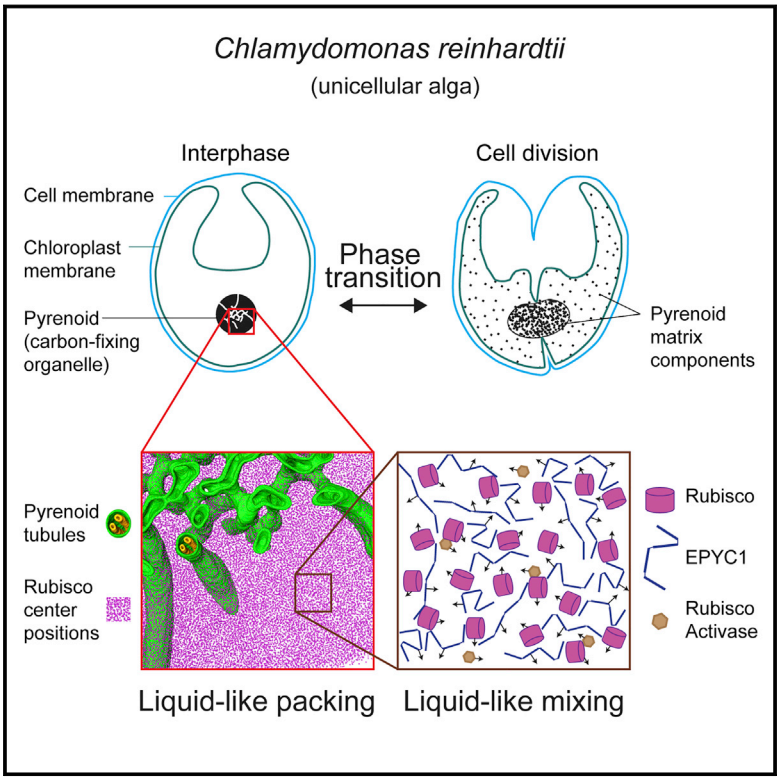
Article

Cell

The Eukaryotic CO_2-Concentrating Organelle Is Liquid-like and Exhibits Dynamic Reorganization

Graphical Abstract



Highlights

- The pyrenoid matrix is not a crystalline solid and instead behaves like a liquid
- The pyrenoid is inherited primarily by fission and can also be assembled *de novo*
- The pyrenoid undergoes a reversible phase transition during cell division
- Modeling reveals a "magic number" effect that governs phase transitions

Authors

Elizabeth S. Freeman Rosenzweig, Bin Xu, Luis Kuhn Cuellar, ..., Benjamin D. Engel, Luke C.M. Mackinder, Martin C. Jonikas

Correspondence

wingreen@princeton.edu (N.S.W.), engelben@biochem.mpg.de (B.D.E.), mjonikas@princeton.edu (M.C.J.)

In Brief

The pyrenoid, a Rubisco-containing organelle that enhances carbon fixation, mixes internally and undergoes phase transitions.

Freeman Rosenzweig et al., 2017, Cell *171*, 148–162
September 21, 2017
http://dx.doi.org/10.1016/j.cell.2017.08.008

The Eukaryotic CO_2-Concentrating Organelle Is Liquid-like and Exhibits Dynamic Reorganization

Elizabeth S. Freeman Rosenzweig,[1,2] Bin Xu,[3,11] Luis Kuhn Cuellar,[4,11] Antonio Martinez-Sanchez,[4] Miroslava Schaffer,[4] Mike Strauss,[5] Heather N. Cartwright,[2] Pierre Ronceray,[6] Jürgen M. Plitzko,[4] Friedrich Förster,[4,9] Ned S. Wingreen,[7,8,*] Benjamin D. Engel,[4,*] Luke C.M. Mackinder,[2,10,12] and Martin C. Jonikas[1,2,8,11,12,13,*]

[1]Department of Biology, Stanford University, Stanford, CA 94305, USA
[2]Department of Plant Biology, Carnegie Institution for Science, Stanford, CA 94305, USA
[3]Department of Physics, Princeton University, Princeton, NJ 08544, USA
[4]Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany
[5]Cryo-EM Facility, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany
[6]Princeton Center for Theoretical Science, Princeton University, Princeton, NJ 08544, USA
[7]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA
[8]Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA
[9]Present address: Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 Utrecht, the Netherlands
[10]Present address: Department of Biology, University of York, York YO10 5DD, UK
[11]These authors contributed equally
[12]These authors contributed equally
[13]Lead Contact
*Correspondence: wingreen@princeton.edu (N.S.W.), engelben@biochem.mpg.de (B.D.E.), mjonikas@princeton.edu (M.C.J.)
http://dx.doi.org/10.1016/j.cell.2017.08.008

SUMMARY

Approximately 30%–40% of global CO_2 fixation occurs inside a non-membrane-bound organelle called the pyrenoid, which is found within the chloroplasts of most eukaryotic algae. The pyrenoid matrix is densely packed with the CO_2-fixing enzyme Rubisco and is thought to be a crystalline or amorphous solid. Here, we show that the pyrenoid matrix of the unicellular alga *Chlamydomonas reinhardtii* is not crystalline but behaves as a liquid that dissolves and condenses during cell division. Furthermore, we show that new pyrenoids are formed both by fission and *de novo* assembly. Our modeling predicts the existence of a "magic number" effect associated with special, highly stable heterocomplexes that influences phase separation in liquid-like organelles. This view of the pyrenoid matrix as a phase-separated compartment provides a paradigm for understanding its structure, biogenesis, and regulation. More broadly, our findings expand our understanding of the principles that govern the architecture and inheritance of liquid-like organelles.

INTRODUCTION

Although pyrenoids mediate approximately one-third of global carbon fixation (Mackinder et al., 2016), the molecular structure and biogenesis of these biogeochemically fundamental organelles remain largely unknown. Pyrenoids are non-membrane-bound, proteinaceous structures that contain a matrix packed with the CO_2-fixing enzyme Rubisco. In many species, the matrix is traversed by membrane tubules that are continuous with the photosynthetic thylakoid membranes (Gibbs, 1962; Ohad et al., 1967; Engel et al., 2015). Pyrenoids are a central feature of the algal CO_2-concentrating mechanism, which supplies Rubisco with a high concentration of its substrate CO_2, enabling more efficient carbon capture than that of most land plants (Badger et al., 1998; Meyer et al., 2012). Pyrenoids are found within the chloroplasts of a diverse array of photosynthetic eukaryotes, including nearly all freshwater and marine algae, as well as a group of non-vascular plants (Brown et al., 1967; Wang et al., 2015). Although pyrenoids were among the first organelles to be scientifically documented (Vaucher, 1803), they have remained largely uncharacterized at a molecular level due to the limited availability of genetic tools for algae.

Based on results in the model unicellular green alga *Chlamydomonas reinhardtii* (*Chlamydomonas* hereafter), we recently proposed that a protein called Essential Pyrenoid Component 1 (EPYC1; also known as LCI5) links Rubisco holoenzymes together to form the pyrenoid matrix (Mackinder et al., 2016). EPYC1 localizes to the pyrenoid matrix, is of similar abundance to Rubisco, and is required for Rubisco's localization in the pyrenoid matrix, supporting a possible structural role. EPYC1 binds Rubisco, and the EPYC1 protein sequence consists of four nearly identical ~60 amino acid repeats, suggesting a model in which each of EPYC1's repeats contains a Rubisco binding site, allowing EPYC1 to link multiple Rubisco holoenzymes together. The molecular details of the EPYC1-Rubisco binding interaction, and the packing organization of EPYC1 and Rubisco in the matrix, remain unknown.

Electron micrographs of several species suggest that the densely packed pyrenoid matrix is crystalline (Holdsworth, 1968; Kowallik, 1969; Bertagnolli and Nadakavukaren, 1970), while micrographs of other species suggest the matrix is amorphous (Griffiths, 1970; Meyer et al., 2012). Our recent cryo-electron

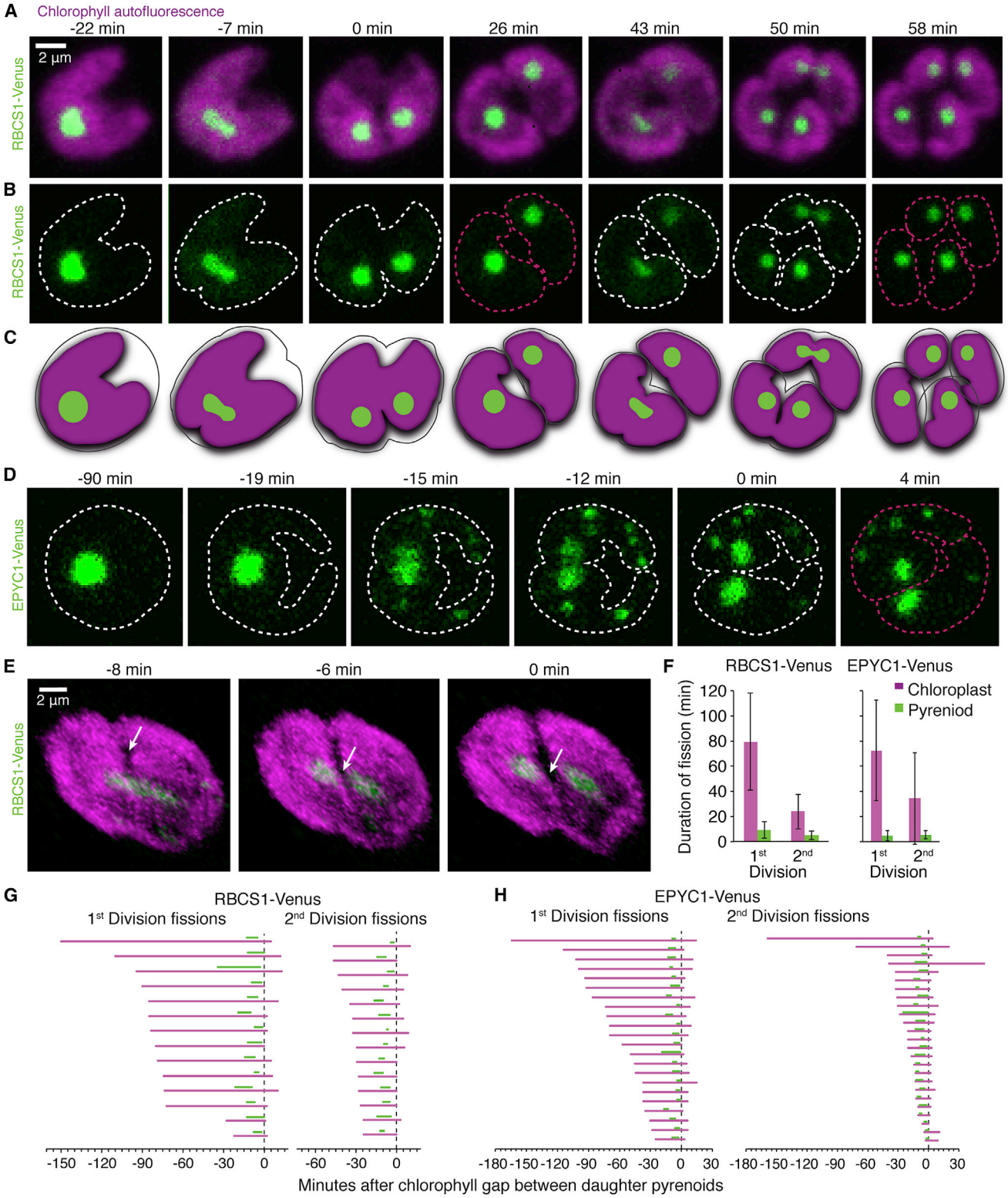


Figure 1. The Pyrenoid Is Inherited Primarily by Fission

(A and B) Confocal *Z* sum images of pyrenoid divisions by fission, with chlorophyll autofluorescence shown in magenta (A) and RBCS1-Venus in green (A and B); $t = 0$ is the first observation of a gap in chlorophyll between the daughter pyrenoids in the first division shown. Dashed curves (B) represent approximate chloroplast outlines in the mother (white) and daughter (pink) cells.

(legend continued on next page)

tomography (cryo-ET) study of *Chlamydomonas* revealed that the average packing of Rubisco molecules within the pyrenoid matrix resembles a hexagonal close-packed lattice (Engel et al., 2015). We have proposed that EPYC1 could link Rubisco holoenzymes together into this arrangement (Mackinder et al., 2016). However, if Rubisco is immobilized within a rigid lattice, it is not clear how Rubisco Activase, which regenerates Rubisco's active sites (Pollock et al., 2003) but is significantly less abundant than Rubisco in pyrenoids (McKay et al.,1991; Wienkoop et al., 2010; Mackinder et al., 2016), could reach enough Rubisco molecules to maintain efficient carbon fixation.

Moreover, pyrenoid biogenesis has not been definitively observed in living cells. Classic electron microscopy studies on fixed cells have suggested that pyrenoids may arise *de novo* (Brown et al., 1967; Retallack and Butler, 1970) in some species and by fission (Brown et al., 1967; Goodenough, 1970) in others.

Here, we describe our study of pyrenoid inheritance in living cells and how our findings inspired us to investigate the *Chlamydomonas* pyrenoid as a liquid. Examination of pyrenoid ultrastructure with improved *in situ* cryo-ET revealed that the pyrenoid matrix is not crystalline, but exhibits liquid-like local order. Further live-cell fluorescence experiments showed that the matrix mixes internally and disperses into the surrounding stroma during cell division, demonstrating that the matrix exhibits liquid-like properties. Finally, using a simple model inspired by the pyrenoid, we predict that phase transitions in some multicomponent, liquid-like biological systems are governed by changes in the ratio of available multivalent binding sites on the constituent macromolecules. This work provides insights into both algal CO_2-concentrating mechanisms and liquid-like biological systems.

RESULTS

The Pyrenoid Exhibits Both Fission and *De Novo* Assembly

To enable the first observations of pyrenoid matrix dynamics in living cells, we expressed pyrenoid matrix proteins tagged with the fluorescent protein Venus and imaged them in 3D with fluorescence time-lapse microscopy during photoautotrophic growth. We tracked inheritance of the pyrenoid matrix by monitoring Venus-tagged Rubisco small subunit 1 (RBCS1) or Venus-tagged EPYC1 and recorded chlorophyll autofluorescence to follow cellular orientation and chloroplast division. The fusion proteins localized to the pyrenoid (Figures 1A and 1D) and did not measurably perturb pyrenoid function (Figures S1A and S1B). During our observation, cells often divided twice in rapid succession, which is typical for *Chlamydomonas* (Cross and Umen, 2015).

We observed that approximately two-thirds of daughter cells inherited their pyrenoid by elongation and then fission of the mother cell's pyrenoid (Figures 1, 2E, 2F, and S2; Movie S1). The duration of pyrenoid fission, defined as the time between visible elongation of the mother pyrenoid and separation of the daughter pyrenoids, was ~7 min, much faster than chloroplast division (~30–80 min, Figure 1F). All pyrenoid fissions occurred during the final minutes of chloroplast division (Figures 1G and 1H).

In all 88 cell divisions in which pyrenoid fission was observed, the completion of pyrenoid division was quickly followed by a gap in the chlorophyll signal between the daughter pyrenoids, generated by the chloroplast cleavage furrow (Figures 1E, 1G, and 1H). Conversely, in the 44 cell divisions in which the furrow did not bisect the pyrenoid, pyrenoid fission did not occur. These findings support the hypothesis that pyrenoid fission is driven by the constrictive force of the chloroplast division furrow (Goodenough, 1970).

In cases where pyrenoid fission did not occur, one of the daughter cells typically inherited the mother cell's whole pyrenoid (~16%–21% of daughter cells; Figures 2A–2F; Movie S2), leaving the other daughter without a pyrenoid. Some of the cells that failed to inherit a pyrenoid remained without a visible pyrenoid for the duration of observation (~8%–9% of all daughter cells; Figures 2A, 2C, 2E, and 2F; Movie S2). In other cases, one or more fluorescent puncta appeared *de novo* and grew or coalesced into an apparent pyrenoid (~6%–7% of all daughter cells; Figures 2B, 2D, and 2E–2I; Movie S3). The proportions of each inheritance pattern in RBCS1-Venus and EPYC1-Venus were not significantly different (Figures 2E and 2F), indicating these patterns are not artifacts related to a particular construct.

Several characteristics of pyrenoid behavior during division are reminiscent of liquids. Toward the end of fission, a "bridge" of matrix material connecting the two lobes of a dividing pyrenoid is briefly visible (Figure 3). After the bridge ruptures, daughter pyrenoids quickly revert to spherical shapes, similar to the behavior of liquid droplets (Stone, 1994; Yanashima et al., 2012). Furthermore, during apparent *de novo* pyrenoid formation, we typically observed that smaller puncta shrank while larger ones grew until the cell contained a single pyrenoid (Figures 2B, 2D, and 2G–2I)—much like Ostwald ripening in systems containing multiple liquid droplets (Hyman et al., 2014)—indicating that components likely exchange between the puncta. Both fission and *de novo* formation have been observed in established liquid-like organelles such as *C. elegans* P granules (Brangwynne et al., 2009; Saha et al., 2016) and *Xenopus* oocyte nucleoli (Brangwynne et al., 2011). The similarity of the pyrenoid's behavior to that of such phase-separated biological liquid droplets suggested the possibility that pyrenoids may also be liquid, rather than static

(C) A cartoon of the approximate locations of the pyrenoid (green), chloroplast (magenta), and cell membrane (black outline).
(D) Example of pyrenoid fission in EPYC1-Venus, annotated as in (B).
(E) Example of the progressing chloroplast cleavage furrow (arrows) appearing to separate daughter pyrenoids. Images are 2D snapshots of 3D *Z* stack reconstructions.
(F) Average and SD of the durations of chloroplast (magenta) and pyrenoid (green) fissions in RBCS1-Venus (left; n = 28 1st and 2nd divisions) and EPYC1-Venus (right; n = 22 1st and 25 2nd divisions).
(G and H) Duration and relative timing of chloroplast (magenta) and pyrenoid (green) division for the pyrenoid fissions plotted in (F). Each bar represents a different division.
See also Figures S1 and S2 and Movie S1.

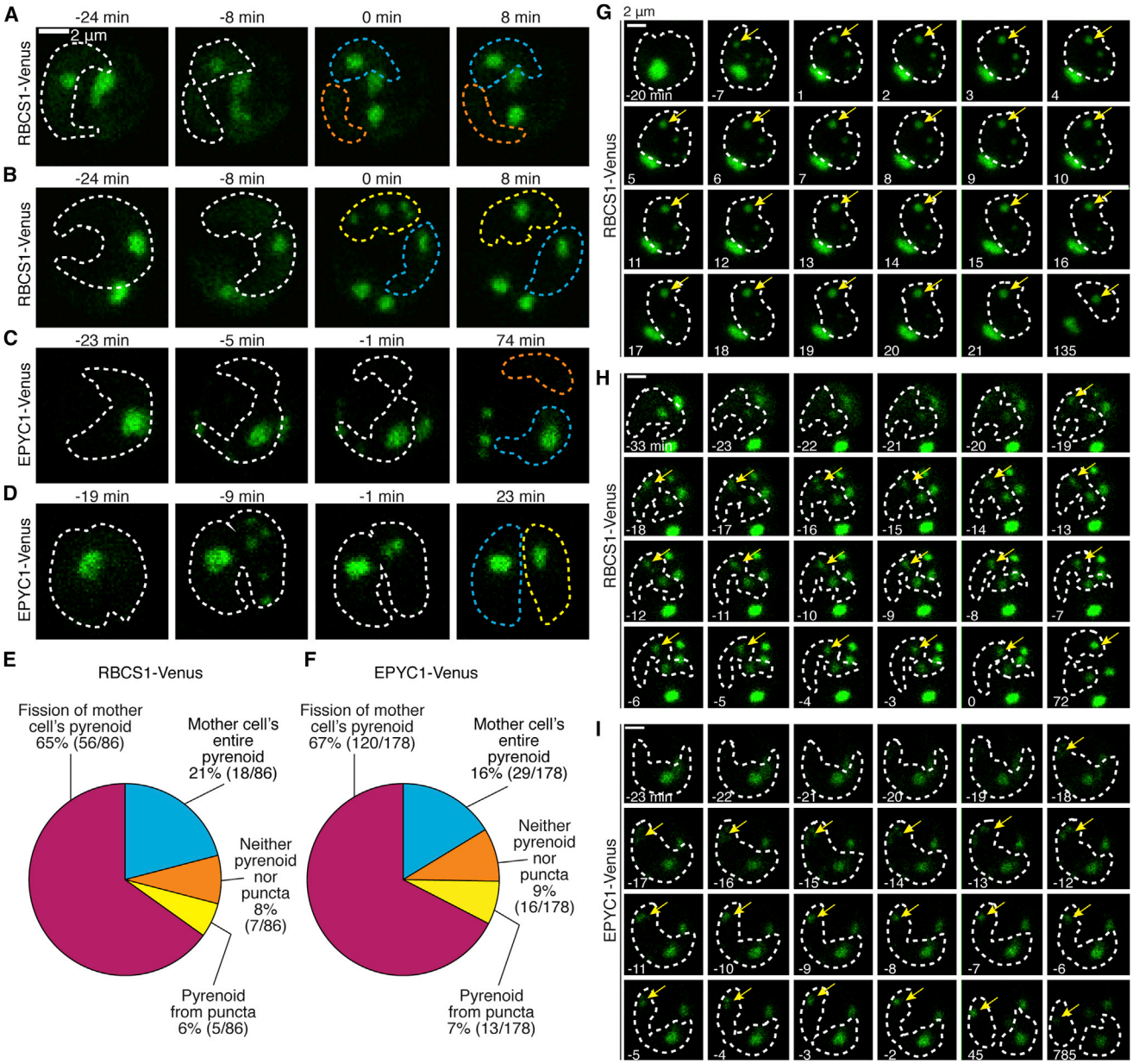


Figure 2. Pyrenoids Can Also Be Inherited by Other Means

(A–D) Examples of other types of pyrenoid inheritance patterns observed in RBCS1-Venus (A and B) and EPYC1-Venus (C and D) cells. (A and C) One daughter (blue) inherits an entire pyrenoid from the mother cell (white) and another daughter (orange) inherits neither a pyrenoid nor puncta. (B and D) One daughter (blue) inherits the entire pyrenoid, and puncta appear in the other daughter (yellow) and coalesce into a new pyrenoid.

