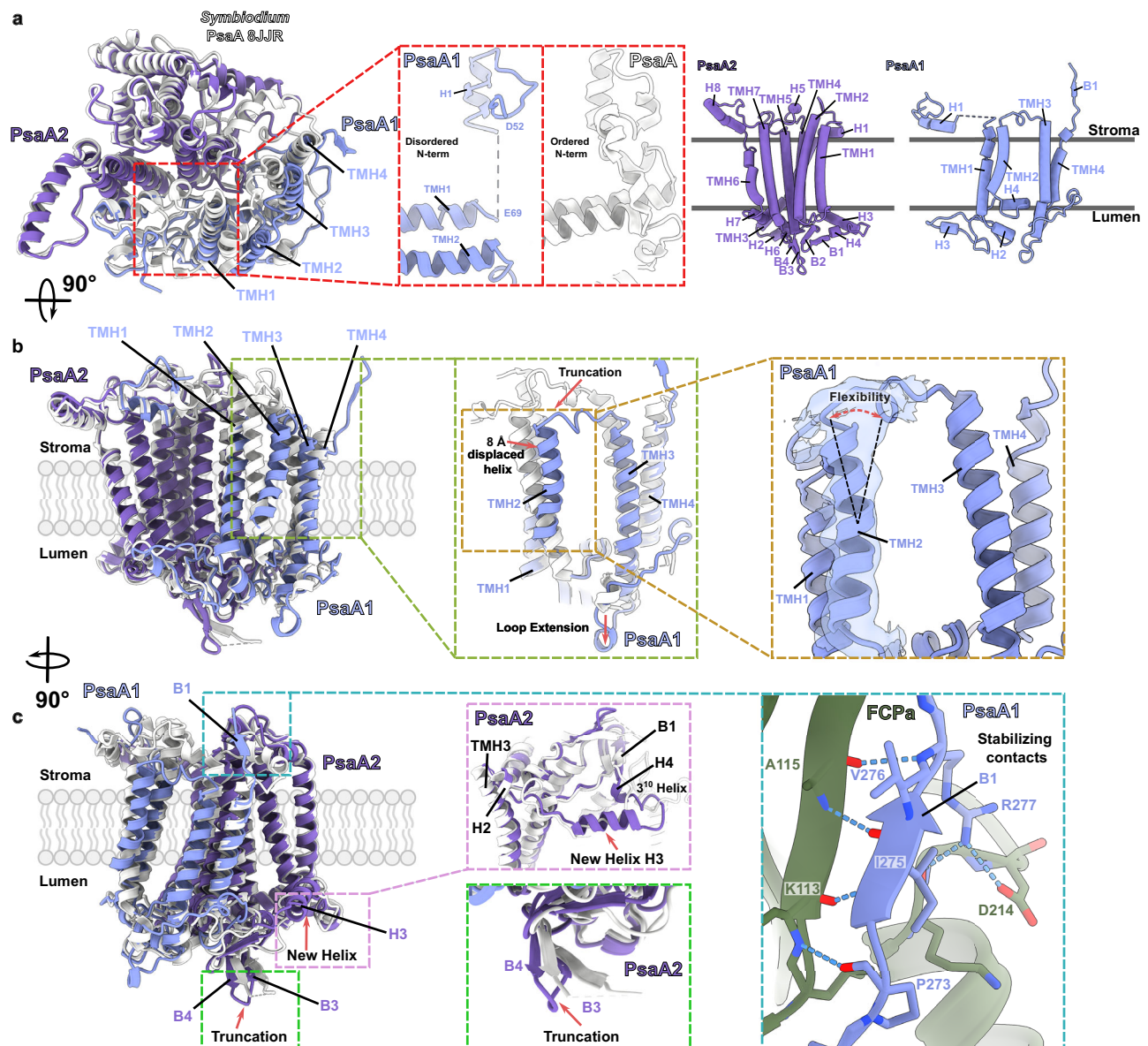
Fig. 3 | Splitting of the core subunit PsaA in *C. velia* and comparative analysis with the PsaA homologue from *Symbiodinium*. a The secondary structure of PsaA1 and PsaA2 is shown and numbered from the N- to the C-terminus: trans-membrane helices (TMH), helices (H), and β-strands (B). Note the partially disordered N-terminus near the ferredoxin-binding site. **b** The shifted TMH2 helix, accompanied by a truncation of the subsequent loop by ~8 Å, is depicted. The lower-resolution density, indicative of local flexibility of this helix, is displayed as a transparent isosurface representation. **c** Several structural remodelings are observed: the loop between β-sheets B3 and B4 is truncated, while the loop region between TMH3 and B1 is remodeled into two newly formed helices, H3 (α) and H4 (3_{10}). The new C-terminus is redirected toward FCPa, where it stabilizes attachment to the PSI core through a newly formed inter-subunit β-sheet and a salt bridge between R277 of PsaA1 and D214 of FCPa.

succeeding beta turn of PsaA1 are surrounded by three helices of PsaV, forming several polar, charged, and hydrophobic interactions with PsaA1 (Fig. 5a). The C-terminus of PsaV attaches to the C-terminal extension of PsaL, to PsaB, and PsaA2 while the N-terminus extends to the light-harvesting antenna FCPb and FCPc, stabilizing their attachment via polar, charged, and hydrophobic interactions (Fig. 5a). PsaV thereby stabilizes the split PsaA1/A2 heterodimer while supporting the attachment of the FCP antenna to the core of PSI.

Superposition of the PSI core subunits with structures of other organisms in the red-algal lineage revealed several protein extensions unique to *C. velia* (Supplementary Fig. 12a). Detailed analysis of these extensions shows that they are most dominant on the stromal side, forming a stabilizing platform for the PsaSOD1/2 heterodimer (Fig. 5b). A large C-terminal extension of PsaD forms the basis of the platform

near PsaSOD1 and is stabilized by several extensions of PsaR and a C-terminal extension of PsaI. The platform for PsaSOD2, in contrast, is formed by a large C-terminal extension of PsaM (Fig. 5b). Despite the significant gain in protein mass due to these extensions, the actual interactions between the PsaSOD1/2 heterodimer and the platform are less extensive than with the linker (Figs. 3 and 5b). Thus, if the primary attachment site to PSI is mediated by the N-terminal linker of PsaSOD1, the platform rather serves as a steric barrier to maintain a gap between active SOD domains and the thylakoid membranes. In theory, the SOD heterodimer could still bind similarly in green algae or plant PSI that lacks the platform and linker, without producing any steric hindrance. However, this interaction would be very weak, supported only by a few van der Waals contacts, and we therefore do not expect a stable PSI-SOD in these organisms (Supplementary Fig. 12b).
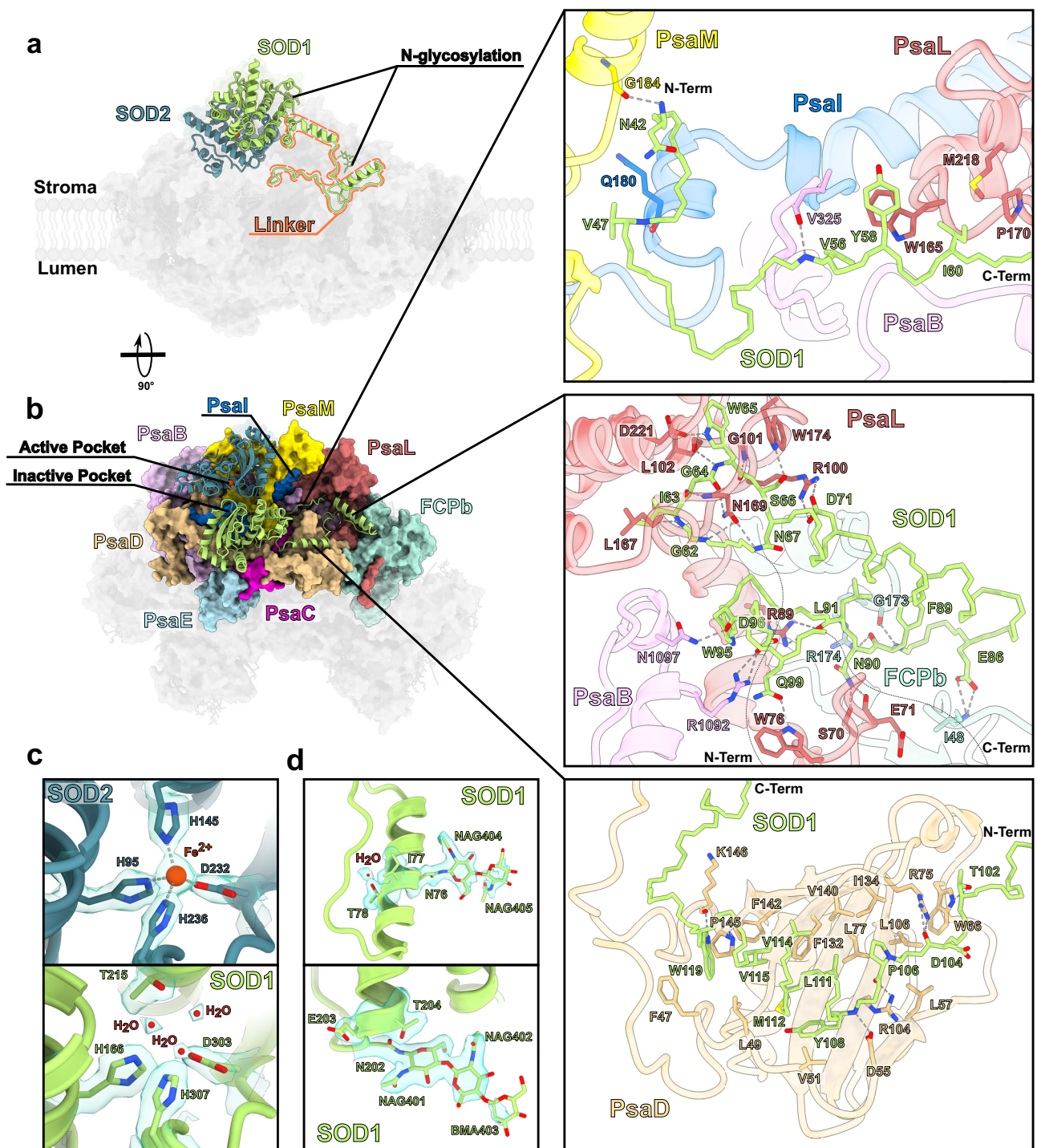
Fig. 4 | The attachment of PsaSOD1/2 subunits to the stromal side of the PSI core. a The PsaSOD1/2 heterodimer stably binds to the PsaCDE triad of PSI via an N-terminal extension (linker, highlighted in orange). **b** The linker of PsaSOD1 is buried in a groove that is formed by stromal parts of PsaD, PsaL, PsaI, PsaB, and FCPb, engaging in hydrophobic, charged, and polar interactions with a total buried surface area of 4675 Å². **c** PsaSOD2 is the catalytically active site with Fe bound to three histidines and one aspartate, while PsaSOD1 has lost its Fe cofactor due to a mutation of one histidine to threonine, resulting in unfavorable coordination for Fe. **d** PsaSOD1 has two N-linked glycosylation sites at the SOD domain and in the N-terminal linker region; both glycosylation sites were confirmed by MS analysis (Supplementary Fig. 10). The conserved NXS/T motif, with the sugar chains, are shown with density. Color coding follows the same scheme as in Fig. 1.

