A Rubisco-binding protein is required for normal pyrenoid number and starch sheath morphology in *Chlamydomonas reinhardtii*

Alan K. Itakura[a,b,1], Kher Xing Chan[c,1,2], Nicky Atkinson[d], Leif Pallesen[b], Lianyong Wang[e], Gregory Reeves[b,c], Weronika Patena[b,e], Oliver Caspari[c,3], Robyn Roth[f], Ursula Goodenough[f], Alistair J. McCormick[d], Howard Griffiths[c,4], and Martin C. Jonikas[b,e,4]

[a]Department of Biology, Stanford University, Stanford, CA 94305; [b]Department of Plant Biology, Carnegie Institution for Science, Stanford, CA 94305; [c]Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, United Kingdom; [d]SynthSys & Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3BF, United Kingdom; [e]Department of Molecular Biology, Princeton University, Princeton, NJ 08544; and [f]Department of Biology, Washington University in St. Louis, St. Louis, MO 63130

Edited by Susan S. Golden, University of California San Diego, La Jolla, CA, and approved August 2, 2019 (received for review March 18, 2019)

A phase-separated, liquid-like organelle called the pyrenoid mediates CO_2 fixation in the chloroplasts of nearly all eukaryotic algae. While most algae have 1 pyrenoid per chloroplast, here we describe a mutant in the model alga *Chlamydomonas* that has on average 10 pyrenoids per chloroplast. Characterization of the mutant leads us to propose a model where multiple pyrenoids are favored by an increase in the surface area of the starch sheath that surrounds and binds to the liquid-like pyrenoid matrix. We find that the mutant's phenotypes are due to disruption of a gene, which we call StArch Granules Abnormal 1 (*SAGA1*) because starch sheath granules, or plates, in mutants lacking SAGA1 are more elongated and thinner than those of wild type. SAGA1 contains a starch binding motif, suggesting that it may directly regulate starch sheath morphology. SAGA1 localizes to multiple puncta and streaks in the pyrenoid and physically interacts with the small and large subunits of the carbon-fixing enzyme Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase), a major component of the liquid-like pyrenoid matrix. Our findings suggest a biophysical mechanism by which starch sheath morphology affects pyrenoid number and CO_2-concentrating mechanism function, advancing our understanding of the structure and function of this biogeochemically important organelle. More broadly, we propose that the number of phase-separated organelles can be regulated by imposing constraints on their surface area.

carbon fixation | phase separation | starch | Rubisco | pyrenoid

Compartmentalization is key to efficiency and specificity of cellular processes. Beyond the classic membrane-bound compartments widely observed in eukaryotic cells, many processes in both eukaryotic and prokaryotic cells are organized in membraneless cellular bodies (1–3). Our understanding of membraneless cellular bodies is undergoing a revolution, enabled by the recent discovery that many of these bodies can exhibit liquid-like behavior such as internal mixing and the ability to dynamically exchange components with the surrounding solution (1). The emerging paradigm is that many of these bodies assemble by phase separation of the constituent proteins and/or nucleic acids, mediated by weak multivalent interactions (3). Many fundamental questions about the principles that underlie the biogenesis and regulation of such phase-separated, liquid-like organelles remain to be addressed. A key outstanding question is how cells regulate whether they have 1 or many droplets of a particular phase-separated body.

The pyrenoid is a phase-separated, liquid-like organelle found in most eukaryotic algae that plays a significant role in global CO_2 fixation (4, 5). The pyrenoid houses the CO_2-fixing enzyme Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase) and enhances the enzyme's operating efficiency by supplying it with CO_2 at a high concentration (5–7). This structure is an essential part of the CO_2-concentrating mechanism (CCM) (5, 8).

Most pyrenoids consist of 3 subcompartments discernible by electron microscopy (Fig. 1*A*) (9). At the core of the pyrenoid is the liquid-like pyrenoid matrix, where Rubisco is densely clustered (6, 10–13). Essential to this clustering is the recently characterized Rubisco linker protein EPYC1 (14, 15). In most species, the pyrenoid matrix is traversed by membrane tubules, which are continuous with the thylakoid membranes of the chloroplast. These thylakoid tubules have been postulated to deliver CO_2 to Rubisco (16–18), at a concentration ~100-fold the concentration outside the cell (19). In many species, the pyrenoid is surrounded by a starch sheath (20) made up of multiple starch granules, also known as starch plates. This starch has been suggested

Significance

Many cellular structures are assembled via phase separation, forming liquid-like droplets in a manner analogous to oil in water. How can the cell control whether there is 1 droplet or many? Here, we provide insights into this question by studying the pyrenoid, an algal liquid-like organelle that mediates a substantial portion of global CO_2 fixation. Whereas most algae have just 1 pyrenoid, we characterize a mutant that has approximately 10 pyrenoids per cell. Our results suggest a model where imposing an increased surface area on the liquid-like pyrenoid matrix favors the formation of multiple pyrenoids, with a concomitant loss of photosynthetic efficiency. Our findings advance our basic understanding of the biogenesis of the pyrenoid and of phase-separated organelles in general.

Author contributions: A.K.I., K.X.C., H.G., and M.C.J. designed research; A.K.I. and K.X.C. conducted growth assays, Western blotting, and confocal microscopy; A.K.I. designed vectors and generated the complemented strain; K.X.C. performed and analyzed data from oxygen evolution and transmission electron microscopy experiments; N.A. and A.J.M. performed the yeast-2-hybrid experiments; L.P. and G.R. performed the mutant screen; L.W. performed time-lapse confocal microscopy; W.P. mapped and indexed insertional mutants; K.X.C. and O.C. performed immunofluorescence staining; R.R. and U.G. performed quick-freeze deep-etch electron microscopy; K.X.C. and A.K.I. analyzed data; and A.K.I., K.X.C., H.G., and M.C.J. formed the model and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]A.K.I. and K.X.C. contributed equally to this work.

[2]Present address: Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana, IL 61801.

[3]Present address: Institut de Biologie Physico-Chimique, CNRS/Sorbonne Universités, 75005 Paris, France.

[4]To whom correspondence may be addressed. Email: hg230@cam.ac.uk or mjonikas@princeton.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1904587116/-/DCSupplemental.

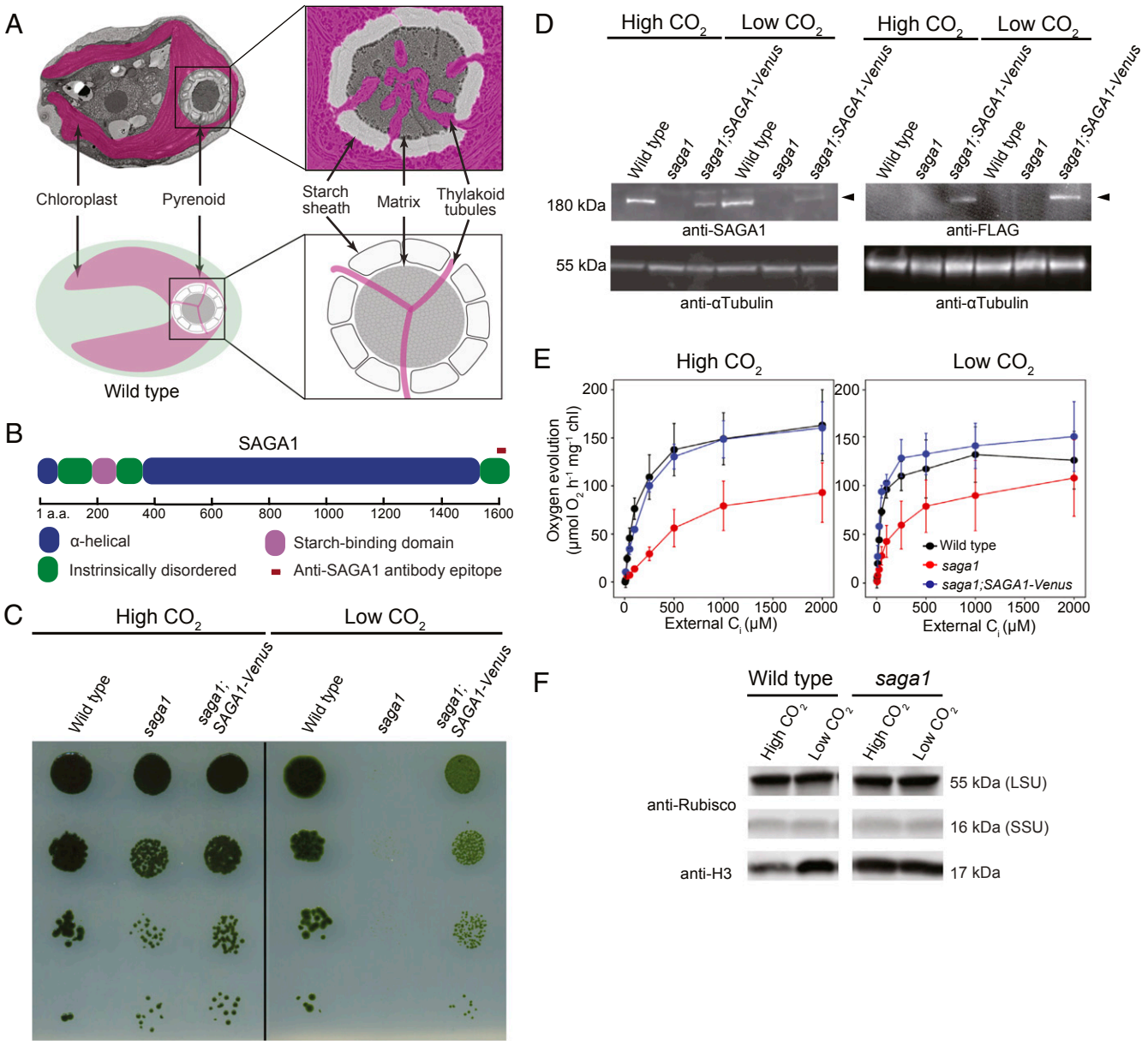


Fig. 1. *SAGA1*, a gene with a predicted starch-binding domain and alpha helical stretch, is necessary for a functional CCM. (*A*) A *Chlamydomonas* cell is shown with the substructures of the pyrenoid labeled. (*B*) SAGA1 has a predicted starch binding domain followed by a long predicted alpha helical region. (*C*) Agar growth phenotypes of wild type, *saga1*, and *saga1;SAGA1-Venus* are shown. Serial 1:10 dilutions were spotted on TP minimal medium and grown at high (4%) and low (0.04%) CO_2 under 500 μmol photons $m^{-2} \cdot s^{-1}$ illumination. (*D*) SAGA1 protein levels in wild-type, *saga1*, and *saga1;SAGA1-Venus* cells grown at low and high CO_2 were probed with an anti-SAGA1 polyclonal antibody and with an anti-FLAG antibody. Anti-tubulin is shown as a loading control. (*E*) Oxygen evolution is shown as a function of external C_i for cells grown in high and low CO_2. (*F*) Rubisco protein levels in wild type and *saga1* grown at low and high CO_2 were probed with a polyclonal antibody raised to Rubisco. Anti-histone H3 is shown as a loading control. LSU: large subunit; SSU: small subunit.

to act as a diffusion barrier that retains CO_2 in the pyrenoid matrix (20). The molecular basis for the biogenesis of the thylakoid tubules and starch sheath, as well as their connections to the pyrenoid matrix, remains unknown in any species.

The model alga *Chlamydomonas reinhardtii* normally has only 1 pyrenoid per cell. Here we describe a pyrenoid-localized protein, SAGA1 (StArch Granules Abnormal 1), in whose absence cells have multiple pyrenoids and abnormally elongated starch sheaths. Our data lead us to propose a model where excessive starch surface area favors the formation of multiple pyrenoids, and where SAGA1 negatively regulates the surface area of the starch sheath to avoid the formation of multiple pyrenoids. Our findings provide a foundation for a molecular understanding of the interactions between the pyrenoid matrix and starch sheath and advance our knowledge of mechanisms that determine the number of phase-separated organelles.

Results

A Screen for CCM Mutants Identified *SAGA1*, a Gene with a Predicted Starch Binding Domain and Alpha Helical Stretch. To identify *Chlamydomonas* mutants with defects in the CCM, we screened an insertional mutant library in search of mutants that require high CO_2 to grow photosynthetically. The screen yielded multiple independent high-CO_2-requiring mutants disrupted in the gene Cre11.g467712, which we call *SAGA1* (StArch Granules Abnormal 1). The predicted gene product of *SAGA1* is a protein

of 1,626 amino acids (expasy.org; Fig. 1*B*). Amino acids 214–280 encode a predicted starch binding motif (CBM20; *SI Appendix*, Fig. S1), a feature conserved in amylolytic enzymes that bind granular starch (21–23). All 5 highly conserved residues necessary for starch binding are present in SAGA1, suggesting that SAGA1 can bind starch (*SI Appendix*, Fig. S1). Following the starch binding motif is a long (~1,000 residues) predicted alpha helical region (Fig. 1*B*) that shares homology with proteins containing coiled-coil domains (*SI Appendix*, Fig. S2). SAGA1 also has a striking amino acid composition: ~36% of SAGA1 consists of only 2 amino acids, alanine (25.3%) and glutamic acid (10.8%) (*SI Appendix*, Fig. S3). These marked features of *SAGA1* are unique among characterized CCM-related proteins. SAGA1 homologs can be found in close *Chlamydomonas* relatives, including *Volvox carteri* and *Tetrabaena socialis*, and in more distant pyrenoid-containing eukaryotic algae, including *Emiliana huxleyi* (9) (*SI Appendix*, Tables S1 and S2).

SAGA1 Is Necessary for a Functional CCM. To confirm the defective CCM growth phenotype, we performed spot assays on the *saga1* mutant and the background strain. The *saga1* mutant was unable to grow in low CO_2 but grew under high CO_2 (Fig. 1*C*). We observed similar growth phenotypes in 2 additional mutant alleles of *SAGA1*, *saga1-2* and *saga1-3* (*SI Appendix*, Fig. S4). We confirmed the presence of the insertional mutagenesis cassette in the *SAGA1* locus by PCR (*SI Appendix*, Fig. S5). We carried out all further experiments using the *saga1* mutant.

To test whether absence of SAGA1 causes the CCM mutant phenotype, we transformed the *saga1* mutant using a construct encoding SAGA1 fused with a C-terminal Venus-3xFLAG tag and a selectable marker (*SI Appendix*, Fig. S6). Two of 9 transformants exhibiting both Venus fluorescence and antibiotic resistance also exhibited a substantial rescue of the low CO_2 growth phenotypes of the *saga1* mutant (Fig. 1*C*). The incomplete rescue of these strains could be due to their lower SAGA1 expression levels than observed in wild type (Fig. 1*D*), possibly resulting from use of an exogenous promoter and/or the removal of several introns in the transformed construct. Alternatively, the addition of a C-terminal Venus-3xFLAG tag could impair the function of SAGA1. We chose 1 of these rescued strains (denoted as *saga1;SAGA1-Venus*; Fig. 1*C*) for use in subsequent experiments. The presence of both the *SAGA1-Venus* cassette and the insertional mutagenesis cassette was confirmed using PCR (*SI Appendix*, Fig. S5).

Using a polyclonal antibody raised against the last 19 amino acids of SAGA1, we detected a ~180-kDa band in the wild type and in the *saga1;SAGA1-Venus-3xFLAG* strain that was completely absent in the *saga1* mutant (Fig. 1*D* and *SI Appendix*, Fig. S7). A band of similar size was detected in the protein sample of *saga1;SAGA1-Venus-3xFLAG* strain when probed with anti-FLAG antibody, confirming that the band of ~180 kDa was the *SAGA1* gene product (Fig. 1*D*). Unlike previously characterized components of the CCM such as EPYC1 (14) and LCIB (24), whose abundances increase under low CO_2, SAGA1 was present at similar abundance per cell under both high and low CO_2 (Fig. 1*D*). This observation suggests that SAGA1 may play a constitutive function under both high- and low-CO_2 growth conditions.

We characterized CCM activity in the *saga1* mutant and in the *saga1;SAGA1-Venus-3xFLAG* strain using photosynthetic O_2 evolution as a proxy for whole-cell affinity for inorganic carbon (C_i; Fig. 1*E*). When acclimated to low-CO_2 conditions, the *saga1* mutant showed a decreased affinity for C_i relative to wild type, indicated by a high photosynthetic $K_{0.5}$ value (~160 μM C_i for *saga1* vs. ~40 μM C_i for wild type; Fig. 1*E* and *SI Appendix*, Table S3). In the *saga1;SAGA1-Venus-3xFLAG* strain, the affinity for C_i uptake was restored (~42 μM C_i; Fig. 1*E*). These results indicate that *SAGA1* is required for a functional CCM.

