ARTICLE

Check for updates

Condensation of Rubisco into a proto-pyrenoid in higher plant chloroplasts

Nicky Atkinson[1], Yuwei Mao[1], Kher Xing Chan [2] & Alistair J. McCormick [1✉]

Photosynthetic CO_2 fixation in plants is limited by the inefficiency of the CO_2-assimilating enzyme Rubisco. In most eukaryotic algae, Rubisco aggregates within a microcompartment known as the pyrenoid, in association with a CO_2-concentrating mechanism that improves photosynthetic operating efficiency under conditions of low inorganic carbon. Recent work has shown that the pyrenoid matrix is a phase-separated, liquid-like condensate. In the alga *Chlamydomonas reinhardtii*, condensation is mediated by two components: Rubisco and the linker protein EPYC1 (Essential Pyrenoid Component 1). Here, we show that expression of mature EPYC1 and a plant-algal hybrid Rubisco leads to spontaneous condensation of Rubisco into a single phase-separated compartment in Arabidopsis chloroplasts, with liquid-like properties similar to a pyrenoid matrix. This work represents a significant initial step towards enhancing photosynthesis in higher plants by introducing an algal CO_2-concentrating mechanism, which is predicted to significantly increase the efficiency of photosynthetic CO_2 uptake.

[1] SynthSys & Institute of Molecular Plant Sciences, School of Biological Sciences, King's Buildings, University of Edinburgh, Edinburgh EH9 3BF, UK. [2] Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 1206 Gregory Drive, Urbana, IL 61801, USA. ✉email: alistair.mccormick@ed.ac.uk

Rubisco (D-ribulose-1,5-bisphosphate carboxylase/ oxygenase) catalyses net CO_2-fixation in all photosynthetic organisms but also reacts with oxygen, leading to a wasteful loss of previously fixed CO_2, nitrogen and energy[1–3]. In plants with the C3 photosynthetic pathway, which includes most major staple crops (e.g. rice, wheat and soybean), Rubisco has a relatively slow turnover rate and operates under sub-optimal CO_2 concentrations of less than half that required for maximal rates of carboxylation[4–6]. Improving the operating efficiency of Rubisco remains a key strategy to increase productivity in C3 crops[7]. This is an important goal given the steep trajectory of crop yield improvements required to meet the demands of our rising global population[8].

Rubisco is typically evenly distributed throughout the chloroplast stroma in plants. However, most unicellular eukaryotic photosynthetic organisms and some non-vascular land plants have evolved highly efficient CO_2-concentrating mechanisms (CCMs) that condense Rubisco into a microcompartment within the chloroplast called the pyrenoid[9,10]. The CCM functions to enrich the pyrenoid with high concentrations of CO_2 to enhance the performance of Rubisco carboxylation and suppress oxygenation[11]. Introduction of such CCMs into C3 plants is predicted to lead to significant increases in CO_2-fixation efficiency and biomass yield[8,12,13]. Recent work in the model green alga *Chlamydomonas reinhardtii* (hereafter Chlamydomonas) has shown that the pyrenoid is a liquid-like condensate that undergoes liquid–liquid phase separation when the CCM is active[14]. Phase separation is facilitated by multivalent interactions between the intrinsically disordered Rubisco-binding protein EPYC1 and the small subunit of Rubisco (SSU)[15–20], of which residues on the α-helices of the native SSU isoforms are critical for binding to EPYC1. Previously we have shown that Arabidopsis can assemble a functional plant-algal hybrid Rubisco complex, with the native large subunit of Rubisco (LSU) and an SSU from Chlamydomonas, which has similar kinetic properties to wild-type (WT) Arabidopsis Rubisco[21]. However, expression of the full coding sequence of EPYC1 in Arabidopsis with plant-algal hybrid Rubisco did not result in Rubisco condensation[18]. Evidence from Chlamydomonas and in vitro studies suggested that an appropriate stoichiometric balance between EPYC1 and Rubisco, and possibly additional regulatory components, might be required for efficient phase separation[9,14,17,22].

Here we show that expression of mature EPYC1 and a plant-algal hybrid Rubisco in the model C3 plant Arabidopsis leads to spontaneous condensation of Rubisco into a single phase-separated compartment in the chloroplast. The condensate has liquid-like properties similar to a pyrenoid matrix. Furthermore, the condensate displaces the thylakoid membranes and is enriched in hybrid Rubisco containing the algal Rubisco small subunit required for phase separation. Promisingly, photosynthetic CO_2 fixation and growth is not impaired in stable transformants compared to azygous segregants.

