A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle

Luke C. M. Mackinder[a], Moritz T. Meyer[b], Tabea Mettler-Altmann[c,1], Vivian K. Chen[a,d], Madeline C. Mitchell[b,2], Oliver Caspari[b], Elizabeth S. Freeman Rosenzweig[a,d], Leif Pallesen[a], Gregory Reeves[a,3], Alan Itakura[a,d], Robyn Roth[e], Frederik Sommer[c,4], Stefan Geimer[f], Timo Mühlhaus[c,4], Michael Schroda[c,4], Ursula Goodenough[e], Mark Stitt[c], Howard Griffiths[b], and Martin C. Jonikas[a,d,5]

[a]Department of Plant Biology, Carnegie Institution for Science, Stanford, CA 94305; [b]Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, United Kingdom; [c]Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam-Golm, Germany; [d]Department of Biology, Stanford University, Stanford, CA 94305; [e]Department of Biology, Washington University in Saint Louis, St. Louis, MO 63130; and [f]Institute of Cell Biology, University of Bayreuth, 95440 Bayreuth, Germany

Edited by Paul G. Falkowski, Rutgers, The State University of New Jersey, New Brunswick, NJ, and approved April 7, 2016 (received for review November 20, 2015)

Biological carbon fixation is a key step in the global carbon cycle that regulates the atmosphere's composition while producing the food we eat and the fuels we burn. Approximately one-third of global carbon fixation occurs in an overlooked algal organelle called the pyrenoid. The pyrenoid contains the CO_2-fixing enzyme Rubisco and enhances carbon fixation by supplying Rubisco with a high concentration of CO_2. Since the discovery of the pyrenoid more that 130 y ago, the molecular structure and biogenesis of this ecologically fundamental organelle have remained enigmatic. Here we use the model green alga *Chlamydomonas reinhardtii* to discover that a low-complexity repeat protein, Essential Pyrenoid Component 1 (EPYC1), links Rubisco to form the pyrenoid. We find that EPYC1 is of comparable abundance to Rubisco and colocalizes with Rubisco throughout the pyrenoid. We show that EPYC1 is essential for normal pyrenoid size, number, morphology, Rubisco content, and efficient carbon fixation at low CO_2. We explain the central role of EPYC1 in pyrenoid biogenesis by the finding that EPYC1 binds Rubisco to form the pyrenoid matrix. We propose two models in which EPYC1's four repeats could produce the observed lattice arrangement of Rubisco in the *Chlamydomonas* pyrenoid. Our results suggest a surprisingly simple molecular mechanism for how Rubisco can be packaged to form the pyrenoid matrix, potentially explaining how Rubisco packaging into a pyrenoid could have evolved across a broad range of photosynthetic eukaryotes through convergent evolution. In addition, our findings represent a key step toward engineering a pyrenoid into crops to enhance their carbon fixation efficiency.

pyrenoid | Rubisco | carbon fixation | *Chlamydomonas reinhardtii* | CO_2-concentrating mechanism

Rubisco, the most abundant enzyme in the biosphere (1), fixes CO_2 into organic carbon that supports nearly all life on Earth (2, 3). Over the past 3 billion y, the enzyme became a victim of its own success as it drew down the atmospheric CO_2 concentration to trace levels (4) and as the oxygen-producing reactions of photosynthesis filled our atmosphere with O_2 (4). In today's atmosphere, O_2 competes with CO_2 at Rubisco's catalytic site, producing the toxic compound phosphoglycolate (5). Phosphoglycolate must be metabolized at the expense of energy and loss of fixed carbon and nitrogen (6). To overcome Rubisco's limitations, many photosynthetic organisms have evolved carbon-concentrating mechanisms (CCMs) (7, 8). CCMs increase the CO_2 concentration around Rubisco, decreasing O_2 competition and enhancing carbon fixation.

At the heart of the CCM of eukaryotic algae is an organelle known as the pyrenoid (9). The pyrenoid is a spherical structure in the chloroplast stroma, discovered more than 130 y ago (10–12). Pyrenoids have been found in nearly all of the major oceanic eukaryotic primary producers and mediate ~28–44% of global carbon fixation (*SI Appendix*, Table S1) (3, 13–17). A pyrenoid typically consists of a matrix surrounded by a starch sheath and traversed by membrane tubules continuous with the photosynthetic thylakoid membranes (18). This matrix is thought to consist primarily of tightly packed Rubisco and its chaperone, Rubisco activase (19). In higher plants and non–pyrenoid-containing photosynthetic eukaryotes, Rubisco is instead soluble throughout the chloroplast stroma. The molecular mechanism by which Rubisco aggregates to form the pyrenoid matrix remains enigmatic.

Two mechanisms for Rubisco accumulation in the pyrenoid have been proposed: (*i*) Rubisco holoenzymes could bind each other directly through hydrophobic residues (20), or (*ii*) a linker protein may link Rubisco holoenzymes together (18, 20). The second model is based on analogy to the well-characterized prokaryotic carbon concentrating organelle, the β-carboxysome, where Rubisco aggregation is mediated by a linker protein consisting of repeats of a

Significance

Eukaryotic algae, which play a fundamental role in global CO_2 fixation, enhance the performance of the carbon-fixing enzyme Rubisco by placing it into an organelle called the pyrenoid. Despite the ubiquitous presence and biogeochemical importance of this organelle, how Rubisco assembles to form the pyrenoid remains a long-standing mystery. Our discovery of an abundant repeat protein that binds Rubisco in the pyrenoid represents a critical advance in our understanding of pyrenoid biogenesis. The repeat sequence of this protein suggests elegant models to explain the structural arrangement of Rubisco enzymes in the pyrenoid. Beyond advances in basic understanding, our findings open doors to the engineering of algal pyrenoids into crops to enhance yields.

Author contributions: L.C.M.M., M.T.M., T.M.-A., M. Schroda, M. Stitt, H.G., and M.C.J. designed research; L.C.M.M., M.T.M., T.M.-A., V.K.C., M.C.M., O.C., E.S.F.R., L.P., G.R., A.I., R.R., F.S., S.G., and T.M. performed research; L.C.M.M., M.T.M., T.M.-A., L.P., M. Schroda, and U.G. contributed new reagents/analytic tools; L.C.M.M., M.T.M., T.M.-A., M.C.M., F.S., S.G., T.M., and M.C.J. analyzed data; and L.C.M.M., M.T.M., T.M.-A., U.G., M. Stitt, H.G., and M.C.J. wrote the paper.