Protein extensions of FCPs antenna stabilize their attachment to the PSI core

Comparison of FCP proteins of *C. velia* and the closely related *Symbiodinium* by superposition showed several structural adaptations in the light-harvesting antenna proteins. The FCPs of *C. velia* can be classified into three types: The first type includes proteins with no extensions in *C. velia* (FCPc, FCPd, and FCPe), while their structural counterparts in *Symbiodinium* have extensions (Supplementary Fig. 13a). The second type includes FCPa with large protein extensions acquired by *C. velia* but absent in *Symbiodinium*, and the third type includes FCPb with extensions that differ from those of their counterparts in *Symbiodinium* (Supplementary Fig. 13b–d).
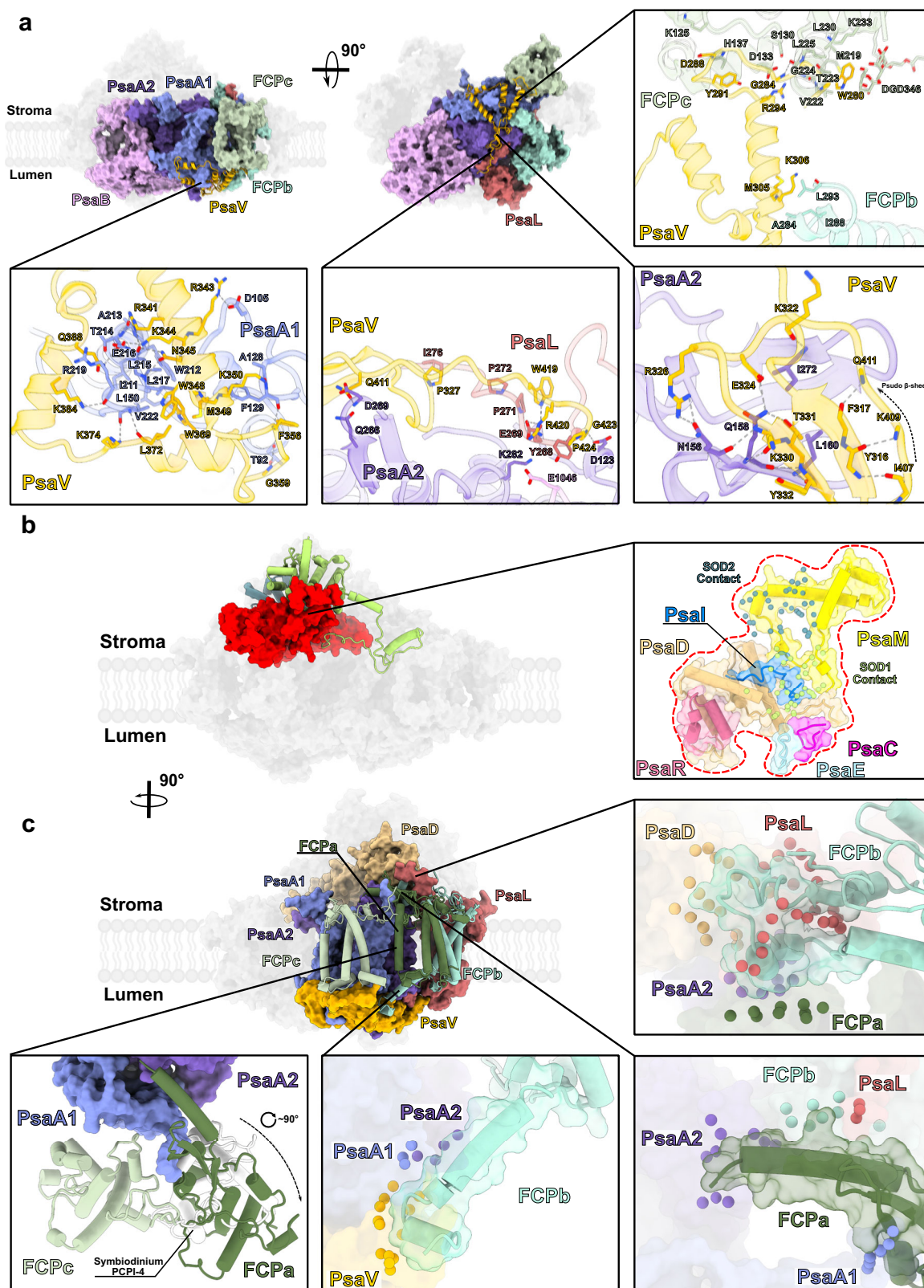
Fig. 5 | PsaV and protein extensions collectively stabilize the split PsaA1/A2, FCPs, and the PsaSOD1/2 heterodimer. a Overall and close-up views showing interactions of PsaV with the PsaA1/A2 dimer, FCPb, FCPc, PsaL, and PsaB. **b** The PsaSOD1/2 platform, formed by stromal protein extensions of several canonical subunits in the PSI core, is highlighted in red. Interacting residues are depicted as spheres centered on Cα atoms, colored according to their respective subunits (< 4 Å distance cut-off for sidechain-to-sidechain interactions). **c** Protein extensions of FCPs stabilize their attachments, with further support from PsaV. Interactions are shown the same way as in (**b**).

Notably, the FCPa is rotated by almost 90° around its transmembrane axis from its position in *Symbiodinium*, resulting in the loss of the canonical interaction with the PSI core via transmembrane helices (Fig. 5c). However, this antenna, organized as a heterodimer with FCPc, possesses a N-terminal extension that forms a long helix and a three-stranded β-sheet (Fig. 5c). This segment interacts with the PSI core on the stromal side, while FCPc anchors the heterodimer on the lumenal side via PsaV, thereby stabilizing the attachment of these antennae (Fig. 5a, c). It is worth noting that the rotated arrangement of the FCPa counterpart was also observed in the diatom structure[38], but not in another dinoflagellate (*Amphidinium carterae*)[29], suggesting that this rotation is a case of convergent evolutionary adaptation rather than divergent evolution with secondary loss (Fig. 5c).

The C-terminal extension of FCPb on the lumenal side interacts with PsaA1/A2 and the novel subunit PsaV (Fig. 5c). Its stromal N-terminal extension wraps around the N-terminal extension of PsaL and interacts with several other subunits, including PsaA2, PsaD, PsaSOD1, and FCPa (Fig. 5c). On the other hand, FCPe and FCPd did not acquire extensions but are attached to the PSI core through PsaD, PsaF, and PsaR for FCPd, while FCPe is connected to PsaF (Supplementary Fig. 13e). Interestingly, FCPe and FCPd do not form any protein-protein contacts, leading to a small gap between the two subunits. Their closest CLA-CLA distance is 26 Å, which indicates a very poor excitonic energy transfer (EET) between these antennas (Supplementary Fig. 13e). Hence, FCPe and FCPd are connected as single subunits to the PSI core and not as dimer.