Results

Expression of maturated EPYC1 in Arabidopsis

We initially sought to test whether high levels of expression of a mature form of EPYC1 could lead to phase separation in a higher plant chloroplast. EPYC1 was truncated according to the predicted transit peptide cleavage site between residues 26 (V) and 27 (A)[18]. A dual GFP expression system was developed to achieve high levels of EPYC1 expression and a favourable stoichiometry with Rubisco. This consisted of a binary vector containing two gene expression cassettes, each encoding mature EPYC1 with an Arabidopsis chloroplastic signal peptide and fused to a different version of GFP (turboGFP (tGFP) or enhanced GFP (eGFP)) to avoid possible recombination events (Fig. 1a and Supplementary Fig. 1). The dual GFP construct (EPYC1-dGFP) was transformed into WT plants or the Arabidopsis *1a3b* Rubisco mutant complemented with an SSU from Chlamydomonas ($S2_{Cr}$)[21]. The resulting transgenic plants (Ep) expressed both EPYC1::eGFP and EPYC1::tGFP, of which the latter was generally more abundant at the protein level (Fig. 1b and Supplementary Table 2). Previously, immunoblots against full-length EPYC1 expressed in $S2_{Cr}$ or WT plants showed additional lower molecular weight bands indicative of proteolytic degradation[18]. Here, degradation products of the mature EPYC1 were relatively reduced when expressed in $S2_{Cr}$ compared to WT plants.

The fluorescence signal for EPYC1-dGFP in WT plants was distributed evenly throughout the chloroplast (Figs. 1c and 2a). In contrast, EPYC1-dGFP in the hybrid $S2_{Cr}$ plants showed only a single dense chloroplastic signal. Transmission electron microscopy confirmed the presence of a single prominent condensed complex in the chloroplast stroma (Fig. 1d). The condensates were spherical in shape and displaced native chlorophyll autofluorescence (Fig. 1e, f and Supplementary Movie 1), indicating that the thylakoid membrane matrix was excluded from the condensate. In protoplasts of leaf mesophyll cells, a condensate was visible in each chloroplast (Supplementary Fig. 2a), and the average size of the condensates was related to the expression level of EPYC1-dGFP (Fig. 1b and Supplementary Fig. 2b, c). We observed that the average diameter of the condensates was 1.6 ± 0.1 μm ($n = 126$; 42 from three individual $S2_{Cr}$ transgenic lines), which was comparable with that of the size range measured for the Chlamydomonas pyrenoid (1.4 ± 0.1 μm; $n = 55$)[23] (Fig. 1g and Supplementary Fig. 2c). The estimated volume of the condensates was 2.7 ± 0.2 μm^3 (ca. 5% of the chloroplast volume). Variations in condensate volume within individual $S2_{Cr}$ transgenic Ep lines were not correlated with chloroplast volume, suggesting that regulation of condensate formation and size was largely independent of chloroplast morphology, and likely limited by the influence of the surrounding network stiffness (i.e. the stroma, which is densely packed with thylakoids) on the dynamics of droplet ripening[24].

Condensates were also observed when EPYC1-dGFP was expressed in the Arabidopsis *1a3b* Rubisco mutant complemented with a native Arabidopsis SSU modified to contain the two α-helices necessary for pyrenoid formation from the Chlamydomonas SSU ($1A_{At}MOD$)[15,21] (Fig. 1c). However, condensates in the $1A_{At}MOD$ background were less punctate (Fig. 2a), which is consistent with the lower affinity of the modified native SSU for EPYC1 observed previously in yeast two-hybrid experiments[18]. Furthermore, we confirmed that condensates were formed when either EPYC1::tGFP or EPYC1::eGFP expression cassettes were transformed into $S2_{Cr}$ individually (Supplementary Fig. 3a). Expression of a full-length (i.e. non-truncated) variant of EPYC1-dGFP in Arabidopsis chloroplasts previously did not result in phase separation[18], which was attributed to low levels of expression and an incompatible stoichiometry between EPYC1 and Rubisco, and possible proteolytic degradation. The results here indicate that the formation of a condensate is primarily due to the expression of the mature variant of EPYC1 rather than the level of expression per se, and that the required stoichiometry with Rubisco is achievable in planta. Furthermore, the observed reduction in proteolytic degradation may be accounted for by sequestration of EPYC1 within a phase-separated compartment (Fig. 1b), which would be less accessible to large protease complexes[25].