Conflict of interest statement: The authors wish to note that the Carnegie Institution for Science has submitted a provisional patent application on aspects of the findings.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]Present address: Cluster of Excellence in Plant Sciences and Institute of Plant Biochemistry, Heinrich-Heine University, 40225 Düsseldorf, Germany.

[2]Present address: Agriculture, Commonwealth Scientific and Industrial Research Organization, Canberra, ACT 2601, Australia.

[3]Present address: Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, United Kingdom.

[4]Present address: Institute of Molecular Biotechnology and Systems Biology, Technical University of Kaiserslautern, 67663 Kaiserslautern, Germany.

[5]To whom correspondence should be addressed. Email: mjonikas@carnegiescience.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1522866113/-/DCSupplemental.

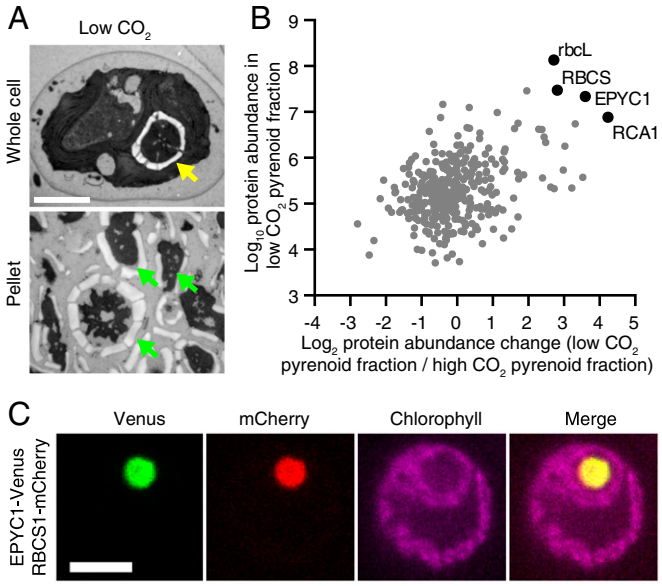

Fig. 1. EPYC1 is an abundant pyrenoid protein. (*A*) TEM images of *Chlamydomonas* whole cells and pyrenoid-enriched pellet fraction from cells grown at low CO_2. The yellow arrow indicates the pyrenoid, and green arrows indicate pyrenoid-like structures. (Scale bar: 2 μm.) (*B*) Mass spectrometry analysis of 366 proteins in pyrenoid-enriched pellet fractions from low- and high-CO_2–grown cells (mean of four biological replicates; raw data are provided in *SI Appendix* and Dataset S1). RbcL, RBCS, EPYC1, and RCA1 (black) are abundant in low-CO_2 pellets, as determined by iBAQ (*y*-axis). In addition, these proteins showed increased abundance in low-CO_2 pellets compared with high-CO_2 pellets, as determined by label-free quantification (LFQ; *x*-axis). (*C*) Confocal microscopy of EPYC1-Venus and RBCS1-mCherry coexpressed in WT cells. (Scale bar: 5 μm.)

domain resembling the Rubisco small subunit (21). Here we find that Rubisco accumulation in the pyrenoid of the model alga *Chlamydomonas reinhardtii* is mediated by a disordered repeat protein, which we term Essential Pyrenoid Component 1 (EPYC1). Our findings suggest a mechanism for aggregation of Rubisco in the pyrenoid matrix, and highlight similarities and differences between the mechanisms of assembly of the eukaryotic and prokaryotic organelles.

Results

EPYC1 Is an Abundant Pyrenoid Component. We hypothesized that the pyrenoid contains unidentified components that are important for its biogenesis. Therefore, we used mass spectrometry to analyze the protein composition of the pyrenoid of *Chlamydomonas*, before and after applying a stimulus that induces pyrenoid growth. When cells are transferred from high CO_2 (2–5% CO_2 in air) to low CO_2 (~0.04% CO_2 in air), the CCM is induced (22) and the pyrenoid increases in size (23). We developed a protocol for isolating largely intact pyrenoids by cell lysis and centrifugation, and applied this protocol to cells before and after a shift from high to low CO_2 (Fig. 1*A* and *SI Appendix*, Fig. S1 *A–C*). Mass spectrometry indicated that the most abundant proteins in the low-CO_2 pyrenoid fraction included the Rubisco large (rbcL) and small (RBCS) subunits, as well as Rubisco activase (RCA1) (Fig. 1*B*, *SI Appendix*, Fig. S1*D*, and Dataset S1).

Strikingly, a fourth protein, previously identified as a low-CO_2–induced nuclear-encoded protein (LCI5; Cre10.g436550) (24), was found in the low-CO_2 pyrenoid fraction with comparable abundance to Rubisco (Fig. 1*B*). Based on the data presented herein, we propose naming this protein Essential Pyrenoid Component 1 (EPYC1). Under low CO_2, the stoichiometry of EPYC1, estimated by intensity-based absolute quantification (iBAQ), was ~1:6 with rbcL and ~1:1 with RBCS (25). Consistent with EPYC1 being a component of the pyrenoid, the abundance of EPYC1 in the pyrenoid fraction was increased by ~12-fold after the shift from high to low CO_2 (Fig. 1*B* and *SI Appendix*, Fig. S1*D* and Dataset S1), an increase similar to that of rbcL (7-fold), RBCS (7-fold), and RCA1 (19-fold). To confirm the pyrenoid localization of EPYC1, we coexpressed fluorescently tagged EPYC1 and RBCS. Venus-tagged EPYC1 showed clear colocalization with mCherry-tagged RBCS in the pyrenoid (Fig. 1*C* and *SI Appendix*, Fig. S1*E*).

EPYC1 Is Essential for a Functional CCM. The high abundance of EPYC1 in the pyrenoid led us to ask whether EPYC1 is required for the CCM. We isolated a mutant in the 5′ UTR of the *EPYC1* gene (*SI Appendix*, Fig. S2*A* and Table S2), which contains markedly reduced levels of *EPYC1* mRNA (*SI Appendix*, Fig. S2*B* and Table S3) and EPYC1 protein (Fig. 2*A*), and lacks transcriptional regulation in response to CO_2 (*SI Appendix*, Fig. S2*B*). Similar to previously described mutants in other components of the CCM, the *epyc1* mutant showed defective photoautotrophic growth in low CO_2, which was rescued by high CO_2 and by reintroducing the *EPYC1* gene (Fig. 2*B* and *SI Appendix*, Fig. S2 *C–E*).