The organization of chlorophylls and the FCP antennas shows a reduced light-harvesting apparatus

In dinoflagellates, the structural changes in the PSI core led to loss of many conserved CLA binding sites compared to diatoms (PDB_ID: 8JJR, 6L4U)[29,30,38,39]. The *C. velia* PSI shares most of these CLA losses (Fig. 6a), with 14 CLA binding sites lost in PsaB, 17 CLA in PsaA, and 1 in PsaF, PsaI, and PsaJ. However, *C. velia* stands out from *Symbiodinium* by absence of additional CLA in the PSI core (Fig. 6b); a detailed structural explanation is shown in Supplementary Fig. 14. Specifically, four clusters can be identified in which CLA sites are lost. Cluster 4 is the largest one and is adjacent to PsaA1, where 8 CLA disappeared in a region where antenna proteins would typically transfer energy to the core (Fig. 6b). The same phenomenon was observed for clusters 1 and 2, while the smaller cluster 3 lacks several CLA despite the adjacent FCPb. Overall, these findings indicate that the absence of the antenna correlates with the absence and shifts of respective CLA connections in the PSI core. On the other hand, two new CLA binding sites appeared in PsaF (CLA403 and CLA404), which are not observed in diatoms or dinoflagellates (Fig. 6a). Specifically, the new CLA403 in PsaF creates an EET pathway between FCPe and the core, likely to compensate the inability of FCPe to transfer excitation energy to FCPd (Supplementary Fig. 13e).

Similar to CLA, there are fewer carotenoids in the *C. velia* PSI core (Fig. 6c) than in *Symbiodinium*. Specifically, we found that 7 carotenoids are missing, while one new BCR1231 was acquired in PsaB. Additionally, one LRX L1236 in PsaB was found in a region where 4 carotenoids were missing in *C. velia*, whereas they are bound to PsaI and PsaM in *Symbiodinium*. Together with the loss, many of the carotenoids appear to be shifted, apparently dispersing the overall distribution of carotenoids within the PSI.

The PSI from *Symbiodinium* possesses an inner antenna belt assembled from 12 antenna monomers (Fig. 6d). In contrast, the PSI in *C. velia* binds only five FCPs; three located near the PsaA1/A2 heterodimer and two near the PsaF region. We cannot rule out that some antenna proteins have been stripped off during the isolation of the complex. Nevertheless, several structural features support the absence of additional antennae in vivo. Superposition with the *Symbiodinium* PSI shows a loop extension on the lumenal side of PsaB that would

produce a steric clash with a hypothetically attached FCP (PCPI-10 in reference structure) (Supplementary Fig. 15a). The two neighboring FCPs (PCPI-9 and PCPI-11) would also produce steric clashes with PsaR and PsaB, respectively (Supplementary Fig. 15b, c). Hence, the binding in *C. velia* of a PCPI-1 equivalent would also be destabilized by the absence of interactions with its neighboring FCP.
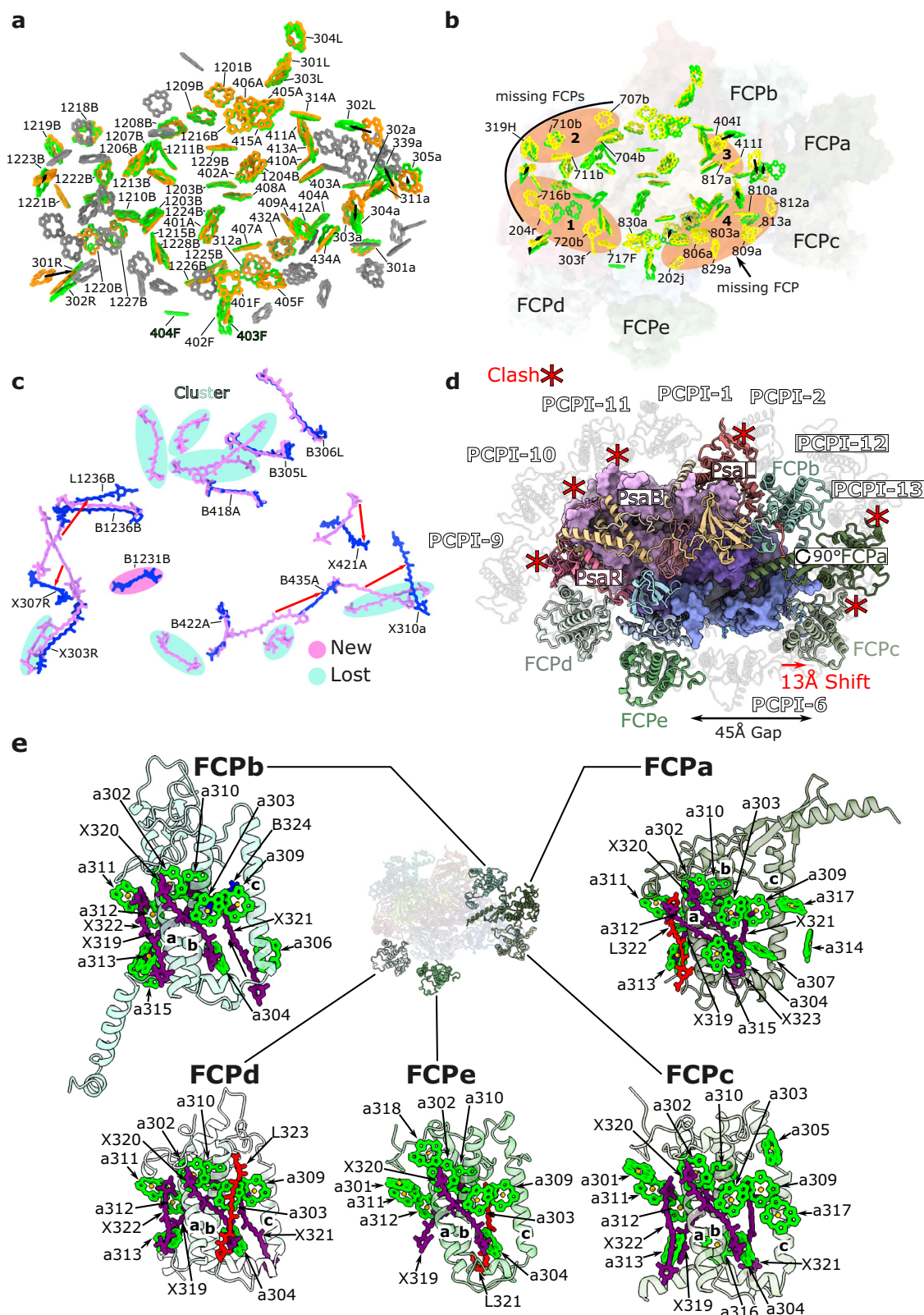
Moreover, the position of FCPe is conserved in *C. velia*, while FCPc is shifted by about 13 Å towards FCPa (Fig. 6d). This shift leaves a gap between FCPe and FCPc, leading to a loss of interaction for an additional FCP binding in this location and the stabilizing protein extension of PsaJ is absent in *C. velia*. Collectively, these structural rearrangements make a FCP (PCPI-6) binding between FCPe and FCPc unlikely (Supplementary Fig. 15d). Likewise, FCPa is also shifted about the same distance as FCPc. In addition, FCPa is rotated around the transmembrane axis by almost 90°. These rearrangements would lead to a steric clash with the neighbor PCPI-13 FCP-equivalent (Supplementary Fig. 15e). Consequently, a hypothetical PCPI-12 would also be destabilized as it loses its interaction site with the missing PCPI-13. Finally, the PCPI-2 equivalent would clash with the loop extension of PsaL, making an antenna binding in this position also unlikely (Supplementary Fig. 15f). We conclude that our structural data support a significant reduction of the PSI antenna size in *C. velia* compared to diatoms and *Symbiodinium*, which may represent a specific adaptation to *C. velia*'s ability to cope with fluctuating and high light conditions.

The function of PsaV and the N-terminal extension of PsaI parallels that of PsaU

The N-terminal extension of PsaI attaches to the lumenal side at PsaB and appears to stabilize the protein extension of PsaB between residues 605–656 (Supplementary Fig. 16). The interactions between PsaB and PsaI are primarily van der Waals and hydrophobic in nature. As the loop extension of PsaB prevents the binding of an antenna protein at this position (Supplementary Fig. 15a), the N-terminal extension of PsaI was likely acquired to support structural changes in PsaB that co-evolved with the minimization of the light-harvesting system. Interestingly, the *Symbiodinium* protein PsaU occupies a similar position to both PsaV and PsaI (Supplementary Fig. 16). Nevertheless, the structural configurations differ markedly. For PsaV compared to PsaU, the N- and C-terminal orientations are reversed, and the primary and secondary structures are different. PsaI also shows large differences in the extent and positioning of its helices. In addition, at the position where PsaU extends toward its C-terminus, the new helix in the PsaB loop (residues 410–421), which specifically evolved in *C. velia*, appears to substitute for PsaU at that location (Supplementary Fig. 16). These findings suggest that PsaV/PsaI evolved independently of PsaU. Nevertheless, these proteins likely arose to stabilize newly evolved structural deviations: PsaV clamps the split PsaA together and supports FCPa attachment, whereas PsaI stabilizes the newly evolved PsaB loop, which prevents antenna attachment in a canonical way. This is reminiscent of PsaU in *Symbiodinium*[30], which also stabilizes changes in its loop region, but in PsaA rather than PsaB.