The *saga1* mutant showed a decreased photosynthetic rate at saturating concentrations of C_i when acclimated to both low and high CO_2 (Fig. 1*E*). The Rubisco protein content of *saga1* was similar to wild type (Fig. 1*F*), suggesting that decreased photosynthetic efficiency of the *saga1* mutant under both low- and high-CO_2 growth conditions may be due to decreased availability of CO_2 to Rubisco or decreased Rubisco activity. We conclude from these observations that SAGA1 is required for a functional CCM and for maximal CO_2 uptake in cells acclimated to both low and high CO_2.

***saga1* Mutants Have Aberrant Starch Sheaths.** Because SAGA1 contains a starch binding motif, we sought to determine if the *saga1* mutant had alterations in the starch sheath surrounding the pyrenoid matrix. Electron microscopy revealed that the *saga1* mutant had an abnormal distribution of starch around the pyrenoid, with fewer starch plates observed per pyrenoid (Fig. 2 *A* and *B*). These abnormal starch plates were elongated: The length of each plate in the *saga1* mutant was ~30% longer that of wild-type starch plates (Fig. 2 *A* and *C*). Furthermore, the starch surrounding each *saga1* pyrenoid appeared thinner, with a decreased total cross-sectional area (Fig. 2 *A* and *D*). We also occasionally observed starch plates that appeared to stack on each other, and single elongated starch plates that entirely or almost entirely surrounded regions of matrix (Fig. 2*E*).

The *saga1;SAGA1-Venus* strain showed partial complementation of the number of starch plates per pyrenoid and of the length of the starch plate–Rubisco interface (Fig. 2 *A*–*C*). The incomplete complementation could be due to the lower expression of SAGA1 relative to wild type (Fig. 1*D*). Interestingly, the pyrenoid starch cross-sectional area in the *saga1;SAGA1-Venus* strain was larger than that of wild type or that of the *saga1* mutant (Fig. 2*D*), suggesting that pyrenoid starch plate biosynthesis was still abnormal in the *saga1;SAGA1-Venus* strain. Based on these results, we conclude that SAGA1 is required for normal starch sheath formation around the pyrenoid and that in the absence of SAGA1 pyrenoid plates are elongated and thinner.

***saga1* Mutants Have Multiple Pyrenoids.** Surprisingly, the electron micrographs revealed that *saga1* mutant cells have multiple pyrenoids, observed as electron-dense bodies identical in texture to the pyrenoid matrix found in wild-type cells (Fig. 3*A* and *SI Appendix*, Fig. S8). These multiple pyrenoids contained Rubisco, as detected by immunogold labeling and immunofluorescence (Fig. 3*B* and *SI Appendix*, Fig. S9 and Table S4). Whereas multiple pyrenoids were rare in wild-type cells, ~40% of *saga1* mutant cross-sections contained multiple pyrenoids (Fig. 3 *C* and *D*). Restoration of *SAGA1* expression in the *saga1;SAGA1-Venus* strain eliminated the multiple pyrenoid phenotype, indicating that disruption of *SAGA1* was responsible for this phenotype (Fig. 3*D*).

Given that *saga1* mutant cells have the same amount of Rubisco per cell as wild type (Fig. 1*F*), and considering that this amount of Rubisco is distributed across multiple pyrenoids, one would expect the average pyrenoid matrix cross-sectional area to be smaller. Indeed, the average cross-sectional area of Rubisco matrix per pyrenoid in the *saga1* mutant was lower than in wild type (Fig. 3*E*). Absolute *saga1* cell size was similar to that of wild-type cells (*SI Appendix*, Fig. S10).

To count the number of pyrenoids per cell in the *saga1* mutant and wild type, we expressed an mCherry-tagged Rubisco small subunit in each strain (14). In nearly all wild-type cells, the *Rbcs1*-mCherry signal was found in a single, large punctum at the base of the chloroplast, consistent with the presence of a single pyrenoid (Fig. 3*F*). Strikingly, in the *saga1* mutant, we observed an average of ~10 *Rbcs1*-mCherry puncta per cell (Fig. 3 *F* and *G* and *SI Appendix*, Fig. S11*A*), and no *saga1* cell was observed to have a singular punctum. This remarkable number of pyrenoids per cell is consistent with the increased incidence of multiple pyrenoids we

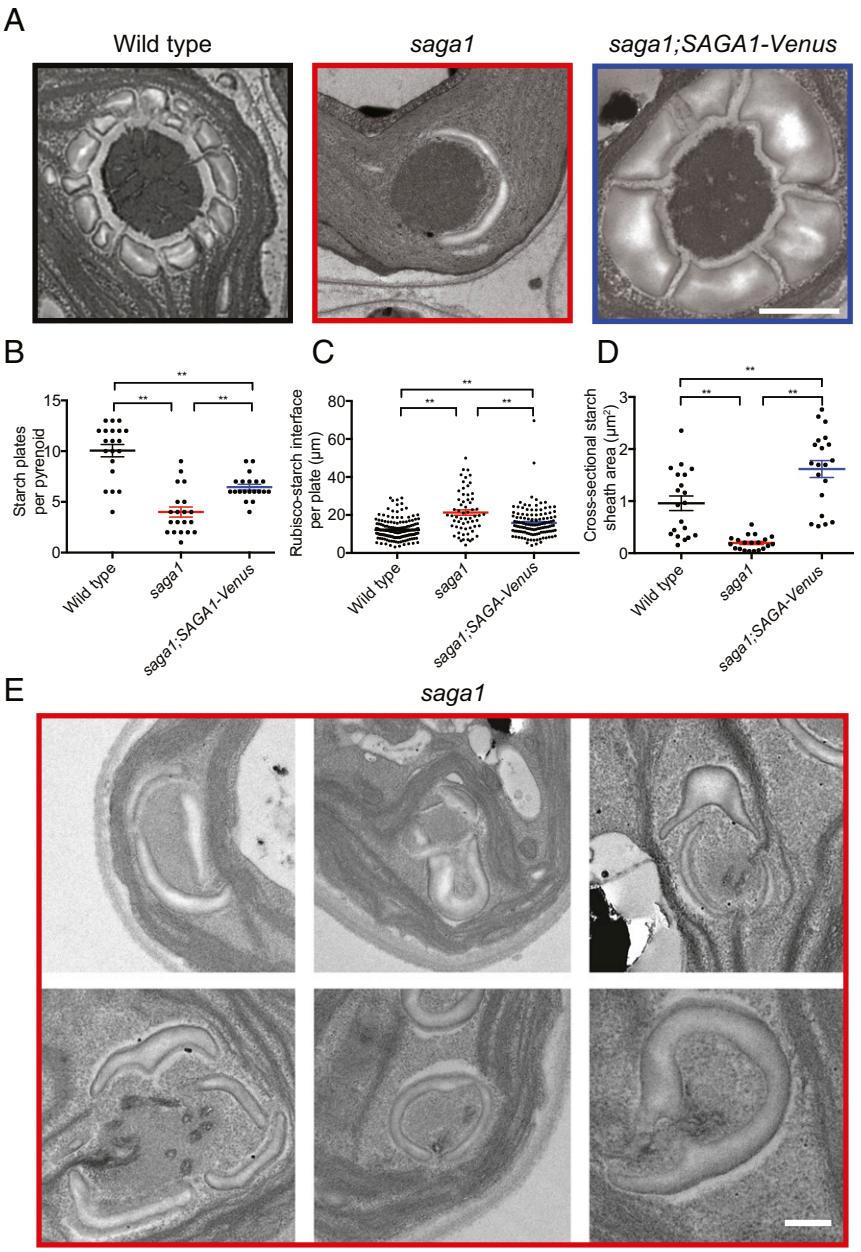


Fig. 2. The pyrenoids of the *saga1* mutant have aberrant starch sheaths. (*A*) Representative TEMs showing pyrenoids of wild-type, *saga1*, and *saga1;SAGA1-Venus* cells grown at low CO_2, illustrating defects in the starch sheath. (*B*) Quantification of the number of starch granules per pyrenoid in wild-type, *saga1*, and *saga1;SAGA1-Venus* cells grown at low CO_2 (n = 20 pyrenoids; Mann–Whitney U test). (*C*) Length of starch granules in wild-type, *saga1*, and *saga1;SAGA1-Venus* cells grown at low CO_2 (n = 162, 114, 126 starch granules; Mann–Whitney U test). (*D*) The cross-sectional area of pyrenoid starch in wild-type, *saga1*, and *saga1;SAGA1-Venus* cells grown at low CO_2 (n = 20 cells; Mann–Whitney U test). Error bars: SEM; $^{**}P < 0.001$. (*E*) Example TEMs showing pyrenoids of *saga1* cells with starch granules either "stacking" on top of each other or appearing to "pinch" away portions of the matrix. (Scale bars, 500 nm.)

observed in the electron micrographs. The positions and sizes of these pyrenoids remained stable over the course of 1 h (*SI Appendix*, Fig. S11*B*). We conclude that the absence of SAGA1 leads to multiple pyrenoids.

Most *Saga1* Mutant Pyrenoids Lack a Visible Pyrenoid Tubule Network. In wild-type cells, thylakoid tubules traverse the pyrenoid matrix, forming a thylakoid tubule network (Fig. 3 *A* and *H* and *SI Appendix*, Fig. S12). However, nearly all pyrenoids of the *saga1* mutant lacked such hallmark thylakoid tubules (Fig. 3 *A* and *H*).

Intriguingly, thylakoid tubules could still be observed in ∼8% of pyrenoids (9/113) in the *saga1* mutant cell sections. In these pyrenoids, the number of observed tubules per pyrenoid was significantly decreased (Fig. 3*I* and *SI Appendix*, Fig. S12). Of the remaining tubuleless pyrenoids in the *saga1* mutant, 48% had thylakoid membranes visible adjacent to the periphery of the pyrenoid matrix, failing to penetrate the matrix (*SI Appendix*, Fig. S12). Single pyrenoids with visible tubules were restored to wild-type levels in the *SAGA1;SAGA1-Venus* strain (Fig. 3*H*), indicating that the absence of pyrenoid tubules in most pyrenoids is due to the disruption of *SAGA1*.

SAGA1-Venus Localizes to the Pyrenoid. To better understand how absence of SAGA1 produces abnormal pyrenoids, we sought to determine the localization of SAGA1. In the *saga1;SAGA1-Venus*

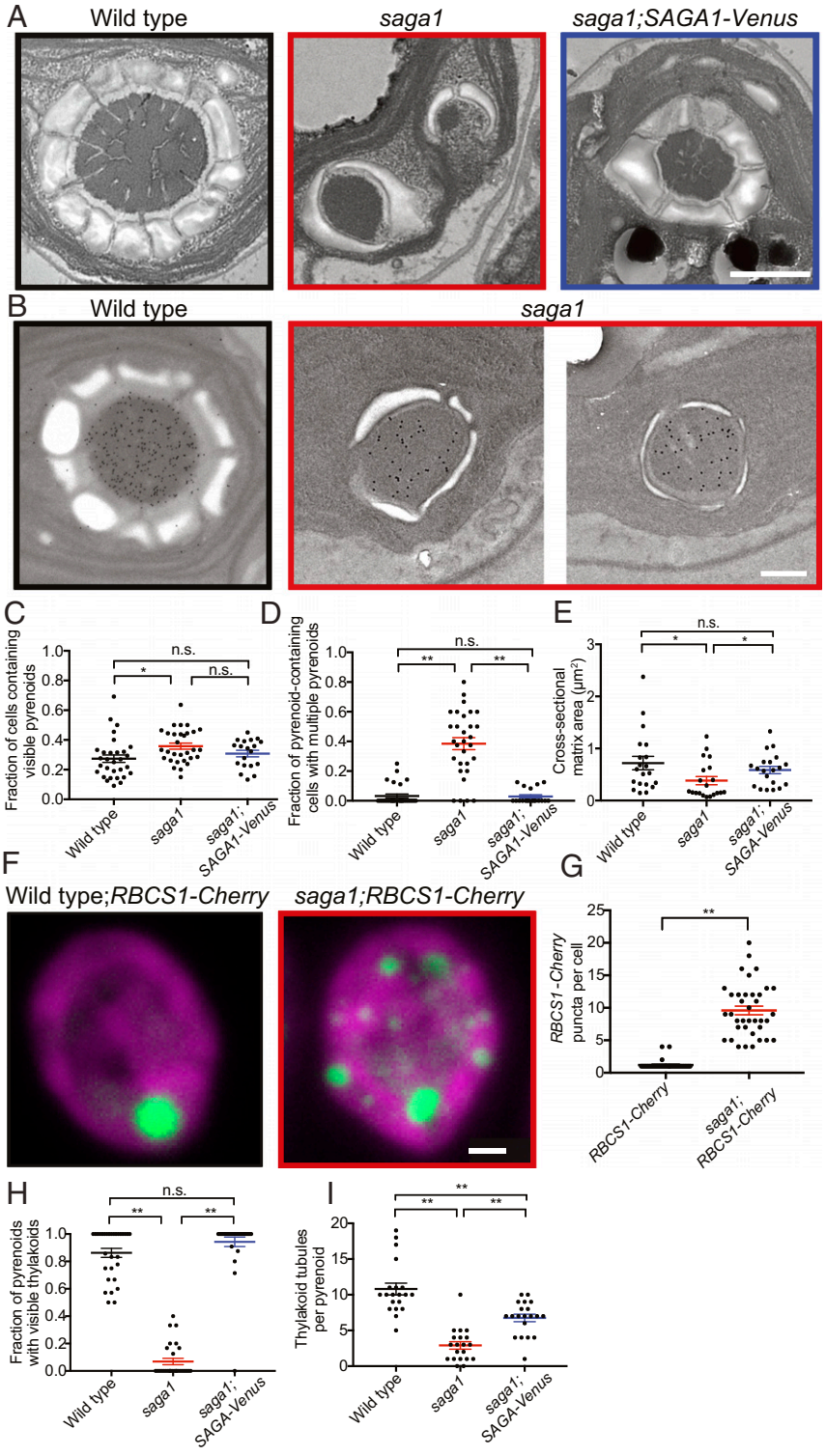


Fig. 3. *saga1* mutant cells have multiple pyrenoids and often lack thylakoid tubules. (*A*) Representative TEMs of wild-type, *saga1*, and *saga1*;*SAGA1-Venus* cells grown at low CO_2, illustrating the frequent observation of multiple pyrenoids in *saga1*. (Scale bar, 500 nm.) (*B*) Representative immunogold staining images of wild type and *saga1* pyrenoids using an anti-Rubisco antibody. (Scale bar, 500 nm.) (*C*) The fraction of cells containing visible pyrenoids in wild-type, *saga1*, and *saga1*;*SAGA1-Venus* cells grown at low CO_2 ($n = 30$ fields of view; Mann–Whitney *U* test). (*D*) The fraction of pyrenoid-containing cells that contain more than 1 pyrenoid in wild-type, *saga1*, and *saga1*;*SAGA1-Venus* cells grown at low CO_2 ($n = 30$ fields of view; Mann–Whitney *U* test). (*E*) Pyrenoid matrix cross-sectional area in wild-type, *saga1*, and *saga1*;*SAGA1-Venus* cells grown at low CO_2 ($n = 20$ pyrenoids). (*F*) Representative summed z-stacks of wild-type and *saga1* cells constitutively expressing RBCS1-mCherry (green) grown in low CO_2. Magenta is chlorophyll autofluorescence. (Scale bar, 1 μm.) (*G*) Quantification of the number of RBCS1-mCherry puncta per cell (wild type: 1.2 ± 0.7 puncta per cell; *saga1:* 9.6 ± 4.1; $n = 35$ cells, Mann–Whitney *U* test). (*H*) Quantification of the fraction of pyrenoids with visible thylakoid tubules ($n = 30$ fields of view; Mann–Whitney *U* test). (*I*) The number of thylakoid tubules per pyrenoid in wild-type, *saga1*, and *saga1*;*SAGA1-Venus* cells grown at low CO_2 ($n = 20$ pyrenoids, Mann–Whitney *U* test). Error bars: SEM; n.s., not significant; $P \geq 0.05$; $*P < 0.05$; $**P < 0.001$.

strain, we observed Venus fluorescence in a spherical region of low chlorophyll fluorescence at the base of the chloroplast, indicating that SAGA1 localizes to the restored pyrenoid in the complemented mutant (Fig. 4). Consistent with this finding, SAGA1 peptides have been detected in the pyrenoid proteome (25). Intriguingly, SAGA1-Venus fluorescence signal was not uniform but formed multiple small puncta and streaks in the pyrenoid. These streaks only partially resemble the peripheral mesh-like localization of LCI9, the star-like thylakoid localization of PSAH, and the punctate thylakoid localization of PSBP4 (26), suggesting that SAGA1 has a distinct localization from these proteins. Some of the SAGA1-Venus signal appeared to partially overlap with chlorophyll autofluorescence at the pyrenoid periphery (*SI Appendix*, Fig. S13). To determine if the overlap between the SAGA1-Venus fluorescence and chlorophyll autofluorescence was significant, we calculated the correlation between the pixel intensities of the 2 fluorescence channels, comparing to random expectation (see *Methods* for more details). We observed a subtle but significant overlap between the 2 fluorescence channels, suggesting that a portion of SAGA1 may colocalize with pyrenoid tubules (*SI Appendix*, Fig. S13 and Table S5). We conclude that SAGA1-Venus localizes to the pyrenoid.