We further tested the CCM activity in the *epyc1* mutant by measuring whole-cell affinity for inorganic carbon, inferred from photosynthetic O_2 evolution. When grown under low CO_2, the *epyc1* mutant showed a reduced affinity for inorganic carbon (increased $K_{0.5}$) relative to WT ($P = 0.0055$, Student's *t* test; $n = 5$) (Fig. 2*C* and *SI Appendix*, Fig. S2*F* and Table S4). The affinity of the *epyc1* mutant under low CO_2 was slightly greater than that of WT at high CO_2, indicating a residual level of CCM activity. This activity may be due to trace levels of EPYC1 in the *epyc1* mutant (*SI Appendix*, Fig. S2 *A* and *B*), or a normal CO_2 concentration followed by inefficient capture by Rubisco.

EPYC1 Is Required for Normal Pyrenoid Size, Number, and Matrix Density. Given that EPYC1 is in the pyrenoid and is required for the CCM, we explored whether the *epyc1* mutant shows any visible defects in pyrenoid structure. Thin-section transmission electron microscopy (TEM) revealed that the *epyc1* mutant had smaller pyrenoids than WT at both low and high CO_2 (low CO_2: $n = 37–79$, $P < 10^{-19}$, Welch's *t* test; high CO_2: $n = 18–22$, $P < 10^{-5}$,

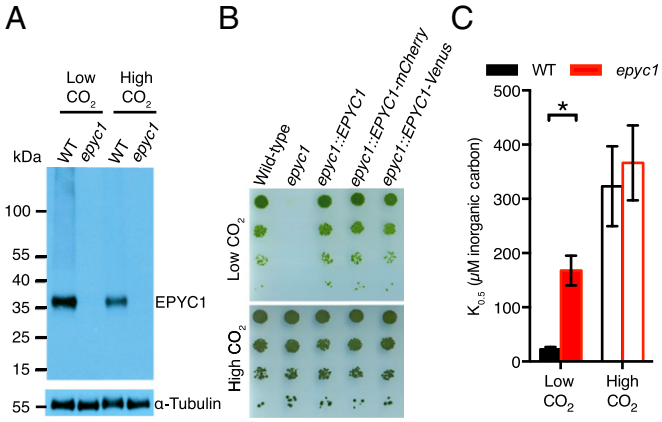

Fig. 2. EPYC1 is an essential component of the carbon-concentrating mechanism. (*A*) EPYC1 protein levels in WT and *epyc1* mutant cells grown at low and high CO_2 were probed by Western blot analysis with anti-EPYC1 antibodies. Anti-tubulin is shown as a loading control. (*B*) Growth phenotypes of WT, *epyc1*, and three lines complemented with *EPYC1*. Serial 1:10 dilutions of WT, *epyc1*, *epyc1::EPYC1*, *epyc1::EPYC1-mCherry*, and *epyc1::EPYC1-Venus* lines were spotted on TP minimal medium and grown at low and high CO_2 under 500 μmol photons m^{-2} s^{-1} illumination. (*C*) Inorganic carbon affinity of WT and *epyc1* cells. Cells were pregrown at low and high CO_2, and whole-cell inorganic carbon affinity was measured as the concentration of inorganic carbon at half-maximal O_2 evolution. Data are a mean of five low-CO_2 or three high-CO_2 biological replicates. Error bars represent SEM. *$P = 0.0055$, Student's *t* test.

Welch's t test) (Fig. 3 *A* and *B* and *SI Appendix*, Fig. S3 *A* and *B* and Fig. S4). *Chlamydomonas* typically has one pyrenoid per cell (17). The *epyc1* mutant showed a higher frequency of multiple pyrenoids; 13% of nondividing *epyc1* cells ($n = 231$) showed multiple pyrenoids, compared with 3% of WT cells ($n = 252$; $P = 0.00048$, Fisher's exact test of independence) (*SI Appendix*, Table S5). Higher-resolution quick-freeze deep-etch electron microscopy (QFDEEM) indicated a lower packing density of granular material in the pyrenoid matrix of the *epyc1* mutant (Fig. 3*C* and *SI Appendix*, Figs. S3*C* and S5). This defect was most noticeable when cells were grown in low CO_2, but was also visible at high CO_2.