Characterisation of the condensate

We next tested if the condensates exhibited internal mixing characteristics consistent with
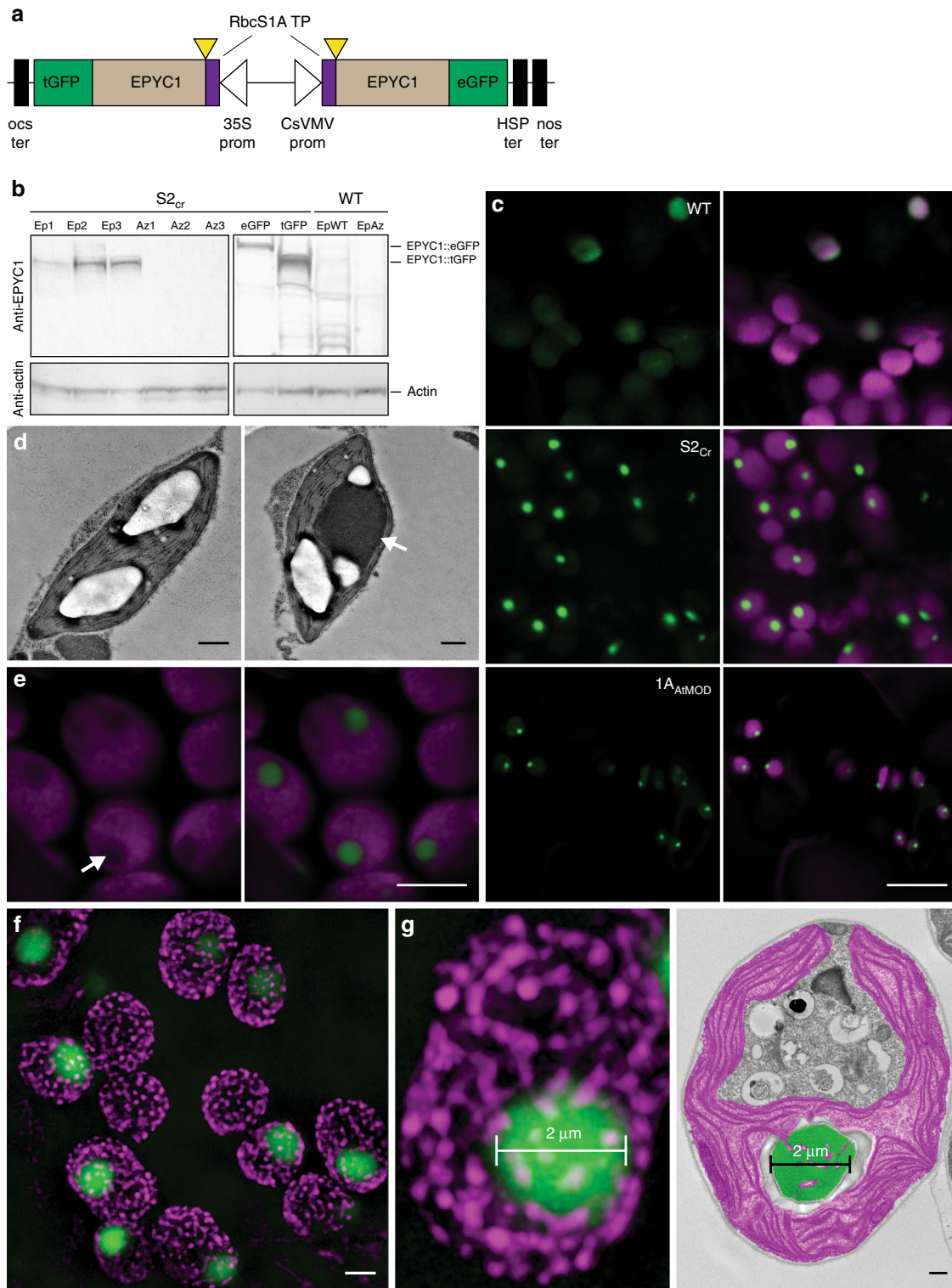
the liquid-like behaviour of pyrenoids (Fig. 2b, c). Fluorescence recovery after photobleaching (FRAP) assays showed that condensates in live leaf cells had similar or increased rates of re-homogenisation of EPYC1-dGFP (i.e. full recovery after 20−40 s) compared to previous in vitro[17] and in alga[14] reports. The more rapid interchange in condensates compared to Chlamydomonas pyrenoids may be due to the relatively reduced availability of EPYC1 binding sites on Rubisco in the S2$_{Cr}$ plant-algal hybrid Rubisco background[19]. In contrast, condensates in leaf tissue chemically cross-linked with formaldehyde showed no recovery

after photobleaching, which is consistent with that observed for Chlamydomonas pyrenoids[14].