The reduced antenna size and specific pigmentation favor direct antenna-to-core EET

A detailed analysis of the five FCPs revealed structural differences in pigment composition (Fig. 6e). FCPa has 12 CLA binding sites and 5 carotenoid binding pockets, FCPb has 10 CLA and 5 carotenoids, FCPc has 12 CLA and 4 carotenoids, FCPd has 8 CLA and 5 carotenoids, and FCPe has 9 CLA and 3 carotenoids. The bindings sites for CLA309, CLA302, CLA303, CLA304, CLA310, CLA311, CLA312, CLA313 (only absent in FCPe), CLA319, and the carotenoid binding pockets at positions 319, 320, and 321, and 322 (absent only in FCPe) are conserved among FCPs (Fig. 6e). *C. velia* specific LRX carotenoids were acquired at position 322 in FCPa, 321 in FCPe, and 323 in FCPd. We observed a

significant reduction in CLAs, especially in LRX containing FCPe and FCPd. Hence, this reduction in pigments appears to correlate with the acquisition of the LRX carotenoid, suggesting co-evolution between CLA reduction and the gain of LRX. We estimated the most probable EET pathways in the PSI using Förster models[40]. We find that *C. velia* connects each subunit separately to the core with relatively strong EET couplings, while the inter-FCP connections are weak (Supplementary

Fig. 17). The CLA-CLA pair between FCPd and FCPe is unlikely to transfer energy (Supplementary Fig. 13e), and a poor connection is also found between FCPa and FCPb. This implies that the direct FCP-to-core transfer is the dominant pathway. Only in the case of FCPa and FCPc, EET transfer takes place between light-harvesting proteins (Supplementary Fig. 17). Thus, the simplification of the light-harvesting apparatus also favors direct EET to the PSI core.

Fig. 6 | Reduced number of CLA and carotenoids in the PSI core and structural changes in the FCP antenna. a Lost CLA in *C. velia* PSI core compared to the diatom structure with PDB_ID: 6L4U. Missing CLAs in *C. velia* compared to 6L4U are shown in gray, all CLAs present in *C. velia* in green, and conserved CLAs shared between 6L4U and *C. velia* in orange. Shifted CLA binding sites are indicated by arrows. All conserved CLA binding sites in *C. velia* are labeled. **b** Comparison of CLA binding sites in *Symbiodinium* (yellow) and *C. velia* (green), with four clusters of decreased CLA sites in orange. Lost CLA sites in *Symbiodinium* are labeled (PDB_ID: 8JJR). Absent and present FCPs are indicated. **c** Comparison of carotenoids in *Symbiodinium* (purple) and *C. velia* (blue) PSI core. Lost carotenoids are in turquoise, new ones in pink. Carotenoids are labeled by type and subunit, with arrows indicating the shifts of carotenoid binding sites. **d** *C. velia* FCP antenna compared to *Symbiodinium* shows steric clashes (red star) and a - 45 Å shift of FCPa and FCPc, resulting in a widened gap (black arrow). **e** Close-up view of all five FCP antennae. Transmembrane helices are labeled, pigments numbered according to the PDB file of the current work.

Discussion

Unprecedented evolutionary novelties observed in *C. velia* include a split PsaA subunit of PSI and a split β-subunit of ATP synthase, the first steps of CLA biosynthesis located in the mitochondrion, a linear plastid genome, and widespread oligouridylation of transcripts[4,41]. In addition, the recruitment of SODs as stable PSI-associated components represents another remarkable innovation. The high-resolution structure of *C. velia* PSI solved in this work, shows how the complex has reshaped to allow attachment of the SOD (PsaSOD2), scaffolded over the inactive PsaSOD1 subunit. Furthermore, the canonical core proteins also evolved to provide a docking site for the SOD enzyme. The specialisation of PSI extended to FCPs, which contain unique extra segments for anchoring to a modified core-complex. A PSI-SOD complex is also present in chromerid alga *V. brassicaformis*, although it possesses an intact PsaA polypeptide[5]. Therefore, the split PsaA is probably not essential for the binding of SODs to PSI but might optimize the function of the complex. The new subunit PsaV and the new N-terminal extension of PsaI likely co-evolved to serve as compensatory elements for modifications of *C. velia* PSI. PsaV clamps the split PsaA more tightly to stabilize it, while it also facilitates the attachment of FCP proteins. In contrast, the N-terminal extension of PsaI stabilizes a newly formed extension in PsaB, which itself likely evolved as a compensatory structural element to replace the original FCP attachment site.

The function of SODs in phototrophs is not limited to ROS scavenging but these enzymes also play a role in redox balance. In the so-called Mehler reaction or water-water cycle (WWC), PSI complexes photoreduce molecular oxygen (O_2) to O_2^-, especially under conditions of highly reduced ferredoxin pool. The toxic O_2^- is rapidly reduced by SOD enzymes, and the resulting peroxide is converted back to water by peroxidases or catalases. In this way, the electrons from split water in PSII are transferred back to water in the stroma, and the process results in ATP synthesis without producing NADPH.

It is not entirely clear what physiological role the WWC plays in different groups of phototrophs because available quantifications of electron flow via WCC are (at best) only approximations. In model plant species (Arabidopsis, tobacco) and green algae, the SOD-mediated WWC has probably a minor role[42,43]. However, in diatoms and *Symbiodinium*, the WWC can contribute up to 50% of total electron flow[44,45]. We postulated a lowered affinity of ferredoxin to *C. velia's* PSI. This decrease in ferredoxin reduction rate should lead to a faster over-reduction of P700 and, consequently, the PSI-driven photoreduction of O_2 into O_2^-. In plants and green algae, this potentially harmful process is mitigated by activating cyclic electron flow[46], which redirects electrons back to plastoquinone/plastocyanin. In *C. velia*, the available physiological data imply an intensive alternative electron sink[23] and, given the strong WWC in *Symbiodinium*[44,45], we hypothesize that the PSI-SOD complex can operate as a "Mehler machine", which favors photoreduction of O_2^- over linear electron flow (see below).

A proposed intensive WWC in *C. velia* chloroplast would also require a relatively high concentration of enzymes scavenging produced quantities of toxic H_2O_2. To address this point, we used proteomics profiling to quantify soluble and membrane-bound proteins from *C. velia* cells grown at low-stress conditions. Notably, 2-Cys peroxiredoxin (PRX) is among the top 50 most abundant soluble proteins (Supplementary Table 4). This enzyme is a thiol-specific PRX in

chloroplast, and is essential for the WWC in Arabidopsis[47]. PRX-driven peroxide reduction contributes to redox balance (re-oxidation of ferredoxin) as an additional electron sink. The soluble extract from *C. velia* also contains high levels of two MnSODs (Supplementary Table 4), one (Cvel_4244) with a predicted signal peptide[5]. While MnSODs in chlorophytes are typically targeted to mitochondria[48], the very high MnSODs levels in *C. velia* suggests plastid localization, similar to diatoms[49]. If so, the concentration of SODs in *C. velia* chloroplast is enormous as FeSODs belong to the most abundant membrane-associated proteins (Supplementary Table 4). These data support our working model (Fig. 7) that, in *C. velia*, a large fraction of electrons generated by PSII can be directed by PSI into O_2^- and immediately detoxified by SODs and PRXs.

The evolutionary splitting of the PsaA subunit was poorly understood so far as it was hard to discern its biological function[4]. Based on our structure, we speculate that the splitting leads to greater conformational flexibility of PsaA1, which is evolutionarily adapted to connect the rotated FCPa antenna to the PSI core (Fig. 4). The unusual attachment of FCPa allows other antennae to shift, creating a large gap between FCPe and FCPc. This gap could serve as a docking site for other proteins or protein complexes. PSI and PSII in *C. velia* are probably tightly interconnected via antennas that enable energy spillover[23] or form a megacomplex[5]. We speculate that the splitting of PsaA is part of a radical redesign that ensures binding of SOD together with the attachment of (oxygen-producing) PSII (Fig. 7). In *C. velia*, PSII also lost several subunits, including PsbM that stabilizes the PSII dimer[5], and the CP47 subunit has a larger unique extension at the lumenal side.

However, this explanation needs to be considered also in light of the newly acquired subunit PsaV, which clamps together the split PsaA subunits at the lumenal side and also stabilizes the attachment of other antenna components, such as FCPb and FCPc. Hence, the FCPa attachment (a formation of a gap) is likely not the only reason for the splitting. An alternative hypothesis could be that possible higher O_2^- concentration in *C. velia* leads to differential local damage. It is conceivable that local O_2^- damage affects the PsaA subunit unevenly, which in turn could have favored the evolutionary splitting of PsaA to minimize the replacement cost by allowing only one subunit to be exchanged at higher rates. This is supported by the higher transcript coverage of PsaA2, which suggests increased synthesis of this subunit[4]. PsaA2 is also located closer to the reaction center, where most O_2^- should form.

The ability of *C. velia* to efficiently adapt to changing light conditions suggests that additional FCPs may dynamically and transiently associate with PSI-5FCP, though likely in a different configuration than in *Symbiodinium* due to steric constraints. Alternatively, no additional FCPs bind, instead, antenna proximity and interaction strength might be modulated to optimize energy transfer. The reduced antenna size could serve as a photoprotective adaptation under high light, balancing over-reduction. Antenna minimization was previously shown to benefit photosynthetic performance under light stress[50,51].