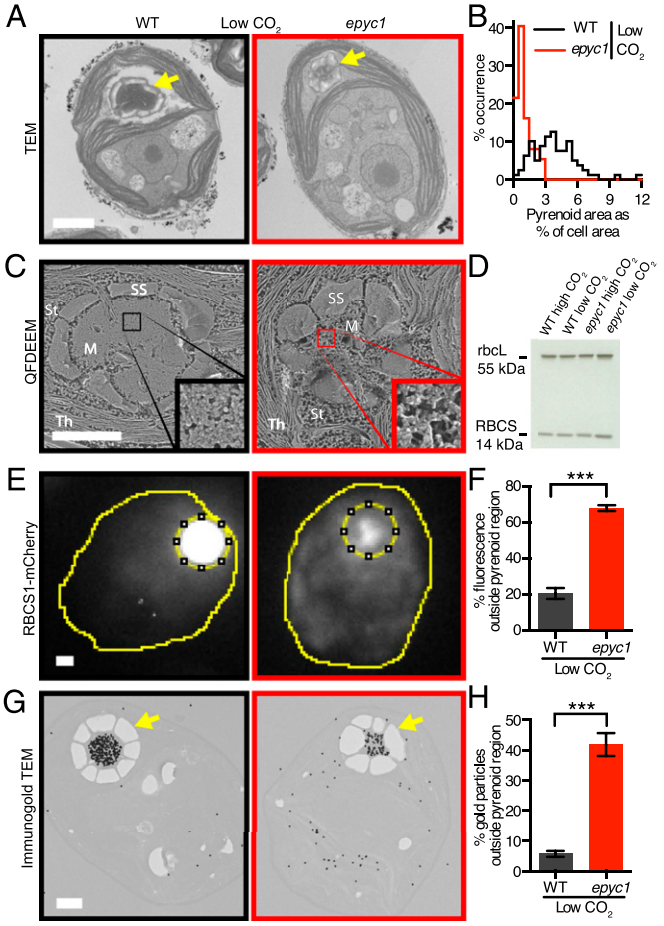

Fig. 3. EPYC1 is essential for Rubisco aggregation in the pyrenoid. (*A*) Representative TEMs of WT and *epyc1* cells grown at low CO_2. (*B*) Quantification of the pyrenoid area as a percentage of cell area of WT and *epyc1* cells grown at low CO_2. Data are from TEM images as represented in *A*. *epyc1*: $n = 37$; WT: $n = 79$. $P < 10^{-19}$, Welch's t test. (*C*) QFDEEM of the pyrenoid of WT and *epyc1* cells grown at low CO_2. M, pyrenoid matrix; St, stroma; Th, thylakoids; SS, starch sheath. (*Inset*) Four hundred percent zoom view of the pyrenoid matrix. (*D*) Rubisco protein levels in WT and *epyc1* cells grown at low and high CO_2 were probed by Western blot analysis. (*E*) Localization of Rubisco was determined by microscopy of WT and the *epyc1* mutant containing RBCS1-mCherry. The sum of the fluorescence signals from Z stacks was used for quantitation. (*F*) The fraction of RBCS1-mCherry signal from outside the pyrenoid region (inner dotted line in *E*) was quantified in WT and *epyc1* cells at low CO_2. *epyc1*: $n = 27$; WT: $n = 27$. ***$P < 10^{-15}$, Student's t test. (*G*) Representative images of anti-Rubisco immunogold labeling of WT and *epyc1* cells grown at low CO_2. Gold particles were enlarged 10× for visibility. (*H*) The fraction of immunogold particles outside the pyrenoid was quantified. WT: $n = 26$ cells, 8,123 gold particles; *epyc1*: $n = 27$ cells, 2,708 gold particles. ***$P < 10^{-15}$, Student's t test. In *F* and *H*, data are mean values, with error bars indicating SEM. Yellow arrows indicate pyrenoids. (Scale bars: 1 μm.)

Interestingly, the *epyc1* mutant retains a number of canonical pyrenoid characteristics (17), including correct localization in the chloroplast, the presence of a starch sheath under low CO_2, and traversing membrane tubules, suggesting that normal levels of EPYC1 are not required for these characteristics. In addition, the *epyc1* mutant showed normal levels of the carbonic anhydrase CAH3, which is thought to be central in delivering CO_2 to Rubisco in the pyrenoid (*SI Appendix*, Fig. S2*G*).

EPYC1 Is Required for Rubisco Assembly into the Pyrenoid. Our observations of decreased pyrenoid size and apparent matrix density in the *epyc1* mutant could be explained by decreased whole-cell levels of Rubisco. However, Western blot analysis revealed no detectable difference in rbcL and RBCS abundance in *epyc1* relative to WT cells or between cells grown at low and high CO_2 levels (Fig. 3*D* and *SI Appendix*, Fig. S3*D*). This result led us to hypothesize that the localization of Rubisco was perturbed in the *epyc1* mutant. To test this hypothesis, we generated WT and *epyc1* cell lines expressing Rubisco tagged with mCherry, and determined the distribution of fluorescence signal by microscopy. Remarkably, a large fraction of Rubisco was found outside the pyrenoid in the *epyc1* mutant. In *epyc1* cells grown in low CO_2, 68% of fluorescence from Rubisco tagged with mCherry was found outside the pyrenoid region, compared with 21% in WT cells ($n = 27$; $P < 10^{-15}$, Student's t test) (Fig. 3 *E* and *F* and *SI Appendix*, Fig. S6). Immunogold-EM confirmed the mislocalization of Rubisco in *epyc1*. In pyrenoid-containing sections of low-CO_2–grown *epyc1* cells, 42% of anti-Rubisco immunogold particles were found outside the pyrenoid, whereas only 6% were found outside the pyrenoid in WT (WT: $n = 26$ cells, 8,123 gold particles; *epyc1*: $n = 27$ cells, 2,708 gold particles; $P < 10^{-15}$, Student's t test) (Fig. 3 *G* and *H* and *SI Appendix*, Fig. S7).

If EPYC1 functions in the recruitment of Rubisco to the pyrenoid solely at low CO_2 (23), then the *epyc1* mutant could be trapped in a "high-CO_2" state of Rubisco localization (23). However, the *epyc1* mutant showed a defect in Rubisco localization even under high CO_2 (*SI Appendix*, Fig. S3 *E* and *F* and Fig. S6), where the fraction of Rubisco-mCherry fluorescence outside the pyrenoid region increased to 80% in *epyc1*, compared with 68% in the WT (WT: $n = 20$; *epyc1*: $n = 20$; $P = 10^{-6}$, Student's t test). We conclude that EPYC1 is required for Rubisco localization to the pyrenoid not only at low CO_2, but also at high CO_2.

EPYC1 and Rubisco Are Part of the Same Complex. EPYC1 could promote the localization of Rubisco to the pyrenoid by a physical interaction. Thus, we immunoprecipitated EPYC1 and Rubisco, and probed the eluates by Western blot analysis (Fig. 4*A* and *SI Appendix*, Fig. S8*A*). Immunoprecipitation of tagged EPYC1 pulled down the Rubisco holoenzyme, and, reciprocally, tagged RBCS1 pulled down EPYC1. We conclude that EPYC1 and Rubisco are part of the same supramolecular complex in the pyrenoid. The high abundance of EPYC1 in the pyrenoid, EPYC1's physical interaction with Rubisco, and the dependence of Rubisco on EPYC1 for localization to the pyrenoid all suggest that EPYC1 plays a structural role in pyrenoid biogenesis.

The EPYC1 Protein Consists of Four Nearly Identical Repeats. To gain insight into how EPYC1 might contribute to pyrenoid biogenesis, we performed a detailed analysis of the EPYC1 protein sequence. This analysis indicated that EPYC1 consists of four nearly identical ~60-aa repeats (Fig. 4 *B*–*D*), flanked by short N and C termini, in contrast to a previous study suggesting only three repeats (26). We found that each repeat consists of a predicted disordered domain and a shorter, less disordered domain containing a predicted alpha helix (Fig. 4*C* and *SI Appendix*, Fig. S8 *B* and *C*). Given that these repeats cover >80% of the EPYC1 protein, it is likely that the Rubisco binding sites are contained within the repeats.