To test for the presence of Rubisco, the condensates were extracted from leaf tissue and examined by immunoblot. The condensates could be pelleted by gentle centrifugation[16], and isolated condensates were shown to be enriched in EPYC1-dGFP and Rubisco (Fig. 3a). Protein samples were probed with either a polyclonal Rubisco antibody with a greater specificity for higher plant LSU and SSUs, or an antibody raised against the RbcS2 SSU isoform of Chlamydomonas (CrRbcS2). Comparison of the

Fig. 1 Expression of EPYC1 in the Arabidopsis line S2$_{Cr}$ results in condensate formation. a Schematic representation of the dual GFP expression system (EPYC1-dGFP) for EPYC1 truncated at amino acid residue 27 (as indicated by yellow triangles) and fused at the N-terminus to the chloroplastic transit peptide (TP) sequence of the Arabidopsis Rubisco small subunit RbcS1A. EPYC1 was further fused at the C-terminus to either enhanced GFP (eGFP) or turboGFP (tGFP), and driven by the 35S CaMV promoter (35S prom) or CsVMV promoter (CsVMV prom), respectively. For the latter expression cassette, a dual terminator system was used to increase expression[53] (Supplementary Fig. 1). **b** EPYC1 protein levels in Arabidopsis plants as assessed by immunoblot analysis with anti-EPYC1 antibodies. Shown are three T2 S2$_{Cr}$ transgenic plants expressing EPYC1-dGFP (Ep1−3) and azygous segregants (Az1−3), and S2$_{Cr}$ plants transformed with only EPYC1::tGFP (55.4 kDa) or EPYC1::eGFP (63.9 kDa). Also displayed are a T2 EPYC1-dGFP WT transformant (EpWT) and azygous segregant (AzWT). Anti-actin is shown as a loading control underneath. **c** Expression of EPYC1-dGFP in WT, S2$_{Cr}$ and 1A$_{At}$MOD backgrounds. Green and purple signals are GFP and chlorophyll autofluorescence, respectively. Overlapping signals are white. Scale bar = 10 µm for all images. **d** TEM images of chloroplasts from S2$_{Cr}$ plants with (right) and without (left) expression of EPYC1. A white arrowhead indicates the dense dark grey area of the EPYC1 condensate. The large white structures are starch granules. Scale bars = 0.5 µm. A representative chloroplast from a wild-type plant expressing EPYC1-dGFP is shown for comparison in Extended Data Fig. 3. **e** Chlorophyll autofluorescence is reduced at the site of EPYC1-dGFP accumulation (white arrow). Scale bar = 5 µm. **f** SIM microscopy showing EPYC1-dGFP condensates inside the chloroplast (see also Supplementary Movie 1). The magenta puncta show the position of grana stacks. Light magenta puncta indicate grana stacks behind the condensate. Scale bar = 2 µm. **g** Example comparison of condensate size (left, 2 µm) with that of a pyrenoid in Chlamydomonas (right, representative TEM image highlighting the pyrenoid in green and chloroplast in purple). Scale bar for TEM image = 0.5 µm. Images of EPYC1-dGFP condensates in the S2$_{Cr}$ background are from line Ep3 (**c**, **f**, **g**). Confocal images are representative of multiple imaging sessions (**c**, **e**). SIM microscopy (**f**, **g**) imaging was performed once. The immunoblots shown were derived from the same experiment and gels/blots were processed in parallel. Immunoblots results were representative of six gels/blots (**a**), TEM images are representative of 26 and 12 images from S2$_{Cr}$ plants with and without EPYC1, respectively (**d**) and 55 images from Chlamydomonas (**g**). Source data are provided as a Source Data file.
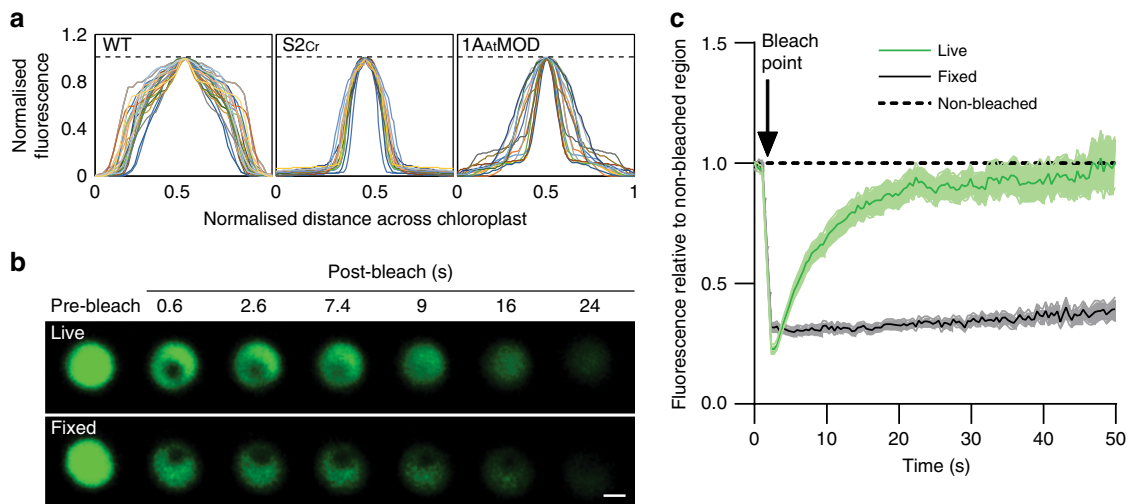