(E and F) Proportion of observed RBCS1-Venus (E) and EPYC1-Venus (F) daughter cells that exhibited each observed inheritance pattern; the distribution of inheritance patterns in EPYC1-Venus cells was not significantly different from that of RBCS1-Venus cells (Chi-square test, $p = 0.8$).

(G–I) Stills from time course image captures in which pyrenoids were observed to grow or coalesce from puncta that appeared in the chloroplast stroma during division. (G and H) RBCS1-Venus. (I) EPYC1-Venus. Images are 2D projections of the sum of pixel values in each channel in a *Z* stack through the whole cell at each time point. The chloroplast of the dividing cell of interest in each series is outlined in white. Arrows point to growing or coalescing pyrenoids. $t = 0$ is defined as the first minute at which the daughter chloroplasts are observed to be distinct in 3D.

See also Figures S1 and S2 and Movies S2 and S3.

solids or crystals, as has been proposed (Holdsworth 1968; Kowallik 1969; Bertagnolli and Nadakavukaren 1970). We therefore decided to investigate the fine-scale structure and potential liquid-like dynamics of the pyrenoid matrix.

The Matrix Is Not Crystalline and Has Liquid-like Organization

The different reports of crystalline and amorphous pyrenoid matrices emerged from classical electron micrographs

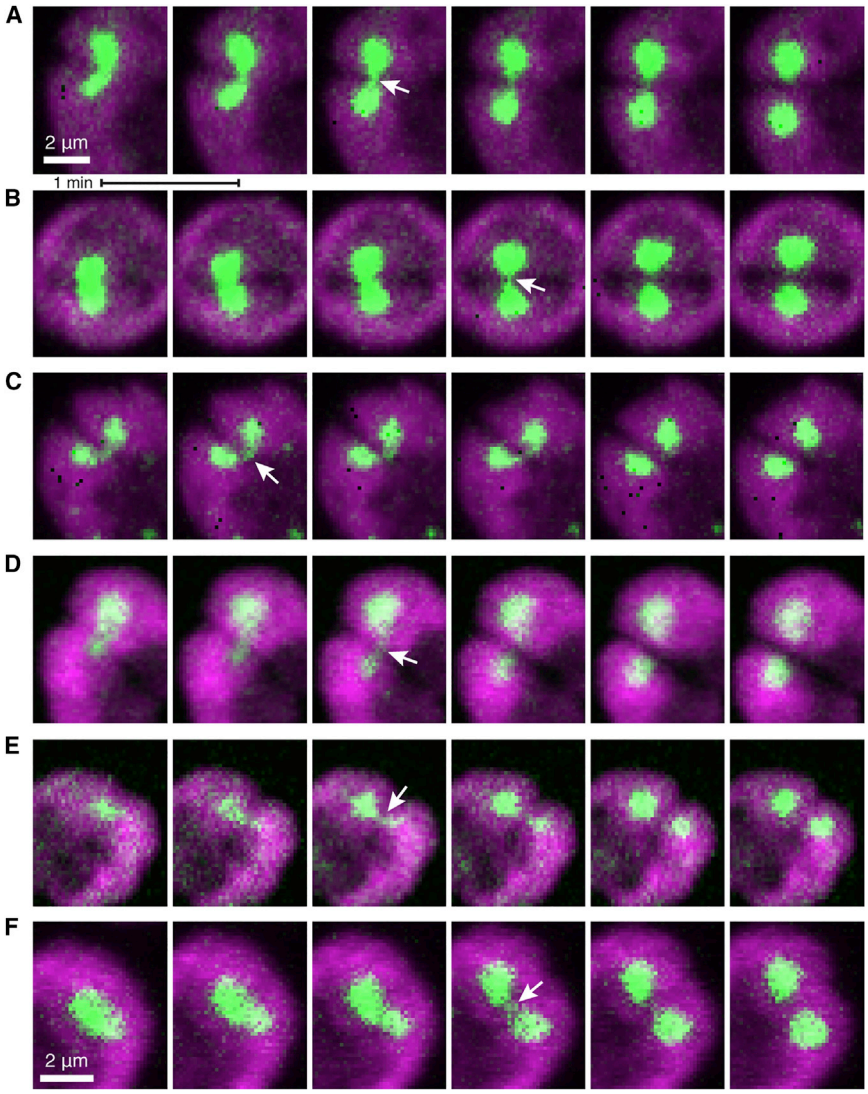


Figure 3. A "Bridge" of Matrix Material Connects Nascent Daughter Pyrenoids during Fission

(A–E) Examples of pyrenoid fissions in RBCS1-Venus cells, with each row corresponding to a separate cell division. Magenta is chlorophyll autofluorescence; green is RBCS1-Venus. Frames are 1 min apart. Scale bar in (A) represents scale in (A)–(E).

(F) Example of pyrenoid fission in an EPYC1-Venus cell. Magenta is chlorophyll autofluorescence and green is EPYC1-Venus.

Images represent the sum of pixel values in each channel from a *Z* stack through the whole cell. See also Movie S1.

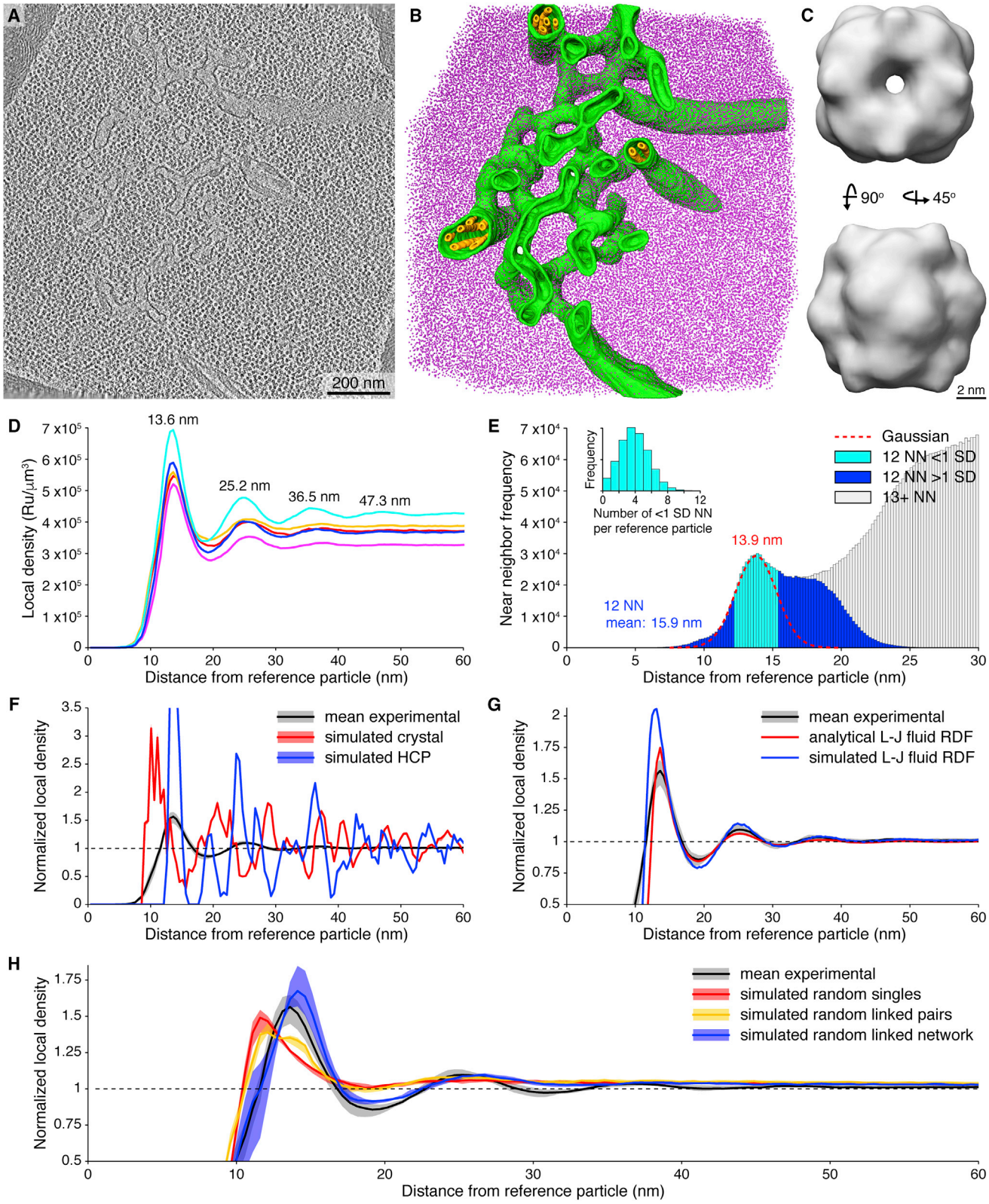
(Holdsworth, 1968; Kowallik, 1969; Bertagnolli and Nadakavukaren, 1970; Griffiths, 1970). However, these micrographs have limited resolution, and the native arrangement of Rubisco may be compromised by artifacts from the sample preparation procedure, which involves chemical fixation, dehydration, plastic embedding, and heavy metal staining (Crang and Klomparens, 1988).

To avoid such artifacts, we analyzed the positions of individual Rubisco holoenzymes within the *Chlamydomonas* pyrenoid matrix by *in situ* cryo-ET (Figure 4A; Movie S4), which provides 3D views of the native cellular environment at molecular resolution (Asano et al., 2016; Schaffer et al., 2017). In order to achieve the complete freezing vitrification required for high-quality cryo-ET, we examined *mat3-4* cells (Umen and Goodenough, 2001), which are smaller due to premature cell division but have normal pyrenoid morphology and a functional CO_2-concentrating mechanism. The high fidelity of direct electron detector imaging enabled us to localize ~97.5% of the Rubisco holoenzymes within five pyrenoid tomograms, totaling 192,100 particles (Figures 4B, 4C, S3, and S4A–S4C).

As a quantitative metric for the packing of Rubisco within the matrix, we calculated the local concentration of holoenzymes in an expanding shell around every Rubisco. Our analysis sampled only the pyrenoid matrix, avoiding potential effects from the traversing membrane tubules and pyrenoid borders. Quasiperiodic near-neighbor peaks of Rubisco density eroded and then vanished as distance increased from the reference particle (Figure 4D), suggesting that the pyrenoid lacks the long-range order characteristic of a crystal.

We previously reported that the average local neighborhood in the pyrenoid matrix resembles a lattice of hexagonal close packing (HCP) with ~15 nm between particle centers (Engel et al., 2015). However, due to limitations in imaging, this conclusion was not reached by determining the positions of individual Rubisco holoenzymes, but rather by averaging together larger volumes that contained neighborhoods of multiple Rubisco particles. The improved quality of our new tomograms enabled us to pinpoint the positions of each Rubisco, revealing that the HCP we previously observed is actually an average of a much more heterogeneous environment. In HCP, the 12 nearest neighbors to any given particle are equidistant. However, when we plotted the center-to-center distances between holoenzymes and their 12 nearest neighbors, we observed a specific peak corresponding to 4.4 ± 1.8 neighbors that were 13.9 ± 1.5 nm away, with most of the remaining neighbors at a range of longer distances, for an average distance of 15.9 ± 2.9 nm (Figure 4E). The diameter of Rubisco is ~10–13 nm, depending on orientation (Taylor et al., 2001); thus, there are ~1–4 nm between a Rubisco and its nearest neighbors in the specific peak.

To directly compare the observed Rubisco packing to that of a crystal, we performed the same local density analysis on



(legend on next page)

simulated data in which the exact volume of the pyrenoid matrix from the tomograms was replaced with one of two crystalline matrix models: one based on HCP with 13.9 nm between particle centers, and the other based on the packing observed in crystallized Rubisco (Taylor et al., 2001) (Figures 4F, S4E, and S4F). Compared to the experimental data, both crystalline simulated datasets produced much taller short-range peaks along with pronounced long-range peaks. Furthermore, the Rubisco concentration observed in the pyrenoid matrix (377,856 ± 37,823 Rubiscos/μm^3; 628 ± 63 μM) was 28% lower than in the HCP simulated data (526,039 ± 709 Rubiscos/μm^3; 874 ± 1 μM) (Figure S3). Therefore, we conclude that the arrangement of Rubisco in the pyrenoid is neither HCP nor that observed in crystallized Rubisco.

To explore whether the pyrenoid matrix has liquid-like organization, we next compared the observed Rubisco packing to the distribution of molecules within known fluids. The local density of Rubisco within the pyrenoid matrix fit well to the radial distribution function of a simple model for liquid molecular interactions, the Lennard-Jones fluid (Johnson et al., 1993) (Figure 4G), determined both analytically and by molecular dynamics simulations (see the STAR Methods). Additionally, similar radial distribution functions have been measured for colloidal liquids (Gu et al., 2010; Thorneywork et al., 2014). Therefore, we conclude that the radial distribution of Rubisco within the matrix resembles that of a liquid.

We sought to determine whether effects at the macromolecular scale could produce the liquid-like organization that we observed in the pyrenoid matrix. We first examined whether random packing of Rubisco particles could yield the local density seen in our experimental data. Using Monte Carlo simulations, we randomly placed the same number of Rubisco particles as the experimental data throughout identical pyrenoid matrix volumes (Figures 4H and S4G). The local density of this simulated data fit our experimental data poorly, showing that the organization of pyrenoid matrix is not random.

We next randomly mapped Rubisco particles into the matrix volumes while imposing specific distance constraints between either linked pairs of particles or interconnected linked networks (Figures 4H and S4H–S4J). As we increased the number of constraints on particle spacing from the single constraint between paired particles to the numerous constraints between particles in a network, the local density of Rubisco much more closely resembled the observed distribution in the pyrenoid matrix. This result suggests that holoenzymes could be linked in a network with a favored Rubisco-to-Rubisco distance. EPYC1 may be the molecular structure that enforces this spacing.

The Pyrenoid Mixes Internally

The Rubisco packing distribution we observed is consistent with either a liquid or an amorphous solid state. In order to discriminate between these states, a dynamic assay is necessary. The most direct test of liquid-like behavior is whether the constituent components exhibit internal mixing (Brangwynne et al., 2009), which is traditionally assayed *in vivo* by fluorescence recovery after photobleaching (FRAP) (van Royen et al., 2009; Hyman et al., 2014).

Strikingly, our FRAP experiments revealed that the pyrenoid matrix mixes internally (Figure 5). After bleaching approximately half the volume of the matrix (Figures S5A–S5C), we observed signal re-homogenization on a timescale of ~20 s (Figures 5A, 5C, 5G, and 5I–5L; Movie S5), similar to that of established liquid-like compartments (Brangwynne et al., 2009; Li et al., 2012; Kroschwald et al., 2015; Saha et al., 2016). During recovery, the signal from the unbleached region decreased as the signal in the bleached region increased, and a "wave" of fluorescence could be seen moving from the unbleached to the bleached region (Figures 5C and 5E), suggesting that recovery is primarily due to internal rearrangements, rather than import of stromal RBCS1-Venus. In contrast, pyrenoids in chemically crosslinked cells showed no recovery (Figures 5B, 5D, 5F, and 5G; Movie S5).

Further FRAP experiments demonstrated that the other major components of the pyrenoid matrix, EPYC1 and Rubisco Activase (McKay et al., 1991) (RCA1), are also mobile (Figures 5G–5L and S5D–S5F). Differences in the initial recovery rates of the three matrix components (Figure 5H) suggest that they may have distinct mobilities within the matrix, apparently inversely

Figure 4. The Pyrenoid Matrix Is Not Crystalline but Exhibits Short-Range Liquid-like Order

(A) Slice through a tomographic volume of the native *Chlamydomonas* pyrenoid.

(B) Segmentation of the tomogram shown in (A) with localized positions of 46,567 Rubisco holoenzymes (magenta) mapped into the volume. Green and yellow: pyrenoid tubule membranes.

(C) *In situ* subtomogram average of Rubisco (16.5 Å resolution; Figure S4A) generated from 30,000 particles extracted from the tomogram shown in (A).

(D) The local density of neighbor Rubisco particles as a function of the distance from the reference particle. Each line represents a separate tomogram, showing the sum of the local densities around every Rubisco. The distances to peaks of high local Rubisco concentration are indicated.

(E) Histogram of distances from reference particles to their nearest neighbors (NN), summed from all five tomograms. Red dashed line, Gaussian distribution fit to the 13.9 nm NN peak; light blue bars, distance to the 12 NN within 1 SD (<1 SD) of the 13.9 nm peak; dark blue bars, distance to the 12 NN beyond 1 SD (>1 SD) from the 13.9 nm peak; gray bars, distance to further (13+ NN) neighbors. Mean distance to the 12 NN = 15.9 nm. Inset: distribution of the number of neighbors per reference particle (mean = 4.4 neighbors) that are <1 SD from the 13.9 nm peak.

(F) The normalized local density of neighbor particles (local density divided by the global density), showing the mean value ± 99% confidence interval (CI) of the experimental data (black) compared to crystalline simulated data generated within the same tomogram volumes (Figures S4E and S4F). Red, crystal structure packing (Taylor et al., 2001); blue, 13.9 nm-spaced HCP.

(G) The mean value ± 99% CI of the experimental data's normalized local density (black) fit with the radial distribution function of a Lennard-Jones fluid generated by an analytical model (Morsali et al., 2005) (red) and by molecular dynamics simulations (Plimpton, 1995) (blue).

(H) The normalized local density of neighbor particles, showing the mean value ± 99% CI of the experimental data (black) compared to random simulated data generated within the same tomogram volumes (Figures S4G–S4J). Red, single particles; yellow, pairs linked by 13.9 ± 1.5 nm; blue, linked networks.

See also Figures S3 and S4 and Movie S4.