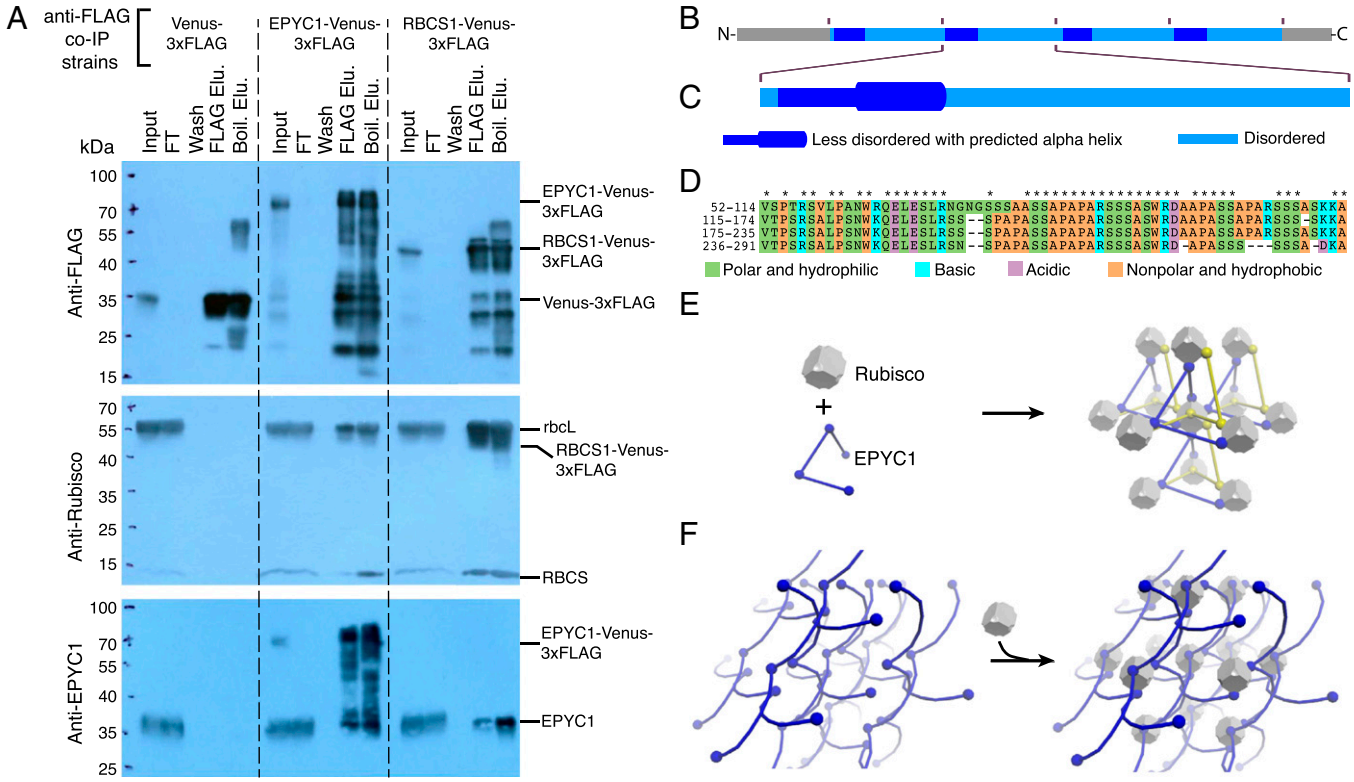

Fig. 4. EPYC1 forms a complex with Rubisco. (*A*) Anti-FLAG coimmunoprecipitation (co-IP) of WT cells expressing Venus-3×FLAG, EPYC1-Venus-3×FLAG, and RBCS1-Venus-3×FLAG. For each co-IP, the input, flow-through (FT), fourth wash (wash), 3×FLAG elution (FLAG Elu.), and boiling elution (Boil. Elu.) were probed with anti-FLAG, anti-Rubisco, or anti-EPYC1. Labels on the right show the expected sizes of proteins. (*B*) Analysis of the EPYC1 protein sequence showing that EPYC1 consists of four nearly identical repeats. (*C*) Each repeat has a highly disordered domain (light blue) and a less-disordered domain (dark blue) containing a predicted alpha-helix (thicker line) rich in charged residues. (*D*) Amino acid alignments of the four repeats. Asterisks indicate residues that are identical in all four repeats. (*E* and *F*) Two models illustrate how EPYC1 could bind the Rubisco holoenzyme in a manner compatible with the observed packing of Rubisco in the pyrenoid. (*E*) EPYC1 and Rubisco could form a codependent network. If each EPYC1 can bind four Rubisco holoenzymes, and each Rubisco holoenzyme can bind eight EPYC1s, eight EPYC1 proteins could connect each Rubisco to twelve neighboring Rubiscos. (*F*) EPYC1 could form a scaffold onto which Rubisco binds. Both arrangements could expand indefinitely in every direction. For clarity, the spacing between Rubisco holoenzymes was increased and EPYC1 is depicted in both yellow and blue.

We Propose Two Models for Rubisco Assembly into the Pyrenoid Matrix by EPYC1. If each repeat of EPYC1 binds Rubisco, then EPYC1 could link multiple Rubisco holoenzymes together to form the pyrenoid matrix. Multiple Rubisco binding sites on EPYC1 could arrange Rubisco into the hexagonal closely packed or cubic closely packed arrangement observed in recent cryoelectron tomography studies of the *Chlamydomonas* pyrenoid (18). EPYC1 and Rubisco could interact in one of two fundamental ways: (*i*) EPYC1 and Rubisco could form a codependent network (Fig. 4*E*), or (*ii*) EPYC1 could form a scaffold onto which Rubisco binds (Fig. 4*F*). Importantly, the 60-aa repeat length of EPYC1 is sufficient to span the observed 2- to 4.5-nm gap between Rubisco holoenzymes in the pyrenoid (18), and a stretched-out repeat could potentially span the observed 15-nm Rubisco center-to-center distance. A promising candidate for an EPYC1-binding site on Rubisco would be the two alpha-helices of the small Rubisco subunit. When these helices are exchanged for higher-plant alpha-helices, pyrenoids fail to form and the CCM does not function, but holoenzyme assembly is normal (20).