C. velia utilizes a form II Rubisco, which has fast kinetics but suffers from low CO_2-O_2 selectivity[3,52]. This fact could imply that *C. velia* concentrates carbon around Rubisco, but so far no such mechanism (e.g., pyrenoid) has been identified. The correlation of O_2 production
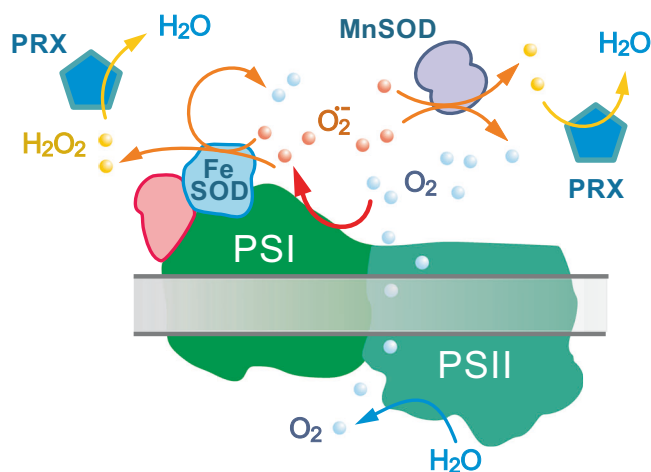
Fig. 7 | The PSI of *C. velia* as an enzymatic machine for oxygen photoreduction. PSI and PSII are co-localized or organized as a supercomplex in *C. velia* and oxygen therefore concentrates around PSI. Due to the weaker affinity of ferredoxin for PSI, oxygen is readily reduced to superoxide (O_2^-) and immediately dismutated into peroxide and oxygen either by PSI-attached FeSOD or by soluble MnSOD. The oxygen molecules from the dismutase half-reaction can undergo another round of reduction by PSI. The peroxide is finally converted back into water by the abundant thiol peroxiredoxins (PRX). As the reducing substrate of PRX are in most cases thioredoxins, this reaction serves as another sink for electrons in the chloroplast (see "Discussion").

with other photosynthetic parameters, however, suggests that an intensive O_2-consuming process(es) keeps the O_2 concentration in *C. velia* stroma low[23]. We hypothesize that the expected large capacity of the WWC in *C. velia* could thus prevent the over-reduction of the photosynthetic electron transport chain while lowering the oxygenase activity of type II Rubisco.

Methods

Culturing of *C. velia* cells
C. velia (RM12) cells were grown photoautotrophically in a rotary shaker in 500 ml Erlenmeyer flasks bubbled with air and under continuous light (day light bulbs) at 27 °C and in artificial seawater medium with f/2 nutrients. The applied light intensity was 100 µmol photons m^{-2} s^{-1}.

Purification of *C. velia* PSI
The resuspended cells in thylakoid buffer (20 mM MES, pH 6.5, 20 mM MgCl$_2$, 10 mM CaCl$_2$) were mixed with 0.5 ml of glass beads (0.1–0.2 mm diameter) in a 2 ml screw capped tube and broken using Mini-BeadBeater (BioSpec; twenty shaking cycles, 30 s each with 2-min breaks for cooling the suspension on ice). Membranes were separated from the cell extract by centrifugation (40,000 × *g*, 20 min) and solubilized with β-DDM (β-DDM/CLA = 15, w/w for 1 h on room temperature). The solubilized membranes (3 mL) were loaded on the top of a step sucrose gradient (38.5 mL in total, Thickwall polycarbonate ultracentrifuge tube, Beckmann Coulter) prepared from 12 ml 20% sucrose, 12 ml of 15% sucrose and 10 mL 10% sucrose in thylakoid buffer containing 0.01% LMNG (Supplementary Fig. 1a). After ultracentrifugation (18 h, 100,000 × *g*) the green fraction containing PSI was collected (5 mL) and concentrated on Amicon 100 kDa microconcentrators to 1 mL. From two independent purifications, one preparation was further separated on a 4–14% clear-native PAGE with 1% (w/v) A8-35 amphipol[53]. The second preparation was injected onto an Agilent-1260 HPLC system and separated on a Yarra SEC-4000 column (Phenomenex) using mobile phase (20 mM MES, pH 6.5, 100 mM NaCl, pH 6.5, containing 0.01% (w/v) LMNG at a flow rate of 0.25 ml min^{-1} at 15 °C. Eluted proteins and complexes were detected using a diode-

array detector and a fluorescence detector set to 440/675 nm (excitation/emission wavelengths) (Supplementary Fig. 1b).

2D electrophoresis
The individual components of *C. velia* PSI were resolved by incubating the gel strip from the first dimension in 25 mM Tris/HCl pH 7 containing 1% (w/v) SDS and 1% (w/v) dithiothreitol for 30 min at room temperature and by subsequent separation in the second dimension by SDS electrophoresis in a denaturing 12–20% (w/v) polyacrylamide gel containing 7 M urea. The separated proteins were stained with Coomassie Brilliant Blue.

Quantitative protein mass spectrometry
Cell lysate of *C. velia*, prepared as described earlier, was centrifuged (40,000 × *g*, 20 min), the supernatant collected as the soluble fraction and the pelleted membrane resuspended in 2 mL of the thylakoid buffer. The centrifugation was repeated, the pellet resuspended in 1 mL and analyzed as membrane fraction. To 20 µL of each fraction, ammonium bicarbonate (Sigma Aldrich; final concentration 50 mM) and Rapigest (Waters) surfactant (final concentration 0.1% v/v) were added to obtain 50 µL of sample. The sediment was resuspended in 50 µL of 50 mM ammonium bicarbonate with 0.1% (v/v) Rapigest and sonicated by 10 consecutive 1-s pulses interrupted by 1 s cooldown using a UP100H ultrasonic processor (Hielscher) on ice. Both mixtures were incubated at 60 °C for 45 min, then cooled to room temperature, proteomic grade trypsin (Pierce) was added to a final concentration of 10 ng/µL and incubated at 37 °C. After 12 h, samples were acidified by mixing with formic acid to a final concentration of 0.1% (v/v) and peptides were isolated by the StageTip procedure[54] to obtain 40 µL of the sample. LC-MS/MS analysis (*n* = 2) was performed on an UltiMate 3000 UHPLC (Thermo Fisher) on-line coupled to a TimsTOF pro (Bruker) mass spectrometer. 2 µL of the sample were trapped for 2 min on a ThermoFisher trap (0.3 × 5 mm, C18, 5 µm) column, then separated by reverse phase liquid chromatography on an Acclaim PepMap RSLC column (75 µm × 15 cm, C18, 2 µm, 100 Å; ThermoFisher). The 30 min gradient started at 3% of acetonitrile in 0.1% formic acid and rose to 50% acetonitrile in 0.1% formic acid. Peptides were ionized by CaptiveSpray, spectra were acquired in data dependent PASEF (Parallel Accumulation/Serial Fragmentation) mode with an accuracy of 0.2 ppm for precursors and 0.5 ppm for fragments. Raw data were processed by the MaxQuant/Andromeda software 1.6.14[55] and[56] compared to *C. velia* protein databases downloaded from CryptoDB (https://cryptodb.org/cryptodb/app) and with a list of plastid-encoded proteins in *C. velia* obtained from Prof. Miloslav Oborník (Biological Centre, Czech Academy of Sciences). Statistical analysis was performed in Perseus 1.6.14.0[57].

Analysis of protein gel spots using protein MS
Selected spots were excised from the gel, chopped into smaller pieces, and destained by sonication in 50 mM Tris-Cl (pH 8.0) containing 50% acetonitrile. Samples (*n* = 1) were then reduced with 10 mM TCEP at 65 °C for 10 min and subsequently alkylated with 20 mM iodoacetamide for 1 h in the dark. Both reactions were performed in 100 mM Tris-Cl (pH 8.0). The samples were further washed by sonication in water, acetonitrile, and 50% acetonitrile, and then dried in a vacuum concentrator (SpeedVac DNA 130, Thermo Fisher Scientific).

For digestion, the gel pieces were rehydrated in 50 mM ammonium bicarbonate containing sequencing-grade trypsin Gold (Promega) at a 1:50 protease:protein ratio and incubated at 37 °C for 18 h. Proteolysis was quenched by adding 0.1% trifluoroacetic acid (TFA) in 50% acetonitrile, followed by sonication, and the supernatants were collected into new tubes. The gel pieces were then extracted a second time with 20 µl of 0.1% TFA in 50% acetonitrile, sonicated, and the supernatants were pooled. Collected peptide solutions were dried in the vacuum concentrator and resuspended in 20 µl of 0.1% TFA in 5%

acetonitrile. For PsaS1, one half of the gel/sample was deglycosylated before digestion by overnight incubation with 500 U PNGase F (New England Biolabs) in 50 mM ammonium bicarbonate.

LC–MS/MS analysis was carried out on an EvoSep One system (EvoSep) coupled online to a timsTOF SCP mass spectrometer (Bruker Daltonics). A 15 µl aliquot of each sample was loaded onto Evotips (Evotip PURE, EvoSep) equilibrated in 0.1% formic acid and 0.015% DDM. Peptides were separated on a PepSep column (150 µm × 15 cm, C18, 1.5 µm, Bruker) using the EvoSep "30 samples per day" method (44 min gradient achieved with 0.1% formic acid in water and 0.1% formic acid in acetonitrile, at a flow rate of 0.5 µl/min). Spectra were acquired in data-dependent mode employing PASEF.