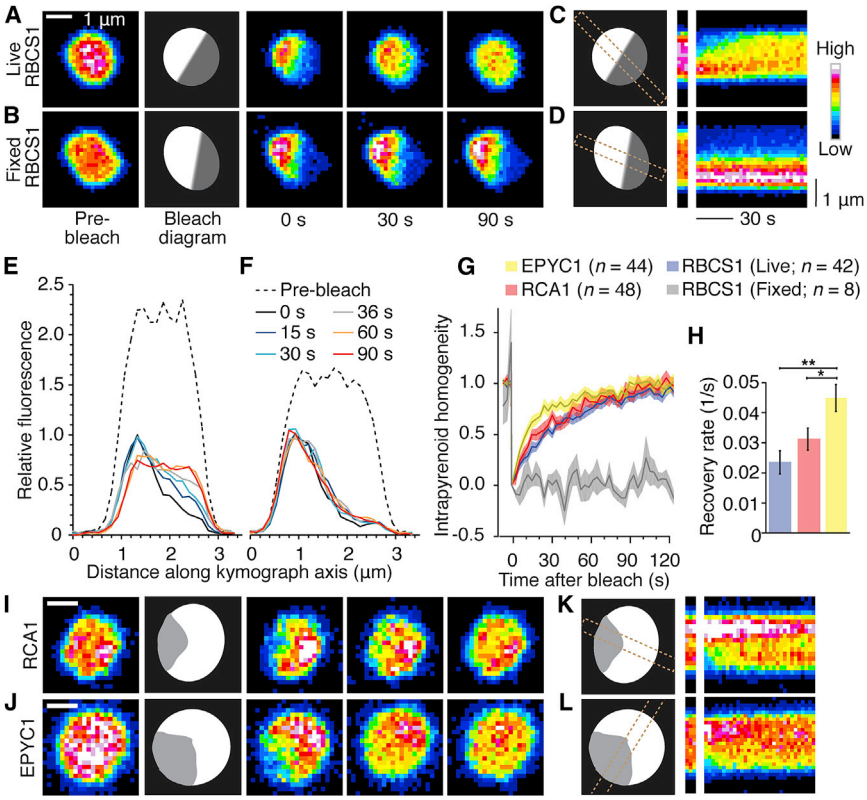


Figure 5. Pyrenoid Matrix Components Mix Internally

(A and B) FRAP in live (A) and fixed (B) RBCS1-Venus pyrenoids. Cartoons depict the approximate bleached region (dark gray). Different intensity display scales are used in the pre- and post-bleach image sets.

(C and D) Kymographs of the live (C) and fixed (D) pyrenoids shown in (A) and (B), respectively. From left to right: pyrenoid cartoon showing the region used to create the kymographs (dashed rectangle), the pre-bleach section of the kymographs, and the post-bleach kymographs.

(E and F) Fluorescence recovery occurs from within the pyrenoid in live pyrenoids (E), but does not occur in fixed pyrenoids (F). The x axis is µm along the dashed regions in (C) and (D).

(G) Average fluorescence recovery profiles ± SEM for pyrenoids in live RBCS1-Venus (blue), RCA1-Venus (red), or EPYC1-Venus (yellow) cells and in fixed RBCS1-Venus cells (gray).

(H) Average recovery rates ± SEM over the first 12 s in (G). $^{**}p < 0.005$; $^{*}p < 0.05$ (one-way ANOVA and post hoc Bonferroni means comparison).

(I–L) Examples of half-pyrenoid FRAP in live RCA-Venus (I and K) and EPYC1-Venus (J and L) cells, with images from the recovery time courses (I and J) and corresponding kymographs (K and L) as shown in (A)–(D). Physical and temporal scales are the same as in (A)–(D).

See also Figure S5 and Movie S5.

proportional to their molecular weights (Rubisco holoenzymes are ~540 kDa, RCA1 hexamers are ~270 kDa, EPYC1 is ~35 kDa). The different mobilities suggest that although these matrix components can bind to each other, the major mobile unit is not a complex between EPYC1 and Rubisco; thus, EPYC1-Rubisco binding interactions (Mackinder et al., 2016) are transient. Alongside our observations of liquid-like fission (Figures 1 and 3) and local order (Figure 4), the internal mixing of matrix components demonstrates that the *Chlamydomonas* pyrenoid matrix behaves like a liquid.

The Pyrenoid Disperses during Cell Division

A remarkable property of liquid-like protein compartments is their ability to transition between an aggregated liquid phase and a dispersed soluble phase (Brangwynne et al., 2009; Li et al., 2012). Surprisingly, our experiments revealed that the *Chlamydomonas* pyrenoid matrix also appears to undergo such a phase transition during division. A portion of the RBCS1-Venus and EPYC1-Venus signals rapidly dispersed from the pyrenoid matrix into the stroma for ~20 min near the end of chloroplast and pyrenoid division (Figures 6 and S6). During this time, the Venus signal in the chloroplast stroma rose dramatically, while the signal in the pyrenoid dropped by ~35%–50%. The Venus intensity in the pyrenoid subsequently recovered, coincident with a reduction in the stromal signal (Figures 6B and 6C). The increased stromal signal during chloroplast division is unlikely to be due to new protein synthesis because these transitions occurred on the timescale of 1–5 min and were not associated with an increase in the total fluorescence signal (Figures 6B and S6). Pyrenoid fission nearly always occurred during the time of increased stromal signal (Figures 6B, 6D, and S6), raising the intriguing possibility that the partial dispersal may be associated with a decrease in matrix surface tension or viscosity, which could facilitate the progression of the cleavage furrow through the pyrenoid.

We observed pyrenoid component dispersal regardless of the mode of pyrenoid inheritance (Figure S6). During dispersal, puncta of matrix material often appeared transiently throughout the stroma (Figures 1D, S1C, and S1D; Movie S3). These *de novo* matrix protein puncta may correspond to the "dense regions" that were observed in electron micrographs of dividing *Chlamydomonas* cells (Goodenough, 1970). In some daughter cells that failed to inherit part or all of the maternal pyrenoid during chloroplast division, we observed that these puncta grew into apparent new pyrenoids (Figures 2B, 2D, and 2G–2I). Our observations suggest a model in which the dispersal of the building blocks of the matrix serves as a redundant mechanism to fission, facilitating equal partitioning of the pyrenoid matrix to daughter chloroplasts. The inheritance of dispersed pyrenoid components enables rapid *de novo* pyrenoid assembly in daughter chloroplasts that fail to inherit a pyrenoid by other means.

A "Magic Number" Effect in Multicomponent Phase Separation

The liquid-like nature of the pyrenoid fits well with the principle that biological liquid compartments are formed by weak

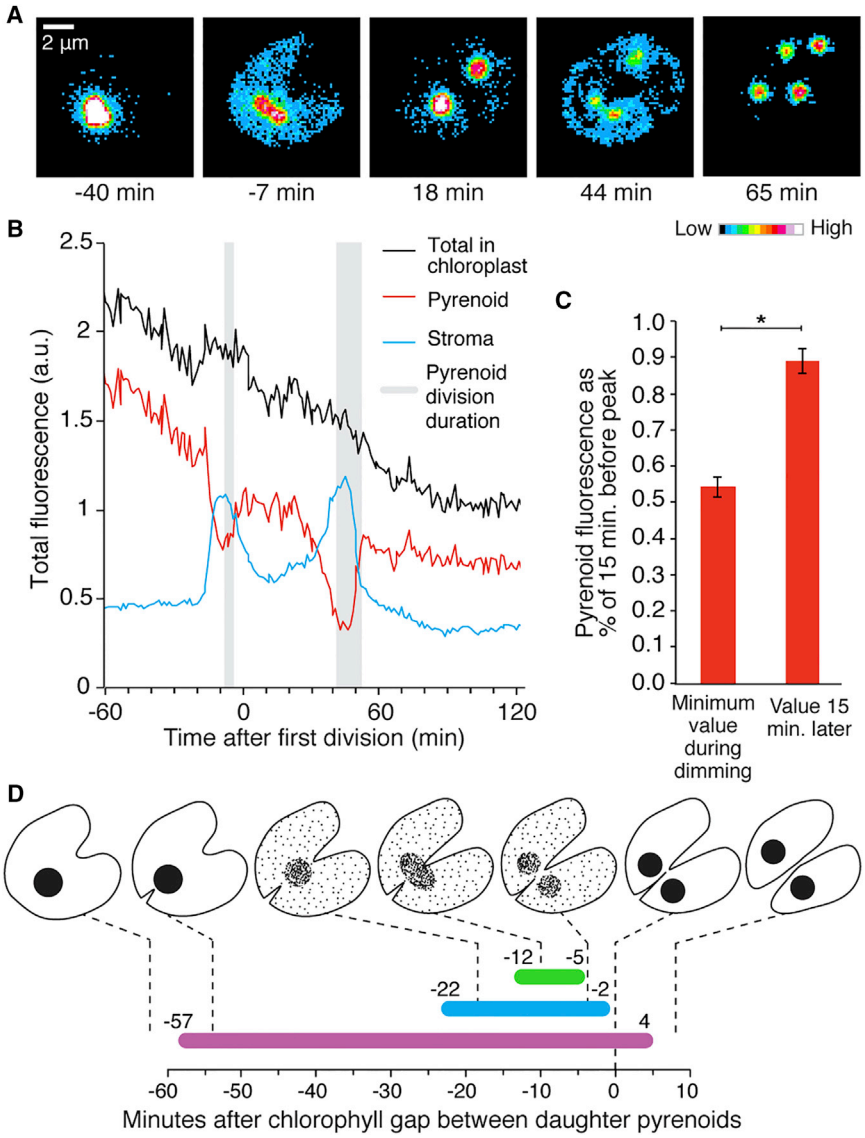


Figure 6. The Pyrenoid Matrix Disperses and Re-aggregates during Division

(A) Heatmaps of the RBCS1-Venus signal during the divisions in Figures 1A–1C. Times are defined as in Figure 1.

(B) Raw signal from (A) plotted over time by regions of interest, representing the sum through the whole z stack in each masked region over time. Times of pyrenoid divisions are highlighted in gray.

(C) The average signal in the pyrenoid during division is significantly lower than that 15 min later (Wilcoxon matched-pairs signed-ranks test; $^{*}p \leq 2 \times 10^{-6}$; n = 31), shown ± SEM.

(D) Timeline of an average cell division with pyrenoid fission. Chloroplast division (magenta), pyrenoid dissolution (cyan), and pyrenoid fission (green) are displayed relative to the moment the chloroplast division furrow passes between the daughter pyrenoids (t = 0). Cartoons depicting the aggregation state of the pyrenoid matrix are shown above each stage, with the chloroplast outlined in black, aggregated matrix components shown as filled black circles, and partially dispersed matrix components as speckles.

See also Figure S6 and Movies S1, S2, and S3.

multivalent binding between constituent proteins (Li et al., 2012). The two major constituents of the pyrenoid, Rubisco and EPYC1, likely have multiple binding sites for each other. The octameric symmetry of the Rubisco holoenzyme makes it plausible that Rubisco has eight binding sites for EPYC1. Additionally, the four repeats of the EPYC1 protein suggest that EPYC1 has four binding sites for Rubisco (Mackinder et al., 2016). We have previously proposed that dynamic regulation of the availability of EPYC1's binding sites could cause a change in the aggregation state of the pyrenoid (Mackinder et al., 2016).

In the model liquid-like system involving binding between the SRC homology 3 (SH3) domain and a proline-rich motif ligand (PRM), it has been established that decreasing the number of binding sites on flexible linker proteins promotes a phase transition from an aggregated to a dispersed state (Li et al., 2012). Surprisingly, modeling of the EPYC1-Rubisco interaction predicts the existence of an additional effect that governs phase transitions in multicomponent liquid-like biological systems.

Because the binding mechanism between Rubisco and EPYC1 is uncharacterized, we developed a simplified computational model to investigate how interactions between these two components may drive pyrenoid aggregation and dissolution. In our model, "Rubisco holoenzymes" and "EPYC1 molecules" occupy a 2D square grid. Each Rubisco holoenzyme is modeled as a 2 × 4 rectangle of 8 binding sites for EPYC1, and each EPYC1 is modeled as a flexible chain of 4 binding sites for Rubisco. Each binding site on Rubisco can bind to a single EPYC1 site and vice versa. Binding occurs when a Rubisco site and an EPYC1 site occupy the same grid position. While our model is simplified and abstract, it allows the observation of certain fundamental behaviors.

Our simulation produced the expected increased aggregation when we increased the number of binding sites on EPYC1 from 4 to 5 (Figures 7B and 7C). However, contrary to the established paradigm for liquid-like systems, we also observed increased aggregation when we decreased the number of binding sites on EPYC1 from 4 to 3 (Figures 7A and 7B).

Further investigation into this behavior revealed a "magic number" effect in the simulated EPYC1-Rubisco system. Magic numbers occur in various contexts in chemistry and physics (Sakurai et al., 1999; Steppenbeck et al., 2013), in which a certain number of particles form an unusually stable state, such as filled electron shells in atoms. However, to our knowledge, magic numbers have not previously been observed in a biological

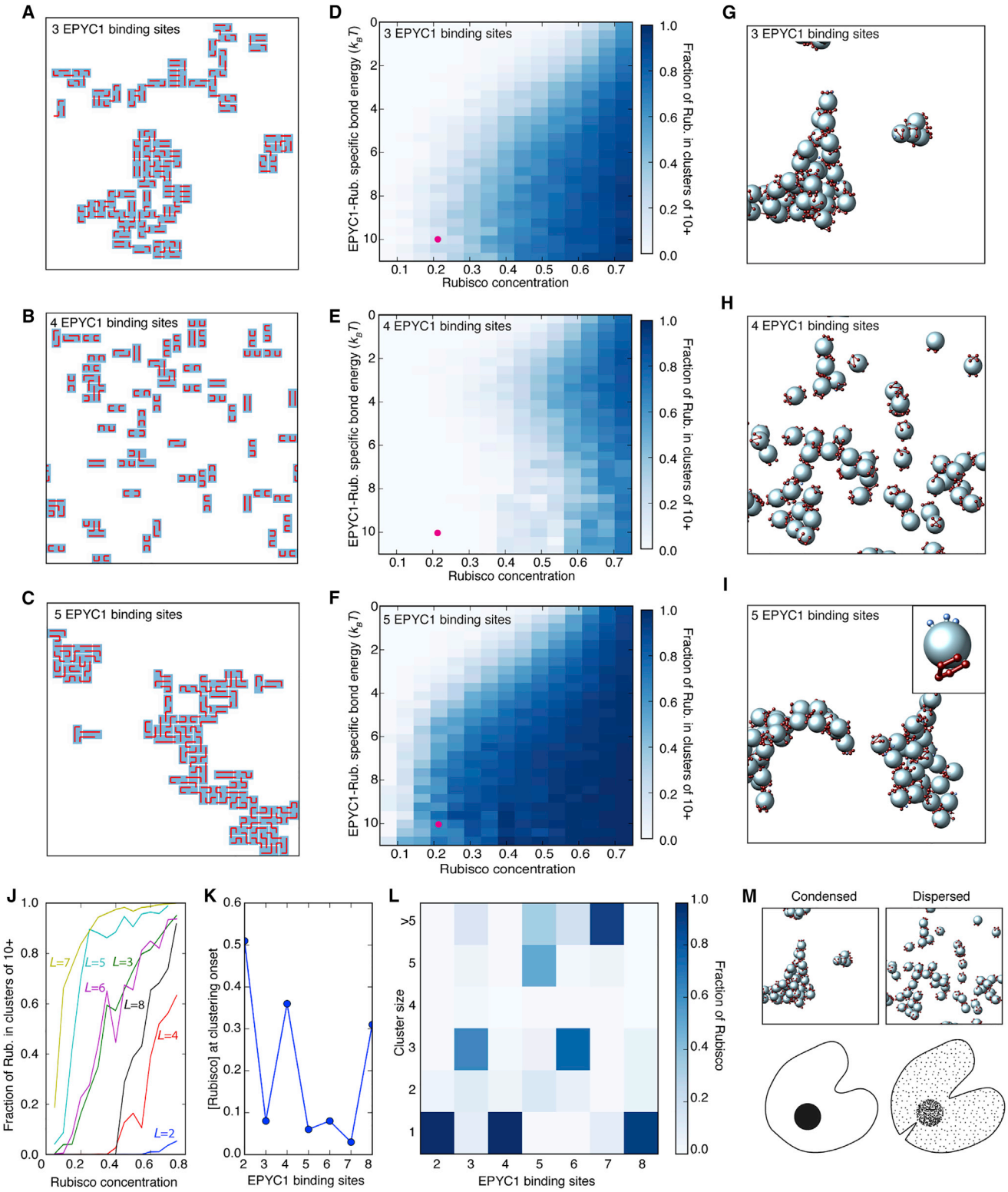


Figure 7. Simulations of an EPYC1-Rubisco System Reveal an Effect of Binding Site Stoichiometry on the Aggregation State

(A–C), Snapshots of simulations with 3 (A), 4 (B), and 5 (C) Rubisco binding sites on EPYC1. "Rubiscos" (blue rectangles) and "EPYC1s" (red polymers) bind when they occupy the same sites in a 2D grid. Snapshots are from simulations with 10 k_BT specific bonds and 0.1 k_BT nearest-neighbor non-specific bonds.

(legend continued on next page)

system. In our model, the magic number effect arises because all 8 binding sites of one Rubisco can be exactly saturated by two EPYC1s with 4 binding sites each to form a stable trimer, resulting in minimal aggregation. When the magic number stoichiometry is broken, either with more or fewer binding sites on EPYC1, larger aggregates form.

To ensure that the magic number effect is not an artifact of the lattice geometry or of two dimensions, we also simulated the system using a more computationally intensive but more realistic three-dimensional, off-lattice representation. Similar to the lattice simulation results, we observed more aggregation when we increased the number of binding sites on EPYC1 from 4 to 5 (Figures 7H and 7I) and when we decreased it from 4 to 3 (Figures 7G and 7H). Additionally, small trimer complexes containing one Rubisco and two EPYC1s were commonly observed when the number of binding sites on EPYC1 was 4, but these complexes were rarely seen when the number of binding sites on EPYC1 was 3 or 5. Thus, the magic number effect persists in three dimensions and off-lattice. Furthermore, the effect is robust for a wide range of protein concentrations and model parameters (Figures 7D–7F and S7; Table S1; Movie S6).

To explore the generality of this effect, we used the 2D lattice model to consider an even wider range of the number of Rubisco binding sites on EPYC1 (Figures 7J–7L) and found additional magic numbers at 2 and 8 binding sites per EPYC1. When the number of binding sites on EPYC1 is a magic number, small heterocomplexes are favored; consequently, higher protein concentrations are required in order to form large aggregates (Figures 7J–7L). These results are robust to constraints such as restricting EPYC1 to only bind one end of each Rubisco (Figure S7). Thus, in addition to the previously established general trend that increasing the number of binding sites in a flexible linker protein promotes phase separation, there are strong magic number effects that impact the phase diagram.

DISCUSSION

Phase Separation Helps Explain Known Pyrenoid Physiology

Phase separation provides a framework for understanding the structure, biogenesis, and inheritance of the pyrenoid. Our observations suggest that the re-localization of Rubisco from the pyrenoid to the stroma in response to high CO_2 (Borkhsenious et al., 1998) or darkness (Mitchell et al., 2014) is a phase transition of the pyrenoid matrix. The use of phase transitions could allow rapid reorganization of Rubisco to enhance CO_2 fixation in fluctuating environmental conditions. Furthermore, the fluidity of the pyrenoid matrix resolves the paradox of how Rubisco Activase chaperones can efficiently access the more abundant Rubisco active sites throughout the matrix.

EPYC1-Rubisco Interactions May Drive Pyrenoid Phase Separation

The pyrenoid's liquid-like nature may be mediated by binding between Rubisco and EPYC1's sequence repeats. Each EPYC1 repeat displays low complexity, with 73% of the repeat regions consisting of just alanine, serine, proline, and arginine (Mackinder et al., 2016). Low complexity domains and repeat regions have been widely implicated in mediating the liquid-like properties and phase separation of nonmembrane-bound organelles (Li et al., 2012; Mitrea and Kriwacki, 2016). The exact nature of the binding between Rubisco and EPYC1 is unknown, but it is likely that surface-exposed α-helices on the Rubisco small subunit play a role in the binding interaction because a specific sequence in these helices is required for Rubisco aggregation in *Chlamydomonas* (Meyer et al., 2012). Weak multivalent interactions have been implicated in mediating biological phase separation in other liquid-like organelles (Li et al., 2012; Hyman et al., 2014). Thus, if EPYC1-Rubisco interactions are weak, then these two binding partners could explain both the structure and the fluidity of the matrix.

A Magic Number Effect Could Facilitate Phase Transitions

We speculate that a magic number effect could help explain the rapid phase transitions of the pyrenoid matrix that we observed experimentally (Figure 6). The rapid partial dissolution and condensation of the pyrenoid suggests regulation at the level of EPYC1-Rubisco interactions. EPYC1 phosphorylation changes in response to CO_2 availability (Turkina et al., 2006; Wang et al., 2014), and phosphorylation could change the affinity of EPYC1 binding sites for Rubisco. As a result of the magic

(D–F) Heatmaps of the fraction of Rubiscos that are in clusters of >10 Rubiscos connected by EPYC1s with 3 (D), 4 (E), or 5 (F) Rubisco binding sites. The fraction of grid sites occupied by Rubiscos (x axis) is varied, with an equal fraction of grid sites occupied by EPYC1s. The specific bond energy (y axis) is varied, while the nonspecific bond energy is fixed at 0.1 k_BT. Red dots indicate the parameters used for the snapshots in (A)–(C).