Proteins with Similar Physicochemical Properties to EPYC1 Are Present in a Diverse Range of Eukaryotic Algae. The primary sequences of disordered proteins like EPYC1 are known to evolve rapidly compared with those of structured proteins, but their physicochemical properties are under selective pressure and are evolutionarily maintained (27). Therefore, we searched for proteins with similar physicochemical properties (i.e., repeat number, length, high isoelectric point, disorder profile, and absence of transmembrane domains) across a broad range of algae (*SI Appendix*, Table S6). Excitingly, proteins with similar properties are found in most pyrenoid-containing algae and appear to be absent from pyrenoid-less algae, suggesting that EPYC1-like proteins may play similar roles in pyrenoids across eukaryotic algae. A thorough assessment of the generality of linker proteins will be enabled by future proteomic analyses of pyrenoids from a diverse range of algae.

Discussion

Our data provide strong support for the concept that Rubisco clustering into the pyrenoid is required for an efficient CCM in eukaryotic algae (9). Current models of the CCM (17, 28) suggest that CO_2 is released at a high concentration from the thylakoid tubules traversing the pyrenoid matrix. The mislocalization of Rubisco to the stroma of the *epyc1* mutant could decrease the efficiency of CO_2 capture by Rubisco, explaining the severe CCM defect observed in this mutant.

The observations presented here suggest that Rubisco packaging to form the matrix of the eukaryotic pyrenoid is achieved by a different mechanism than that used in the well-characterized prokaryotic β-carboxysome. In the β-carboxysome, aggregation of Rubisco is mediated by the protein CcmM (CO_2 concentrating mechanism protein M). CcmM contains multiple repeats of a domain resembling the Rubisco small subunit, and incorporation of these domains into separate Rubisco holoenzymes is thought to produce a link between Rubisco holoenzymes (21). Given that the

EPYC1 repeats show no homology to Rubisco and are highly disordered, it is likely that they bind to the surface of Rubisco holoenzymes rather than becoming incorporated in the place of small subunits. The simplicity of such a surface-binding mechanism potentially explains how Rubisco packaging into a pyrenoid could have evolved across a broad range of photosynthetic eukaryotes through convergent evolution (17, 29), leading to the dominant role of pyrenoids in aquatic CO_2 fixation. Such a surface-binding mechanism may even organize Rubisco in prokaryotic α-carboxysomes, where the intrinsically disordered Rubisco-binding repeat protein CsoS2 plays a poorly understood role in assembly (30).

In addition to being a key structural component, EPYC1 could regulate Rubisco partitioning to the pyrenoid or Rubisco kinetic properties. The Rubisco content of the pyrenoid changes in response to CO_2 (23 and our data), whereas total cellular Rubisco remains constant (Fig. 3*D*). Given that EPYC1 is required for Rubisco localization to the pyrenoid, changes in EPYC1 abundance and/or Rubisco-binding affinity could affect Rubisco partitioning to the pyrenoid. Consistent with this hypothesis, EPYC1 was previously found to be up-regulated at both the transcript and protein levels in response to light and low CO_2 (26), and our data further support this finding (Fig. 2*A* and *SI Appendix*, Fig. S2*A*). Moreover, previous studies have shown that EPYC1 becomes phosphorylated at multiple sites in response to low CO_2 (26, 31), potentially affecting its binding affinity for Rubisco.

Another mode of regulation of EPYC1–Rubisco binding could be through the methylation of Rubisco. Rubisco is methylated in multiple residues (32), and in *Chlamydomonas*, the predicted methyltransferase CIA6 is required for Rubisco localization to the pyrenoid (33). It is also possible that EPYC1 binding to Rubisco alters the kinetic properties of Rubisco to fine-tune its performance in the pyrenoid.

Along with advancing our understanding of the molecular mechanisms underlying global carbon fixation, our findings may contribute to the future engineering of crops with enhanced photosynthesis. There is great interest in introducing a CCM into C_3 plants, given that this enhancement is predicted to increase yields by up to 60% and to improve the efficiency of nitrogen and water use (34). Although much remains to be done to improve our understanding of the algal CCM, recent work suggests that algal components may be relatively easy to engineer into higher plants (35). Our discovery of a possible mechanism for Rubisco assembly to form the pyrenoid is a key step toward engineering an algal CCM into crops.

Materials and Methods

Strains and Culture Conditions. WT *Chlamydomonas* CC-1690 (36) was used for pyrenoid enrichment and proteomics. WT *Chlamydomonas* cMJ030 (CC-4533) (37) was used for all other experiments. The *epyc1* mutant was isolated from a collection of high-CO_2–requiring mutants generated by transformation of the pMJ016c mutagenesis cassette into cMJ030 (37). All experiments were performed under photoautotrophic conditions supplemented with high CO_2 (3% or 5% vol/vol CO_2-enriched air) or low CO_2 (air, ∼0.04% vol/vol CO_2).

Proteomics. Pyrenoid enrichment was performed using a modified protocol based on previous studies (38, 39). In brief, cells were harvested by centrifugation (3,220 × *g* for 2 min at 4 °C), lysed by sonication, and then centrifuged again at 500 × *g* for 3 min at 4 °C to obtain a soluble fraction and a pellet fraction. Shotgun proteomics on the soluble and pellet fractions was performed as described by Mühlhaus et al. (40). Raw MS data files were processed with MaxQuant version 1.5.2.8 (41).

Cloning. *EPYC1* (Cre10.g436550) and *RBCS1* (Cre02.g120100) ORFs were amplified from gDNA and cloned into pLM005 (Venus) or pLM006 (mCherry) by Gibson assembly (42).

Transformation of *Chlamydomonas*. Constructs were transformed into the nuclear genome of WT and *epyc1* strains by electroporation as described by Zhang et al. (37). To screen for Venus- and mCherry-expressing colonies, transformation plates were imaged with a Typhoon Trio fluorescence scanner (GE Healthcare).

Microscopy. TEM images of the enriched pyrenoid fraction and whole cells before pyrenoid enrichment were prepared and taken according to Nordhues et al. (43). TEM imaging for pyrenoid area analysis and immunogold localization of Rubisco was based on methods described by Meyer et al. (20). QFDEEM was performed as described by Heuser (44). Fluorescence microscopy was performed using a spinning-disk confocal microscope (Leica DMI6000) with the following settings: Venus, 514 nm excitation with 543/22 nm emission; mCherry, 561 nm excitation with 590/20 nm emission; and chlorophyll, 561 nm excitation with 685/40 nm emission.

Quantitative Real-Time PCR. EPYC1 gene transcript levels were determined by qRT-PCR. CDNA was synthesized from total RNA, and relative gene expression was measured in real time using SYBR Green. Gene expression was calculated according to the method of Livak and Schmittgen (45), relative to RCK1 (Cre06.g278222) (46). The primers used are listed in *SI Appendix*, Table S2.