For protein identification, data were peak-picked in DataAnalysis (v6.1, Bruker Daltonics), exported to Mascot generic files, and searched with Mascot (v2.7.0, Matrix Science, London, UK) against databases containing *C. velia* protein sequences as described above with manually added sequence of PsaM and cRAP contaminants. Search parameters were as follows: enzyme specificity, Trypsin/P with up to two missed cleavages; carbamidomethylation (Cys) as a fixed modification; N-terminal acetylation and Met oxidation as variable modifications; and Asn deamidation as an additional variable modification for the deglycosylated sample. Mass tolerances were set to 15 ppm for precursors and 0.05 Da for fragments. Decoy searches were enabled with FDR < 1%, IonScore cutoff set to 25, and minimum peptide length > 5 amino acids. For glycan identification of PsaSOD1, raw data were also searched in PEAKS Studio 12.5 (Bioinformatics Solutions Inc.) using the Glycan workflow, GlycanScore cutoff set to 20 and identical search parameters as described above.

Pigment analysis

To identify pigments associated with isolated *C. velia* PSI, we excised a piece of CN gel (~ 3 × 2 mm) containing the separated complex and cut it further into several smaller pieces. Gel pieces were incubated in 200 µl of 25 mM MES buffer, pH 6.5 with 0.04% β-DDM overnight to elute PSI from the gel. After centrifugation, 50 µl of the supernatant were injected into the Agilent-1260 HPLC system equipped with a diode-array detector. Pigments were separated on a reverse-phase column (Zorbax Eclipse C18, 5 µm particle size, 3.9 × 150 mm; Agilent) with 35% (v/v) methanol and 15% (v/v) acetonitrile in 0.25 M pyridine (solvent A) and 20% (v/v) methanol, 20% (v/v) acetone, 60% (v/v) acetonitrile as solvent B. Pigments were eluted with a linear gradient of solvent B (30–95% (v/v) in 25 min) in solvent A followed by 95% of solvent B in solvent A at a flow rate of 0.8 ml min^{-1} at 40 °C. CLA and carotenoids were detected at 450 nm; the obtained peaks were integrated and molar stoichiometries calculated from calibration curves prepared using authentic standards. Since no chemical standard for LRX was available, we used a standard of echinenone (Sigma-Aldrich). This xanthophyll has a similar absorption maximum as LRX (~ 450 nm)[17] and, although the quantification of LRX is not absolutely precise, this approach provides a good estimate of the LRX content.

Grid preparations

The grids for the two datasets were prepared as follows. For dataset 1 (SEC purified sample), Quantifoil R 1.2/1.3 grids with 300 copper mesh (100 pieces, Thermo Fisher Scientific, USA) were glow-discharged for 45 s at 35 mA using a GloQube Plus instrument (Quorum Technologies, UK). The purified PSI sample, with a CLA concentration of ~0.3 mg/ml, was spun down at 22,000 × g for ten minutes to pellet larger aggregates. Three microliters of the supernatant were applied to 1.2/1.3 Cu Quantifoil 300 mesh grids. The sample was incubated on the grid for 30 s in the humidified chamber (100% humidity) at 4 °C using a Vitrobot (FEI Company, USA) plunge-freezing machine. The grids were then blotted for 2.5 s and plunge-frozen into liquid ethane.

For dataset 2, the sample eluted from clear-native PAGE was used. The greenish band in the gel was excised, cut and smashed into smaller

fragments using a scalpel. Gel fragments (~ 1 ml) were placed into a 15 ml Falcon tube, and the PSI was eluted by diffusion into 3 ml sample buffer (20 mM MES pH 6.5, 0.15 M NaCl, and 0.02% β-DDM) by incubating it using a rotary mixer at 4 °C overnight. The resuspension was then centrifuged for 2 h at 300,000 × g to pellet gel particles. The resulting supernatant was concentrated using a spin concentrator (Vivaspin 500, 100 kDa cutoff) and further used for grid preparations. Quantifoil R 1.2/1.3 grids with 300 copper mesh were glow-discharged for 30 s at 35 mA using a GloQube Plus instrument. After incubating the native gel eluted sample for 5 s, the grids were blotted for 3 s and plunge-frozen in liquid ethane. The blotting chamber temperature was set to 4 °C, and the relative humidity to 100%. All grids were then clipped and loaded into a 300-kV Titan Krios G4 microscope (Thermo Fisher Scientific, USA).

Cryo-EM data collections

Both datasets were collected on a Titan Krios G4 microscope, fitted with a Falcon 4i camera and Selectris × Energy Filter (Thermo Fisher Scientific, USA). Data collection was carried out using EPU 3.5.1.6034 software at a magnification of 165 k, which corresponds to a nominal pixel size of 0.729 Å/px. A total of 20,980 micrographs for dataset 1 and 11,917 micrographs for dataset 2 were acquired from grids that showed a good particle distribution. The dose rate was 6.73 electrons per physical pixel per second, and images were recorded for 3.16 s, corresponding to a total dose of 40 electrons per Å². Defocus was set to −0.5, −0.8, −1.1, −1.5, −1.8, and −2.1 µm.

Cryo-EM single particle analysis

Supplementary Fig. 2 provides an overview of the data processing workflow. Two datasets were independently processed using cryoSPARC[58] v4.5.3, as detailed below. Initially, motion correction and CTF estimation were performed on the raw data in cryoSPARC's Live session. Particle picking was conducted using blob picking, followed by an initial 2D classification. Promising 2D classes were selected to generate templates for subsequent particle picking. From these, 2,446,533 particles from Dataset 1 and 3,360,609 particles from Dataset 2 were extracted using a 500-pixel box size, which were subsequently binned to a 200-pixel box size (dataset 1) or 96-pixel box size (dataset 2). The particle stacks and processed exposures were then exported to cryoSPARC's workspace for further processing.

Both datasets underwent two rounds of 2D classification followed by Ab initio reconstruction to create initial maps for heterogeneous refinement. For Dataset 1, two rounds of heterogeneous refinement were performed using five classes (two Ab initio and three junk classes), starting with all 2,446,533 particles. After an additional Ab initio job, non-uniform refinement with 348,740 unbinned, re-extracted particles produced a 2.36 Å map. Dataset 2 underwent eight rounds of heterogeneous refinement using five classes (one Ab initio and four junk classes for the first seven rounds, two Ab initio and three junk classes for the final round). Non-uniform refinement of 184,142 cleaned and re-extracted particles yielded a 2.48 Å map.

The particles from both datasets were converted into star files using *csparc2star* from the PyEM suite (v0.5 Zenodo) (55) for further processing in RELION v5.0.0b[59]. Motion correction was repeated using RELION's built-in implementation. A total of 532,822 particles were extracted with a 500-pixel box size, and Datasets 1 and 2 were merged for subsequent steps. A reference map was generated via 3D refinement without a mask, employing a global angular search, and used to create a solvent mask. Several iterative steps were then performed. First, a 3D refinement with a solvent mask and local angular search resulted in a 2.48 Å map, which was sharpened using the postprocessing routine in RELION. Second, following the initial postprocessing step, CTF refinement (without fourth-order aberration correction) and a subsequent 3D refinement with local angular search improved the map to a 2.03 Å resolution. Third, another postprocessing step was

followed by particle polishing (separately for each dataset) and another 3D refinement, yielding a 1.92 Å map. Fourth, after a third postprocessing step, additional CTF refinement with fourth-order aberration correction and another 3D refinement produced a 1.85 Å map. Finally, following a fourth postprocessing step, a second round of polishing (separately for each dataset) and subsequent 3D refinement yielded a map with a 1.86 Å resolution. A final CTF refinement led to a consensus map with an overall resolution of 1.84 Å.

To improve local resolution in the FCPa, FCPe, and SOD regions, three individual local masks were created using ChimeraX's map eraser tool[60]. 3D refinements with these masks and local angular searches produced maps with resolutions of 2.17 Å (FCPa), 2.05 Å (FCPe), and 1.95 Å (SOD), as shown in Supplementary Fig. 2.

To enhance the local resolution of the N-terminal region of PsaI, particles from datasets 1 and 2 were manually separated using a text editor. 3D refinements with a solvent mask and local angular search were performed on each dataset individually. While dataset 1 yielded a map with poor PsaI density, dataset 2 produced a refined PsaI region map with an overall resolution of 1.94 Å.

Sequence alignment, BLAST searches, and protein modeling

Amino acid sequences for alignments were retrieved from either PDB structure files or NCBI BLAST/GenBank. The alignment of Acylated Galactolipid Associated Protein 1 (AGAP1) from *Arabidopsis thaliana* and *C. velia* was performed with the MUSCLE algorithm in Geneious Prime (Supplementary Fig. 4c). To find the amino acid sequence of AGAP1, we conducted a BLAST search with the homologous enzyme sequence from *Arabidopsis thaliana*. However, conventional BLASTp, even when restricted to sequences of *C. velia*, failed to yield results. We therefore used tBLASTn with the amino acid sequence of *Arabidopsis thaliana* in the Transcriptome Shotgun assembly library and restricted results again to *C. velia*. The resulting transcript (GenBank ID: JO804824) was shorter than that of *Arabidopsis thaliana* and is potentially truncated. As has been pointed out previously, sequence data of photosynthetic active organisms is often poorly annotated or data bases suffer from other issues, such as redundant entries[61,62]. We then retrieved the full-length sequence with a BLAST search performed in Geneious Prime, using Patent Data Bases as source. Both the truncated and full version of this sequence and resulting protein models were oddly similar to distantly related plant species. The deposited *C. velia* sequences[63] were indeed found previously to be contaminated with other sequences and an erratum was published[64]. Using the sequence data from Woehle et al. as custom BLAST data base in Geneious Prime and *Arabidopsis thaliana* AGAP1, we then found another *C. velia*, putative AGAP1 sequence. We could confirm this sequence in the online data base CryptoDB (https://cryptodb.org/cryptodb/app/) with the ID Cvel_10729, as it is annotated as putative mono- and diacylglycerol lipase. We used this sequence to model a potential *C. velia* AGAP1 fold with AlphaFold (Supplementary Fig. 4d).