(G–I), Snapshots of off-lattice 3D simulations with 3 (G), 4 (H), and 5 (I) Rubisco binding sites on EPYC1 for Rubisco. "Rubiscos" (blue spheres with 4 binding sites on each end) and "EPYC1s" (red polymers) bind when their binding sites overlap. Inset in (I): zoom-in with one Rubisco and one EPYC1 with 5 binding sites; 4 of the 5 binding sites of the EPYC1 are in specific bonds with the Rubisco. Snapshots are from simulations with 10 k_BT specific bonds and a Lennard-Jones nonspecific interaction with ϵ = 0.1 k_BT. The molecules occupy ~2% of the total space in these simulations, with equal total numbers of EPYC1 and Rubisco binding sites.

(J) Fraction of Rubiscos in clusters of >10 Rubiscos for EPYC1s with different numbers of binding sites in the 2D model. The specific bond energy is 10 k_BT and the nonspecific bond energy is 0.1 k_BT.

(K) The concentration of Rubisco at which clustering begins for systems with different numbers of EPYC1 binding sites in the 2D model. The onset is determined from the curves in (J) (see Figure S7).

(L) Heatmap of the distribution of Rubisco cluster sizes for different numbers of EPYC1 binding sites in the 2D model. Each column depicts the normalized cluster-size distribution at [Rubisco] = 0.15, with 10 k_BT specific bond energy and 0.1 k_BT nonspecific bond energy.

(M) Schematic of a possible mechanism by which magic numbers could regulate the formation and dissolution of pyrenoids: EPYC1s with 3 binding sites favor Rubisco clustering in the pyrenoid (left), while EPYC1s with 4 binding sites form stable 2:1 complexes of EPYC1:Rubisco that dissolve into the chloroplast stroma (right).

See also Figure S7, Table S1, and Movie S6.

number effect, only a single binding site on each EPYC1 may need to be modified to trigger a pronounced phase change (Figure 7).

Both lattice and off-lattice models indicate that clustering can be strongly suppressed if the number of binding sites on one species (e.g., Rubisco) is an integer multiple of the number of binding sites on the other (e.g., EPYC1), as this favors the assembly of small oligomers in which all binding sites are saturated. In general, we expect this magic number effect to pertain to multicomponent, multivalent binding systems where bonds are one-to-one and saturable. Moreover, the effect requires these specific bonds to have an energy of several times the product of the Boltzmann constant and temperature (k_BT), strong enough for most small oligomers to be fully bonded without defects (Figures 5D–5F). Because such binding energies are quite common, magic number effects can be predicted broadly for interacting protein pairs such as SH3-PRM (Li et al., 2012) and SIM-SUMO (Banani et al., 2016) as well as for RNA-protein droplets (Lin et al., 2015). Additional insights could be gained in the future by investigating the influence of polymer flexibility, stoichiometry, and mixed valency populations on the magic number effect described in this work.

Other Species May Also Have Liquid-like Pyrenoids

Our results suggest that the pyrenoid "regression" (reduction in size and disappearance) observed in some algae during chloroplast division in zoosporogenesis (Brown and Arnott, 1970) may be a phase transition from an aggregated to a soluble phase. Previous reports suggested that different species of algae undergo either pyrenoid fission or pyrenoid regression during cell division (Brown et al., 1967; Griffiths, 1970). However, our observations demonstrate that both phenomena can occur simultaneously in the same cell.

The observation of pyrenoid fission and regression in other species of algae, combined with the amorphous appearance of the matrix of many species (Griffiths, 1970), suggests that the liquid-like nature may be a general property of all pyrenoids. This hypothesis can be tested as new genetic tools become available in a broad range of algae.

Pyrenoids are not only broadly distributed in the Chloroplastida, but are also found in five out of the seven supergroups of the eukaryotic tree of life (Excavates, Stramenopiles, Alveolates, Rhizaria, and Archaeplastids) (Burki, 2014), and it is thought that pyrenoids have evolved multiple times (Villarreal and Renner, 2012; Meyer and Griffiths, 2013). Intriguingly, across sequenced algae, the presence of a pyrenoid is correlated with the presence of a protein with EPYC1-like physiochemical properties, such as a low-complexity repeat sequence, a high pI, and no transmembrane domains (Mackinder et al., 2016). If the only requirement for aggregating diffuse Rubisco into a liquid-like pyrenoid matrix is the evolution of a linker protein with multiple weak binding sites for Rubisco, pyrenoids could be relatively simple to evolve.

Pyrenoid Behavior Is Distinct from that of Carboxysomes

Cyanobacteria have a functionally similar structure to the pyrenoid called the carboxysome. Like the pyrenoid, carboxysomes contain aggregated Rubisco and are not surrounded by a membrane. There are two classes of carboxysomes, α and β, which are thought to have arisen through convergent evolution (Rae et al., 2013).

Rubisco in β-carboxysomes is thought to be linked together by CcmM, a protein containing multiple Rubisco small subunit-like domains, each of which is thought to be incorporated into a different Rubisco holoenzyme (Long et al., 2010). Such incorporation would be expected to preclude internal mixing. Indeed, fluorescently tagged Rubisco in β-carboxysomes does not recover after photobleaching (Chen et al., 2013).

Less is known about the arrangement of Rubisco in α-carboxysomes. The proposed α-carboxysome Rubisco linker protein CsoS2 (Cai et al., 2015) is a disordered repeat protein, raising the possibility that the α-carboxysome may have liquid-like characteristics. Whereas we have observed pyrenoid biogenesis by both fission and de novo assembly, biogenesis of new α- and β-carboxysomes has only been observed to occur *de novo* (Iancu et al., 2010; Chen et al., 2013). It will be interesting to see whether α-carboxysomes do exhibit liquid-like behavior or whether the underlying structural and biogenesis principles are a general distinguishing feature between carboxysomes and pyrenoids.

The Pyrenoid Provides Insights into Liquid-like Organelle Biology

To our knowledge, the pyrenoid is the only example so far of a liquid-like organelle specific to photosynthetic organisms. Like other previously characterized liquid-like organelles (Brangwynne et al., 2011; Saha et al., 2016), new pyrenoids can be generated by either *de novo* aggregation or fission into two daughters. The first characterized liquid-like organelle, the *C. elegans* P granule, leverages the properties of phase transitions to promote its own asymmetric inheritance by preferentially condensing at one end of the cell before cell division (Brangwynne et al., 2009; Saha et al., 2016). In contrast, the pyrenoid appears to leverage the liquid-like properties of partial dissolution and fission to promote symmetric inheritance, allowing both daughter cells to inherit a pyrenoid.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - RBCS1-, RCA1-, *epyc1* EPYC1-, and EPYC1-Venus *Chlamydomonas reinhardtii* strain generation and culture conditions
  - Cells used for cryo-electron tomography
- METHOD DETAILS
  - Division microscopy and analysis
  - Vitrification, cryo-focused ion beam milling, cryo-electron tomography, and tomogram reconstruction
  - Localization of Rubisco within tomograms and subtomogram averaging
  - Local density analysis and generation of simulated data

- Fitting Lennard-Jones fluid models to the experimental data
- Spot tests to assay growth of Venus-expressing strains
- Fluorescence Recovery After Photobleaching microscopy and analysis
- "Magic number" modeling

- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Division image analysis
  - Tomogram analysis
  - FRAP image analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, one table, and six movies and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2017.08.008.

AUTHOR CONTRIBUTIONS

E.S.F.R. performed the FRAP and division experiments and data analysis. L.C.M.M. generated the fluorescent constructs and made the initial FRAP observation. E.S.F.R., L.C.M.M., and M.C.J. designed and interpreted the FRAP and division work. H.N.C. assisted confocal data acquisition and analysis. M. Schaffer and B.D.E. performed cryo-ET data acquisition. L.K.C. and M. Strauss performed the cryo-ET image analysis. A.M.-S. performed the cryo-ET statistical analysis and simulated data generation. A.M.-S. and B.D.E. performed the Lennard-Jones analytical fit. B.X. performed the Lennard-Jones simulation fit. J.M.P., F.F., and B.D.E. supervised the cryo-ET work. B.X. and N.S.W. designed and interpreted the on-lattice EPYC1-Rubisco interaction modeling work. B.X. performed the on-lattice EPYC1-Rubisco interaction modeling work. B.X. and P.R. performed the 3D off-lattice simulation. E.S.F.R., B.X., N.S.W., B.D.E., and M.C.J. wrote the text with input from all authors.

ACKNOWLEDGMENTS

We thank W. Baumeister, J. Berry, D. Ehrhardt, J. Feldman, W. Frommer, A. Grossman, Y. Meir, M.B. Mudgett, M. Scott, H. Stone, V. Walbot, and members of the Jonikas laboratory for helpful discussion and suggestions, and we thank U. Goodenough, A. Grossman, A. McCormick, and M.T. Meyer for critical reading of the manuscript. This work was supported by the National Science Foundation (EF-1105617 and IOS-1359682 to M.C.J. and PHY-1305525 to N.S.W.), the Carnegie Institution for Science (to L.C.M.M. and M.C.J.); NIH (T32GM007276 to E.S.F.R. and 7DP2GM119137-02 to M.C.J.), the Simons Foundation and HHMI (55108535), and Princeton University (to M.C.J.), a CONACyT-DAAD Graduate Scholarship (to L.K.C), a Fundación Séneca Postdoctoral Fellowship (to A.M.-S.), an Alexander von Humboldt Foundation Postdoctoral Fellowship (to B.D.E), and the Deutsche Forschungsgemeinschaft (FO 716/4-1 to F.F.).

Received: March 9, 2017
Revised: June 12, 2017
Accepted: August 4, 2017
Published: September 21, 2017

REFERENCES

Asano, S., Engel, B.D., and Baumeister, W. (2016). In situ cryo-electron tomography: a post-reductionist approach to structural biology. J. Mol. Biol. *428* (2 Pt A), 332–343.

Badger, M.R., Andrews, T.J., Whitney, S.M., Ludwig, M., Yellowlees, D.C., Leggat, W., and Price, G.D. (1998). The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae. Can. J. Bot. *76*, 1052–1071.

Banani, S.F., Rice, A.M., Peeples, W.B., Lin, Y., Jain, S., Parker, R., and Rosen, M.K. (2016). Compositional control of phase-separated cellular bodies. Cell *166*, 651–663.

Bertagnolli, B.L., and Nadakavukaren, M.J. (1970). An ultrastructural study of pyrenoids from *Chlorella pyrenoidosa*. J. Cell Sci. *7*, 623–630.

Bharat, T.A.M., and Scheres, S.H.W. (2016). Resolving macromolecular structures from electron cryo-tomography data using subtomogram averaging in RELION. Nat. Protoc. *11*, 2054–2065.

Borkhsenious, O.N., Mason, C.B., and Moroney, J.V. (1998). The intracellular localization of ribulose-1,5-bisphosphate Carboxylase/Oxygenase in *chlamydomonas reinhardtii*. Plant Physiol. *116*, 1585–1591.

Brangwynne, C.P., Eckmann, C.R., Courson, D.S., Rybarska, A., Hoege, C., Gharakhani, J., Jülicher, F., and Hyman, A.A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science *324*, 1729–1732.

Brangwynne, C.P., Mitchison, T.J., and Hyman, A.A. (2011). Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. Proc. Natl. Acad. Sci. USA *108*, 4334–4339.

Brown, R.M., Jr., and Arnott, H.J. (1970). Structure and function of the algal pyrenoid. I. Ultrastructure and cytochemistry during zoosporogenesis of *Tetracystis excentrica*. J. Phycol. *6*, 14–22.

Brown, R.M., Jr., Arnott, H.J., Bisalputra, T., and Hoffman, L.R. (1967). The pyrenoid: Its structure, distribution, and function. J. Phycol. *3* (Supp), 5–6.

Burki, F. (2014). The eukaryotic tree of life from a global phylogenomic perspective. Cold Spring Harb. Perspect. Biol. *6*, a016147.

Cai, F., Dou, Z., Bernstein, S.L., Leverenz, R., Williams, E.B., Heinhorst, S., Shively, J., Cannon, G.C., and Kerfeld, C.A. (2015). Advances in understanding carboxysome assembly in *Prochlorococcus* and *Synechococcus* implicate CsoS2 as a critical component. Life (Basel) *5*, 1141–1171.

Chen, A.H., Robinson-Mosher, A., Savage, D.F., Silver, P.A., and Polka, J.K. (2013). The bacterial carbon-fixing organelle is formed by shell envelopment of preassembled cargo. PLoS ONE *8*, e76127.

Cheng, S., Cetinkaya, M., and Gräter, F. (2010). How sequence determines elasticity of disordered proteins. Biophys. J. *99*, 3863–3869.

Crang, R.F.E., and Klomparens, K.L. (1988). Artifacts in Biological Electron Microscopy (Plenum Press).

Cross, F.R., and Umen, J.G. (2015). The *Chlamydomonas* cell cycle. Plant J. *82*, 370–392.

Engel, B.D., Schaffer, M., Kuhn Cuellar, L., Villa, E., Plitzko, J.M., and Baumeister, W. (2015). Native architecture of the *Chlamydomonas* chloroplast revealed by in situ cryo-electron tomography. eLife *4*, e04889.

Feric, M., Vaidya, N., Harmon, T.S., Mitrea, D.M., Zhu, L., Richardson, T.M., Kriwacki, R.W., Pappu, R.V., and Brangwynne, C.P. (2016). Coexisting liquid phases underlie nucleolar subcompartments. Cell *165*, 1686–1697.

Förster, F., Pruggnaller, S., Seybert, A., and Frangakis, A.S. (2008). Classification of cryo-electron sub-tomograms using constrained correlation. J. Struct. Biol. *161*, 276–286.

Förster, F., Han, B.G., and Beck, M. (2010). Visual proteomics. Methods Enzymol. *483*, 215–243.

Gibbs, S.P. (1962). The ultrastructure of the pyrenoids of green algae. J. Ultrastruct. Res. *7*, 262–272.

Goodenough, U.W. (1970). Chloroplast division and pyrenoid formation in *Chlamydomonas reinhardtii*. J. Phycol. *6*, 1–6.

Griffiths, D.J. (1970). The pyrenoid. Bot. Rev. *36*, 29–58.

Gu, L., Xu, S., Sun, Z., and Wang, J.T. (2010). Brownian dynamics simulation of the crystallization dynamics of charged colloidal particles. J. Colloid Interface Sci. *350*, 409–416.

Hofmann, H., Soranno, A., Borgia, A., Gast, K., Nettels, D., and Schuler, B. (2012). Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. Proc. Natl. Acad. Sci. USA *109*, 16155–16160.

Holdsworth, R.H. (1968). The presence of a crystalline matrix in pyrenoids of the diatom, *Achnanthes brevipes*. J. Cell Biol. *37*, 831–837.

Hrabe, T., Chen, Y., Pfeffer, S., Cuellar, L.K., Mangold, A.V., and Förster, F. (2012). PyTom: a python-based toolbox for localization of macromolecules in cryo-electron tomograms and subtomogram analysis. J. Struct. Biol. *178*, 177–188.

Hyman, A.A., Weber, C.A., and Jülicher, F. (2014). Liquid-liquid phase separation in biology. Annu. Rev. Cell Dev. Biol. *30*, 39–58.

Iancu, C.V., Morris, D.M., Dou, Z., Heinhorst, S., Cannon, G.C., and Jensen, G.J. (2010). Organization, structure, and assembly of α-carboxysomes determined by electron cryotomography of intact cells. J. Mol. Biol. *396*, 105–117.

Johnson, J.K., Zollweg, J.A., and Gubbins, K.E. (1993). The Lennard-Jones equation of state revisited. Mol. Physiol. *78*, 591–618.

Keown, J.R., Griffin, M.D., Mertens, H.D., and Pearce, F.G. (2013). Small oligomers of ribulose-bisphosphate carboxylase/oxygenase (Rubisco) activase are required for biological activity. J. Biol. Chem. *288*, 20607–20615.

Klammt, C., Novotná, L., Li, D.T., Wolf, M., Blount, A., Zhang, K., Fitchett, J.R., and Lillemeier, B.F. (2015). T cell receptor dwell times control the kinase activity of Zap70. Nat. Immunol. *16*, 961–969.

Kowallik, K. (1969). The crystal lattice of the pyrenoid matrix of *Prorocentrum micans*. J. Cell Sci. *5*, 251–269.

Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. *116*, 71–76.

Kropat, J., Hong-Hermesdorf, A., Casero, D., Ent, P., Castruita, M., Pellegrini, M., Merchant, S.S., and Malasarn, D. (2011). A revised mineral nutrient supplement increases biomass and growth rate in *Chlamydomonas reinhardtii*. Plant J. *66*, 770–780.

Kroschwald, S., Maharana, S., Mateju, D., Malinovska, L., Nüske, E., Poser, I., Richter, D., and Alberti, S. (2015). Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. eLife *4*, e06807.

Li, P., Banjade, S., Cheng, H.-C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J.V., King, D.S., Banani, S.F., et al. (2012). Phase transitions in the assembly of multivalent signalling proteins. Nature *483*, 336–340.

Lin, Y., Protter, D.S., Rosen, M.K., and Parker, R. (2015). Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. Mol. Cell *60*, 208–219.

Long, B.M., Tucker, L., Badger, M.R., and Price, G.D. (2010). Functional cyanobacterial beta-carboxysomes have an absolute requirement for both long and short forms of the CcmM protein. Plant Physiol. *153*, 285–293.

Mackinder, L.C.M., Meyer, M.T., Mettler-Altmann, T., Chen, V.K., Mitchell, M.C., Caspari, O., Freeman Rosenzweig, E.S., Pallesen, L., Reeves, G., Itakura, A., et al. (2016). A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. Proc. Natl. Acad. Sci. USA *113*, 5958–5963.

Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51.

McKay, R.M.L., Gibbs, S.P., and Vaughn, K.C. (1991). RuBisCo activase is present in the pyrenoid of green algae. Protoplasma *162*, 38–45.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., and Teller, E. (1953). Equation of state calculations by fast computing machines. J. Chem. Phys. *21*, 1087–1092.

Meyer, M., and Griffiths, H. (2013). Origins and diversity of eukaryotic CO_2-concentrating mechanisms: lessons for the future. J. Exp. Bot. *64*, 769–786.

Meyer, M.T., Genkov, T., Skepper, J.N., Jouhet, J., Mitchell, M.C., Spreitzer, R.J., and Griffiths, H. (2012). Rubisco small-subunit α-helices control pyrenoid formation in *Chlamydomonas*. Proc. Natl. Acad. Sci. USA *109*, 19474–19479.

Mitchell, M.C., Meyer, M.T., and Griffiths, H. (2014). Dynamics of carbon-concentrating mechanism induction and protein relocalization during the dark-to-light transition in synchronized Chlamydomonas reinhardtii. Plant Physiol. *166*, 1073–1082.

Mitrea, D.M., and Kriwacki, R.W. (2016). Phase separation in biology; functional organization of a higher order. Cell Commun. Signal. *14*, 1.

Morsali, A., Goharshadi, E.K., Mansoori, G.A., and Abbaspour, M. (2005). An accurate expression for radial distribution function of the Lennard-Jones fluid. Chem. Phys. *310*, 11–15.

Ohad, I., Siekevitz, P., and Palade, G.E. (1967). Biogenesis of chloroplast membranes. I. Plastid dedifferentiation in a dark-grown algal mutant (*Chlamydomonas reinhardi*). J. Cell Biol. *35*, 521–552.

Plimpton, S. (1995). Fast parallel algorithms for short-range molecular dynamics. J. Comp. Physiol. *117*, 1–19.

Pollock, S.V., Colombo, S.L., Prout, D.L.J., Jr., Godfrey, A.C., and Moroney, J.V. (2003). Rubisco activase is required for optimal photosynthesis in the green alga *Chlamydomonas reinhardtii* in a low-CO_2 atmosphere. Plant Physiol. *133*, 1854–1861.

Rae, B.D., Long, B.M., Whitehead, L.F., Förster, B., Badger, M.R., and Price, G.D. (2013). Cyanobacterial carboxysomes: microcompartments that facilitate CO_2 fixation. J. Mol. Microbiol. Biotechnol. *23*, 300–307.

Retallack, B., and Butler, R.D. (1970). The development and structure of pyrenoids in *Bulbochaete hiloensis*. J. Cell Sci. *6*, 229–241.

Saha, S., Weber, C.A., Nousch, M., Adame-Arana, O., Hoege, C., Hein, M.Y., Osborne-Nishimura, E., Mahamid, J., Jahnel, M., Jawerth, L., et al. (2016). Polar positioning of phase-separated liquid compartments in cells regulated by an mRNA competition mechanism. Cell *166*, 1572–1584.e16.