SAGA1 Binds to Rubisco Large and Small Subunits. In our recent high-throughput immunoprecipitation–mass spectrometry study, we observed that SAGA1 coimmunoprecipitates with EPYC1, an abundant pyrenoid component that links Rubisco holoenzymes together to form the pyrenoid matrix (26). In order to validate this coimmunoprecipitation, we performed a yeast 2-hybrid assay to evaluate SAGA1's binding interactions with EPYC1 as well as the large and small subunits of Rubisco. In this assay, SAGA1 interacted with the large and small subunits of Rubisco, but not with EPYC1 (Fig. 5*A*). These findings suggest that our previously observed coimmunoprecipitation of SAGA1 and EPYC1 was mediated by the binding of both proteins to Rubisco (Fig. 5*A*). The structural homology of SAGA1 to coiled-coil proteins raised the possibility that SAGA could dimerize (27, 28); however, SAGA1 did not appear to interact with itself in a yeast 2-hybrid assay (Fig. 5*A*). Western blots confirmed the presence and molecular weight of SAGA1 and the different preys (Fig. 5*B*). We conclude that SAGA1 binds to Rubisco large and small subunits.

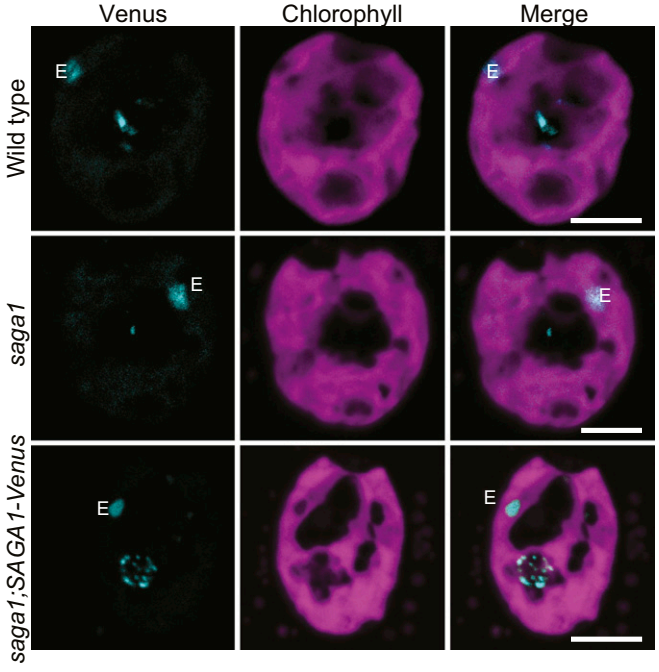


Fig. 4. SAGA1 was localized within the pyrenoid of the rescued mutant *saga1;SAGA1-Venus* after 24-h acclimation to ambient air (low CO_2) to induce the CCM with wild type and *saga1* as controls. All cells were grown at room temperature with illumination of about 50 μmol photons $m^{-2} \cdot s^{-1}$. E: eyespot. White arrows indicated the location of the pyrenoid. (Scale bars, 5 μm.)

Discussion

We Propose a Starch-Centric Model for SAGA1's Function in Pyrenoid Biogenesis. The *saga1* mutant is puzzling because it has 3 prominent phenotypes: abnormal starch plates, multiple pyrenoids, and frequent lack of pyrenoid tubule networks. Occam's razor suggests that SAGA1 likely has a single molecular function whose absence explains all 3 of these phenotypes. We propose a model where SAGA1 guides the formation of pyrenoid starch plates and, in the absence of SAGA1, defective pyrenoid starch plates lead to the formation of multiple pyrenoids, most of which lack tubules (Fig. 6).

We were drawn to this model because the multiple pyrenoid phenotype is at odds with the recent description of the pyrenoid matrix as a phase-separated compartment (4). One would expect that the multiple pyrenoids of the *saga1* mutant would rapidly resolve into a single pyrenoid due to Ostwald ripening, a physical phenomenon whereby large droplets grow by assimilating building blocks from smaller droplets (29). Indeed, Ostwald ripening appears to occur when the pyrenoid matrix is reconstituted in vitro by mixing the Rubisco-binding linker protein EPYC1 with Rubisco (15): Many small droplets initially form, but they rapidly resolve into a smaller number of larger droplets through growth of the large droplets and shrinkage of the small droplets. Some of the large droplets grow to over 5 μm in diameter, corresponding to a volume 2 orders of magnitude larger than that of a pyrenoid in a wild-type cell (Fig. 2*A*), suggesting that the natural tendency of a mixture of EPYC1 and Rubisco is to incorporate all available material into a single large pyrenoid. The presence of multiple stable pyrenoids thus suggests that in the *saga1* mutant additional forces not present in the in vitro system act to counteract Ostwald ripening.

We propose that the excessive surface area of the abnormal starch plates in *saga1* promotes the formation of multiple pyrenoids by counteracting Ostwald ripening. Central to our model is the assumption that starch plates have an affinity for pyrenoid matrix, such that the pyrenoid matrix material "wets" or adheres to 1 side of each starch plate. In wild-type cells, each cell has the appropriate amount of starch plate surface area to coat a single spherical pyrenoid. However, in the *saga1* mutant, the starch plates elongate to reach a total surface area much greater than the surface area of a single pyrenoid. In our model, this mismatch leads to starch plates "pinching" portions of matrix off, which results in multiple smaller pyrenoids (Fig. 6). Indeed, some of the electron micrographs we have observed may represent intermediates in this process, including overlapping starch plates and "pinching" of matrix by single starch plates (Fig. 2*E*). Therefore, the multiple pyrenoids of the *saga1* mutant represent a minimal-energy solution for putting a larger surface area of starch plates in contact with the same volume of liquid-like pyrenoid matrix.

The model raises the question of what mediates binding between starch plates and the matrix. We hypothesize that this binding is mediated by 1 or more proteins with starch-binding domains that also bind to Rubisco or EPYC1 (Fig. 6 and *SI Appendix*, Fig. S14). Furthermore, our model suggests that the starch sheath has an inherent asymmetric property that positions Rubisco to 1 side of the starch sheath plates. This asymmetry is unperturbed in the *saga1* mutant; otherwise, one would expect abnormal starch plates to be surrounded by Rubisco matrix.

Previous studies have found that in the absence of pyrenoid starch cells contain a single pyrenoid and a functional CCM (30). This is in agreement with our proposed model where the multiple

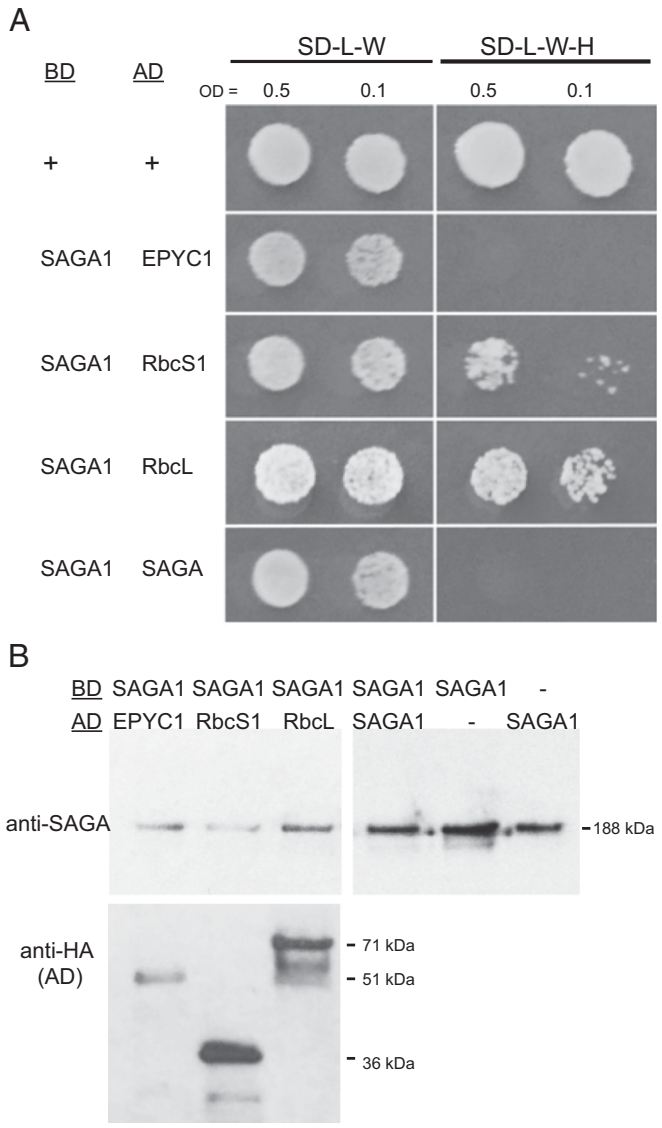


Fig. 5. SAGA1 binds Rubisco small and large subunits. (*A*) Yeast 2-hybrid assays of SAGA1 as the bait (BD; DNA-binding domain) and either EPYC1, Rbcs1, RbcL1, or SAGA1 as the prey (AD; activation domain). *Saccharomyces cerevisiae* expressing the bait were mated with cells expressing an AD and were grown to an OD of 0.5 and 0.1 and spotted onto either nonselective (SD-L-W) or selective (SD-LWH) media. (*B*) Protein levels of the bait and prey proteins in the mated strains were detected by Western blotting using either the anti-SAGA1 antibody or an anti-HA antibody, respectively.

pyrenoid phenotype of the *saga1* mutant is due to the presence of abnormal pyrenoid starch plates, as opposed to the absence of normal pyrenoid starch plates.

Excessive Starch Plate Surface Area:Matrix Volume Ratio Explains Multiple Pyrenoids in Other Mutants. Our proposal that the formation of multiple pyrenoids is caused by an imbalance between the surface area and volume of the pyrenoid matrix is consistent with, and provides an explanation for, the previously unexplained increased incidence of multiple pyrenoids in other mutants, including *epyc1* (14) and *cia6* mutants (31). In these mutants, the pyrenoid matrix volume is decreased because much of the Rubisco is dissolved in the stroma. If pyrenoid starch biosynthesis is largely unchanged, one would expect that these mutants would have an excessive starch sheath surface area relative to pyrenoid matrix volume, which would favor multiple pyrenoids as the matrix adopts a configuration with a higher surface-to-volume ratio, just as we observe in the *saga1* mutant.

The Lack of Tubules Observed in Most Mutant Pyrenoids May Be a Consequence of Cells Retaining a Single Pyrenoid Tubule Network. While most pyrenoids of the *saga1* mutant lacked the canonical thylakoid tubule network, a limited thylakoid tubule network was still observed in ∼10% of pyrenoids (Fig. 3*H* and *SI Appendix*, Fig. S12). We hypothesize that this observation is due to each *saga1* mutant cell's retaining a single, diminished pyrenoid tubule network, with only 1 of the ∼10 pyrenoids per cell containing the network. This proposed model explains the absence of thylakoid tubules in the majority of the pyrenoids of the *saga1* mutant. This model is also consistent with previous work suggesting that the biogenesis of the pyrenoid tubule network is independent of matrix biogenesis, as mutants that entirely lack a pyrenoid matrix still contain a singular tubule network (14, 32).

The pyrenoid tubule networks observed in pyrenoids of the *saga1* mutant have a reduced number of tubules and appear somewhat deformed (Fig. 3*I*). These abnormal pyrenoid tubule networks could be due to altered availability of spaces between starch plates where thylakoid tubules can enter the pyrenoid. Alternatively, the abnormal tubule networks could be explained by a regulatory response of the extent of the thylakoid tubule network to the size of the Rubisco matrix, which is decreased in *saga1* mutants.

The absence of thylakoid tubule networks from most pyrenoids in the *saga1* mutant indicates that most of these pyrenoids are not supplied with CO_2 via the thylakoid network. Consequently, the majority of the Rubisco does not benefit from concentrated CO_2, explaining the decreased affinity to CO_2 and lower maximum photosynthetic rates we observed in the *saga1* mutant (Fig. 1*E*). The absence of a singular pyrenoid and the diminished thylakoid tubule network of the *saga1* mutant may also explain the decreased affinity of the mutant for C_i under high CO_2 (*SI Appendix*, Table S3).

Recent work suggests that the pyrenoid tubules may serve as an "anchor" that localizes the pyrenoid to the base of the chloroplast, as the tubules still retain their canonical localization even in the absence of pyrenoid matrix (14). The scattering of pyrenoids across the stroma in the *saga1* mutant (Fig. 3*F*) could thus be due to their lack of thylakoid tubules.

A Proposed Molecular Role for SAGA1 in Shaping Starch Plates. While the mechanisms that control the shape of starch plates are poorly understood in any organism (33), a direct role for SAGA1 in regulation of starch plate shape is plausible given that the *saga1* mutant exhibited abnormal starch plates, SAGA1 has a CBM20 starch binding domain, and other CBM20 motif-containing proteins have been demonstrated to influence starch biogenesis. Moreover, other coiled-coiled proteins are important for starch granule initiation in *Arabidopsis* (34) and pyrenoid starch synthesis in *Chlamydomonas* (35). The observation that SAGA1-Venus forms streaks near the periphery of the pyrenoid (Fig. 4) is consistent with a possible localization to the edges of starch plates, where SAGA1 could inhibit pyrenoid starch plate elongation (Fig. 6*A*). SAGA1 could influence starch-plate morphology by recruiting starch-remodeling enzymes, or by preventing access of starch-synthesis enzymes. Alternatively, SAGA1 itself could possess its own enzymatic activity; however, it shows no clear homology to any known starch-modifying enzymatic domain. The larger, but less numerous, starch granules we observed in the rescued strain suggest that SAGA1 may also play a role in creating the interfaces between starch plates for thylakoid tubules to enter and exit.

SAGA1's binding interaction with Rubisco could serve to localize SAGA1 to the periphery of the pyrenoid matrix (Fig. 4). A

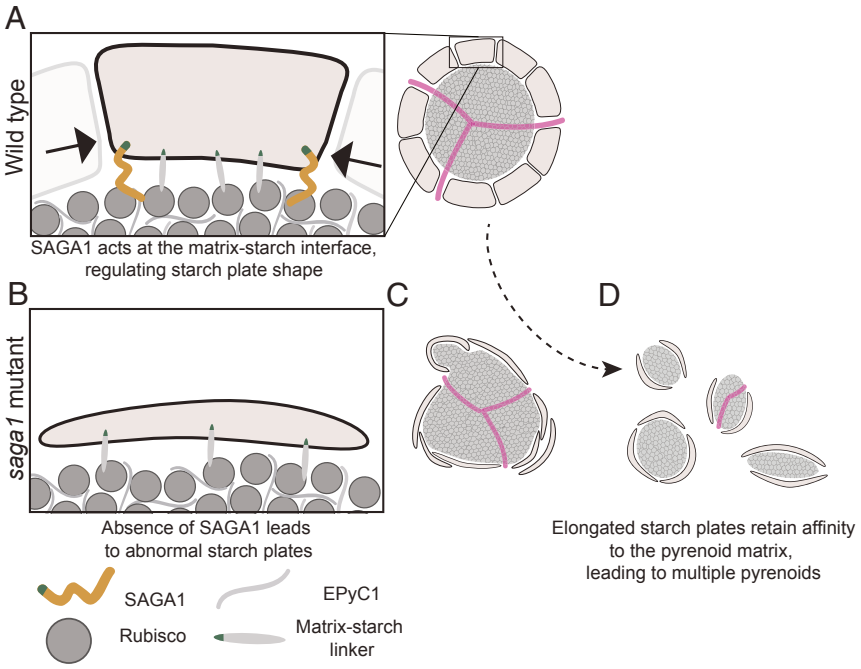


Fig. 6. A proposed model for the function of SAGA1. (*A*) Multiple proteins link starch and the Rubisco matrix. SAGA1 plays a role at the periphery of starch granules in restricting starch granule elongation. (*B*) Loss of SAGA1 results in elongated starch granules. (*C*) The elongated starch granules have increased surface area and their affinity for Rubisco favors a configuration of the Rubisco matrix that has increased surface area, resulting in multiple pyrenoids. (*D*) There is only 1 "knot" of thylakoid tubules per cell and, as a consequence, the extra pyrenoids in the *saga1* mutant lack thylakoid tubules.

size-exclusion effect may prevent the 180-kDa SAGA1 from entering inside the matrix. Indeed, we previously observed that most chloroplast proteins that can enter the pyrenoid matrix are smaller than ∼80 kDa (26). Other phase-separated bodies, such as P-granules, can act as size-exclusion barriers due to the effective mesh size of the condensed phase (36–38). SAGA1 contains a long alpha helical region (*SI Appendix*, Fig. S2); thus, if the Rubisco-binding region is on SAGA1's C terminus, this alpha helical region could span the ∼30-nm gap between the starch sheath and matrix that has been observed in native electron microscopy studies (*SI Appendix*, Fig. S8) (14) and would allow the CBM20 domain to interact with the starch sheath. Our observation of a weak colocalization of SAGA1-Venus with chlorophyll autofluorescence (*SI Appendix*, Fig. S13) could indicate that excess SAGA1 may accumulate along the surface of the pyrenoid tubules, where it could still remain excluded from the matrix.