To compare sequences of canonical subunits (PsaA, B, C, D, E, F, L, M, I, R) of *C. velia* to a closely related species, *Symbiodinium*, we retrieved *Symbiodinium* amino acid sequences from the PDB structure 8JJR and aligned them with the Clustal Omega algorithm in Geneious Prime.

For sequence alignments in Supplementary Fig. 7, sequences were retrieved from PDB file 8JJR and aligned in Geneious Prime with the software's own alignment option (cost matrix Blosum62). To avoid fragmentation of the two shorter sequences (*C. velia* PsaA1 and −2) when aligning to the longer sequence (*Symbiodinium* PsaA), we increased the gap open penalty to 20, the gap extension penalty was left as the default (= 3). PsaF/J sequences of *C. velia* and *Symbiodinium* were aligned with the same settings. SOD sequences of different species (*Chaetoceros tenuissimus* 1, NCBI ID: GFH51814; *Porphyridium purpureum*, NCBI ID and query cover: KAA8495090.1:64-261; *Chaetoceros tenuissimus* 2 NCBI ID: GFH47345; *C. velia*: sequence from current

work; *V. brassicaformis*, NCBI ID and query cover: CEM36361.1:51-331; *Symbiodinium*, NCBI ID and query cover: AAX99422.1:3-197; *Physcomitrium patens*, NCBI ID and query cover: XP_024379755.1:60-263; *Chlamydomonas reinhardtii*, NCBI ID and query cover: AAB04944.1:47-228; *Pisum sativum*, NCBI ID and query cover: XP_050888148.1:65-271) in Supplementary Fig. 8 were aligned using Clustal Omega in Geneious Prime with default settings. FCP sequences from the PSI complex structure of this current manuscript were used for a Clustal Omega alignment with default settings in Geneious Prime. Protein models in Supplementary Fig. 4d (*C. velia* AGAP1), Supplementary Fig. 8 (*C. velia* ferredoxin), and Supplementary Fig. 12 *(Pisum sativum* SODs) were modelled using the SWISS-MODEL server[65].

Model building and refinement

The atomic model building was carried out using Coot 9.8.98[66] in combination with ISOLDE[67] in ChimeraX 1.7.1[60]. The known subunits were predicted using SWISS MODEL[65] and then rigid-body fitted into the map as a starting model using UCSF Chimera v1.17[68]. Real-space refinement in Coot was then carried out to rebuild the model, deleting residues that were absent in the density and correcting the rotamers and deviations in the main chain tracings. The unknown proteins were first modeled manually by tracing about 100 amino acids and estimating the sidechains from the density. The predicted sequences were then exported and BLASTed[27] against genomic and transcriptomic databases of *C. velia*[28]. The hits were unambiguously linked to only one protein. The sequences were then used to manually rebuild the unknown proteins and extend the missing parts using Coot. For model building, all maps (consensus, SOD-region, FCPe-region, FCPa-region) were first aligned in Chimera, resampled, and then imported to Coot to be used simultaneously for model building. Target maps were changed when necessary, so that model fitting was always done in maps with the best local quality. Model-map weights were adjusted accordingly. In lower quality maps of the FCP regions, more weight was given to the model restraints, while in the core regions of PSI, the map weight was increased and model restraint was lowered. All ligands, like CLAs, carotenoids, lipids, detergents, and other cofactors, were manually placed using restraints files generated from the Grade2 server (http://grade.globalphasing.org). The modelling of ligand identities was based on high-resolution density map. If the density was located in a well-resolved region of the map, then inspection of both sharpened and unsharpened maps was sufficient to unambiguously identify the ligand. For the LRX ligands, the density was unambiguous for at least one head group (one side of the molecule showed high-resolution density and the other side was less well-resolved). Nevertheless, the quality of the density on the well-resolved side was sufficient to confidently determine the ligand's identity. Similarly, the RRL ligand shows well-resolved density for the core moiety and sugar group, including correct sugar-ring puckering, whereas the aliphatic chains are less defined. After fitting the model, PHENIX software suite (1.21-5207)[69] was used for an initial round of model refinement. Ready-set was utilized to add hydrogens and determine the link records for the coordination of the chlorophylls. The overall refinement weight was optimized and applied in subsequent refinement steps. For the refinement, a composite map was created with PHENIX, which averaged all maps weighted by their local correlation coefficient with the model. This map was first used to refine the model using the *minimization_global* options. After the initial refinement with PHENIX real-space refine, the model was imported into ISOLDE (v1.7)/ChimeraX, to fix clashes and rotameric outliers of the sidechains. Then, a second round of real-space refinement was conducted, with several iterations of rebuilding, fixing outliers, and model validation using MolProbity[70]. The final refinement was done by refining only the atomic B-factors using the consensus map instead of the composite map. Finally, validation was conducted using the validation server of the Protein Data Bank (PDB) (https://validate-rcsb-1.wwpdb.org/). All maps and the

model were deposited in the PDB databank. All structural figures were generated with Chimera X[60] and UCSF Chimera[68].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Atomic coordinates of the PSI supercomplex from *C. velia* have been deposited in the PDB under accession code 9HYU. The composite density map, including the core region, SOD region, FCPa region, PsaI N-terminal region, and FCPe region, has been deposited in the Electron Microscopy Data Bank (EMDB) under accession code EMD-52518 (composite map). The consensus map is deposited under accession code EMD-52508, along with EMD-52501 (SOD region), EMD-52524 (FCPa region), EMD-52517 (FCPe region), and EMD-52507 (PsaI N-terminal region). Whole-cell proteomics data and the MS analysis of protein spots form the 2D gel have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers and PXD060022 and PXD068416. Source data are provided with this paper.