Sakurai, M., Watanabe, K., Sumiyama, K., and Suzuki, K. (1999). Magic numbers in transition metal (Fe, Ti, Zr, Nb, and Ta) clusters observed by time-of-flight mass spectrometry. J. Chem. Phys. *111*, 235–238.

Schaffer, M., Engel, B.D., Laugks, T., Mahamid, J., Plitzko, J.M., and Baumeister, W. (2015). Cryo-focused ion beam sample preparation for imaging vitreous cells by cryo-electron tomography. Bio Protoc. *5*, e1575.

Schaffer, M., Mahamid, J., Engel, B.D., Laugks, T., Baumeister, W., and Plitzko, J.M. (2017). Optimized cryo-focused ion beam sample preparation aimed at in situ structural studies of membrane proteins. J. Struct. Biol. *197*, 73–82.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Steppenbeck, D., Takeuchi, S., Aoi, N., Doornenbal, P., Matsushita, M., Wang, H., Baba, H., Fukuda, N., Go, S., Honma, M., et al. (2013). Evidence for a new nuclear "magic number" from the level structure of 54Ca. Nature *502*, 207–210.

Stone, H.A. (1994). Dynamics of drop deformation and breakup in viscous fluids. Annu. Rev. Fluid Mech. *26*, 65–102.

Taylor, T.C., Backlund, A., Bjorhall, K., Spreitzer, R.J., and Andersson, I. (2001). First crystal structure of Rubisco from a green alga, *Chlamydomonas reinhardtii*. J. Biol. Chem. *276*, 48159–48164.

Thévenaz, P., Ruttimann, U.E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. IEEE Trans. Image Process. *7*, 27–41.

Thorneywork, A.L., Roth, R., Aarts, D.G.A.L., and Dullens, R.P.A. (2014). Communication: radial distribution functions in a two-dimensional binary colloidal hard sphere system. J. Chem. Phys. *140*, 161106.

Turkina, M.V., Blanco-Rivero, A., Vainonen, J.P., Vener, A.V., and Villarejo, A. (2006). CO_2 limitation induces specific redox-dependent protein phosphorylation in *Chlamydomonas reinhardtii*. Proteomics *6*, 2693–2704.

Umen, J.G., and Goodenough, U.W. (2001). Control of cell division by a retinoblastoma protein homolog in *Chlamydomonas*. Genes Dev. *15*, 1652–1661.

van Royen, M.E., Farla, P., Mattern, K.A., Geverts, B., Trapman, J., and Houtsmuller, A.B. (2009). Fluorescence recovery after photobleaching (FRAP) to study nuclear protein dynamics in living cells. Methods Mol. Biol. *464*, 363–385.

Vaucher, J.-P. (1803). Histoire des conferves d'eau douce, contenant leurs différens modes de reproduction, et la description de leurs principales espèces, suivie de l'histoire des trémelles et des ulves d'eau douce (J.J. Paschoud).

Villarreal, J.C., and Renner, S.S. (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*, 18873–18878.

Wales, D.J., and Doye, J.P.K. (1997). Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. J. Phys. Chem. *101*, 5111–5116.

Wang, H., Gau, B., Slade, W.O., Juergens, M., Li, P., and Hicks, L.M. (2014). The global phosphoproteome of *Chlamydomonas reinhardtii* reveals complex organellar phosphorylation in the flagella and thylakoid membrane. Mol. Cell. Proteomics *13*, 2337–2353.

Wang, Y., Stessman, D.J., and Spalding, M.H. (2015). The CO_2 concentrating mechanism and photosynthetic carbon assimilation in limiting CO_2: how *Chlamydomonas* works against the gradient. Plant J. *82*, 429–448.

Wiegand, T., and Moloney, K.A. (2004). Rings, circles, and null-models for point pattern analysis in ecology. Oikos *104*, 209–229.

Wienkoop, S., Weiss, J., May, P., Kempa, S., Irgang, S., Recuenco-Munoz, L., Pietzke, M., Schwemmer, T., Rupprecht, J., Egelhofer, V., and Weckwerth, W. (2010). Targeted proteomics for Chlamydomonas reinhardtii combined with rapid subcellular protein fractionation, metabolomics and metabolic flux analyses. Mol. Biosyst. *6*, 1018–1031.

Yanashima, R., García, A.A., Aldridge, J., Weiss, N., Hayes, M.A., and Andrews, J.H. (2012). Cutting a drop of water pinned by wire loops using a superhydrophobic surface and knife. PLoS ONE *7*, e45893.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
In situ subtomogram average of Rubisco within the *Chlamydomonas* pyrenoid	This paper; EMDataBank	EMDB: EMD-3694
Experimental Models: Organisms/Strains		
C. reinhardtii: wild-type: cMJ030	Chlamydomonas Resource Center	CC-4533 cw15
C. reinhardtii: epyc1 mutant: cMJ030 background	Mackinder et al., 2016; Chlamydomonas Resource Center	CC-5360
C. reinhardtii: RBCS1-Venus: cMJ030 background	Mackinder et al., 2016; Chlamydomonas Resource Center	CC-5357
C. reinhardtii: EPYC1-Venus: cMJ030 background	Mackinder et al., 2016; Chlamydomonas Resource Center	CC-5359
C. reinhardtii: RCA1-Venus: cMJ030 background	This paper; Chlamydomonas Resource Center	CC-5358
C. reinhardtii: epyc1 EPYC1-Venus: *epyc1* background	Mackinder et al., 2016; Chlamydomonas Resource Center	CC-5361
C. reinhardtii: mat3-4	Umen and Goodenough, 2001; Chlamydomonas Resource Center	CC-3994 *mat3-4* mt+
Recombinant DNA		
pLM005-RBCS1-Venus	Mackinder et al., 2016; GenBank	GenBank: KY550376
pLM005-EPYC1-Venus	Mackinder et al., 2016; GenBank	GenBank: KX077944
pLM005-RCA1-Venus	Mackinder et al., 2016; GenBank	GenBank: KY550375
Software and Algorithms		
LASX	Leica	http://www.leica-microsystems.com/products/microscope-software/details/product/leica-las-x-ls/
FIJI	Schindelin et al., 2012	https://imagej.net/Fiji/Downloads
Imaris	Bitplane	http://www.bitplane.com/imaris/imaris?gclid=Cj0KEQjwioHIBRCes6nP56Ti1IsBEiQAxxb5G4wQ2n13WrbWgP9nQXa11cQ5KVq94tzi6J5jBKZwoa0aArD68P8HAQ
SlideBook	Intelligent Imaging Innovations	https://www.intelligent-imaging.com/slidebook
StackReg	Thévenaz et al., 1998	http://bigwww.epfl.ch/thevenaz/stackreg/
OriginPro	OriginLab	http://www.originlab.com/Origin
SerialEM	Mastronarde, 2005	http://bio3d.colorado.edu/SerialEM/
IMOD	Kremer et al., 1996	http://bio3d.colorado.edu/imod/
K2Align	Dimitry Tegunov	https://github.com/dtegunov/k2align
PyTom	Hrabe et al., 2012	http://pytom.org/
Amira	FEI	https://www.fei.com/software/amira-3d-for-life-sciences/
MATLAB	MathWorks	https://www.mathworks.com/
RELION	Bharat & Scheres, 2016	http://www2.mrc-lmb.cam.ac.uk/relion/index.php/Main_Page/
LAMMPS	Plimpton, 1995	http://lammps.sandia.gov/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Modeling Rubisco-EPYC1 aggregation	This paper	https://github.com/binarybin/RubiscoSimulation
Fitting the Lennard-Jones RDF to experimental data	This paper	https://github.com/anmartinezs/LJRDFfit/
Other		
Typhoon Trio fluorescence scanner	GE Healthcare	Part No. 63005587 (Discontinued)
LumiBar LED lights	LumiGrow	Cat#8100-5502
TCS SP8 confocal microscope	Leica	http://www.leica-microsystems.com/products/confocal-microscopes/details/product/leica-tcs-sp8/
LED Angel Eye Headlight Accent Light Kit	SuperBrightLeds.com	AE80-RGB12
Countess II Automated Cell Counter	ThermoFisher Scientific	Cat# AMQAF1000
Poly-L-lysine-coated 8 well μ-slides	ibidi	Cat#80824
Spinning Disk Confocal Microscope	Leica	DMI6000B

CONTACT FOR REAGENT AND RESOURCE SHARING

Please contact the Lead Contact, Martin C. Jonikas (mjonikas@princeton.edu), with any requests regarding reagents used in this study.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

RBCS1-, RCA1-, *epyc1* EPYC1-, and EPYC1-Venus *Chlamydomonas reinhardtii* strain generation and culture conditions

All strains expressing fluorescent proteins were generated in Mackinder et al. (2016), as follows: the DNA encoding the protein of interest (RBCS1, RCA1, or EPYC1) was amplified from *Chlamydomonas reinhardtii* gDNA and cloned behind the PsaD promoter with a Venus-3xFLAG on the C terminus, in a construct containing the *aphVIII* gene for paromomycin resistance. Vector sequences have been deposited at GenBank under accession numbers KY550376 (pLM005-RBCS1-Venus), KX077944 (pLM005-EPYC1-Venus), and KY550375 (pLM005-RCA1-Venus). Linearized constructs were transformed into wild-type (cMJ030; also known as CC-4533 cw15; mating type minus) or *epyc1* (also in the cMJ030 background) (Mackinder et al., 2016) *Chlamydomonas reinhardtii* by electroporation, which results in random integration into the nuclear genome (Mackinder et al., 2016). Strains expressing the construct were selected by growth on agar plates containing Tris-Acetate-Phosphate (TAP) plus paromomycin, and subsequent fluorescence screening for Venus on a Typhoon Trio fluorescence scanner (GE Healthcare) with excitation at 532 nm and emission at 555/20 nm. Strains containing RBCS1- and EPYC1-Venus were authenticated by western blots in Mackinder et al. (2016). The above strains have been deposited at the Chlamydomonas Resource Center under accession numbers CC-5357 (RBCS1-Venus), CC-5359 (EPYC1-Venus), CC-5358 (RCA1-Venus), CC-5360 (*epyc1*), and CC-5361 (*epyc1* EPYC1-Venus).

For microscopy, photoheterotrophic pregrowth cultures were inoculated from ~two-week-old TAP plates into 50 mL liquid TAP media in 250 mL glass Erlenmeyer flasks, and grown under continuous white light (~40 μmol photons m^{-2} s^{-1}) while shaking at 125 rpm until cells reached log phase growth (2-5 × 10^6 cells/mL; 3-4 days after inoculation). Photoautotrophic cultures were inoculated by centrifuging the pregrowth cultures at 1000 g for 5 min at room temperature and then resuspending the pellet in 50 mL liquid minimal Tris-Phosphate (TP) media (Kropat et al., 2011). Photoautotrophic cultures were grown in 50 mL liquid TP in glass tubes bubbled with air (~0.04% CO_2) under continuous light (150 μmol photons m^{-2} s^{-1} red and blue LumiGrow LumiBar LED lights). After several rounds of dilution and growth in TP, cells were imaged when they reached a density of ~1-2x10^6 cells/mL.

Cells used for cryo-electron tomography

Chlamydomonas mat3-4 cells (Umen and Goodenough, 2001), acquired from the Chlamydomonas Resource Center (Univ. of Minnesota, CC-3994, mt+), were used because they vitrify better than wild-type cells due to their smaller size (~5 μm) but have normal pyrenoid morphology and a functional carbon-concentrating mechanism, shown by growth in low-CO_2 conditions (Umen and Goodenough, 2001), which inhibit carbon-concentrating-deficient strains such as the *epyc1* mutant (Mackinder et al., 2016). Cells were grown in TAP media with constant light (~90 μmol photons m^{-2} s^{-1}) and aeration with normal atmosphere until mid-log phase.

METHOD DETAILS

Division microscopy and analysis

Image acquisition

Strains were grown and prepared for microscopy as described above. Cells were imaged at room temperature every minute for 8-16 hr on a Leica TCS SP8 laser scanning confocal microscope in resonant scanning mode with LASX software and a 63X NA 1.4 oil objective, while illuminated with a 626 nm red LED circle affixed to the condenser (LED Angel Eye Headlight Accent Light Kit; SuperBrightLeds.com, AE80-RGB12). Venus fluorescence was imaged by exciting with a white light laser at 488 nm and collecting emission from 499-551 nm on a HyD SMD hybrid detector (Leica) with lifetime gate filter (1-6 ns) to reduce background due to chlorophyll autofluorescence. Chlorophyll autofluorescence was imaged simultaneously with the same 488 nm excitation, and emission was collected from 680-750 nm on a PMT (Leica). Images were collected with 8-line averaging, 0.3 μm steps through the entire cell volume, and adaptive focus control. Displayed division analysis data represent images collected over at least two independent biological replicates for each strain. Sample sizes were chosen based on similar published studies (Chen et al., 2013; Feric et al., 2016).

Image analysis

4D TIFFs were exported to FIJI as virtual stacks and dividing cells were manually marked in MaxZ projections over time. The pyrenoid inheritance pattern and timing of division was manually noted for each dividing cell in 3D in Imaris (Bitplane). If a cell exhibited more than one mode of inheritance (for instance, both pyrenoid fission and puncta), then fission took precedence for categorization. The duration of chloroplast division and pyrenoid fission (as defined in the text) were measured for all observed RBCS1-Venus pyrenoid fission events (n = 14 first divisions and 14 second divisions) and randomly selected EPYC1-Venus fission events (n = 22 first divisions and 25 second divisions). The extent of pyrenoid dissolution and increase in stromal Venus signal was measured in FIJI as follows: for each dividing cell, a SumZ projection was created through the whole cell. Then, any saturated pixels were masked out, and intensity-based masks were created from SumZ projections of the chlorophyll and Venus channels to mask out any signal in the Venus channel from outside the chloroplast or outside the pyrenoid, respectively. Finally, for each masked region of each cell, the total intensity (sum of the unmasked pixels in the entire image; RawIntDen) was measured at each time point and exported to Excel. The Wilcoxon Matched-Pairs Signed-Ranks Test was used to quantify the difference in the average pyrenoid signal between maximal dimming and 15 min beforehand in 31 divisions (with the error bars denoting SEM) as described below.

Vitrification, cryo-focused ion beam milling, cryo-electron tomography, and tomogram reconstruction

Using a Vitrobot Mark 4 (FEI), cells were blotted onto carbon-coated EM grids (Quantifoil Micro Tools; 4 μL cell culture per grid) and plunge-frozen into a liquid ethane/propane mixture. Cryo-focused ion beam (cryo-FIB) milling was performed on a dual-beam Quanta3D FIB/SEM microscope (FEI) by scanning the cells with Gallium ions, as previously described (Schaffer et al., 2015, 2017). Thinned samples were transferred to a Titan Krios 300 kV transmission electron microscope (FEI) equipped with a 968 Quantum post-column energy filter (Gatan), and imaged with a K2 Summit direct detector (Gatan) operated in movie mode at 12-17 frames per second. Using SerialEM software (Mastronarde, 2005), single-axis tilt-series were recorded from -60° to $+60^\circ$ (in two halves separated at 0°) at 2° increments, with an object pixel size of 3.42 Å, a target defocus of -4 or -5 μm, and a cumulative electron dose of $\sim$100 electrons/Å^2. Image frames were aligned using in-house developed software to correct for beam-induced motion. Using IMOD software (Kremer et al., 1996), tilt-series were aligned by patch-tracking and reconstructed by weighted back projection. The biggest gain in image quality compared to the previous study (Engel et al., 2015) was the result of using a direct electron detector instead of a CCD camera.

Localization of Rubisco within tomograms and subtomogram averaging

Tomogram masking and template matching

Tomograms were binned to a pixel size of 13.68 Å and subjected to template matching using the PyTom software (Hrabe et al., 2012). The template was generated from the X-ray crystallography structure of *Chlamydomonas* Rubisco (Taylor et al., 2001) (PDB entry: 1GK8) by lowpass filtering the electron density map to a resolution of 33 Å, a value determined by the approximate first zero of the contrast transfer function in the tomograms (Förster et al., 2010). Tomogram masks were manually segmented in Amira software (FEI), restricting the search area to only the pyrenoid matrix (Figure S3A). For each tomogram, the template matching cross-correlation function was filtered by the tomogram mask, and the remaining cross-correlation peaks were exhaustively extracted with a mask radius of 9.5 nm. Histograms of cross-correlation scores from the extracted peaks revealed Gaussian distributions of true-positive particles at the high values. A cutoff was set to two standard deviations toward the low-valued tail of each Gaussian, and all particles with scores below this cutoff were discarded (Figure S3B).

Subtomogram averaging and classification

Subvolumes corresponding to the extracted peaks were binned to a pixel size of 6.84 Å and aligned in PyTom using a real-space expectation maximization method that implements gold-standard alignment (Hrabe et al., 2012). This alignment procedure did not impose D_4 symmetry. The number of iterations was set to 10, the initial angular increment to 3°, and the angular shells to 3. Next, subtomogram classification was performed using constrained principal component analysis (CPCA) (Förster et al., 2008). PyTom was used to calculate similarity matrices and for hierarchical clustering, while CPCA and k-means clustering were performed in MATLAB

(MathWorks). Subtomograms were lowpass filtered to 38 Å, 5 eigenvectors were used, and the number of classes was set between 99 and 165, depending on the number of particles in each tomogram. Classes were then merged by hierarchical clustering, using constrained cross-correlation as a distance measure, yielding a "positive" and "negative" class for each tomogram (Figure S3C).

High-resolution subtomogram averaging

Unbinned subtomogram volumes (3.42 Å pixel size) from "tomogram 1" corresponding to the 30,000 highest template matching cross-correlation scores in the "positive" class were used to generate the average shown in Figure 4C. The subvolumes were corrected for the contrast transfer function by phase-flipping in IMOD, then aligned with imposed D_4 symmetry using gold-standard alignment in RELION (Bharat and Scheres, 2016). This procedure was restricted to a local angular search, using initial angle and offset sampling rates of 3.7° and 0.34 nm, respectively, while the maximal offset value was set to 1.7 nm. The initial reference was a subtomogram average filtered to 4 nm resolution. In the final average, the B-factor variable was set to 4 $Å^2$ for each 1 electron/$Å^2$, and the cumulative electron dose of the subtomograms was restricted by only using the central 60° of the tilt-series (−30° to +30°).

Local density analysis and generation of simulated data

Measurement of radial local Rubisco density

Pyrenoid matrix organization was quantified by Radial Local Density (RLD) estimation, defined as:

$$\lambda(r) = \frac{\sum_{i=0}^{n} C\left(S_h\left(\vec{x}_{p_i}, r\right)\right)}{V\left(S_h\left(\vec{x}_{p_i}, r\right)\right)}$$

The coordinates of the reference particle are represented by the vector $\vec{x}_{p_i}$. Function C counts the number of particles contained within the S_h subvolume, which is a radial shell subset of the S_P total masked matrix volume $(S_h \subset S_P)$. Function V computes the volume of S_h. Since all particles in S_P are used, then $\{\vec{x}_{p_0}, \ldots, \vec{x}_{p_n}\} \in S_P$. The definition of the local subvolume S_h associated to a particle $\vec{x}_{p_i}$ and a radius r is:

$$S_h\left(\vec{x}_{p_i}, r\right) = \left\{\vec{x} \in S_P \middle| \left\{d\left(\vec{x}, \vec{x}_{p_i}\right) \geq (r - \Delta r/2)\right\} \text{i} \left\{d\left(\vec{x}, \vec{x}_{p_i}\right) \leq (r + \Delta r/2)\right\}\right\}$$

where d is the Euclidean distance function. In an unbounded space, S_h would correspond with a spherical shell centered at $\vec{x}_{p_i}$, with radius r (distance to reference particle p_i) and a shell thickness of $\Delta r > 0$. However, because the pyrenoid matrix volume S_P is actually a finite irregularly bounded space, $V(S_h)$ cannot be accurately estimated by analytical formulas. Thus, we used a numerical estimator analogous to Wiegand and Moloney (2004), but adapted for 3D volumes instead of 2D areas. This approach restricts S_h to the masked pyrenoid matrix volume, eliminating edge effects that would have arisen from the inclusion of the membrane tubules and areas outside of the pyrenoid. For all RLD measurements in Figure 4 (experimental and simulated data in panels D, F and H, experimental data in panel G), we set Δr = 1.4 nm as a compromise between precision and graph smoothness.

As the global pyrenoid matrix Rubisco density, ρ, is slightly different for each pyrenoid (Figures 4D and S3D), we normalized each RLD by the global density so that RLD shape could be directly compared between multiple tomograms: $\lambda' = \lambda/\rho$.