Western Blot Analysis. Protein levels of EPYC1 and CAH3 in WT and the *epyc1* mutant were measured according to Heinnickel et al. (47). Rubisco levels were measured as described by Meyer et al. (20).

O_2 Evolution and Spot Tests. Apparent affinity for inorganic carbon was determined using the oxygen evolution method described by Badger et al. (48). Spot tests were performed by spotting serially diluted WT, *epyc1*, and complemented cell lines onto Tris-phosphate (TP) plates. Plates were incubated in low or high CO_2 under 100 or 500 μmol photons m^{-2} s^{-1} of light for 7 d.

Coimmunoprecipitation. Cell lysate from WT cells expressing the bait proteins (Venus-3×FLAG, EPYC1-Venus-3×FLAG, or RbcS1-Venus-3×FLAG) was incubated with anti-FLAG M2 antibody (Sigma-Aldrich) bound to protein G Dynabeads (Life Technologies). Bait proteins with interaction partners were eluted by 3×FLAG competition, followed by boiling in 1× Laemmli buffer.

EPYC1 Sequence Analysis. To understand the intrinsic disorder of EPYC1, the full-length amino acid sequence was run through several structural disorder prediction programs, including VL3, VLTX (49), and GlobPlot 2 (50). To look for regions of secondary structure, the full-length and repeat regions of the EPYC1 amino acid sequence were analyzed by PSIPRED v3.3 (51) and Phyre2 (52).

Proteins with EPYC1-Like Physicochemical Properties in Other Algae. Complete translated genomic sequences from pyrenoid and non-pyrenoid algae were analyzed for tandem repeats using Xstream (53). The isoelectric point, disorder profile (54), and presence of transmembrane domains (55) of Xstream hits were calculated. Proteins with three or more repeats, a pI >8, an oscillating disorder profile with a frequency between 40 and 80 amino acids, and no transmembrane domains were classified as potential EPYC1-like Rubisco linker proteins.

More detailed information on the materials and methods used in this study is provided in *SI Appendix, SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank J. N. Skepper, L. Carter, and M. Rütgers for TEM support, discussions on immunogold optimization, and ultramicrotomy; H. Cartwright for confocal microscopy support; S. Ramundo for technical advice with coimmunoprecipitation; W. Patena for help with data analysis; and W. Frommer, V. Walbot, P. Walter, and T. Cuellar for comments on the manuscript. The project was funded by National Science Foundation Grants EF-1105617 and IOS-1359682 (to L.C.M.M., L.P., G.R., and M.C.J.); the Carnegie Institution for Science (L.C.M.M. and M.C.J.); National Institutes of Health Grant T32GM007276 (to E.S.F.R., V.K.C., and A.I.); Biotechnology and Biological Sciences Research Council Grant BB/M007693/1 (to M.T.M. and H.G.); the Federal Ministry of Education and Research, Germany, within the frame of the *Go*FORSYS Research Unit for Systems Biology (Grant FKZ 0313924, to T.M.-A., F.S., M. Schroda, and M. Stitt); and the International Max Planck Research School of the Max Planck Society (T.M.-A. and T.M.).

1. Ellis RJ (1979) The most abundant protein in the world. *Trends Biochem Sci* 4(11): 241–244.
2. Falkowski PG, Barber RT, Smetacek V (1998) Biogeochemical controls and feedbacks on ocean primary production. *Science* 281(5374):200–207.
3. Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* 281(5374):237–240.
4. Dismukes GC, et al. (2001) The origin of atmospheric oxygen on Earth: The innovation of oxygenic photosynthesis. *Proc Natl Acad Sci USA* 98(5):2170–2175.
5. Somerville CR, Ogren WL (1979) A phosphoglycolate phosphatase-deficient mutant of *Arabidopsis*. *Nature* 280(5725):833–836.
6. Bauwe H, Hagemann M, Fernie AR (2010) Photorespiration: Players, partners and origin. *Trends Plant Sci* 15(6):330–336.