In addition to the model we presented above, alternative models can partially explain our data. SAGA1 could be implicated in thylakoid biogenesis or maintenance, for example by serving as a linker between the thylakoid tubules and the matrix. In this model, an absence of SAGA1 leads to fewer thylakoid tubules, which then liberates the matrix from its canonical location and leads to the formation of multiple pyrenoids. While also a compelling model, SAGA1's starch binding domain, the elongated starch plates of the *saga1* mutant, and the liquid-like nature of the pyrenoid matrix motivate our preference for a model centered around the ratio of starch plate surface area to matrix volume. We hope that our data and model will help guide future work on the molecular function of SAGA1.

Implications for Regulation of Phase-Separated Organelles. Depending on the species of phase-separated organelle, the number of distinct droplets per cell can range from a single body, for example the Balbiani body (39), to dozens of bodies, for example stress granules (40). Moreover, many of these organelles change in number in response to environmental or developmental cues (4, 41). Illuminating work has demonstrated that the cytoskeleton (F-actin) regulates stress granule formation in *Xenopus* (42), while the control of the number of other phase separated bodies, such as P bodies and nucleoli, remains enigmatic (3).

Our model for how starch influences the number of pyrenoids suggests a basic principle for how cells could regulate the number of a species of phase-separated organelle, by constraining their surface-to-volume ratio through the interaction of a liquid-phase component with a peripheral component. More broadly, this principle could be useful for the regulation of number of phase-separated organelles in synthetic biology systems (43, 44).

Methods

Cloning. The *SAGA1* (Cre11.g467712) open reading frame (ORF) was assembled from 3 synthesized fragments, the first 2 containing introns and the third lacking introns. Twenty-one introns in total were included in the final sequence. Next, this assembled *SAGA1* ORF was cloned into pRAM118 (GenBank accession no. MK357711). This cassette is identical to pLM005 (14), but the *AphIII* cassette for paromomycin resistance cassette is replaced with an *AphII* cassette for hygromycin resistance. This allows for transformation of the *saga1* mutant.

Strains and Culture Conditions. Wild-type *C. reinhardtii* cMJ030 was used for all experiments. The *saga1* and *saga1-2* mutants were isolated from a collection of high-CO_2-requiring mutants generated by transformation of the pMJ016c mutagenesis cassette into cMJ030 (45). These mutants can be found at the Chlamydomonas Resource Center (*saga1*: CC5420; *saga1*-2: CC5421). The *saga1-3* mutant was obtained from the CLiP collection (46) and can be found at the Chlamydomonas Resource Center (ID: LMJ.RY0402.060639). The *saga1;SAGA1-Venus* strain (CC5422) was generated by transformation of pRAM118 containing a Venus-tagged *SAGA1* into the *saga1* mutant. The *saga1;RbcS1-mCherry* strain (CC5423) was generated by transforming *saga1* with a pLM006 construct containing *Rbcs1* (14). Cells were grown to a concentration ∼2×10^6 cells per mL prior to experiments. All experiments were performed under photoautotrophic conditions in Tris-phosphate medium (47) with high CO_2 (3% [vol/vol] or 5% [vol/vol] CO_2-enriched air) or low CO_2 (air, ∼0.04% [vol/vol] CO_2).

Spot Tests. Cells were pregrown in liquid Tris-acetate-phosphate medium (47). The cells were washed with Tris-phosphate medium and resuspended to a concentration of 6×10^5 cells per mL. Cells were then diluted by factors of 1, 10, 100, and 1,000. Fifteen microliters were spotted onto solid Tris-phosphate plates. The cells were grown under high- and low-CO_2 concentrations for 7 d with 500 μmol photons $m^{-2} \cdot s^{-1}$.

O_2 Evolution. The apparent affinity for C_i was determined using the oxygen evolution method as previously described in ref. 14 using an OXYVY1 Hansatech Oxyview System (Hansatech Instruments).

Microscopy. Cell samples used for transmission electron microscopy (TEM) were fixed with 2% (vol/vol) glutaraldehyde and embedded in epoxy resin mix. TEM was performed using a Tecnai G2 80- to 200-kV transmission electron microscope (FEI Company) and imaged with AMT Image Capture Engine software (Advanced Microscopy Techniques). Measurement of the section area and perimeter of the cells, pyrenoids, and the surrounding starch sheath was performed using ImageJ2 (Fiji) (48–50).

Fluorescence microscopy on wild type, *saga1*, and *saga1;SAGA1-Venus* to observe the localization of SAGA1 was performed using Leica TCS SP5 confocal microscope (Leica Microsystems) and imaged with the Leica Application Suite Advanced Fluorescence software (Leica Microsystems). Fluorescence microscopy for *saga1;RbcS1-mCherry* was performed with a spinning-disk confocal microscope (3i custom-adapted Leica DMI6000). Samples were mounted on poly-L-lysine–coated plates (Ibidi) and covered with low-melting-temperature agarose (Invitrogen) to prevent desiccation during imaging.

The following excitation and emission settings were used: Venus, 514 nm excitation with 543/22 nm emission; mCherry, 561 nm excitation with 590/20 nm emission; and chlorophyll, 561 nm emission with 685/40 nm emission.

Analysis of Colocalization of SAGA1-Venus Fluorescence and Chlorophyll Autofluorescence. Pyrenoids were visually demarcated using SAGA1-Venus fluorescence and intensity values of each pixel in the demarcated pyrenoid were extracted in the chlorophyll and Venus channel. The Spearman's rank correlation coefficient was calculated between the chlorophyll and Venus channel for every pixel.

In order to compare to the expected Spearman's rank coefficient that we would observe by chance, we independently scrambled the pixel intensity values of the chlorophyll and Venus channels in 100-pixel blocks and then calculated the Spearman's rank coefficient. This was repeated 100 times to create a distribution of rho values specific to each image. The observed rho value was compared to this distribution to determine a *P* value. We performed the same analysis on a nonpyrenoid region of the chloroplast to control for bleed-through of the Venus channel into the chlorophyll channel. All rho values and *P* values are found in *SI Appendix*, Table S5.

Transformation of *Chlamydomonas*. Constructs were transformed into the nuclear genome of *saga1* mutant by electroporation as described previously (45). Transformants were first screened for hygromycin resistance then screened for Venus- and mCherry-expressing colonies using a Tecan Infinite M1000 PRO plate reader (Tecan Austria GmbH). Excitation and emission settings were as follows: Venus, 532-nm excitation with 555/20 emission; mCherry, 532-nm excitation with 610/30 nm emission; and chlorophyll, 633-nm excitation with 670/30 nm emission.

Western Blot Analysis. Protein levels of SAGA1, Rubisco, alpha-tubulin, and histone H3 in wild type, *saga1*, and *saga1;SAGA1-Venus* were measured as described previously (51).

Analysis of Protein Interactions Using Yeast 2-Hybrid. The 2-hybrid vectors pGBKT7 and pGADT7 were used to detect interactions between proteins of interest. Genes were amplified and cloned into each vector to create fusions with either the GAL4 DNA binding or activation domain, respectively. Yeast cells were then cotransformed with binding and activation domain vectors. Successful transformants were grown in liquid, harvested, diluted to an optical density at 600 nm (OD_{600}) of 0.5 or 0.1, and plated onto SD-L-W and SD-L-W-H containing increasing concentrations of the HIS3 inhibitor triaminotriazole (3-AT) and incubated for 3 d before assessing growth. Protein extraction for Western blots was carried out by resuspending cells to an OD_{600} of 1 from an overnight liquid culture in a lysis buffer (50 mM Tris·HCl [pH 6], 4% [vol/vol] SDS, 8 M urea, 30% [vol/vol] glycerol, 0.1 M dithiothreitol, and 0.005% [wt/vol] bromophenol blue), incubating 65 °C for 30 min, and loading directly onto a 10% Bis-Tris protein gel (Expedeon). Proteins were transferred to a nitrocellulose membrane then probed with anti-SAGA (1:2,000) or anti-HA (1:5,000) primary antibody and HRP-linked goat anti-rabbit secondary antibody (1:10,000).

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

ACKNOWLEDGMENTS. We thank F. Fauser, R. Jinkerson, and J. Vilarrasa-Blasi for sharing phenotypic data for CLiP mutants; L. C. Mackinder and M. T. Meyer for advice and method support; J. N. Skepper, L. Carter, and the Cambridge Advanced Imaging Centre for advice and sample preparation for transmission electron microscopy; S. Ramundo for the pRAM118 vector; H. Cartwright, K. Geisler, and M. Szecówka for confocal microscopy support; and the H.G. and M.C.J. laboratories for helpful discussions. The project was funded by National Science Foundation (EF-1105617, IOS-1359682, and MCB-1146621); the National Institutes of Health (DP2-GM-119137); the Simons Foundation and Howard Hughes Medical Institute (55108535) grants to M.J.; and Biotechnology and Biological Sciences Research Council grants (BB/M007693/1) to H.G. and A.J.M. A.K.I. was supported by a National Institutes of Health Cell and Molecular Biology training grant (5T32GM 7276-42). K.X.C. was supported by the Cambridge-IDB International Scholarship, the Cambridge Philosophical Society, the Cambridge Trust, the Lundgren Fund, and the Houston Putnam Lowry Prize (Fitzwilliam College).

1. C. P. Brangwynne *et al*., Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729–1732 (2009).
2. A. K. Itakura, R. A. Futia, D. F. Jarosz, It pays to be in phase. *Biochemistry* **57**, 2520–2529 (2018).
3. S. Boeynaems *et al*., Protein phase separation: A new phase in cell biology. *Trends Cell Biol.* **28**, 420–435 (2018).
4. E. S. Freeman Rosenzweig *et al*., The eukaryotic CO_2-concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell* **171**, 148–162.e19 (2017).
5. M. T. Meyer, C. Whittaker, H. Griffiths, The algal pyrenoid: Key unanswered questions. *J. Exp. Bot.* **68**, 3739–3749 (2017).
6. K. Kuchitsu, M. Tsuzuki, S. Miyachi, Characterization of the pyrenoid isolated from unicellular green alga Chlamydomonas reinhardtii: Particulate form of RuBisCO protein. *Protoplasma* **144**, 17–24 (1988).
7. A. Küken *et al*., Effects of microcompartmentation on flux distribution and metabolic pools in *Chlamydomonas reinhardtii* chloroplasts. *eLife* **7**, e37960 (2018).
8. S. Miyachi, M. Tsuzuki, I. Maruyoma, M. Gantar, S. Miyachi, Effects of CO2 concentration during growth on the intracellular structure of Chlorella and Scenedesmus (Chlorophyta). *J. Phycol.* **22**, 313–320 (1986).
9. M. Meyer, H. Griffiths, Origins and diversity of eukaryotic CO2-concentrating mechanisms: Lessons for the future. *J. Exp. Bot.* **64**, 769–786 (2013).
10. O. N. Borkhsenious, C. B. Mason, J. V. Moroney, The intracellular localization of ribulose-1,5-bisphosphate carboxylase/oxygenase in chlamydomonas reinhardtii. *Plant Physiol.* **116**, 1585–1591 (1998).
11. G. Lacoste-Royal, S. P. Gibbs, Immunocytochemical localization of ribulose-1,5-bisphosphate carboxylase in the pyrenoid and thylakoid region of the chloroplast of chlamydomonas reinhardtii. *Plant Physiol.* **83**, 602–606 (1987).
12. E. Morita, H. Kuroiwa, T. Kuoriwa, N. Hisayoshi, High localization of ribulose-1,5-bisphosphate carboxylase/oxygenase in the pyrenoids of Chlamydomonas reinhardtii (Chlorophyta), as revealed by cryofixation and immunogold electron microscopy. *J. Phycol.* **33**, 68–72 (1997).
13. K.-H. Süss, I. Prokhorenko, K. Adler, In situ association of Calvin cycle enzymes, ribulose-1,5-bisphosphate carboxylase/oxygenase activase, ferredoxin-NADP+ reductase, and nitrite reductase with thylakoid and pyrenoid membranes of Chlamydomonas reinhardtii chloroplasts as revealed by immunoelectron microscopy.*Plant Physiol.* **107**, 1387–1397 (1995).
14. L. C. M. Mackinder *et al*., A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5958–5963 (2016).
15. T. Wunder, S. L. H. Cheng, S.-K. Lai, H.-Y. Li, O. Mueller-Cajar, The phase separation underlying the pyrenoid-based microalgal Rubisco supercharger. *Nat. Commun.* **9**, 5076 (2018).
16. J. A. Raven, CO2-concentrating mechanisms: A direct role for thylakoid lumen acidification? *Plant Cell Environ.* **20**, 147–154 (1997).
17. B. D. Engel *et al*., Native architecture of the Chlamydomonas chloroplast revealed by in situ cryo-electron tomography. *eLife* **4**, 1–29 (2015). Correction in: *eLife* **4**, e11383 (2015).
18. N. A. Pronina, S. Avramova, D. Georgiev, V. E. Semenenko, A pattern of carbonic anhydrase activity in Chlorella and Scenedesmus on cell adaptation to high light and low CO2 concentration. *Sov. Plant Physiol.* **28**, 32–40 (1981).
19. M. R. Badger *et al*., The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO2-concentrating mechanisms in algae. *Can. J. Bot.* **76**, 1052–1071 (1998).
20. Z. Ramazanov *et al*., The induction of the CO2-concentrating mechanism is correlated with the formation of the starch sheath around the pyrenoid of Chlamydomonas reinhardtii. *Planta* **195**, 210–216 (1994).
21. M. Goto, T. Semimaru, K. Furukawa, S. Hayashida, Analysis of the raw starch-binding domain by mutation of a glucoamylase from Aspergillus awamori var. kawachi expressed in Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* **60**, 3926–3930 (1994).