Generation of simulated tomogram volumes

For direct comparison to the experimental data, all of the simulated tomograms were created by placing Rubisco particles into the same pyrenoid matrix volumes as those in the real tomograms (Figures S3A and S4D–S4I), as defined by the manual masking step described above.

I. Simulated crystal

These data were generated by propagating the unit cell of crystalized *Chlamydomonas* Rubisco (Taylor et al., 2001) throughout the matrix volumes. Despite their "noisy" appearance, the crystalline profiles have less variance than the experimental data – their 99% confidence intervals are almost too small to see in Figure 4F. Two factors cause the simulated data to look "noisy": 1) unlike the heterogeneous organization of the experimental data, the crystalline packing is exact, leading to dramatic peaks. 2) RLD is sampled with a spherical shell (of 1.4 nm width), whereas the crystalline organization does not propagate in a spherical pattern. This mismatch between the spherical RLD measurement and the crystalline packing causes the "noisy" jagged appearance of the plots.

II. Simulated hexagonal close packing

Within the matrix volumes, an HCP lattice was generated with 13.9 nm between the centers of all nearest neighbors, a distance based on the preferred nearest neighbor spacing measured in Figure 4E.

III. Simulated random singles

A binary Rubisco template was generated by thresholding our Rubisco subtomogram average (Figure 4C) to allow the minimum near neighbor distance found in our experimental data (Figure 4E). Using Monte Carlo simulations, these templates were sequentially mapped into the matrix volumes, using random positions and orientations but forbidding overlap between templates. The procedure was stopped upon reaching the same number of Rubisco particles as the experimental data.

IV. Simulated random linked pairs

Pairs of two Rubisco templates, each with random orientation, were placed together with an inter-center distance defined by the experimental data's distribution (mean = 13.9 nm, standard deviation = 1.5 nm; Figure 4E). Pairs were sequentially mapped into

the matrix volumes, using random positions and orientations, while avoiding overlap with previously inserted particles. These steps were repeated until the same number of particles as the experimental data were reached.

V. Simulated random linked network

First, a seed Rubisco template was randomly placed into the matrix volume, while avoiding overlap with already inserted particles. Second, a zone for the potential placement of neighbors was defined as a shell around the seed template in a range of 5.5-7 nm from the seed particle surface (Figure S4J). Third, a neighbor Rubisco particle was randomly placed with its center inside the zone. Fourth, the zone for potential neighbor placement was updated to encompass a range of 5.5-7 nm around both particles. The rounds of random neighbor placement and zone updating were sequentially repeated until a maximum of 8 networked particles were placed. After 8 particles were placed or the network failed to place a particle due to lack of space, a new seed was randomly created within the matrix volume and the procedure was repeated. When two networks encountered each other, they merged their neighbor placement zones. The procedure was stopped when the same number of Rubisco particles as the experimental data were reached.

Fitting Lennard-Jones fluid models to the experimental data

RLD is a robust estimator of the Radial Distribution Function (RDF) for proteins in a finite, bounded, and irregularly shaped space like the pyrenoid matrix. Thus, analytical models for RDF can also be applied to RLD results. We fit our experimental data with the Lennard-Jones fluid analytical model proposed in Morsali et al. (2005). This RDF model is constrained by 65 constants that Morsali et al. (2005) calculated from 353 molecular dynamics simulations of argon atoms interacting via the Lennard-Jones potential under a range of state variables. We set the Lennard-Jones length parameter (σ = 13.9 nm) to scale the x axis to reduced r^*, and optimized the variables for reduced temperature (T^* = 17.296) and reduced density (ρ^* = 1.143) by minimizing the mean squared error between the model and our data with the Python SciPy library's global optimization algorithm described in Wales and Doye (1997). With these values for σ, T^* and ρ^*, the Morsali RDF model fit our data well (root mean squared error: 0.0935, mean squared error: 0.0087, maximum squared error: 0.5308, standard deviation of squared error: 0.05819). Our optimized value for T^* fell outside the range of molecular dynamics data that Morsali et al. (2005) used to develop their model; therefore, we also determined the Lennard-Jones RDF for these conditions using a molecular dynamics simulation (Plimpton, 1995).

The simulation was performed using the LAMMPS Molecular Dynamics Simulator (Plimpton, 1995). The simulation space was a 10x10x10 box with hard-wall boundary conditions and N = 1000 particles. We employed a Lennard-Jones interaction between particles, defined by:

$$V_{LJ} = \epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right].$$

We set the interaction strength at $\epsilon = 1$, and the Lennard-Jones length scale σ was chosen to reproduce the reduced density via $\rho^* = N\sigma^3/V$, where V is the system volume. The reduced density (ρ^* = 1.143) and reduced temperature (T^* = 17.296) were set to the values determined by the analytical fit of the experimental RLD to the Lennard-Jones fluid RDF formula from Morsali et al. (2005) as described above. To avoid boundary artifacts, the RDF was computed from the particle-to-particle distances around a set of reference particles occupying a small 2x2x2 box at the center of the full simulation volume. To compare the simulated RDF with the experimental RLD curve, we set the vertical scale by matching the far distance amplitude, and then set the horizontal scale by a least-squares fit to the experimental data.

There is an important geometrical difference between the Lennard-Jones RDF models and our experimental measurements of the pyrenoid's RLD. Lennard-Jones models consider spherical particles [such as argon (Morsali et al., 2005)] that always have one minimum-energy distance between particle centers (r_m). However, Rubisco particles are not spherical, and have a minimum diameter of ~10 nm and a maximum diameter of ~13 nm. Thus, unlike spherical particles, a single mimimum-energy distance between the surfaces of neighboring Rubiscos (as shown in our linked-network model, Figure S4J) yields a range of minimum-energy distances between particle centers (a distribution of r_m values) instead of a single, discrete distance. This may explain why our experimental data had a broader first peak than that observed in Lennard-Jones models.

The Lennard-Jones potential describes the attractive and repulsive forces between small molecules (our models used argon), the balance of which results in a single preferred distance between neighbors. Rubisco particles exist on a much larger size scale than argon atoms, and are thus subject to different molecular forces such as protein interactions. However, we reasoned that this simple Lennard-Jones fluid might nonetheless serve as an informative analogy for liquid organization. Interestingly, the Lennard-Jones distribution appears to be robust to changes in scale, matching well not only to ~12 nm pyrenoid Rubisco and 120 nm colloidal particles (Gu et al., 2010), but also to 4 μm colloidal particles (Thorneywork et al., 2014). Due to differences in the underlying molecular forces, the analogy between the pyrenoid and a Lennard-Jones fluid is limited to the interpretation that the pyrenoid matrix may be liquid; the analogy is not intended for the application of other Lennard-Jones descriptors to properties of the pyrenoid.

Spot tests to assay growth of Venus-expressing strains

Pre-growth cultures were inoculated in TAP as described above. After 4 days, 50 mL of each culture were washed of acetate twice by centrifuging for 5 min at 2000 g and resuspending the pellet in 20 mL TP. After washing, cells were counted three times (Countess II Automated Cell Counter; ThermoFisher Scientific) and diluted serially to contain 10^4, 10^3, 10^2, or 10 cells per 15 μL. 15 uL of each

strain in each concentration were spotted in replicate onto TAP and TP plates and left to dry in the dark for one hour. TAP plates were then wrapped in Parafilm and aluminum foil and kept at room temperature in the dark for 12 days before imaging; TP plates were incubated in custom containers with filtered air flow (~0.04% CO_2) with 40 hr of acclimation at 50 μmol photons m^{-2} s^{-1} red and blue LED light, and then 3 days in ~100 μmol photons m^{-2} s^{-1} before imaging. Representative samples out of 4 total replicates (2 biological replicates, each with two technical replicates) per condition are shown.

Fluorescence Recovery After Photobleaching microscopy and analysis

Slide preparation

For live cell imaging, 300 μL of photoautotrophic culture at ~$2x10^6$ cell/mL were plated onto poly-L-lysine-coated 8 well μ-slides (ibidi, 80824) and allowed to settle for ~5 min. The supernatant was removed, and the remaining cells were coated with 300 μL of TP in 1.5% UltraPure Low Melting Point Agarose (Invitrogen) at ~40°C and allowed to cool for ~20 min. For fixed cell imaging, cells were fixed in 10% glutaraldehyde as follows: ~70 × 10^6 cells were pelleted by centrifugation for 10 min at 1000 g at room temperature, resuspended in 6 mL of 10 mM HEPES-KOH (pH 7) at 4°C, and transferred to a cold 20 mL glass beaker. 1.5 mL of 10% glutaraldehyde solution (300 uL of 50% glutaraldehyde in 1.2 mL of 10 mM HEPES-KOH, pH 7) was added in successive 100 uL drops every five seconds while swirling vigorously. The mixture was incubated on ice and agitated every 10 min for one hour, after which the fixed cells were centrifuged (1000 g, 5 min, 4°C), resuspended in 6 mL cold 10 mM HEPES-KOH (pH 7), centrifuged again, and resuspended in 1 mL of 10 mM HEPES-KOH (pH 7). 300 μL were plated for microscopy, as described above.

FRAP image acquisition

Pyrenoids were imaged at mid-plane at room temperature on a spinning disk confocal microscope (Leica DMI6000B custom-adapted with a Yokogawa CSU-X1 spinning disk head; a Photometrics Evolve 512 camera; and Intelligent Imaging Innovations SlideBook software, Vector FRAP, LaserStack, and mSAC spherical aberration systems), with a 100X oil objective (HCX PI APO, 1.4-0.7 NA; Leica). Venus fluorescence was imaged by excitation at 514 nm and emission with a YFP 540/15 filter (Semrock) and 445/514/561 nm Yokogawa dichroic beamsplitter (Semrock) under the following conditions: 100 ms exposure every 3 s for 100 images, with < 2 mW laser power (measured at the fiber), and a camera gain of 3 and EM gain of 300, with Adaptive Focus Control active before every acquisition. FRAP experiments were conducted one by one on individual pyrenoids centered in the field of view and far from previously bleached pyrenoids on the slide. Vector was used to direct the 514 nm laser at full power for photobleaching (~18 mW, measured at the Vector fiber), which took place between the third and fourth image captures. on a sub-resolution region of interest positioned at the edge of each selected pyrenoid; this bleached ~$^1/_3$-$^1/_2$ of the cross-section of the pyrenoid. For the comparative recoveries of RBCS1-, RCA1-, and EPYC1-Venus graphed in Figure 5, the photobleaching event consisted of one repetition of a 4 ms exposure on a 2x2 pixel region of interest. Due to alterations in the light path of the microscope, subsequent acquisition and bleach conditions were altered to achieve the same cross-sectional proportion of bleaching: for the images of live and fixed RBCS1-Venus FRAP pyrenoids shown in Figure 5, the photobleaching event consisted of 2 repetitions of a 10 ms exposure directed to a 4x4 pixel region of interest; for *epyc1* EPYC1-Venus FRAP, 1 repetition of a 10 ms exposure on a 2x2 pixel region of interest was used. Displayed results are from independent experiments conducted at least three (live RBCS1-, RCA1-, and EPYC1-Venus), two (fixed RBCS1-Venus), or one (*epyc1* EPYC1-Venus) times. Sample sizes were chosen based on similar published studies (Klammt et al., 2015).

FRAP quantitative image analysis

FRAP images were analyzed in FIJI software (Schindelin et al., 2012). The StackReg translation registration plug-in (Thévenaz et al., 1998) was used to align image sets that shifted in XY during imaging. For each bleached pyrenoid that was analyzed, a kymograph was generated using the "Reslice" function on a 3-pixel-wide, 3.5 μm-long rectangle that spanned the bleached and unbleached regions (shown in Figure 5), avoiding any saturated pixels. To plot the fluorescence recoveries, the signal along 3-pixel-wide rectangles over the bleached and unbleached regions of each kymograph were measured, respectively; these signals were exported to Excel and compared to calculate fluorescence recovery, shown as "intra-pyrenoid homogeneity" over time for each bleached pyrenoid, as follows: for each time point, the signal from the bleached region of the kymograph was divided by that from the unbleached region.

The signal ratio at the fourth time point ($t = 0$; immediately post-bleach) was defined as $y = 0$ for each pyrenoid by subtraction. Each recovery plot was then normalized to the average of the ratios of the three pre-bleach time points, which was defined as $y = 1$. This normalized recovery was averaged over the stated number of pyrenoids and displayed with the standard error of the mean for each strain.

Thus, it is important to note that our FRAP curves represent homogeneity over the bleached and unbleached regions, correcting for signal loss due to repeated measurements. In our plots, therefore, a return to $y = 1$ is not a return to the initial signal intensity, but a return to the initial signal homogeneity. This is different from how FRAP recovery is often plotted, and results in seemingly "high" recoveries. However, our FRAP images and kymographs show that the final signal intensities are less than the initial intensities, but that the signal disparity between the bleached and unbleached regions relaxes, and our FRAP kymographs resemble those of other liquid-like organelles (Kroschwald et al., 2015).

Additionally, plotting "homogeneity" allows us to control for potential variations in the proportion of the pyrenoid that was bleached in each event. Because the fluorescence recovery in the pyrenoid cannot be said to draw from an infinite pool, variation in the ratio of the bleached to unbleached pools could produce artifacts in the observed final recovery.

To calculate the fraction of the total pyrenoid volume that was bleached, Z stacks of glutaraldehyde-fixed RBCS1-Venus pyrenoids were captured before and after a bleaching event; in live cells, fluorescence recovered too quickly to section through the entire pyrenoid. 23 planes were imaged with a 0.23-μm step size (spanning 5.06 μm) before and after bleaching, and pre- and post-bleach 3D volumes were reconstructed for 10 pyrenoids in Imaris (Bitplane), using the same intensity thresholds for pre- and post-bleach volumes. The total fluorescence intensity and volume of each reconstructed pre- and post-bleach pyrenoid was exported and analyzed in Excel.

"Magic number" modeling

Model parameters

Simulations were performed using a square grid system of 50 × 50 sites with periodic boundary condition. In the model, each Rubisco is represented as a 4 × 2 rectangle, and each EPYC1 occupies several connected (nearest neighbor) grid sites. Rubisco and EPYC1 are self-avoiding. However, each grid site can be simultaneously occupied by EPYC1 and Rubisco, and if so, the two are considered to form a specific bond. This scheme allows for stoichiometric bonding between EPYC1s and Rubiscos without the artifacts due to crowding that would occur if the two were prevented from occupying the same sites. The simulations also include a weak non-specific attractive interaction between all occupied nearest-neighbor sites, EPYC1-EPYC1, EPYC1-Rubisco, and Rubisco-Rubisco.

We performed Markov-Chain Monte Carlo simulations using the Metropolis algorithm (Metropolis et al., 1953). Briefly, in each simulation step we randomly propose a move of the EPYC1-Rubisco configuration. The move is always accepted if it reduces system energy, and accepted with probability $e^{-(E_f-E_i)/k_BT}$, where E_f and E_i are the final and initial energies, if the move increases system energy. Three categories of moves are proposed: single-EPYC1 moves, single-Rubisco moves and EPYC1-Rubisco joint moves (Figure S7). Single-EPYC1 moves are standard lattice-polymer local moves: the end-point move, the corner move, and the reptation move. Single-Rubisco moves consist of one-step translations in the four cardinal directions and a 90-degree rotation around the Rubisco's center. In the regime of strong specific bonds, EPYC1s and Rubiscos are typically held together by multiple specific bonds, which leads to dynamical freezing. To enable the system to better explore configuration space, we include EPYC1-Rubisco joint moves such that connected clusters of EPYC1s and Rubiscos move together, without breaking any specific bonds. The joint moves consist of translating a connected cluster of EPYC1s and Rubiscos together or rotating the whole cluster by 90-degrees around any point. To obtain thermalized ensembles, we follow a simulated two-step procedure: we keep k_BT constant and gradually increase bond strength. We first increase the non-specific bond from 0 to 0.1 k_BT in 0.005 k_BT increments, keeping the specific bond energy at 0 k_BT. Then the specific bond energy is increased from 0 to 11 k_BT in 0.04 k_BT increments, while the non-specific bond energy is kept at 0.1 k_BT. Each step of annealing is simulated with 50,000-150,000 Monte-Carlo steps and results are averaged over 20-100 of the resulting thermalized snapshots.

Analysis of clustering in the 2D simulations

To assess the extent of clustering of Rubiscos, we consider a cluster to be a group of Rubiscos that are connected by EPYC1s via specific bonds. To quantify the fraction of Rubiscos in large clusters, we employ a cluster-size cutoff of 10+ Rubiscos (Figures 7D–7F). This avoids ambiguities due to smaller clusters that form independent units with all specific bonds satisfied (e.g., 3 Rubiscos with 8 EPYC1s each of 3 binding sites). To determine the onset of clustering, we fit the fraction of Rubiscos in large clusters with a degree 4 spline (Figure S7); for the case of EPYC1s with 8 binding sites, because of the delayed onset of clustering we only use the data for concentrations > 40% for the fit.

Three-dimensional off-lattice model

Molecular dynamics simulations were performed using the LAMMPS Molecular Dynamics Simulator (Plimpton, 1995). We model each Rubisco as a sphere with 4 small spherical binding sites on each end, and each EPYC1 molecule as a polymer with 3, 4, or 5 binding sites (Figures 7G–7I; Movie S6). When the EPYC1 and Rubisco binding sites overlap, specific bonds are formed (see below). We simulate a cubic box of 120 nm in each dimension, with periodic boundary conditions. In the simulation, particles representing Rubisco and EPYC1 interact both non-specifically and via specific EPYC1-Rubisco bonds. Particles denoted by A and B interact non-specifically with each other through the Lennard-Jones potential with a cutoff

$$V(r) = 4\epsilon_{AB}\left[\left(\frac{\sigma_{AB}}{r}\right)^{12} - \left(\frac{\sigma_{AB}}{r}\right)^{6}\right] - 4\epsilon\left[\left(\frac{\sigma_{AB}}{r_c}\right)^{12} - \left(\frac{\sigma_{AB}}{r_c}\right)^{6}\right], r < r_c$$

where $\sigma_{AB} = \sigma_A + \sigma_B$ is the sum of the effective radii of the two particles, ϵ_{AB} is the interaction strength between the two particles, and r_c is the cutoff of the interaction range.

Each Rubisco in the simulation is represented by a sphere of radius $R_R = 5$ nm, which is the radius of gyration of Rubisco (Keown, et al., 2013). Since the specific Rubisco binding sites for EPYC1 are not yet known, we employ a simple model in which each Rubisco has 4 binding sites, forming a rigid square of edge length 3.4 nm, on each of its two ends. The radius of each binding site is 0.9 nm and the center of the binding site is 0.9 nm away from the Rubisco sphere (i.e., 5.9 nm from the center of the Rubisco sphere). Each EPYC1 is represented by 3, 4, or 5 connected binding sites: the binding sites are spheres of radius $R_E = 0.9$ nm, which is the radius of a compact region of 18 amino acids representing the repeat region of EPYC1. We model the unstructured chain of 34 amino acids separating these repeats as harmonic springs with zero rest length and stiffness 0.24 k_BT/nm^2, reflecting the entropic elasticity of

a worm-like polymer chain consisting of 34 units of size 0.35 nm (the approximate size of an amino acid) with a persistence length of 0.5 nm (a rough consensus for polypeptide chains [Hofmann et al., 2012, Cheng et al., 2010]).

We set the non-specific Lennard-Jones interactions between two EPYC1 binding sites, between two Rubisco cores, and between one EPYC1 binding site and one Rubisco core to be the Lennard-Jones potential with a cutoff with interaction energy $\epsilon = 0.1\ k_BT$ and σ the sum of the radii of the two interacting particles. We set $r_c = 1.4\sigma$ as a cutoff so that the system does not include long-range interactions.

In addition to the non-specific interaction, we model the attractive specific interaction between Rubisco and EPYC1 binding sites as a soft potential

$$V(r) = \epsilon_b\left(1 + cos\left(\frac{\pi r}{r_c}\right)\right), r < r_c \ \#(1)$$

with binding energy $\epsilon_b = -10\ k_BT$, r the distance between binding site centers, and a cutoff distance $r_c = 0.45$ nm. The minimum of energy for this interaction corresponds to fully overlapping binding sites. The one-to-one character of EPYC1-Rubisco specific bonds is guaranteed by the Lennard-Jones repulsion between EPYC1 binding sites, which prevents them from binding to the same Rubisco binding site, and by an additional soft repulsion (with the same functional form as Equation (1) and $\epsilon = +50\ k_BT$) between two Rubisco binding sites, which prevents them from binding to the same EPYC1 binding site. There is no Lennard-Jones interaction between a Rubisco binding site and an EPYC1 binding site, in order to allow them to overlap and form a specific bond.