References

1. Nelson, N. & Ben-Shem, A. The complex architecture of oxygenic photosynthesis. *Nat. Rev. Mol. Cell Biol.* **5**, 971–982 (2004).
2. Iwai, M., Patel-Tupper, D. & Niyogi, K. K. Structural diversity in eukaryotic photosynthetic light harvesting. *Annu. Rev. Plant Biol.* **75**, 119–152 (2024).
3. Moore, R. B. et al. A photosynthetic alveolate closely related to apicomplexan parasites. *Nature* **451**, 959–963 (2008).
4. Janouškovec, J. et al. Split photosystem protein, linear-mapping topology, and growth of structural complexity in the plastid genome of *Chromera velia*. *Mol. Biol. Evol.* **30**, 2447–2462 (2013).
5. Sobotka, R. et al. Extensive gain and loss of photosystem I subunits in chromerid algae, photosynthetic relatives of apicomplexans. *Sci. Rep.* **7**, 13214 (2017).
6. Pilon, M., Ravet, K. & Tapken, W. The biogenesis and physiological function of chloroplast superoxide dismutases. *Biochim. Biophys. Acta* **1807**, 989–998 (2011).
7. Okamoto, O. K., Robertson, D. L., Fagan, T. F., Hastings, J. W. & Colepicolo, P. Different regulatory mechanisms modulate the expression of a dinoflagellate iron-superoxide dismutase. *J. Biol. Chem.* **276**, 19989–19993 (2001).
8. Perry, J. J., Shin, D. S., Getzoff, E. D. & Tainer, J. A. The structural biochemistry of the superoxide dismutases. *Biochim. Biophys. Acta* **1804**, 245–262 (2010).
9. Myouga, F. et al. A heterocomplex of iron superoxide dismutases defends chloroplast nucleoids against oxidative stress and is essential for chloroplast development in Arabidopsis. *Plant Cell* **20**, 3148–3162 (2008).
10. Gallie, D. R. & Chen, Z. Chloroplast-localized iron superoxide dismutases FSD2 and FSD3 are functionally distinct in Arabidopsis. *PLoS ONE* **14**, e0220078 (2019).
11. Belgio, E. et al. High photochemical trapping efficiency in Photosystem I from the red clade algae *Chromera velia* and *Phaeodactylum tricornutum*. *Biochim. Biophys. Acta Bioenerg.* **1858**, 56–63 (2017).
12. Pan, H., Šlapeta, J., Carter, D. & Chen, M. Phylogenetic analysis of the light-harvesting system in *Chromera velia*. *Photosynth Res.* **111**, 19–28 (2012).
13. Tichy, J. et al. Light harvesting complexes of *Chromera velia*, photosynthetic relative of apicomplexan parasites. *Biochim. Biophys. Acta* **1827**, 723–729 (2013).
14. Mann, M. et al. Unusual features of the high light acclimation of *Chromera velia*. *Photosynth Res.* **122**, 159–169 (2014).
15. Bina, D. et al. Novel type of red-shifted chlorophyll a antenna complex from *Chromera velia*: II. Biochemistry and spectroscopy. *Biochim. Biophys. Acta* **1837**, 802–810 (2014).
16. Kotabová, E. et al. Novel type of red-shifted chlorophyll a antenna complex from *Chromera velia*. I. Physiological relevance and functional connection to photosystems. *Biochim. Biophys. Acta* **1837**, 734–743 (2014).
17. Takacs, Z. H. et al. NMR structure of chromeraxanthin - a unique allenic carotenoid of the apicomplexan alga *Chromera velia* https://doi.org/10.26434/chemrxiv-2025-5b5vd (2025).
18. Janouškovec, J., Horák, A., Barott, K. L., Rohwer, F. L. & Keeling, P. J. Environmental distribution of coral-associated relatives of apicomplexan parasites. *ISME J.* **7**, 444–447 (2013).
19. Cumbo, V. R. et al. *Chromera velia* is endosymbiotic in larvae of the reef corals Acropora digitifera and *A. tenuis*. *Protist* **164**, 237–244 (2013).
20. Strassert, J. F. H., Irisarri, I., Williams, T. A. & Burki, F. A molecular timescale for eukaryote evolution with implications for the origin of red algal-derived plastids. *Nat. Commun.* **12**, 1879 (2021).
21. Mohamed, A. R. et al. Deciphering the nature of the coral-Chromera association. *ISME J.* **12**, 776–790 (2018).
22. Oborník, M. Photoparasitism as an intermediate state in the evolution of apicomplexan parasites. *Trends Parasitol.* **36**, 727–734 (2020).
23. Quigg, A. et al. Photosynthesis in *Chromera velia* represents a simple system with high efficiency. *PLoS ONE* **7**, e47036 (2012).
24. Kotabová, E., Kaňa, R., Jarešová, J. & Prášil, O. Non-photochemical fluorescence quenching in *Chromera velia* is enabled by fast violaxanthin de-epoxidation. *FEBS Lett.* **585**, 1941–1945 (2011).
25. Nilsson, A. K. et al. Acylated monogalactosyl diacylglycerol: prevalence in the plant kingdom and identification of an enzyme catalyzing galactolipid head group acylation in *Arabidopsis thaliana*. *Plant J.* **84**, 1152–1166 (2015).
26. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
27. Johnson, M. et al. NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36**, W5–W9 (2008).
28. Puiu, D., Enomoto, S., Buck, G. A., Abrahamsen, M. S. & Kissinger, J. C. CryptoDB: the Cryptosporidium genome resource. *Nucleic Acids Res.* **32**, D329–D331 (2004).
29. Li, X. et al. Structures and organizations of PSI-AcpPCI supercomplexes from red tidal and coral symbiotic photosynthetic dinoflagellates. *Proc. Natl. Acad. Sci. USA* **121**, e2315476121 (2024).
30. Zhao, L. S. et al. Architecture of symbiotic dinoflagellate photosystem I-light-harvesting supercomplex in Symbiodinium. *Nat. Commun.* **15**, 2392 (2024).
31. Woo, Y. H. et al. Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. *Elife* **4**, e06974 (2015).
32. Muhlenhoff, U., Haehnel, W., Witt, H. & Herrmann, R. G. Genes encoding eleven subunits of photosystem I from the thermophilic cyanobacterium Synechococcus sp. *Gene* **127**, 71–78 (1993).
33. Ohta, N. et al. Complete sequence and analysis of the plastid genome of the unicellular red alga Cyanidioschyzon merolae. *DNA Res.* **10**, 67–77 (2003).
34. Kato, K. et al. Structural basis for molecular assembly of fucoxanthin chlorophyll a/c-binding proteins in a diatom photosystem I supercomplex. *Elife* https://doi.org/10.7554/eLife.99858 (2024).
35. Zheng, H. et al. CheckMyMetal: a macromolecular metal-binding validation tool. *Acta Crystallogr. D. Struct. Biol.* **73**, 223–233 (2017).
36. Kozuleva, M. et al. Phylloquinone is the principal Mehler reaction site within photosystem I in high light. *Plant Physiol.* **186**, 1848–1858 (2021).
37. Favier, A. et al. The plastid-encoded RNA polymerase-associated protein PAP9 is a superoxide dismutase with unusual structural features. *Front. Plant Sci.* **12**, 668897 (2021).

38. Nagao, R. et al. Structural basis for assembly and function of a diatom photosystem I-light-harvesting supercomplex. *Nat. Commun.* **11**, 2481 (2020).

39. Xu, C. et al. Structural basis for energy transfer in a huge diatom PSI-FCPI supercomplex. *Nat. Commun.* **11**, 5081 (2020).

40. Sener, M. et al. Forster energy transfer theory as reflected in the structures of photosynthetic light-harvesting systems. *Chemphyschem* **12**, 518–531 (2011).

41. Kořený, L., Sobotka, R., Janouškovec, J., Keeling, P. J. & Oborník, M. Tetrapyrrole synthesis of photosynthetic chromerids is likely homologous to the unusual pathway of apicomplexan parasites. *Plant Cell* **23**, 3454–3462 (2011).

42. Asada, K. The water-water cycle as alternative photon and electron sinks. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **355**, 1419–1431 (2000).

43. Huang, W., Yang, Y. J. & Zhang, S. B. The role of water-water cycle in regulating the redox state of photosystem I under fluctuating light. *Biochim. Biophys. Acta Bioenerg.* **1860**, 383–390 (2019).

44. Waring, J., Klenell, M., Bechtold, U., Underwood, G. J. C. & Baker, N. R. Light-induced responses of oxygen photoreduction, reactive oxygen species production and scavenging in two diatom species 1. *J. Phycol.* **46**, 1206–1217 (2010).

45. Roberty, S., Bailleul, B., Berne, N., Franck, F. & Cardol, P. PSI Mehler reaction is the main alternative photosynthetic electron pathway in Symbiodinium sp., symbiotic dinoflagellates of cnidarians. *N. Phytol.* **204**, 81–91 (2014).

46. Jokel, M., Johnson, X., Peltier, G., Aro, E. M. & Allahverdiyeva, Y. Hunting the main player enabling Chlamydomonas reinhardtii growth under fluctuating light. *Plant J.* **94**, 822–835 (2018).

47. Awad, J. et al. 2-cysteine peroxiredoxins and thylakoid ascorbate peroxidase create a water-water cycle that is essential to protect the photosynthetic apparatus under high light stress conditions. *Plant Physiol.* **167**, 1592–1603 (2015).

48. Wolfe-Simon, F., Grzebyk, D., Schofield, O. & Falkowski, P. G. The role and evolution of superoxide dismutases in algae. *J. Phycol.* **41**, 453–465 (2005).

49. Wolfe-Simon, F., Starovoytov, V., Reinfelder, J. R., Schofield, O. & Falkowski, P. G. Localization and role of manganese superoxide dismutase in a marine diatom. *Plant Physiol.* **142**, 1701–1709 (2006).

50. Friedland, N. et al. Fine-tuning the photosynthetic light harvesting apparatus for improved photosynthetic efficiency and biomass yield. *Sci. Rep.* **9**, 13028 (2019).

51. Mussgnug, J. H. et al. Engineering photosynthetic light capture: impacts on improved solar energy to biomass conversion. *Plant Biotechnol. J.* **5**, 802–814 (2007).

52. Prywes, N., Phillips, N. R., Tuck, O. T., Valentin-Alvarado, L. E. & Savage, D. F. Rubisco function, evolution, and engineering. *Annu. Rev. Biochem.* **92**, 385–410 (2023).

53. Konert, M. M., Wysocka, A., Koník, P. & Sobotka, R. High-light-inducible proteins HliA and HliB: pigment binding and protein-protein interactions. *Photosynth Res.* **152**, 317–332 (2022).

54. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).

55. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).

56. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).

57. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).

58. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

59. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S. H. W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.* **478**, 4169–4185 (2021).

60. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

61. Bolger, M. E., Arsova, B. & Usadel, B. Plant genome and transcriptome annotations: from misconceptions to simple solutions. *Brief. Bioinform.* **19**, 437–449 (2018).

62. Steuernagel, B. et al. The NLR-annotator tool enables annotation of the intracellular immune receptor repertoire. *Plant Physiol.* **183**, 468–482 (2020).

63. Woehle, C., Dagan, T., Martin, W. F. & Gould, S. B. Red and problematic green phylogenetic signals among thousands of nuclear genes from the photosynthetic and apicomplexa-related *Chromera velia*. *Genome Biol. Evol.* **3**, 1220–1230 (2011).

64. Woehle, C., Dagan, T., Martin, W. F. & Gould, S. B. Erratum: Red and problematic green phylogenetic signals among thousands of nuclear genes from the photosynthetic and apicomplexa-related *Chromera velia*. *Genome Biol. Evol.* **4**, 617 (2012).

65. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).

66. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).

67. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D. Struct. Biol.* **74**, 519–530 (2018).

68. Pettersen, E. F. et al. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

69. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D. Struct. Biol.* **74**, 531–544 (2018).

70. Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).

Acknowledgements

Author contributions

A.N. and R.S. designed the study. R.S., L.K., and M.L. purified the PSI sample and analyzed pigments; P.K., T.S., and P.M. carried out protein mass spectrometry. A.N. and X.Y. prepared the sample for cryo-EM analysis. A.A., A.N., and X.Y. collected the cryo-EM data and processed the data. A.N. and X.Y. built the atomic model. A.N. and QC analyzed the structure and interpreted the structural data. Q.P. revised and refined the atomic model and deposited the data. E.H. carried out BLAST searches and protein modeling. All authors jointly analyzed the data, prepared the figures, and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at
https://doi.org/10.1038/s41467-025-67637-0.

Correspondence and requests for materials should be addressed to Roman Sobotka or Andreas Naschberger.

Peer review information *Nature Communications* thanks Tanai Cardona who co-reviewed with Emma ChalonerJianing Mi, Dmitry Semchonok, and the other anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at
http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.