22. L. Chen, P. M. Coutinho, Z. Nikolov, C. Ford, Deletion analysis of the starch-binding domain of Aspergillus glucoamylase. *Protein Eng.* **8**, 1049–1055 (1995).
23. C. Christiansen *et al*., A CBM20 low-affinity starch-binding domain from glucan, water dikinase. *FEBS Lett.* **583**, 1159–1163 (2009).
24. T. Yamano *et al*., Light and low-CO2-dependent LCIB-LCIC complex localization in the chloroplast supports the carbon-concentrating mechanism in Chlamydomonas reinhardtii. *Plant Cell Physiol.* **51**, 1453–1468 (2010).
25. Y. Zhan *et al*., Pyrenoid functions revealed by proteomics in *Chlamydomonas reinhardtii*. *PLoS One* **13**, e0185039 (2018).
26. L. C. M. Mackinder *et al*., A spatial interactome reveals the protein organization of the algal CO_2-concentrating mechanism. *Cell* **171**, 133–147.e14 (2017).
27. F. Uhlmann, SMC complexes: From DNA to chromosomes. *Nat. Rev. Mol. Cell Biol.* **17**, 399–412 (2016).
28. A. Matityahu, I. Onn, A new twist in the coil: Functions of the coiled-coil domain of structural maintenance of chromosome (SMC) proteins. *Curr. Genet.* **64**, 109–116 (2018).
29. J. D. Ng *et al*., The crystallization of biological macromolecules from precipitates: Evidence for Ostwald ripening. *J. Cryst. Growth* **168**, 50–62 (1996).
30. A. Villarejo, F. Martinez, M. D. P. Plumed, Z. Ramazanov, The induction of the CO2 concentrating mechanism in a starch-less mutant of Chlamydomonas reinhardtii. *Physiol. Plant.* **98**, 798–802 (1996).
31. Y. Ma, S. V. Pollock, Y. Xiao, K. Cunnusamy, J. V. Moroney, Identification of a novel gene, CIA6, required for normal pyrenoid formation in Chlamydomonas reinhardtii. *Plant Physiol.* **156**, 884–896 (2011).
32. O. D. Caspari *et al*., Pyrenoid loss in Chlamydomonas reinhardtii causes limitations in CO2 supply, but not thylakoid operating efficiency. *J. Exp. Bot.* **68**, 3903–3913 (2017).
33. S. C. Zeeman, J. Kossmann, A. M. Smith, Starch: Its metabolism, evolution, and biotechnological modification in plants. *Annu. Rev. Plant Biol.* **61**, 209–234 (2010).
34. D. Seung, T. B. Schreier, L. Bürgy, S. Eicke, S. C. Zeeman, Two plastidial coiled-coil proteins are essential for normal starch granule initiation in Arabidopsis. *Plant Cell* **30**, 1523–1542 (2018).
35. J. Findinier *et al*., Deletion of BSG1 in Chlamydomonas reinhardtii leads to abnormal starch granule size and morphology. *Sci. Rep.* **9**, 1990 (2019).
36. L. P. Bergeron-Sandoval, N. Safaee, S. W. Michnick, Mechanisms and consequences of macromolecular phase separation. *Cell* **165**, 1067–1079 (2016).
37. D. L. Updike, S. J. Hachey, J. Kreher, S. Strome, P granules extend the nuclear pore complex environment in the C. elegans germ line. *J. Cell Biol.* **192**, 939–948 (2011).
38. M.-T. Wei *et al*., Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. *Nat. Chem.* **9**, 1118–1125 (2017).
39. E. Boke *et al*., Amyloid-like self-assembly of a cellular compartment. *Cell* **166**, 637–650 (2016).
40. M. D. Panas, P. Ivanov, P. Anderson, Mechanistic insights into mammalian stress granule dynamics. *J. Cell Biol.* **215**, 313–323 (2016).
41. X. J. Lian, I.-E. Gallouzi, Oxidative stress increases the number of stress granules in senescent cells and triggers a rapid decrease in p21waf1/cip1 translation. *J. Biol. Chem.* **284**, 8877–8887 (2009).
42. M. Feric, C. P. Brangwynne, A nuclear F-actin scaffold stabilizes ribonucleoprotein droplets against gravity in large cells. *Nat. Cell Biol.* **15**, 1253–1259 (2013).
43. M. Castellana *et al*., Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nat. Biotechnol.* **32**, 1011–1018 (2014).
44. B. S. Schuster *et al*., Controllable protein phase separation and modular recruitment to form responsive membraneless organelles. *Nat. Commun.* **9**, 2985 (2018).
45. R. Zhang *et al*., High-throughput genotyping of green algal mutants reveals random distribution of mutagenic insertion sites and endonucleolytic cleavage of transforming DNA. *Plant Cell* **26**, 1398–1409 (2014).
46. X. Li *et al*., A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis. *Nat. Genet.* **51**, 627–635 (2019).
47. R. J. Spreitzer, L. Mets, Photosynthesis-deficient mutants of chlamydomonas reinhardii with associated light-sensitive phenotypes. *Plant Physiol.* **67**, 565–569 (1981).
48. J. Schindelin *et al*., Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
49. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
50. C. T. Rueden *et al*., ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* **18**, 529 (2017).
51. M. L. Heinnickel *et al*., Novel thylakoid membrane greenCut protein CPLD38 impacts accumulation of the cytochrome b6f complex and associated regulatory processes. *J. Biol. Chem.* **288**, 7024–7036 (2013).

Supplementary Information

A Rubisco-binding protein is required for normal pyrenoid number and starch sheath morphology in *Chlamydomonas reinhardtii*

Alan K. Itakura[a,b1], Kher Xing Chan[c,1], Nicky Atkinson[d], Leif Pallesen[a], Gregory Reeves[a], Weronika Patena[a,f], Oliver Caspari[c], Robyn Roth[f], Ursula Goodenough[f], Alistair J. McCormick[d], Howard Griffiths[c,2], Martin C. Jonikas[a,2]

[1]A.K.I. and K.X.C. contributed equally to this work.

[2]To whom correspondence should be addressed.
Martin Jonikas (mjonikas@princeton.edu) or Howard Griffiths (hg230@cam.ac.uk)

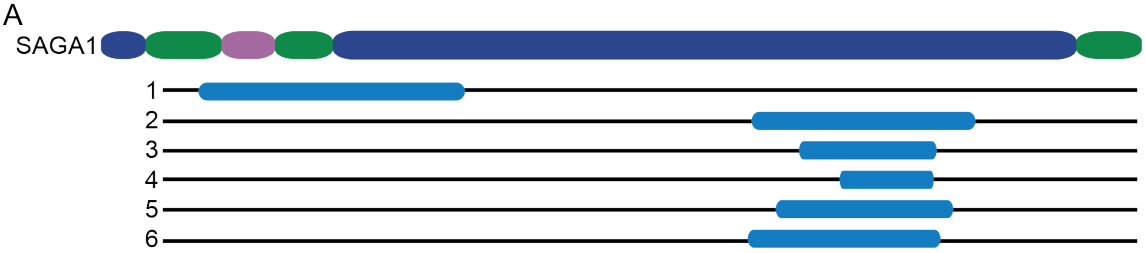
This PDF file includes:

- **Figs S1 to S13**
- **Tables S1 to S6**
- **SI Materials and Methods**
- **SI References**

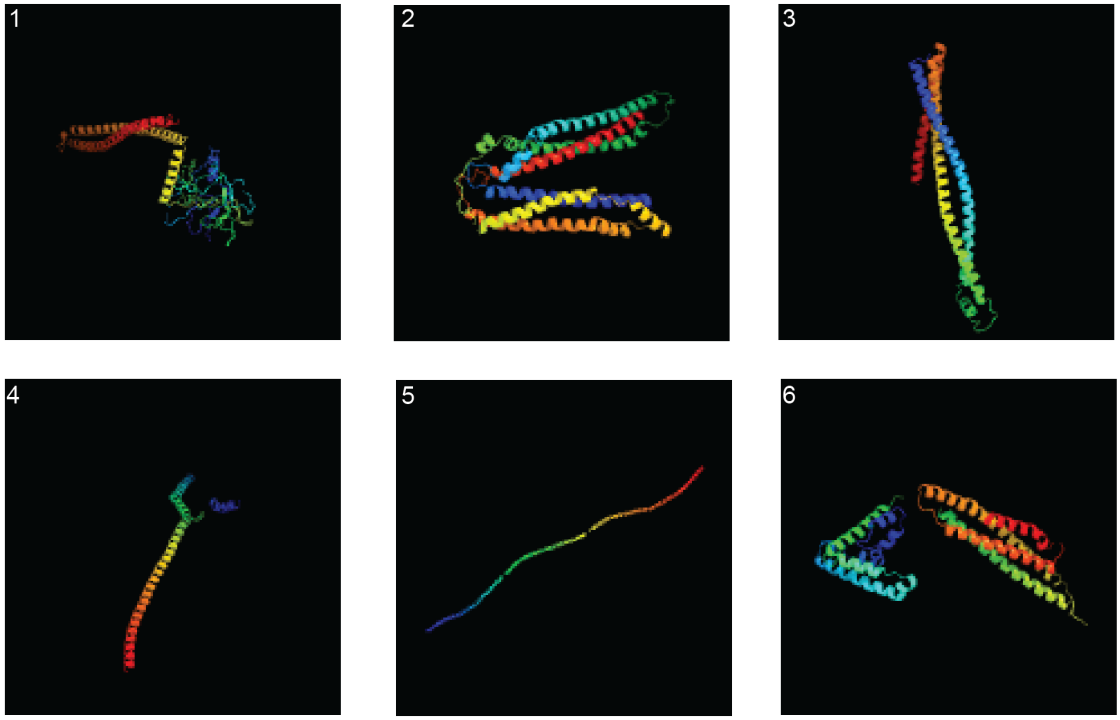
SI Figures

Fig. S1. SAGA1 contains a predicted starch binding domain. SAGA1 (query) was aligned to the conserved starch-binding site 1 of a range of proteins from several species (CDD, NCBI (52, 53); CBM20; E-value = 1.48e^{-08}). This site is suggested to act as an initial starch recognition site. 1CYG_A: Chain A, cyclodextrin glucanotransferase (E.C.2.4.1.19); 1DTU_A: Chain A, *Bacillus circulans* strain 251 cyclodextrin glycosyltransferase; 1ACZ_A: Chain A, glucoamylase; gi 23127960: Lysophospholipase L1 and related esterases [*Nostoc punctiforme* PCC73102]; gi 67926159 Glycoside hydrolase, starch-binding [*Crocosphaera watsonii* WH 8501]; gi 17227665: cyclomaltodextrin glucanotransferase [*Nostoc* sp. PCC 7120]; gi 87123854: Glycoside hydrolase, starch-binding [*Synechococcus* sp. RS9917]; gi 118364918 trehalose-6-phosphate synthase, putative [*Tetrahymena thermophila* SB210]; gi 145502819: hypothetical protein (macronuclear) [*Paramecium tetraurelia* strain d4-2]. Hash mark (#) with yellow highlights: amino acids involved in the starch recognition feature (Feature 1); Grey lower case: unaligned residues; Upper case: aligned residues used to generate PSSM (position-specific scoring matrix); Red to blue color scale: degree of conservation, with red designating highly conserved.

A



B



C

Panel	PDB Template	PDB Header	PDB Molecule	PDB Title	Confidence (%)	Identity (%)	Query range(aa)
1	c1jchC_	ribosome inhibitor, hydrolase	colicin e3	crystal structure of colicin e3 in complex with its immunity protein	98.6	18	36-450
2	c2oevA_	protein transport	programmed cell death 6-interacting protein	crystal structure of alix/aip1	98.4	10	893-1228
3	c4cgkA_	cell cycle	secreted 45 kda protein	crystal structure of the essential protein pcsb from streptococcus2 pneumoniae	98.3	11	950-1157
4	c3ojaB_	protein binding	anopheles plasmodium-responsive leucine-rich repeat protein	crystal structure of lrim1/apl1c complex	98.3	10	1024-1167
5	c1c1gA_	contractile protein	tropomyosin	crystal structure of tropomyosin at 7 angstroms resolution2 in the spermine-induced crystal form	98.2	16	920-1198
6	c1yvlB_	signalling protein	signal transducer and activator of transcription	structure of unphosphorylated stat1	98	9	875-1170

Fig. S2. SAGA1 is predicted to have structural homology to proteins with long alpha helical regions. (*A*) Using Phyre2, the last 1500 amino acids of SAGA1 (due to sequence length

submission limitation) were run in intensive mode. Shown is a schema of the top 6 structural homology results. In blue are the regions of SAGA1 that showed homology to PDB (Protein DataBase) templates. (*B*) The tertiary structure of the top 6 structural homology results are displayed. (*C*) Details of the top 6 structural homology results. PDB template refers to the template that SAGA1 was found to have structural homology with. Provided for each result is the PDB header, the PDB molecule, PDB title, the confidence (the probability that the match between SAGA1 and the PDB template is a true homology) and identity (coverage) of the query sequence that has this particular homology.

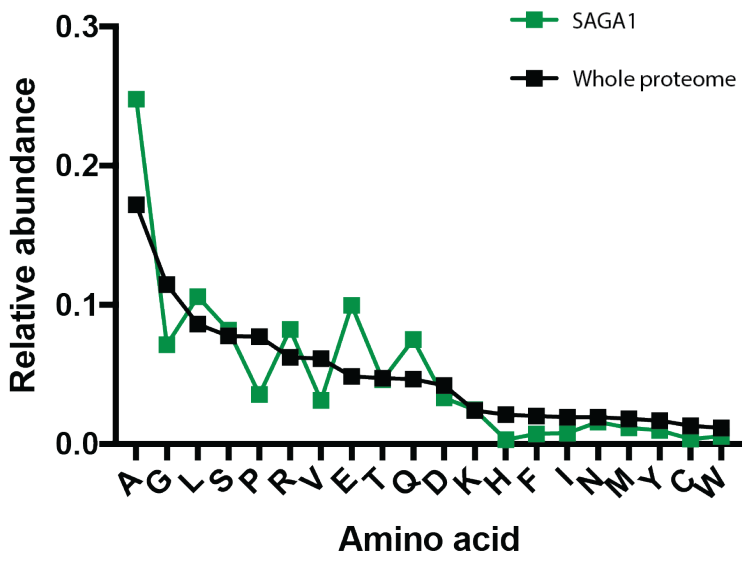


Fig. S3. SAGA1 has an unusual amino acid composition. The relative abundances of each amino acid of SAGA1 and the *Chlamydomonas* proteome were calculated as a fraction of the total number residues in each. Amino acids are ordered from highest to lowest abundance in the *Chlamydomonas* proteome.

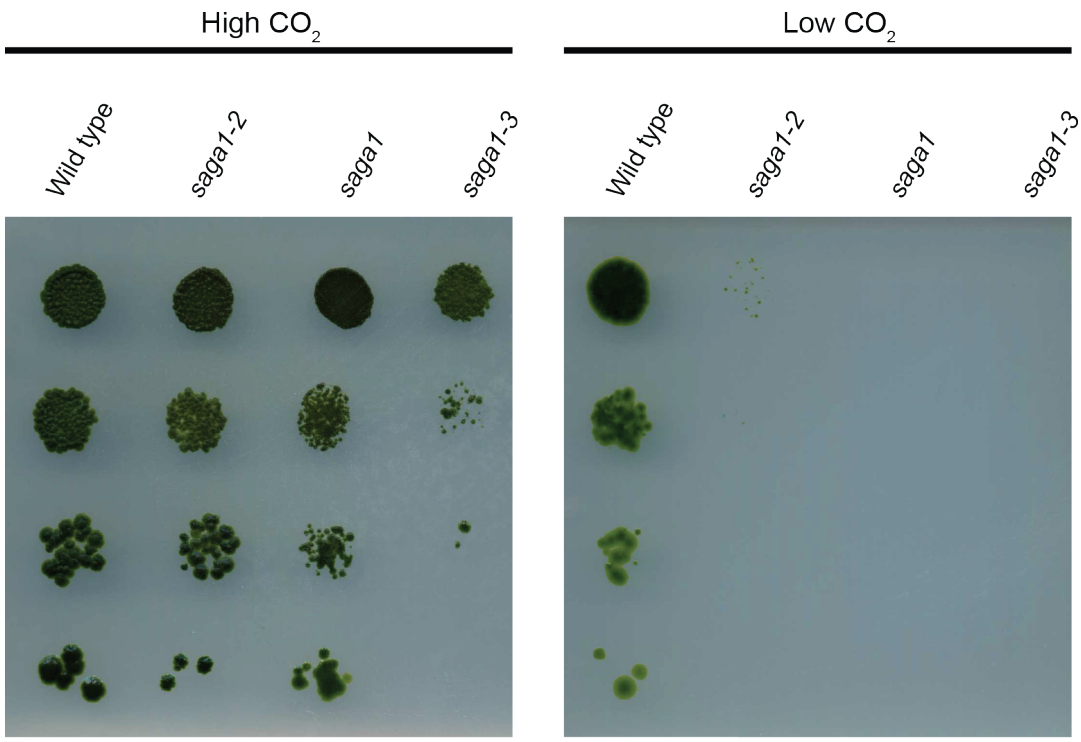


Fig. S4. Other mutant alleles of *saga1* also exhibit a growth defect in low CO_2. Serial 1:10 dilutions of wild type and 3 independent mutant alleles of *saga1* were spotted on TP minimal medium and grown at high and low CO_2 under 500 μmol photos $m^{-2}s^{-1}$ illumination.