We performed molecular dynamics simulations with Langevin dynamics in the NVT ensemble. For convenience, we attributed equal mobilities to all particles, chose a simulation unit length of 1 nm and a time step of 0.008 in LJ units. Each simulation consisted of a total of 150,000,000 time steps. The first 100,000,000 steps were used to let the system reach thermal equilibrium and a snapshot was taken every 10,000 steps after that for clustering analysis.

Analysis of clustering in the 3D simulations

To identify specific bonds between EPYC1 and Rubisco binding sites, we first computed the distance distribution histogram between all EPYC1 and Rubisco binding sites, and found a clear gap around 0.5 nm. Below this gap, the EPYC1 and Rubisco binding sites form a specific bond; above this gap, the binding sites are farther away from each other than the binding site diameter, and thus do not form a specific bond. As in the 2D case, for each snapshot we tabulated the Rubiscos connected by EPYC1s through specific bonds and plotted the fraction of Rubiscos in clusters with 10+ Rubiscos (Figure S7K).

Robustness of the magic-number effect

To confirm that the magic-number effect is robust with respect to our choice of interaction parameters, we performed additional 3D simulations in which we set the EPYC1 and Rubisco binding-site radii to be the same, and equal to either 0.8 nm, 1 nm, or 1.25 nm (i.e., factor of two volume changes), with the distance between the center of the Rubisco core and the Rubisco binding site as 5.8 nm, 6 nm, or 6.25 nm, correspondingly. We also varied the non-specific Lennard-Jones interaction energy and the EPYC1 inter-binding-site spring stiffness by a factor of 2 (increase or decrease) for a fixed concentration of Rubisco (2% volume fraction) and an equal number of Rubisco and EPYC1 binding sites. In Table S1, we report the resulting percentage of Rubiscos in clusters with 10+ Rubiscos for EPYC1s with 3, 4, or 5 binding sites for each of these parameter sets.

QUANTIFICATION AND STATISTICAL ANALYSIS

Division image analysis

Chi-square statistic calculations on the relative proportions of each observed inheritance pattern in RBCS1- and EPYC1-Venus were performed using an online calculator (http://www.socscistatistics.com/tests/chisquare2/Default2.aspx).

The Wilcoxon Matched-Pairs Signed-Ranks Test was used to quantify the difference in the pyrenoid signal during division using an online calculator (http://www.fon.hum.uva.nl/rob/SignedRank/Signed_Rank_Test.html). This nonparametric test assumes data are paired, but that pairs are independent of each other. In this case, each pair was made up of the signal from the pyrenoind of a single mother cell at 1) the minimal value during dimming, and 2) the value 15 min after that point (including both daughters); thus, different pairs represent different mother cells. Sample size and error definitions can be found in the relevant figure legends.

Tomogram analysis

Nine tomograms were screened for their reconstruction quality based on IMOD patch-tracking scores and the resolution of Rubisco subtomogram averages produced from each tomogram's extracted particles. The final dataset used for quantitative analysis was five tomograms, representing five cells from three separate liquid cultures.

Following the localization of Rubisco holoenzymes within pyrenoid tomograms by volume masking, template matching, subtomogram averaging and classification (described above), Rubisco concentrations for each tomogram were calculated by dividing the number of particles in the "positive" class by the masked volume of the pyrenoid matrix. Error values displayed in the text are standard deviations, unless marked otherwise.

FRAP image analysis

Images were screened for quality before quantitative analysis: any pyrenoids with indistinct bleach regions, many saturated pixels, or loss of focus during recovery were discarded. If a pyrenoid contained 1-2 saturated pixels, those pixels were not included in quantification. Average recovery rates for n = 44 (EPYC1-Venus), 48 (RCA1-Venus), 42 (live RBCS1-Venus), 8 (fixed RBCS1-Venus), or 28 (*epyc1* EPYC1-Venus) pyrenoids are shown with error bars or shading representing standard error of the mean (SEM). FRAP data were collected over 1 (*epyc1* EPYC1-Venus) or 3 (RBCS1-, EPYC1-, and RCA1-Venus) biological replicates. For comparison of recovery rates between strains, a one-way ANOVA with post hoc Bonferroni means comparison was performed using OriginPro software; the variances for these recovery rates were 6.36×10^{-4} (RBCS1-Venus), 6.58×10^{-4} (RCA1-Venus), and 8.88×10^{-4} (EPYC1-Venus). A one-way ANOVA test assumes normal distributions, independent samples, and equal variances, all of which appeared to be met in our data

DATA AND SOFTWARE AVAILABILITY

The computer code used in modeling Rubisco-EPYC1 aggregation and fitting the Lennard-Jones RDF to experimental data are available through https://github.com/binarybin/RubiscoSimulation and https://github.com/anmartinezs/LJRDFfit/, respectively.

The datasets generated during and/or analyzed during the current study are available from the corresponding authors upon reasonable request. Sequence information for the fluorescent constructs is available from GenBank under accession numbers GenBank: KY550376 (pLM005-RBCS1-Venus), GenBank: KX077944 (pLM005-EPYC1-Venus), and GenBank: KY550375 (pLM005-RCA1-Venus). The in situ subtomogram average of *Chlamydomonas* Rubisco has been deposited in the EMDataBank with accession code EMD-3694.

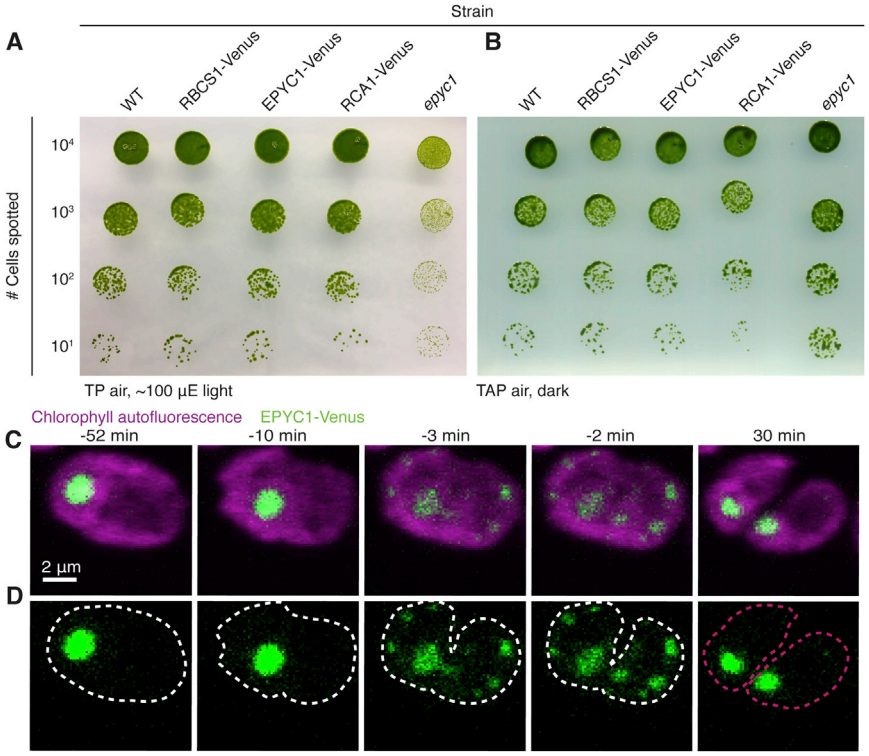


Figure S1. Fluorescently Tagged Lines Grow Normally and Specks Can Appear Transiently Even during Fission, Related to Figures 1, 2, 3, 5, and 6

(A and B) Spot tests on cells grown in photoautotrophic conditions that require a functional pyrenoid (A) versus heterotrophic conditions that do not require a functional pyrenoid (B) show that tagging pyrenoid matrix proteins (RBCS1, EPYC1, RCA1) with the fluorescent protein Venus results in wild-type-like growth. In comparison, the *epyc1* mutant, which has perturbed pyrenoid structure (Mackinder et al., 2016), displays a growth defect under conditions that require a functional pyrenoid. TP air cells were imaged ~5 days after spotting; TAP dark cells were imaged ~12 days after spotting.

(C and D) Additional example of pyrenoid fission in an EPYC1-Venus cell, with transient puncta appearing in the stroma. Images are summed through *Z*, with chlorophyll autofluorescence shown in magenta, and RBCS1-Venus in green; $t = 0$ is the first observation of a gap in chlorophyll between the daughter pyrenoids in the first division shown. Dashed curves represent approximate chloroplast outlines in the mother (white) and daughter (pink) cells.

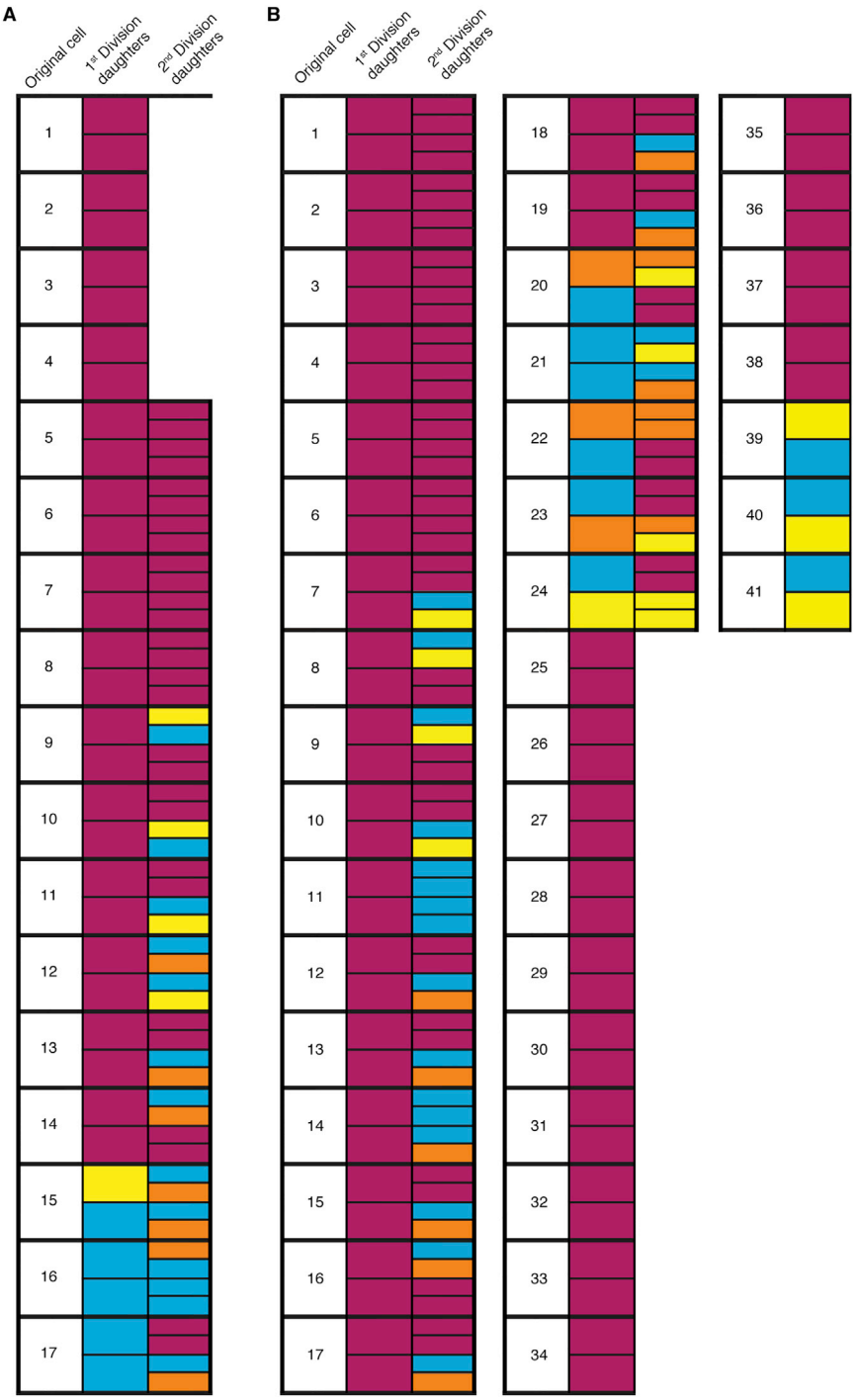


Figure S2. The Mode and Lineage of Pyrenoid Inheritance Were Tracked for RBCS1-Venus and EPYC1-Venus Daughter Cells, Related to Figures 1, 2, 3, and 6

(A and B) Pyrenoid inheritance was observed in daughter cells resulting from the first division (middle columns) or second division (right columns) of 17 original RBCS1-Venus cells, resulting in 86 daughters (A) and 41 original EPYC1-Venus cells, resulting in 178 daughters (B). Colors correspond to the pattern of pyrenoid inheritance in that cell. As in Figures 1 and 2, pyrenoid fissions are displayed in magenta, and inheritance of an entire previously existing pyrenoid is blue. Cases in which no pyrenoid was inherited and puncta appeared are shown in yellow; cases in which neither a pyrenoid nor puncta were inherited are shown in orange. In some lineages, the mother cell contained two pyrenoids at the start of observation, allowing both daughters to be labeled in blue. Some observed cells only divided once. The observed proportions of these inheritance patterns are displayed in Figure 2. Images from selected divisions in this figure are presented elsewhere as follows: Figures 1A–1C, 6A, and 6B, and Movie S1 are RBCS1-Venus original cell 7 divisions 1 and 2; Figure 1D is EPYC1-Venus original cell 9

(legend continued on next page)

division 1; Figure 1E is RBCS1-Venus cell 9 division 1; Figure 2A and Movie S2 are RbcS1-Venus cell 14 division 2; Figure 2B is RBCS1-Venus cell 9 division 2; Figure 2C is EPYC1-Venus cell 19 division 2; Figure 2D is EPYC1-Venus cell 40 division 1; Figures 2G–2I are RBCS1-Venus cell 15 division 2, RBCS1-Venus cell 10 division 2, and EPYC1-Venus cell 8 division 2, respectively; Figure 3A is RBCS1-Venus cell 12 division 1, B is RBCS1-Venus cell 6 division 1, C is RBCS1-Venus cell 5 division 1, D is RBCS1-Venus cell 13 division 1, E is RBCS1-Venus cell 3 division 1, and F is EPYC1-Venus cell 19 division 1; Figures S1C and S1D are EPYC1-Venus cell 29 division 1; Figure S6A is RBCS1-Venus cell 5 divisions 1 and 2; Movie S3 and Figure S6B are RBCS1-Venus cell 9 divisions 1 and 2; Figure S6C is RBCS1-Venus cell 14 divisions 1 and 2, D is RBCS1-Venus cell 13 divisions 1 and 2, E is RBCS1-Venus cell 12 divisions 1 and 2, and H is EPYC1-Venus cell 5 divisions 1 and 2.

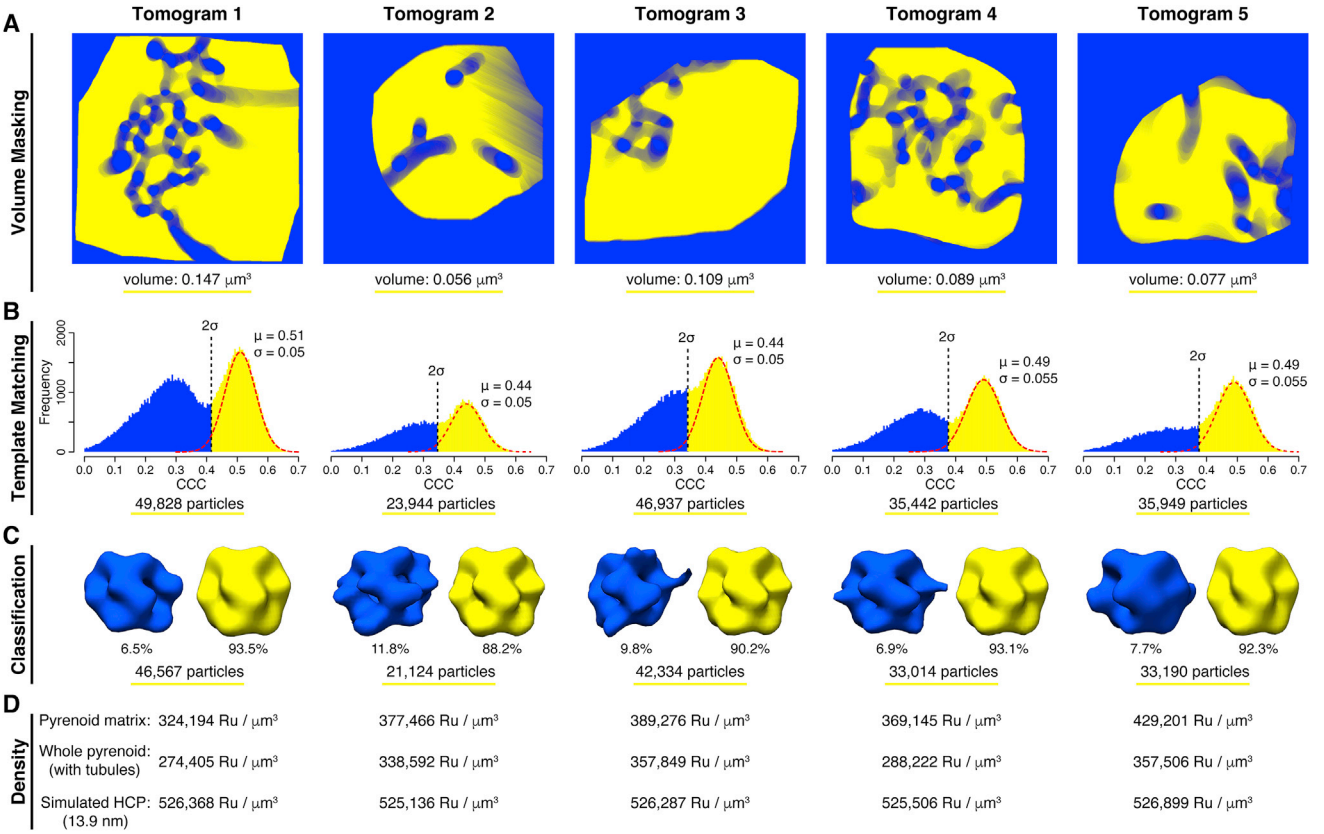


Figure S3. In Situ Localization of Rubisco Holoenzymes within Tomograms, Related to Figure 4

(A) Tomogram volumes were masked to separate the pyrenoid matrix (yellow), which was template matched for Rubisco particles, from the rest of the volume (blue), which was excluded from the search. 2D projection images of the 3D masks are shown. The volumes contained within the matrix masks are underlined in yellow.

(B) Following template matching, particles were exhaustively extracted and plotted in histograms by their cross-correlation coefficient (CCC) to the Rubisco template. Clear true-positive peaks with higher CCC values than the false-positive noise were apparent for all tomograms. Each true-positive peak was fit with a Gaussian distribution (red dashed line), and a two standard deviation (2σ) cutoff was used to separate particles to be processed further (yellow, ~97.5% of true positives) from those to be discarded (blue). The numbers of particles that were kept are underlined in yellow.

(C) Following 3D subvolume alignment, the particles were subjected to hierarchical classification to remove contaminating false positives. Between 6.5% and 11.8% of particles were removed per tomogram (blue averages), leaving only clean true-positives (yellow averages). The averages were all filtered to 30 Å resolution for clear comparison. The numbers of particles remaining after classification are underlined in yellow.

(D) Calculation of the concentration of Rubisco particles within each masked pyrenoid matrix region, adjusted for the ~97.5% sampling of true positives in (B). For comparison to lower-resolution light microscopy measurements, concentrations that do not omit the volume of the pyrenoid tubules were also calculated. Concentrations for the HCP simulated data generated within the same masked matrix volumes (Figure S4E) are also listed.