7. Sage RF, Sage TL, Kocacinar F (2012) Photorespiration and the evolution of C4 photosynthesis. *Annu Rev Plant Biol* 63:19–47.
8. Giordano M, Beardall J, Raven JA (2005) CO_2 concentrating mechanisms in algae: Mechanisms, environmental modulation, and evolution. *Annu Rev Plant Biol* 56: 99–131.
9. Badger MR, et al. (1998) The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae. *Can J Bot* 76(6): 1052–1071.
10. Schmitz F (1882) *Die Chromatophoren der Algen: Vergleichende Untersuchungen über Bau und Entwicklung der Chlorophyllkörper und der analogen Farbstoffkörper der Algen* (M. Cohen & Sohn, Bonn, Germany).
11. Vaucher J-P (1803) *Histoire des Conferves D'eau Douce: Contenant Leurs Différents Modes De Reproduction, Et La Description De Leurs Principales Espèces, Suivie De L'histoire Des Trémelles Et Des Ulves D'eau Douce* (JJ Paschoud, Geneva, Switzerland).
12. Brown R (1967) Pyrenoid: Its structure distribution and function. *J Phycol* 3(Suppl 1): 5–7.
13. Behrenfeld MJ, et al. (2001) Biospheric primary production during an ENSO transition. *Science* 291(5513):2594–2597.
14. Rousseaux CS, Gregg WW (2013) Interannual variation in phytoplankton primary production at a global scale. *Remote Sens* 6(1):1–19.
15. Mann GD (1996) Chloroplast morphology, movements and inheritance in diatoms. *Cytology, Genetics and Molecular Biology of Algae*, eds Chaudhary BR, Agrawal SB (SPB Academic Publishing, Amsterdam), pp 249–274.
16. Thierstein HR, Young JR, eds (2004) *Coccolithophores: From Molecular Processes to Global Impact* (Springer, Heidelberg, Germany).
17. Meyer M, Griffiths H (2013) Origins and diversity of eukaryotic CO_2-concentrating mechanisms: Lessons for the future. *J Exp Bot* 64(3):769–786.
18. Engel BD, et al. (2015) Native architecture of the *Chlamydomonas* chloroplast revealed by in situ cryo-electron tomography. *eLife* 4:e04889.
19. McKay RML, Gibbs SP (1991) Composition and function of pyrenoids: Cytochemical and immunocytochemical approaches. *Can J Bot* 69(5):1040–1052.
20. Meyer MT, et al. (2012) Rubisco small-subunit α-helices control pyrenoid formation in *Chlamydomonas*. *Proc Natl Acad Sci USA* 109(47):19474–19479.
21. Long BM, Badger MR, Whitney SM, Price GD (2007) Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. *J Biol Chem* 282(40):29323–29335.
22. Berry J, Boynton J, Kaplan A, Badger M (1976) Growth and photosynthesis of *Chlamydomonas reinhardtii* as a function of CO_2 concentration. *Carnegie Institution of Washington Year Book* (Carnegie Institution of Washington, Washington, DC), Vol 75, pp 423–432.
23. Borkhsenious ON, Mason CB, Moroney JV (1998) The intracellular localization of ribulose-1,5-bisphosphate carboxylase/oxygenase in *Chlamydomonas reinhardtii*. *Plant Physiol* 116(4):1585–1591.
24. Miura K, et al. (2004) Expression profiling-based identification of CO_2-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant Physiol* 135(3):1595–1607.
25. Schwanhäusser B, et al. (2011) Global quantification of mammalian gene expression control. *Nature* 473(7347):337–342.
26. Turkina MV, Blanco-Rivero A, Vainonen JP, Vener AV, Villarejo A (2006) CO_2 limitation induces specific redox-dependent protein phosphorylation in *Chlamydomonas reinhardtii*. *Proteomics* 6(9):2693–2704.
27. Dunker AK, Silman I, Uversky VN, Sussman JL (2008) Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* 18(6):756–764.
28. Wang Y, Stessman DJ, Spalding MH (2015) The CO_2-concentrating mechanism and photosynthetic carbon assimilation in limiting CO_2: How *Chlamydomonas* works against the gradient. *Plant J* 82(3):429–448.
29. Villarreal JC, Renner SS (2012) Hornwort pyrenoids, carbon-concentrating structures, evolved and were lost at least five times during the last 100 million years. *Proc Natl Acad Sci USA* 109(46):18873–18878.
30. Cai F, et al. (2015) Advances in understanding carboxysome assembly in *Prochlorococcus* and *Synechococcus* implicate CsoS2 as a critical component. *Life (Basel)* 5(2):1141–1171.
31. Wang H, et al. (2014) The global phosphoproteome of *Chlamydomonas reinhardtii* reveals complex organellar phosphorylation in the flagella and thylakoid membrane. *Mol Cell Proteomics* 13(9):2337–2353.
32. Taylor TC, Backlund A, Bjorhall K, Spreitzer RJ, Andersson I (2001) First crystal structure of Rubisco from a green alga, *Chlamydomonas reinhardtii*. *J Biol Chem* 276(51): 48159–48164.
33. Ma Y, Pollock SV, Xiao Y, Cunnusamy K, Moroney JV (2011) Identification of a novel gene, CIA6, required for normal pyrenoid formation in *Chlamydomonas reinhardtii*. *Plant Physiol* 156(2):884–896.
34. Long SP, Marshall-Colon A, Zhu XG (2015) Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* 161(1):56–66.
35. Atkinson N, et al. (2015) Introducing an algal carbon-concentrating mechanism into higher plants: Location and incorporation of key components. *Plant Biotechnol J*, 10.1111/pbi.12497.
36. Sager R (1955) Inheritance in the green alga *Chlamydomonas reinhardi*. *Genetics* 40(4):476–489.
37. Zhang R, et al. (2014) High-throughput genotyping of green algal mutants reveals random distribution of mutagenic insertion sites and endonucleolytic cleavage of transforming DNA. *Plant Cell* 26(4):1398–1409.
38. Kuchitsu K, Tsuzuki M, Miyachi S (1988) Changes of starch localization within the chloroplast induced by changes in CO_2 concentration during growth of *Chlamydomonas reinhardtii*: Independent regulation of pyrenoid starch and stroma starch. *Plant Cell Physiol* 29(8):1269–1278.
39. Kuchitsu K, Tsuzuki M, Miyachi S (1991) Polypeptide composition and enzyme activities of the pyrenoid and its regulation by CO_2 concentration in unicellular green algae. *Can J Bot* 69(5):1062–1069.
40. Mühlhaus T, Weiss J, Hemme D, Sommer F, Schroda M (2011) Quantitative shotgun proteomics using a uniform 15N-labeled standard to monitor proteome dynamics in time course experiments reveals new insights into the heat stress response of *Chlamydomonas reinhardtii*. *Mol Cell Proteom* 10(9):M110.004739.
41. Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies, and proteome-wide protein quantification. *Nat Biotechnol* 26(12):1367–1372.
42. Gibson DG, et al. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6(5):343–345.
43. Nordhues A, et al. (2012) Evidence for a role of VIPP1 in the structural organization of the photosynthetic apparatus in *Chlamydomonas*. *Plant Cell* 24(2):637–659.
44. Heuser JE (2011) The origins and evolution of freeze-etch electron microscopy. *J Electron Microsc (Tokyo)* 60(Suppl 1):S3–S29.
45. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 25(4):402–408.
46. Schloss JA (1990) A *Chlamydomonas* gene encodes a G protein β subunit-like polypeptide. *Mol Gen Genet* 221(3):443–452.
47. Heinnickel ML, et al. (2013) Novel thylakoid membrane GreenCut protein CPLD38 impacts accumulation of the cytochrome b6f complex and associated regulatory processes. *J Biol Chem* 288(10):7024–7036.
48. Badger MR, Kaplan A, Berry JA (1980) Internal inorganic carbon pool of *Chlamydomonas reinhardtii*: Evidence for a carbon dioxide-concentrating mechanism. *Plant Physiol* 66(3):407–413.
49. Obradovic Z, et al. (2003) Predicting intrinsic disorder from amino acid sequence. *Proteins* 53(Suppl 6):566–572.
50. Linding R, Russell RB, Neduva V, Gibson TJ (2003) GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res* 31(13):3701–3708.
51. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res* 41(Web Server issue, W1):W349–W357.
52. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10(6):845–858.
53. Newman AM, Cooper JB (2007) XSTREAM: A practical algorithm for identification and architecture modeling of tandem repeats in protein sequences. *BMC Bioinformatics* 8(1):382.
54. Romero P, et al. (2001) Sequence complexity of disordered protein. *Proteins* 42(1): 38–48.
55. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J Mol Biol* 305(3):567–580.