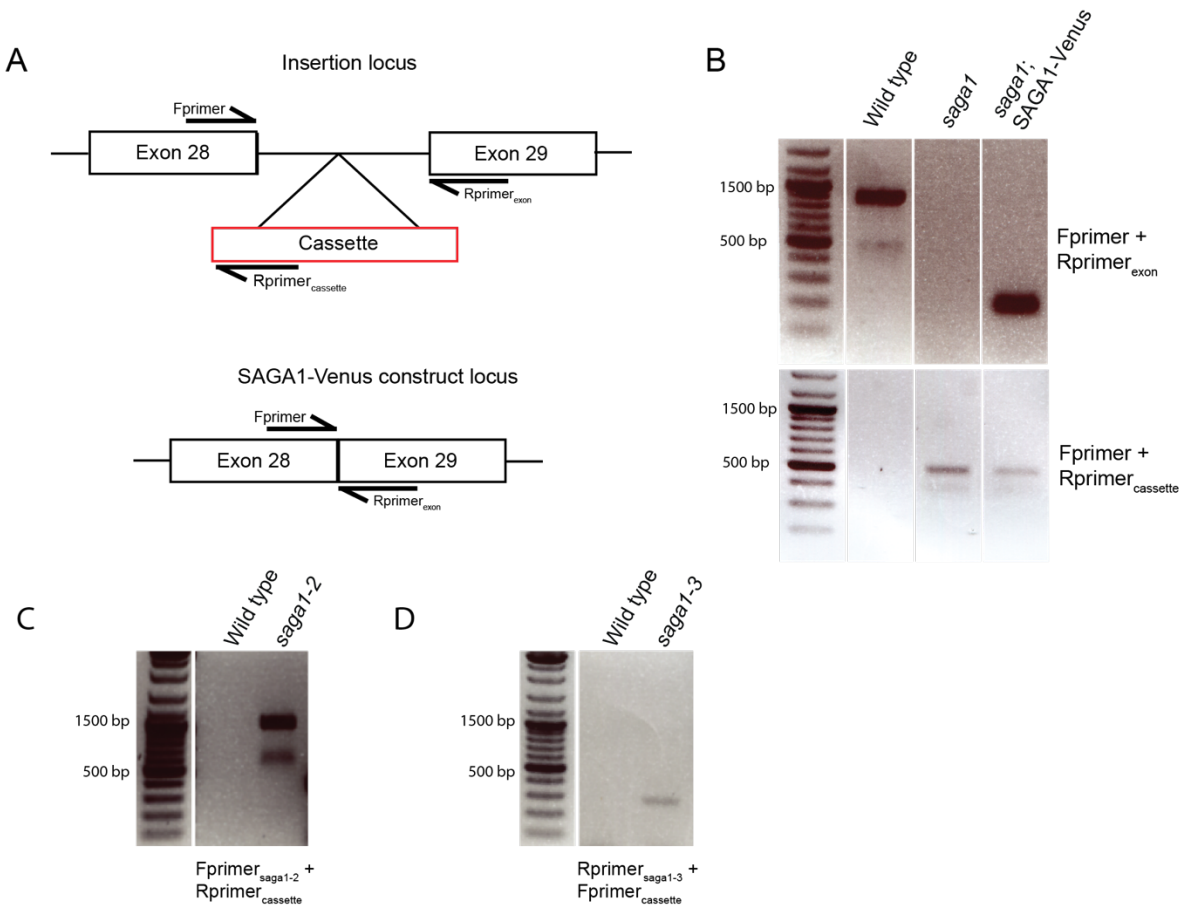


Fig. S5. The complemented *saga1* mutant contains both the insertion cassette and the SAGA1-Venus construct. (*A*) Schema depicting the PCR confirmation strategy of the *saga1* mutant and *saga1;SAGA1-Venus* complement. (*B*) The Fprimer and $Rprimer_{exon}$ yielded a 900 bp product in the wild type but did not produce a product in the *saga1* mutant due to the presence of the insertion cassette. Fprimer and $Rprimer_{cassette}$ yielded a 500 bp product in the *saga1* mutant and the *saga1;SAGA1-Venus* strain due to the presence of the insertion cassette. In the *saga1;SAGA1-Venus* strain, Fprimer and $Rprimer_{exon}$ yielded a 250 bp product, smaller than in the wild type because the complementation cassette lacks intron 28 (Fig. S13). (*C*) PCR confirming the presence and location of the insertion cassette in the *saga1-2* mutant. $Fprimer_{saga1\text{-}2}$ is upstream of the insertion cassette. The double bands indicate the presence of tandem cassette. The *saga1-2* insertion was mapped to intron 28. (*D*) PCR confirming the presence and location of the insertion cassette in the *saga1-3* mutant. $Rprimer_{saga1\text{-}3}$ is downstream of the insertion cassette. $Fprimer_{cassette}$ lies at the 3' end of the cassette. The insertion was mapped to intron 25. Primer sequences can be found in Table S6.

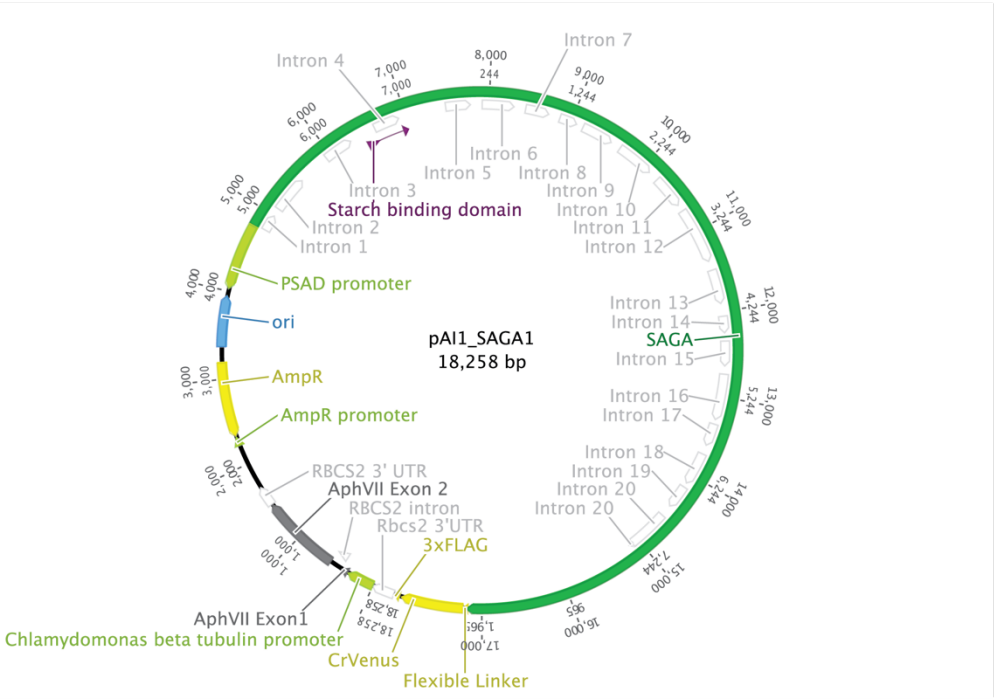


Fig. S6. The pRAM118-SAGA1 construct. The *SAGA1* gene includes the first 20 introns; the remaining introns were omitted during gene synthesis. *SAGA1* is driven by a *PSAD* promoter and is followed by a *CrVenus* tag and a *3xFLAG* tag. The construct contains *AmpR* cassette for ampicillin resistance in *E. coli* and an *AphVII* cassette for hygromycin selection in *C. reinhardtii*. The construct is derived from pLM005, but with an *AphVII* for hygromycin resistance instead of *AphIII* for paromomycin resistance. The construct and sequence can be found at the Chlamydomonas Resource Center as pRAM118_SAGA1_Venus_3xFLAG.

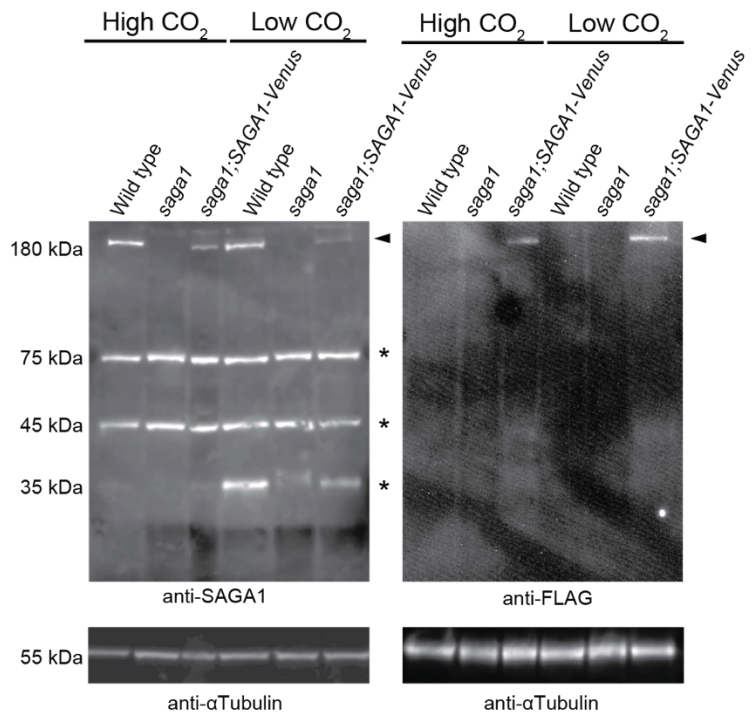


Fig. S7. Full membranes from Fig. 1D. SAGA1 protein levels in wild-type, *saga1*, and *saga1*;*SAGA1-Venus* cells grown at low and high CO_2 were probed with an anti-SAGA1 polyclonal antibody and with an anti-FLAG antibody. Arrow indicates the *SAGA1* protein product. The asterisks indicate non-specific bands. Anti-tubulin is shown as a loading control.

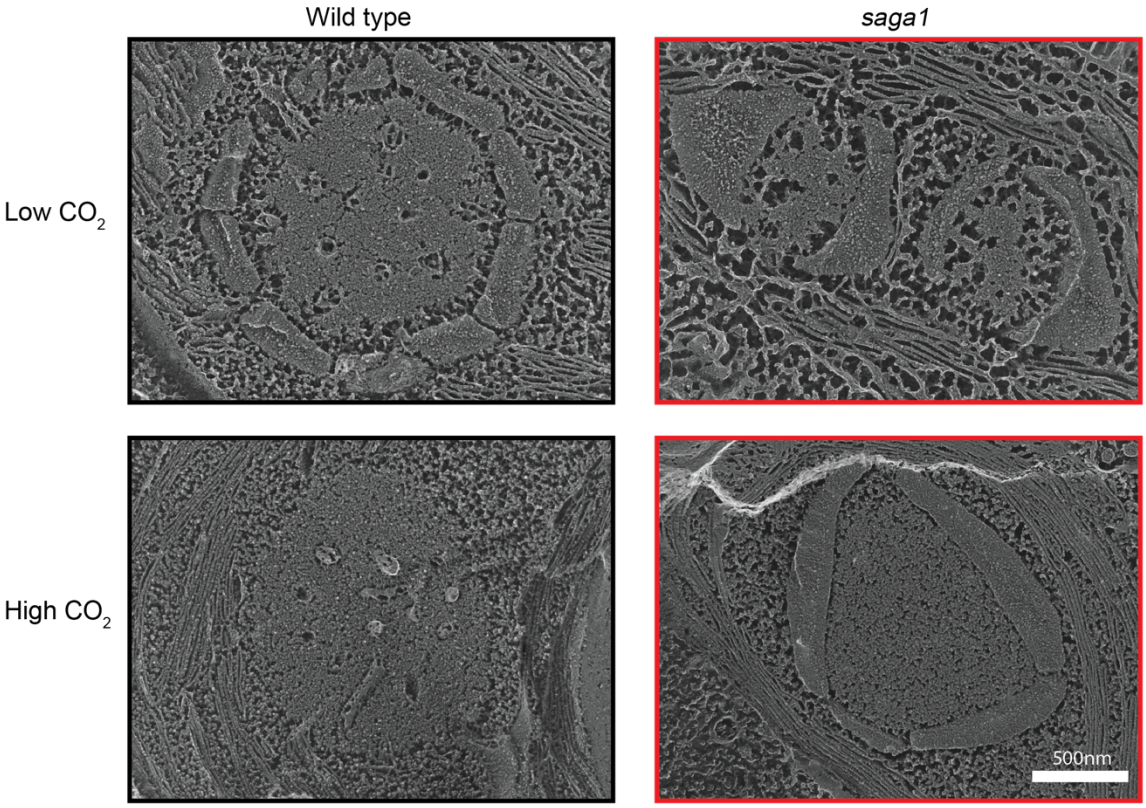


Fig. S8. Quick-freeze deep-etch cryo-electron microscopy reveals abnormal pyrenoid structure in the *saga1* mutant. Representative quick-freeze deep-etch cryo-electron microscopy images of cells grown in high and low CO_2 highlighting multiple pyrenoids, abnormal starch sheaths, and a decreased number of thylakoid tubules in the *saga1* mutant. Samples were fixed with glutaraldehyde prior to electron microscopy. Scale bar = 500nm.

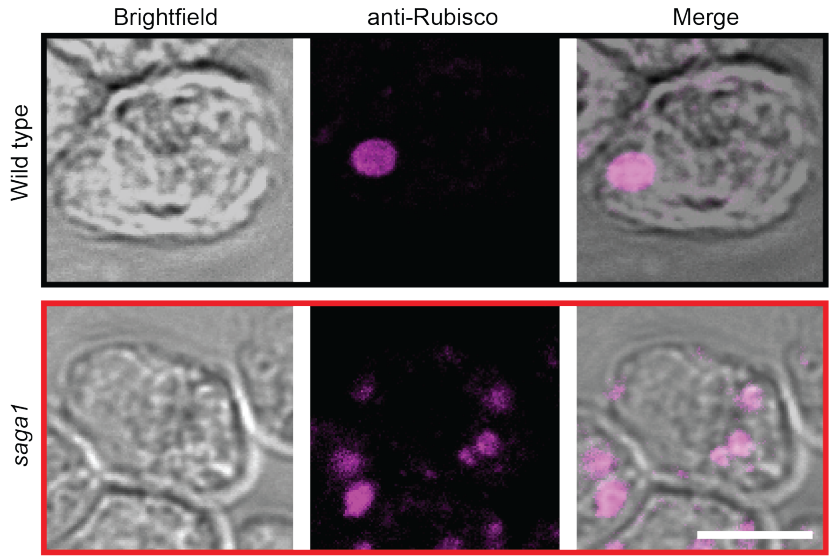


Fig. S9. Rubisco is localized to the pyrenoids of wild-type and *saga1* mutant *Chlamydomonas reinhardtii* cells grown in low CO_2. Subcellular localization of Rubisco in wild-type and *saga1* mutant cells shown by indirect immunofluorescence assay using an anti-Rubisco antibody. Multiple Rubisco localization sites could be observed in each *saga1* mutant cell, in contrast to a single site in each wild-type cell. Scale bar = 5 μm.

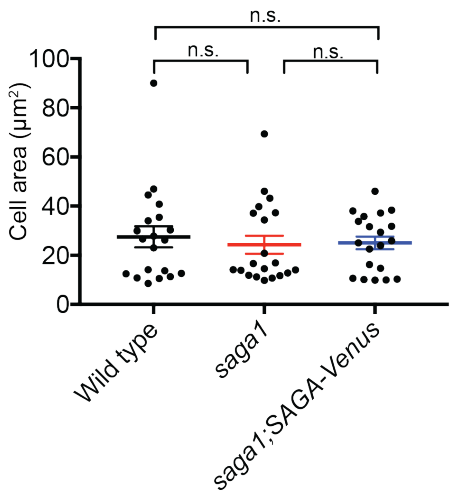


Fig. S10. *saga1* mutant cells are of similar size to wild-type cells. The areas of wild-type, *saga1*, and *saga1*;*SAGA1-Venus* cells imaged using TEM were quantified (N=10 cells; Mann Whitney *U* test; n.s., not significant). Error bars: SEM.