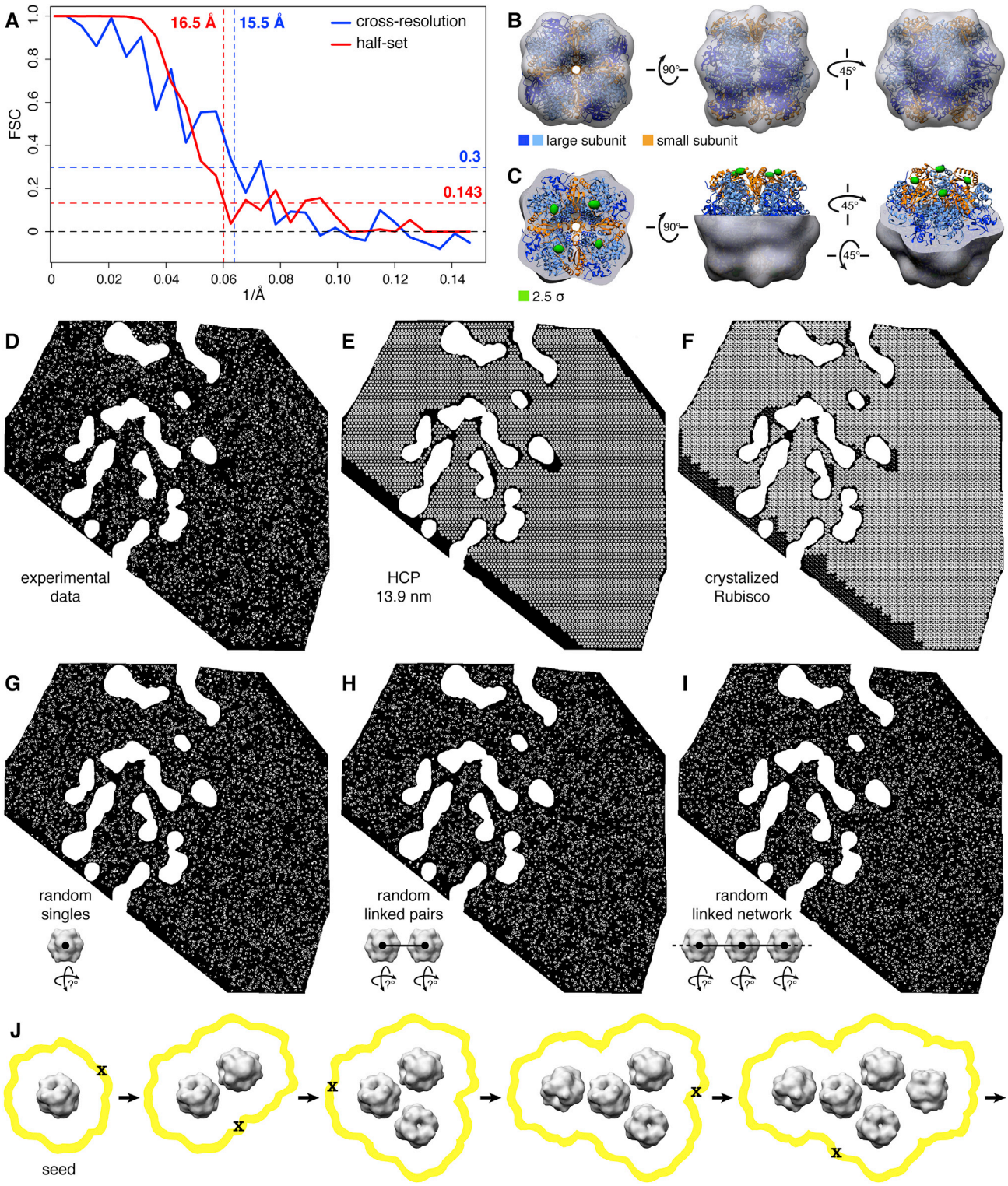


Figure S4. Assessment of the Rubisco In Situ Subtomogram Average and Generation of the Crystalline and Randomly Packed Simulated Pyrenoid Data, Related to Figure 4

(A) Fourier shell correlation (FSC) resolution estimates for the subtomogram average shown in Figure 4C. Resolution was calculated both by cross-resolution of the full dataset to the crystal structure (Taylor et al., 2001) (blue line, 15.5 Å, 0.3 cutoff) and by randomly splitting the dataset and comparing the consistency of the two half-sets (red line, 16.5 Å, 0.143 cutoff).

(legend continued on next page)

(B) The Rubisco crystal structure (Taylor et al., 2001) fitted into the subtomogram average.

(C) A difference map between the crystal structure and the subtomogram average. There is only a very small significant difference (green volume: 2.5σ variance) near the Rubisco small subunits. Half of the subtomogram average has been cut away to more clearly show the position of the difference density relative to the crystal structure.

(D–I) Experimental Rubisco positions (D), HCP with 13.9 nm spacing between Rubisco centers (E), Rubisco packing within a Rubisco crystal (F), randomly positioned single Rubisco particles (G), randomly positioned linked pairs of Rubisco with 13.9 ± 1.5 nm between pair centers (H), and a randomly positioned network of linked Rubisco (I), each mapped into the same tomogram volume (corresponding slices are shown for each volume). In (G – I), the rotation symbols with "?°" indicate that the particles have random orientations in addition to their random positions.

(J) Schematic for the generation of the random linked network in (I), illustrated from left to right. A seed Rubisco template is randomly placed into the matrix volume, and a zone for the potential placement of neighbors (yellow) is defined in a shell occupying the volume that is 5.5-7 nm from the seed particle surface. A second particle is randomly placed within the neighbor zone (center position of particle indicated by "X"), and the zone for potential neighbor placement is then updated to occupy a shell of 5.5-7 nm around both particles. The rounds of random neighbor placement and zone updating are iteratively repeated to grow the random linked network. After a maximum of 8 particles have been added to a network, a new seed is randomly placed elsewhere in the matrix volume and another network is formed.

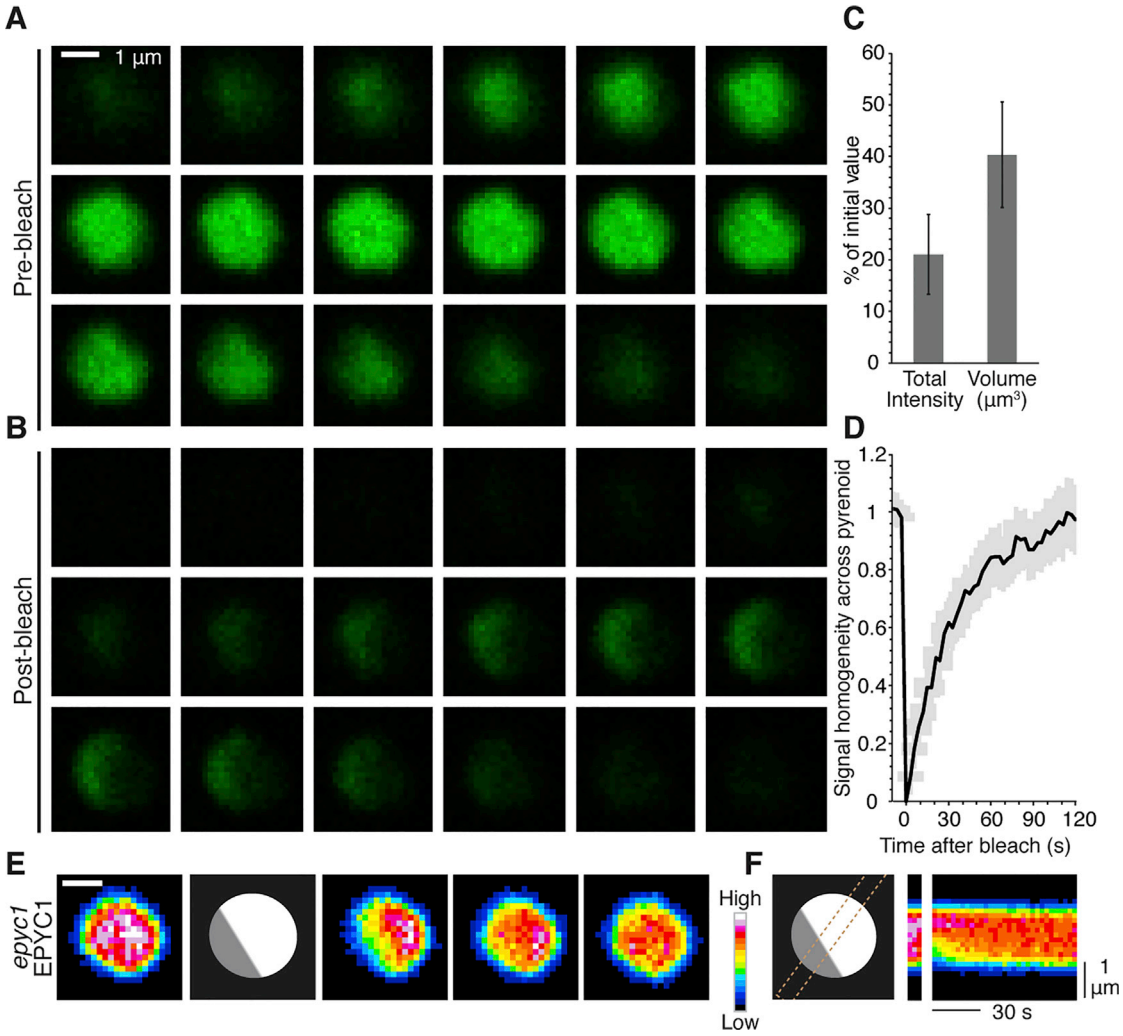


Figure S5. The Bleached Region during FRAP Experiments Extends Vertically through the Pyrenoid Volume, and RCA1-Venus and EPYC1-Venus Are Also Mobile within the Pyrenoid, Related to Figure 5

(A and B) Serial sections through a fixed RBCS1-Venus pyrenoid before (A) and after (B) a bleach event. Sections are 0.23 µm apart.

(C) Average total intensity and volume of the Venus signal after bleaching, as a percent of the initial value, in 3D reconstructions of 10 fixed RBCS1-Venus pyrenoids before and after a bleach event. Error bars are standard deviation.

(D) Average FRAP recovery profiles ± SEM for 28 pyrenoids in live *epyc1* mutants complemented with EPYC1-Venus. The y axis is intra-pyrenoid homogeneity.

(E and F) Examples of FRAP in an *epyc1* EPYC1-Venus pyrenoid. Cartoons depict the approximate bleached region (dark gray). Image stills from the recovery time-course (E) and corresponding kymograph (F) as shown in Figure 5.

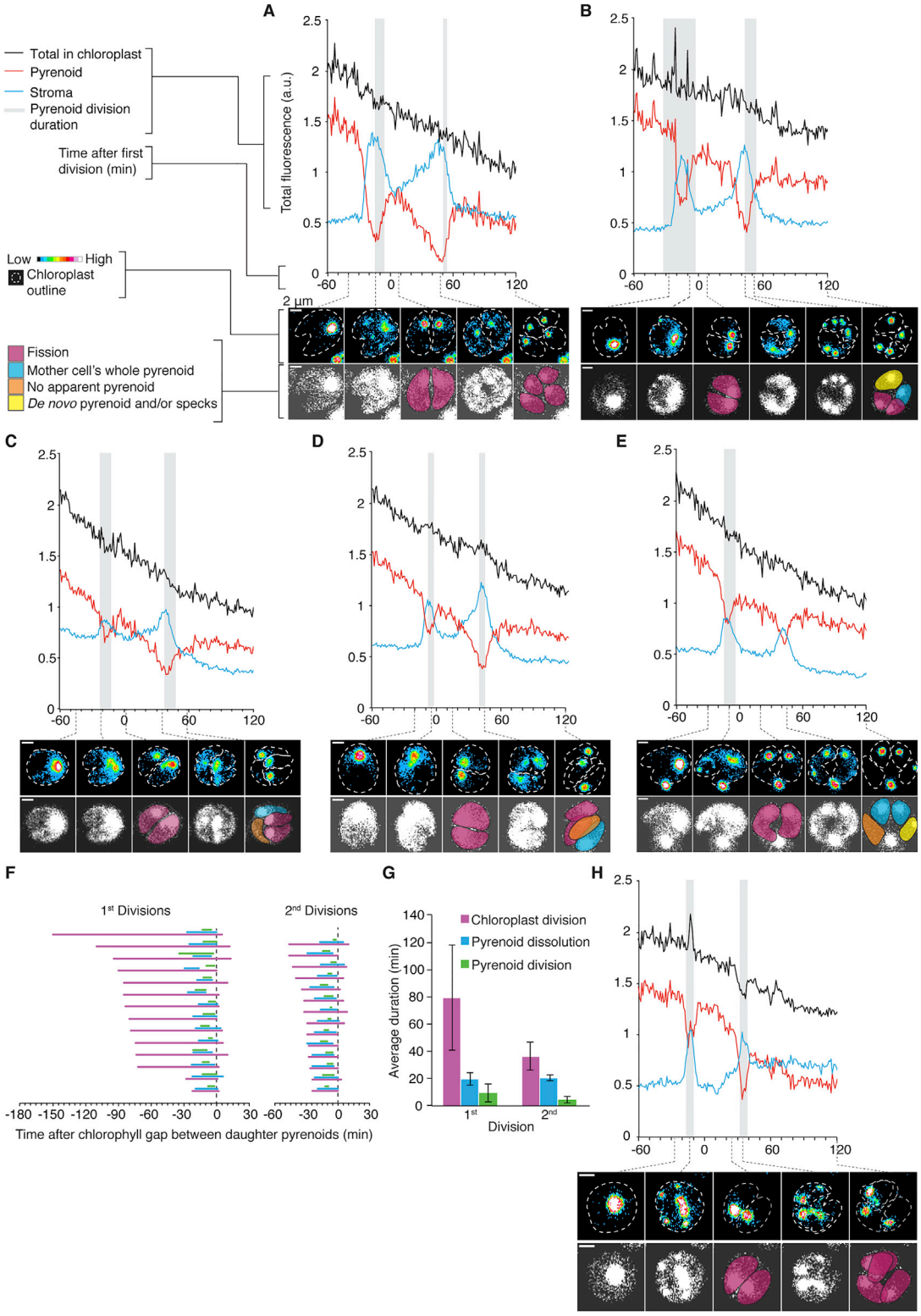


Figure S6. Pyrenoid Matrix Components Undergo Relocalization out of and into the Pyrenoid during Cell Division, Related to Figures 1, 2, and 6

(A–E) Examples of the changing localization of the RBCS1-Venus signal during divisions in five lineages. As in Figures 6A and 6B, the plots represent the sum of the fluorescence through the whole Z stack in each masked region over time. The time window during which the pyrenoid is undergoing fission is highlighted in

(legend continued on next page)

gray; note that the pyrenoids in the second division of (E) do not undergo fission, and thus lack a gray highlight. $t = 0$ is the minute in which a chlorophyll gap first appears between daughter pyrenoids in the first division. Underneath each plot, snapshots of the RBCS1-Venus signal at several different points are displayed in two ways: on the top, the RBCS1-Venus signal is false-colored based on intensity, with the approximate chloroplast boundaries outlined in white, and signal from outside the chloroplast masked out. Beneath that, an overexposed version of each image is displayed, with the daughter cells colored to match the mode of pyrenoid inheritance in that cell. The second division in (C) is also shown in Figure 2A; the second division in (E) is also shown in Figure 2B.

(F) Duration and relative timing of the RBCS1-Venus chloroplast (magenta) and pyrenoid (green) division for the pyrenoid fissions, as in Figure 1G, with the duration and relative timing of pyrenoid dissolution superimposed in blue. Left: 1st divisions; right: 2nd divisions. $t = 0$ is the first observation of a gap in chlorophyll between the daughter pyrenoids. We observed only one instance in which pyrenoid dissolution and fission appeared to be temporally distinct; in this case, the mother cell contained two pyrenoids (which is rare, but does occur), and one daughter cell inherited both an entire pyrenoid and a pyrenoid by fission.

(G) Average and standard deviation of the durations of chloroplast (magenta) and pyrenoid (green) fissions, as well as pyrenoid dissolution (blue) for the RBCS1-Venus divisions shown in (F) ($n = 28$).

(H) Example of the changing localization of Venus signal during a series of divisions in EPYC1-Venus cells, displayed as in (A – E).

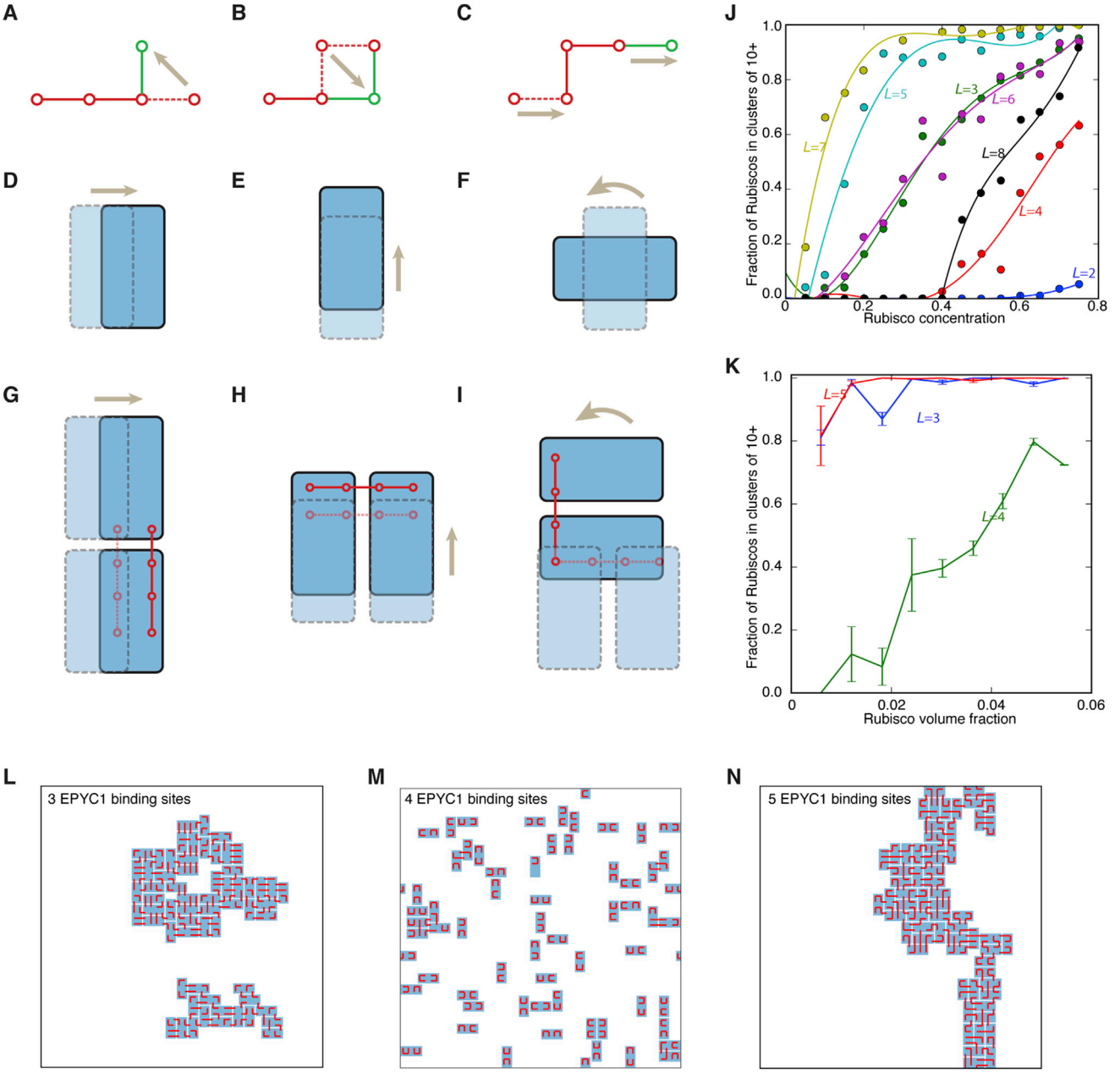


Figure S7. Moves in Monte Carlo Simulations, Determination of Onset of Clustering, and Snapshots of Simulations with Binding of Each EPYC1 Restricted to One End of a Rubisco, Related to Figure 7

(A–I) Schematics of Rubisco and EPYC1 moves in Monte Carlo simulations.

(J) Determination of clustering onsets in Monte Carlo simulations. Data points are from the simulation data in Figure 7J. Each curve for a fixed number of EPYC1 interacting sites is fitted with a 4th order polynomial, and the highest zero-crossing is taken as the onset of clustering in Figure 7H.

(K) Fraction of Rubiscos in clusters of > 10 Rubiscos for EPYC1s with 3, 4, or 5 binding sites in the off-lattice 3D simulation. The specific bond energy is 10 k_BT and the Lennard-Jones nonspecific interaction energy is $\epsilon = 0.1\ k_BT$.

(L–N) Snapshots of an EPYC1-Rubisco system with 8 binding sites on Rubisco holoenzymes for EPYC1 with four binding sites, with the constraint that a given EPYC1 cannot bind to both ends of one Rubisco. Percentage of occupied grid sites is 20% in each layer.