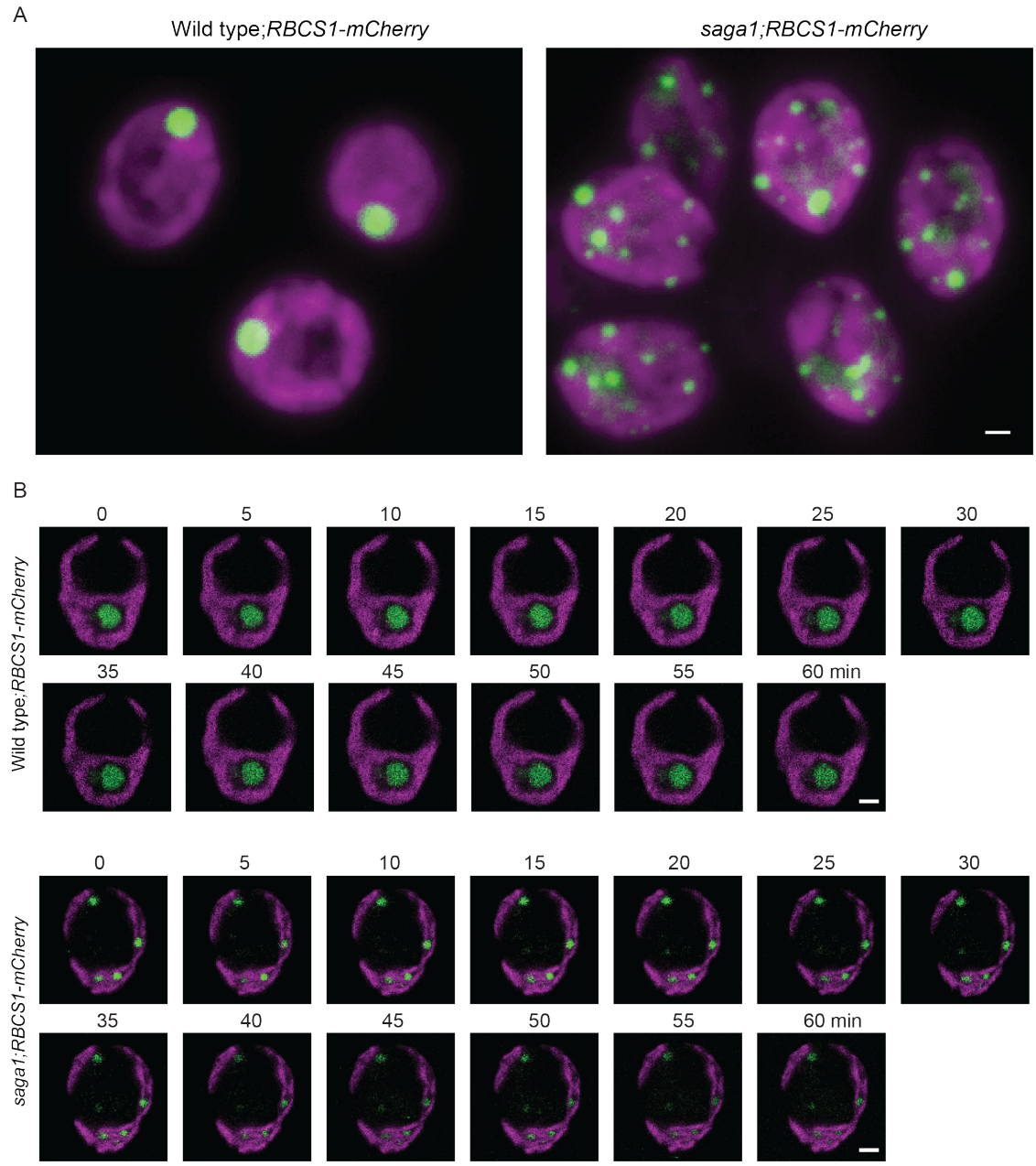
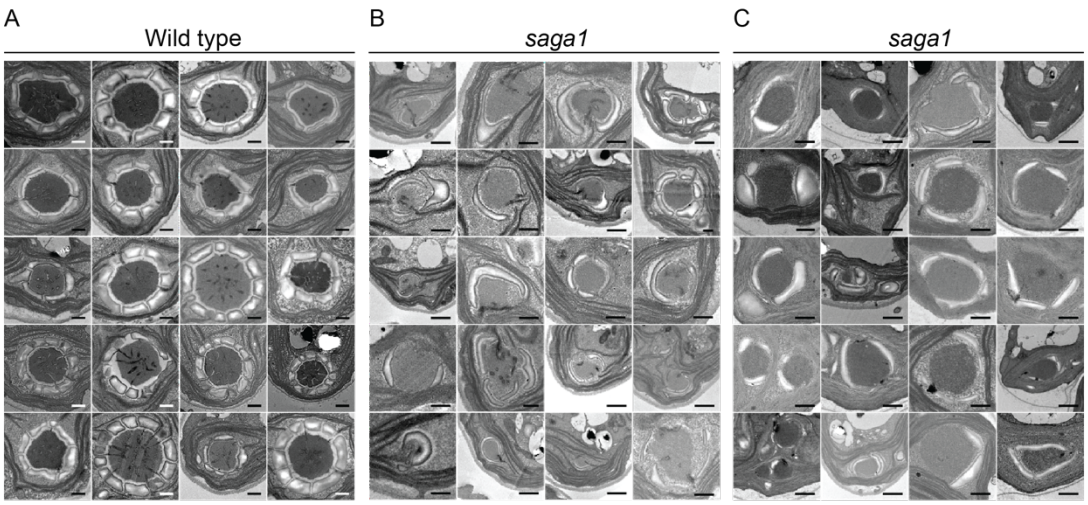


Fig. S11. The *saga1* mutant has multiple stable pyrenoids. (*A*) Representative summed z-stacks of a field of *saga1* and wild-type cells constitutively expressing RBCS1-mCherry grown in low CO_2. One cell from each field view is also shown in Fig. 3. (*B*) A 1-hour time course of *saga1* and wild-type cells expressing RBCS1-mCherry. Green is RBCS1-mCherry fluorescence and magenta is chlorophyll autofluorescence. Scale bar = 1μm.



107

108 **Fig. S12.** Pyrenoids in the *saga1* mutant occasionally contain thylakoid tubules. (*A*) TEM
109 images of wild-type pyrenoids. (*B*) TEM images of *saga1* mutant pyrenoids with thylakoid
110 tubules. (*C*) TEM images of *saga1* mutant pyrenoids without visible thylakoid tubules. Scale
111 bar = 500nm.

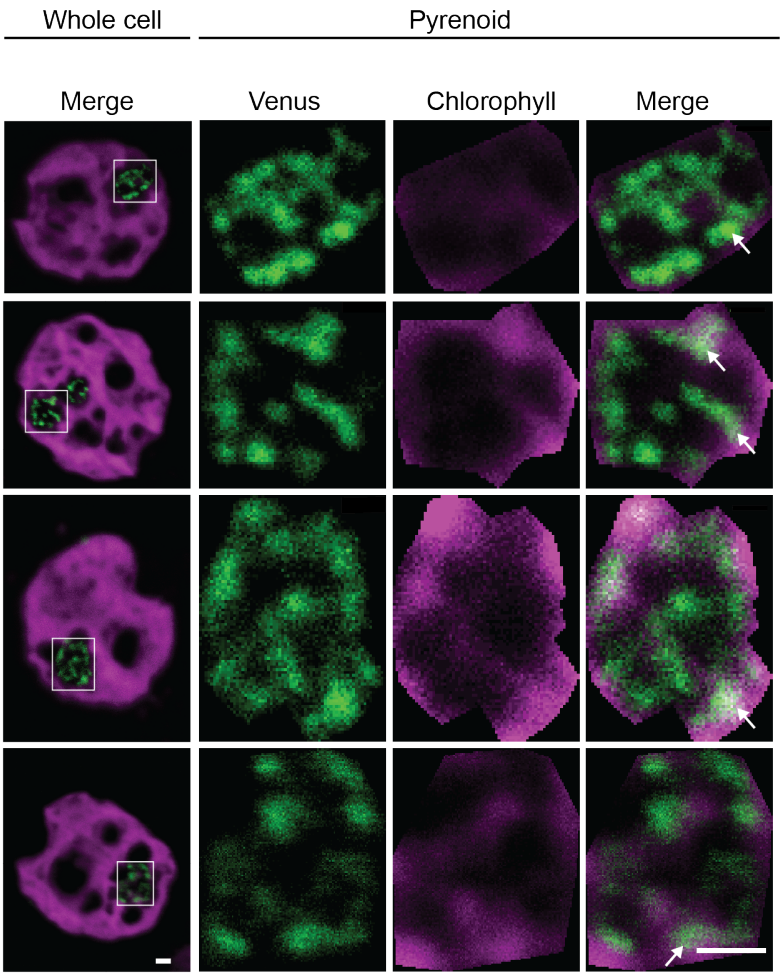


Fig. S13. SAGA1 partially colocalizes with chlorophyll autofluorescence in the pyrenoid. Representative confocal fluorescent microscopy images of SAGA1-Venus (green) constitutively expressed in *saga1* mutant cells grown in low CO_2. Magenta is chlorophyll autofluorescence. White arrows indicate white coloration arising from overlap of SAGA1-Venus and chlorophyll autofluorescence. Spearman rank rho values calculated from pixel intensities of the SAGA1-Venus fluorescence and the chlorophyll autofluorescence are listed in Table S5. Left: scale bar = 1 μm. Right: scale bar = 500nm.

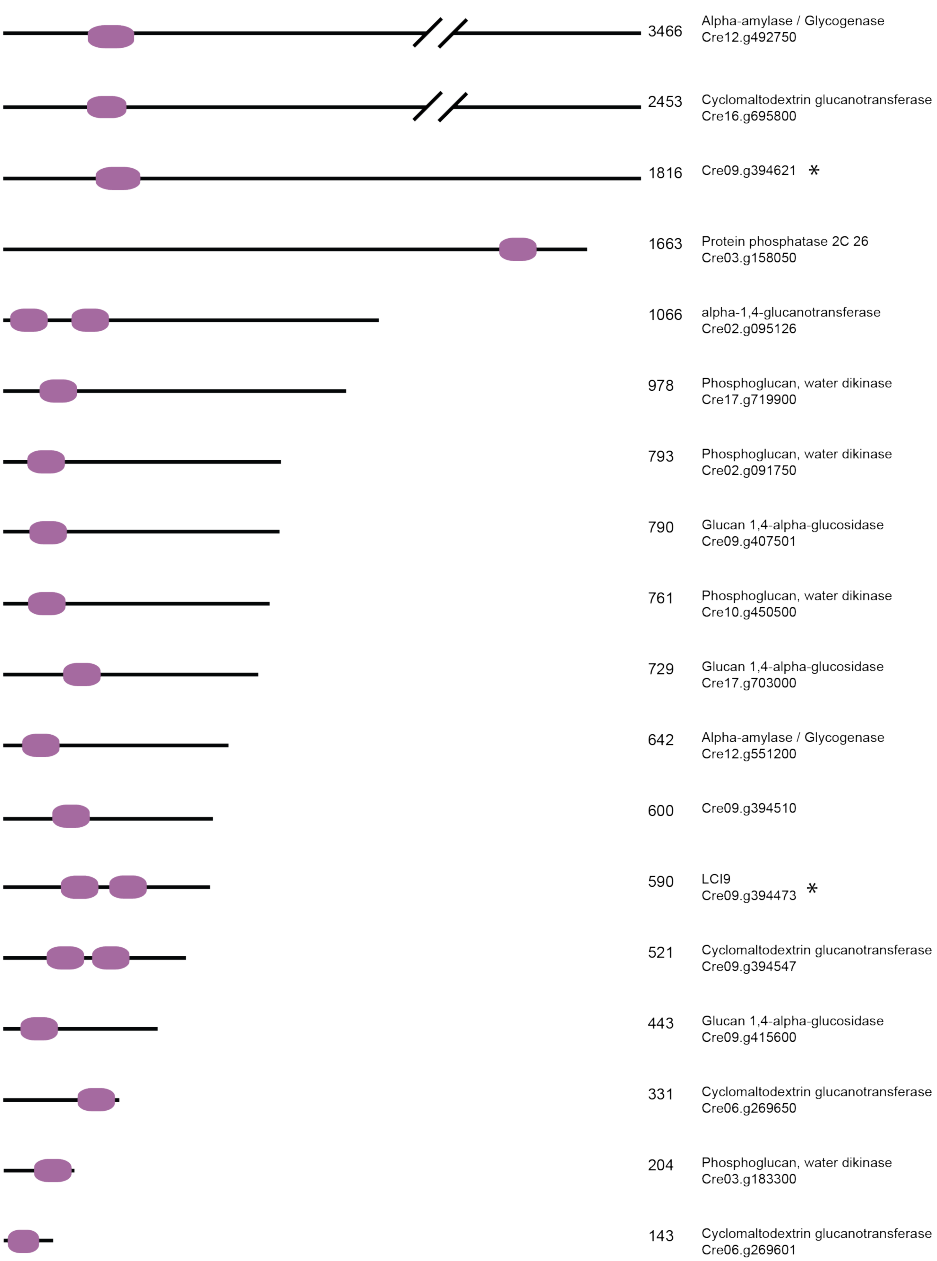


Fig. S14. *Chlamydomonas reinhardtii* proteins with putative starch binding domains. A cartoon depiction of the 18 *C. reinhardtii* proteins that have the CBM_20 motif (PFAM00686; Phytozome). Length of the protein is relative to Cre09.g394621. Included is the gene name and the description of the gene. The asterisks indicate proteins that interacted with components of the pyrenoid matrix by immunoprecipitation-mass spectrometry (26)

SI Tables

Table S1. Protein BLAST results using the SAGA1 starch binding domain (residues 212-280) as a query sequence.

Species	Description	Uniprot ID	Gene ID	Starch binding domain	Coiled-coil	Length	Query sequence	Identity	E-value	Score
Chlamydomonas reinhardtii	Uncharacterized protein	A0A2K3D7T6	CHLRE_11g467712v5	178-308	678	1626	SAGA1 SBD	100.00%	9.60E-47	389
Tetrabaena socialis	Uncharacterized protein	A0A2J7ZZR0	TSOC_007957	65-171	562	1462	SAGA1 SBD	59.40%	1.40E-19	207
Volvox carteri f. nagariensis	Uncharacterized protein	D8TYS9	VOLCADRAFT_105158	98-206, 232-354	403	1552	SAGA1 SBD	41.30%	2.90E-09	138
Gonium pectorale	Uncharacterized protein	A0A150GJ96	GPECTOR_19g330	78-203, 226-334	974	3,273	SAGA1 SBD	45.50%	6.40E-08	129
Gonium pectorale	Uncharacterized protein	A0A150G832	GPECTOR_49g527	14-117	167	672	SAGA1 SBD	39.10%	1.20E-07	127
Volvox carteri f. nagariensis	Uncharacterized protein	D8TPI1	VOLCADRAFT_88626	186-296, 325-432, 509-617	694	2801	SAGA1 SBD	41.50%	5.00E-07	123
Raphidocelis subcapitata	Uncharacterized protein	A0A2V0PDF5	Rsub_08056	42-143	449	1201	SAGA1 SBD	46.20%	7.00E-07	122
Chlamydomonas reinhardtii	alpha amylase?	A0A2K3D1W5	CHLRE_12g492750v5	210-323, 344-468, 491-599	1306	3466	SAGA1 SBD	42.20%	7.10E-07	122
Chlamydomonas reinhardtii	Uncharacterized protein	A0A2K3DEF8	CHLRE_09g394621v5	96-206, 221-343	410	1748	SAGA1 SBD	43.10%	9.9E-7	121
Gracilariopsis chorda	Kinesin-like protein KIN-14R	A0A2V3IYN9	BWQ96_03906	204-325, 669-996	99	1357	SAGA1 SBD	56.40%	2.8E-6	118
Chlorella variabilis	Uncharacterized protein	E1ZG60	CHLNCDRAFT_52636	119-233	49	541	SAGA1 SBD	45.30%	3.70E-06	117

Top 10 results from a Protein BLAST using the Uniprot BLAST portal (unitprot.org). Target database was Eukaryota. Predicted starch binding domains (CBM20) were identified using Interpro; listed are the corresponding a.a. residues. The total length of coiled-coil secondary structures was determined using UniProt's Automatic Annotation pipeline. Identity is a measurement of how similar the query sequence is to the target sequence. E-value is a statistical measure to estimate the number of expected matches in a random database. Score is a normalized metric for how

similar the sequences are, independent of sequence length and database size. E-Threshold was set at 10. BLOSUM62 matrix was used for alignment. No filtering was used. Gaps were permitted.

Table S2. Protein BLAST results using SAGA1 as a query sequence.

Species	Description	Uniprot ID	Gene ID	Starch binding domain	Coiled-coil (total a.a.)	Length (a.a.)	Query sequence	Identity	E-value	Score
Chlamydomonas reinhardtii	Uncharacterized protein	A0A2K3D7T6	CHLRE_11g467712v5	178-308	678	1,626	SAGA1	100.00%	0	7,807
Tetrabaena socialis	Uncharacterized protein	A0A2J7ZZR0	TSOC_007957	65-171	562	1,462	SAGA1	54.50%	0	3,603
Chlamydomonas reinhardtii	Uncharacterized protein	A8J9W7	CHLREDRAFT_151809	N/A	373	577	SAGA1	78.60%	0	1,992
Gonium pectorale	Uncharacterized protein	A0A150FZL5	GPECTOR_106g128	N/A	507	1,080	SAGA1	46.10%	0	1,746
*Volvox carteri**	Mitotic checkpoint protein MAD1	N/A	Vocar.0009s0363	192-305	~1000	1,639	SAGA1	37.60%	1.50E-97	1414
Chlorella sorokiniana	TPR repeat-containing protein	A0A2P6TEU7	C2E21_8179	N/A	343	4,188	SAGA1	27.10%	2.80E-82	801
Porphyra umbilicalis	Uncharacterized protein	A0A1X6PGC9	BU14_0071s0060	N/A	154	2,312	SAGA1	27.60%	7.10E-82	796
Chlorella sorokiniana	Kinesin K39 isoform A	A0A2P6TPQ7	C2E21_5045	N/A	956	2,826	SAGA1	27.70%	8.80E-82	796
Chlamydomonas reinhardtii	Uncharacterized protein	A0A2K3DKA0	CHLRE_07g336750v5	N/A	2450	3,869	SAGA1	28.50%	2.30E-80	785
Emiliania huxleyi	Uncharacterized protein	R1EUH6	EMIHUDRAFT_113389	N/A	1856	4,388	SAGA1	27.80%	3.20E-80	784
Gonium pectorale	Uncharacterized protein	A0A150G2C5	GPECTOR_75g720	N/A	1374	1,898	SAGA1	26.30%	1.30E-79	776

Top 10 results from a Protein BLAST using the Uniprot BLAST portal (unitprot.org). Target database was Eukaryota. Column titles are as in Table S1. * indicates that this result was identified using Phytozome (phytozome.jgi.doe.gov) Protein BLAST.

Table S3. Photosynthetic $K_{0.5}$ (Ci) of wild-type, *saga1* and *saga1;SAGA1-Venus-3xFLAG* cultures.

Sample	Conditions	$K_{0.5}$ (C_i) (μM)	p-value
Wild type	HC	115.8 ± 24.2	-
	LC	38.7 ± 11.8	-
saga1	HC	775.5 ± 88.7	1.0×10^{-5}
	LC	160.0 ± 19.1	0.00015
saga1;SAGA1-Venus-3xFLAG	HC	267.4 ± 84.3	0.20929
	LC	41.9 ± 26.3	0.91083

Data are shown as average ± standard error, which were obtained from at least three independent experiments. The paired t-test was used to determine the significance compared to wild type. C_i, inorganic carbon; HC, high CO_2; LC, low CO_2.

Table S4. Rubisco fraction in the pyrenoid or stroma of wild-type cells and the *saga1* mutant.

Sample	Average area of pyrenoid cross section (nm^2)	Average area of chloroplast cross section (nm^2)	Normalized immunogold particle count (particles nm^{-2}), N		Rubisco in pyrenoid (%)	Comparison to wild type (adjusted p-value)
			Pyrenoid, N_i	Stroma, N_j		
Wild type	9.25×10^6	120.11×10^6	140 ± 17.0	6 ± 1.3	93.8	-
saga1	7.91×10^6	113.21×10^6	50 ± 8.6	4 ± 0.8	90.3	n.s.

The percentage of Rubisco in the pyrenoid was calculated as described in the Supplementary Methods. The data shown are the means ± SEM.

Table S5. Spearman rank rho values calculated from pixel intensities of the SAGA1-Venus fluorescence and chlorophyll autofluorescence.

	Spearman rank rho	*p*-value	Expected mean ± SD
Image 1	0.0336	<0.0001	0.0028±0.0042
Image 2	0.038	<0.0001	0.0027±0.0034
Image 3	0.1128	<0.0001	$6.6577e^{-04}$±$9.2904e^{-04}$
Image 4	0.0699	<0.0001	0.0015±0.0022
Control	0.0023	ns	0.0016±0.0024

Expected Spearman-rank rho distributions (a nonparametric measure of the statistical dependence between the two rankings of two variables, ranging from 0 (no dependence) to 1 (complete dependence)) were created for each image using 100 independently scrambled pixel intensity values of the chlorophyll and Venus channels. The expected means and standard deviations were calculated from these simulated rho distributions. *p*-values were calculated by comparing the observed Spearman rank rho to the distribution of simulated Spearman-Rank rho values (Z-score test).

157 **Table S6.** Primers used to confirm the location of the *saga1* mutants.

158

Primer Name	Sequence
Fprimer (*saga1)*	GCATTGAGATCCGAGATGGT
Rprimer cassette	GCACCAATCATGTCAAGCCT
Rprimer exon (*saga1*)	AGTCCAGGCCGACTACTCC
Rprimer exon (*saga1-3*)	GTGTGAGTGGGATCGCATTCAT
Fprimer cassette (*saga1-3)*	GACGTTACAGCACACCCTTG
Fprimer (*saga1-2)*	CCCACCCCTCACATAAACAC
Rprimer cassette (*saga1-2)*	GCACCAATCATGTCAAGCCT

159

SI Materials and Methods

Screening for the *saga1* mutant

~7,500 mutants on 79 plates, each with 96 colonies were grown on solid TP media in high and low CO_2 in 100 µmol photons m^{-2} s^{-1} light. Mutants that required high CO_2 were sequenced as described in (54). Three mutant alleles mapped to the SAGA1 locus. *saga1:* ~12,590 nt (Intron 26); *saga1-2:* ~13,370 nt (Intron 28)**;** *saga1-3***:** ~12,070 nt (Intron 25). The mutants were confirmed by PCR using a primer that originated from the mutagenesis cassette and a primer in the SAGA1 gene flanking the insertion site. All primers used are in SI Appendix (Table S6).

SAGA1 Sequence Analysis

The full length amino acid sequence of SAGA1 was subjected to PSIPRED (55), NCBI CDD (52) and Phyre2 (56) analyses to identify predicted secondary structures, regions of disordered protein and putative domains.

Generation of the *saga1;SAGA1-Venus* line

Because of the challenge of amplifying across the entire gene, 3 fragments with ~40 nt overlap for Gibson assembly were synthesized (NeoScientific and GeneWiz). These 3 fragments and the resulting construct contained all the exons but only the first 20 introns of SAGA1. Introns 21-34 were removed due to highly repetitive sequences that made them difficult to synthesize and amplify. The 3 fragments were Gibson-assembled into an expression vector that includes a C-terminus Venus and FLAG tag, along with *AphVII* gene for hygromycin resistance under a beta2-tubulin promoter.

Using this construct, we transformed the *saga1* mutant by electroporation as in (45). For each transformation, 14.5 ng kb^{-1} of *Eco*RV-digested construct was mixed with 250 µL of 2 x 10^8 cells mL^{-1} and transformed immediately into *saga1* strains. Cells were selected for hygromycin resistant colonies on hygromycin TAP plates (25 µg mL^{-1}). These colonies were picked into 180 µl of TAP with hygromycin (25 µg mL^{-1}) in a 96-well plate and screened for fluorescence using a Tecan Infinite M1000 PRO plate reader. Excitation and emission settings were: Venus, 532 excitation with 555/20 emission; chlorophyll autofluorescence, 633 excitation with 670/30 emission.

Of the 9 colonies that exhibited a high Venus/chlorophyll fluorescence ratio, 2 showed pyrenoid localization. These 2 complements with pyrenoid Venus fluorescence were the only that rescued the CCM phenotype.

Western Blotting

Chlamydomonas cells were grown to mid-log phase (1 x 10^6 cells mL^{-1}) and then harvested by centrifugation at 5,000 x g. Total protein extraction, detection and quantification of SAGA1 in wild type and *saga1* mutant were performed using method described in (51). Protein was extracted from flash-frozen cells, normalized to chlorophyll absorbance, separated by SDS-PAGE, and western blots were performed with different antibodies.

The primary anti-SAGA1 antibody was used at a 1:10,000 dilution and the secondary horseradish-peroxidase (HRP) conjugated goat anti-rabbit (Life Technologies) at a 1:10,000 dilution. The primary anti-Rubisco antibody was used at a 1:50,000 dilution and the secondary horseradish-peroxidase (HRP) conjugated goat anti-rabbit (Life Technologies) or a goat anti-Rabbit IgG (1:10,000 dilution). To ensure even loading, membranes were stripped (Restore PLUS western blot stripping buffer, Thermo Scientific) and re-probed with anti-tubulin (1:25,000; Sigma) or anti-Histone H3 (1:10,000; abcam) followed by HRP conjugated goat anti-mouse (1:10,000; Life Technologies) or goat anti-Rabbit IgG (H+L) 800 CW (1:5,000; LiCOR).

The anti-SAGA1 antibody was raised in rabbit to the last 19 amino acids of C-terminal tail of SAGA1 (RGTGDSPTRRAFGDWRKNL-cooh) by Yenzym Antibodies (South San Francisco, California, USA). The anti-Rubisco polyclonal antibody was a gift from Prof. John Gray, University of Cambridge.

Oxygen Evolution

Chlamydomonas cells grown in Tris-phosphate medium were harvested by centrifugation at 3,000 x g for 3 min at 20°C and resuspended in 25 mM HEPES-KOH (pH 7.3) to the concentration of about 200 µg chlorophyll ml^{-1} using the equations from (57). Aliquots of cells (1 mL) were added to an OXYV1 Hansatech Oxyview System (Hansatech Instruments, King's Lynn, UK) maintained at 23°C using a circulating water bath. The chamber was sealed and illuminated with 200-300 µmol photons m^{-2} s^{-1} until the cells had depleted the internal inorganic carbon storage. When net oxygen evolution ceased, 10-µl aliquots of sodium bicarbonate were added to the cells at 30-s intervals. The cumulative concentrations of HCO_3^- after each addition were as follows: 2.5, 5, 10, 25, 50, 100, 250, 500, 1000 and 2000 µM. The rate of oxygen evolution was recorded per second using the PicoLog 1216 Data Logger (Pico Technologies, St. Neots, UK). $K_{0.5}$ values were calculated using the Michaelis-Menten equation.

Fixing and Embedding for TEM

Chlamydomonas cells were harvested by centrifugation at 5,000 x *g* and fixed for 1 h at room temperature with fixation buffer (2% (v/v) glutaraldehyde, 0.01% (v/v) hydrogen peroxide in TP medium)). Samples were then fixed and osmicated for 1 h at room temperature in an osmium mix (1% (v/v) OsO_4, 1.5% (w/v) $K_3[Fe(CN)_6]$, 2 mM $CaCl_2$). Bulk staining was performed with 2% (w/v) uranyl acetate for 1 h at room temperature. At the interval of each of the subsequent step, three 5 min washes with distilled water were performed. Samples were dehydrated progressively in 70% (v/v) and 95% (v/v) ethanol, followed by two washes in 100% ethanol and 100% acetonitrile.

The samples were then embedded in epoxy resin (34% (w/v) Quetol 651, 44% (w/v) nonenyl succinic anhydride, 20% (w/v) methyl-5-norbonene-2,3-dicarboxylic anhydride, and 2% (w/v) catalyst dimethylbenzylamine (Agar Scientific, Essex, UK)). The fixed cells were first mixed with a mixture of acetonitrile and epoxy resin at a ratio of 1:1 and left to settle overnight. Over the next two days, the cells were refreshed with 100% epoxy resin. Lastly, the samples were cured at 60°C for at least 24 h.

Sections of 50 nm thickness were prepared with a Leica Ultracut UCT (Leica Microsystems, Milton Keynes, UK), mounted on 300 mesh copper grids and counterstained with uranyl acetate followed by lead citrate (Ms. Lyn Carter, Cambridge Advanced Imaging Centre, UK).

Immunogold labeling

Samples mounted on 300 mesh nickel grids were treated with 4% (w/v) Na-meta-periodate for 15 min and 1% (w/v) periodic acid for 5 min to remove superficial osmium and unmask the epitopes. Then, the samples were blocked for 5 min in blocking buffer of 1% BSA (w/v) in high-salt Tris-buffered saline (HSTBSTT) (0.05% (v/v) Tween 20, 0.05% (v/v) Triton X-100 and 500 mM NaCl). The sections were then incubated overnight at room temperature with diluted antibody (1:1000 dilutions of rabbit anti-Rubisco antibody in blocking buffer) followed by two 5-min washes in HSTBSTT and two 5-min washes in dH_2O. To detect bound antibody, the sections were incubated with a secondary antibody (1:200 dilutions of goat anti-rabbit 15-nm gold conjugates (BBI Solutions, Cardiff, UK) in blocking buffer) for 1 h at room temperature followed by two 5-min washes in HSTBSTT and two 5-min washes in dH_2O. The sections were dried and kept for further observation.

Sections were examined using TEM and evaluation of immunogold labeling was made by using the cell counter and measurement functions of ImageJ2 (Fiji). The percentage of

aggregation of gold particles in pyrenoid and chloroplast stromal area was calculated using these equations:

Particle density in the pyrenoid,

$$d = \frac{n_i}{A_i} \quad (1)$$

Normalized number of particles in the pyrenoid,

$$N_i = (d_i - d_k) \times n_i \quad (2)$$

Calculated percentage of Rubisco aggregation in the pyrenoid,

$$\%N_i = \frac{N_i}{N_i + N_j} \times 100 \quad (3)$$

Where d = density, n = number of particles counted, N = normalized number of particles, A = measured area, i = pyrenoid, j = chloroplast stroma, k = background.

Calculation of particle density of the background of every image:

Non-pyrenoid/chloroplast area,

$$A_k = A_{all} - (A_i + A_j) \quad (4)$$

Particle density of the background of every image,

$$d_k = \frac{n_{other}}{A_k} \quad (5)$$

Where n_{other} is the number of particles counted in the non-pyrenoid/chloroplast area.

Normalized number of particles in the chloroplast stroma, j, was derived from equation (2):

$$N_j = (d_j - d_k) \times n_j$$

Indirect Immunofluorescence

10^6 cells were concentrated to a volume of 500 µl and fixed on the poly-L-lysine-coated slides for 10 min and the excess media was removed from the slides. The slides were placed in cold methanol and incubated at -20°C for about 10 min. After then, the dehydrated cells were rehydrated by incubating the slides in 1X PBS for 5 min. This step was repeated twice. After that, Permeating Solution (2% (v/v) of Triton X-100 in 1X PBS) was added to the jar and incubated for 10 min at room temperature followed by two washes in PBS-Mg solution (5 mM $MgCl_2$ in 1X PBS) for 10 min each then air-dried. About 80 µl of Blocking Solution (1% BSA, 1% cold water fish gelatin, 0.05% Triton X-100 and 0.05% Tween20 in 1X PBS) was pipetted onto each cover slip and the slides were inverted onto the cover slips followed by 30 min incubation at room temperature in the humid chambers. After that, about 80 µl of primary

antibody was added to the cells as described previously. The slides were then placed in the humid chambers and incubated overnight at 4°C.

The next day, the slides were washed thrice with 1X PBS for 5 min at room temperature. Secondary antibody was added to the slides as described previously and incubated at room temperature for 1 h. After the three washes with 1X PBS for 5 min each, the slides were then air-dried and mounted. First, about 25 µl of ProLong® Gold Antifade Reagent (Life Technologies) was added onto the slide and a cover slip was then inverted over the droplet. The mounted slides were left to dry overnight in the dark at 4°C.

Primary antibodies used for indirect immunofluorescence were anti-Rubisco (1:2,000) and anti-SAGA antibodies (1:2,000 dilution) while the secondary antibody used was the Alexa Fluor 488 Goat anti-Rabbit IgG (H+L) antibody (Invitrogen; 1:1,000).

Pyrenoid quantification using Rbcs1-mCherry fluorescence

15 Z-sections were taken per cell at 100x magnification. Cell boundaries were defined using chlorophyll autofluorescence. 3D pyrenoid reconstructions were generated from Z-stacks sections using Imaris software (Bitplane). Pyrenoids were defined as 'vesicle like objects' using mCherry fluorescence, and then quantified.

Quick-freeze deep-etch EM (QFDEEM)

150 mL of each of air-bubbled cultures and 75 mL of high CO_2- bubbled cultures were pelleted at 1,000 g for 10 min at RT to produce pellets of ~200 µL. The pellets were resuspended in 6 mL of ice-cold 10 mM HEPES buffer (pH 7) and transferred to a cold 25 mL glass flask. A freshly prepared solution of 4% glutaraldehyde (Sigma-Aldrich G7651) in 10 mM HEPES (pH 7) was added 100 µL at a time, swirling between drops, until 1.5 mL in total had been added. The mixture was then left on ice for 1 hour, with agitation every 10 min. The mixture was pelleted (1000 g, 5 min, 4° C), washed in cold HEPES buffer, pelleted again, and finally resuspended in 6 mL fresh HEPES. Samples were shipped overnight to St. Louis in 15 mL conical screw cap tubes maintained at 0-4° C.

Microscopy QFDEEM was performed as previously described in (14, 58).

SI References

14. Mackinder LCM, et al. (2016) A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proc Natl Acad Sci* 113(21):5958–5963.
45. 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* 26:1398–1409.
48. Schindelin J, et al. (2012) Fiji: An open-source platform for biological-image analysis. *Nat Methods* 9(7):676–682.
49. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675.
50. Rueden CT, et al. (2017) ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* 18(529):1–26.
51. 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.
52. Marchler-Bauer A, et al. (2017) CDD/SPARCLE: Functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 45(D1):D200--D203.
53. Sorimachi K, Le Gal-Coëffet MF, Williamson G, Archer DB, Williamson MP (1997) Solution structure of the granular starch binding domain of Aspergillus niger glucoamylase bound to β-cyclodextrin. *Structure* 5(5):647–661.
54. Li X, et al. (2016) An Indexed, Mapped Mutant Library Enables Reverse Genetics Studies of Biological Processes in *Chlamydomonas reinhardtii*. *Plant Cell* 28(2):367–387.
55. Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 292:195–202.
56. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modelling, prediction, and analysis. *Nat Protoc* 10(6):845–858.
57. Porra RJ, Thompson WA, Kriedemann PE (1989) Determination of accurate extinction coefficients and simulteneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochim Biophys Acta* 975:384–394.
58. Heuser JE (2011) The origins and evolution of freeze-etch electron microscopy. *J Electron Microsc (Tokyo)* 60(Supplement 1):3–29.