Fig. 2 In planta condensates behave like liquid−liquid phase-separated microcompartments. a Fluorescence distribution plots of EPYC1-dGFP across the chloroplast. The intensity of the GFP fluorescence signal over the cross-section of a chloroplast is shown in WT ($n = 28$), S2$_{Cr}$ (line Ep3, $n = 17$) and 1A$_{At}$MOD ($n = 22$) backgrounds. Both GFP fluorescence and cross-section values have been normalised to 1 (as indicated by the dashed line), with the highest value in the centre. **b** Fluorescence recovery after photobleaching (FRAP) assays. Condensates are shown from live (top) and fixed (bottom) leaf tissue from S2$_{Cr}$ transgenic line Ep3 expressing EPYC1-dGFP. Scale bar = 1 µm for both images. **c** Fluorescence recovery of the bleached area in relation to the non-bleached area of condensates. The mean ± SEM are shown for live ($n = 13$) and fixed ($n = 16$) chloroplasts. Source data are provided as a Source Data file.

resulting immunoblot data showed enrichment of the Chlamydomonas SSU in the pelleted condensate compared to native SSUs. Subsequent Coomassie staining of denatured, gel-separated extracts revealed that nearly half (49%) of Rubisco in the initial extract contained Chlamydomonas SSU, while 82% of Rubisco in the pelleted condensate contained Chlamydomonas SSU (Fig. 3b). Pelleted condensates coalesced into larger droplets when resuspended, as expected for the liquid-like behaviour of EPYC1-Rubisco interactions (Fig. 3c)[17]. Immunogold analysis of chloroplast TEM images from S2$_{Cr}$ plants expressing EPYC1-dGFP showed approximately half (54%) of all Rubisco was contained within the condensate (Figs. 3d, e and Supplementary Fig. 3b). Consistent with Coomasie staining, 81% of Rubisco containing Chlamydomonas SSU was located in the condensate. Thus, condensation of Rubisco is strongly associated with Rubisco complexes bearing the Chlamydomonas SSU, which

constituted approximately 50% of the Rubisco pool. The latter is consistent with the expected expression levels of plant-algal hybrid Rubisco in S2$_{Cr}$[21]. It is currently unclear if Rubisco can form a heterogenous L8S8 complex with different SSU isoforms, or if only a single SSU isoform is favoured during assembly[26]. Thus, it remains unclear whether the Rubisco pool within the condensate was comprised of a mixture of homogeneous Rubisco complexes, or those containing both Arabidopsis and Chlamydomonas SSUs.

Growth phenotype of EPYC1 transformants. We compared the growth performance of three separate T2 EPYC1-dGFP S2$_{Cr}$ transgenic lines (Ep1−3) screened for the presence of condensates with their matching T2 azygous segregant S2$_{Cr}$ lines where the EPYC1-dGFP transgene was absent (Az1−3). Plants
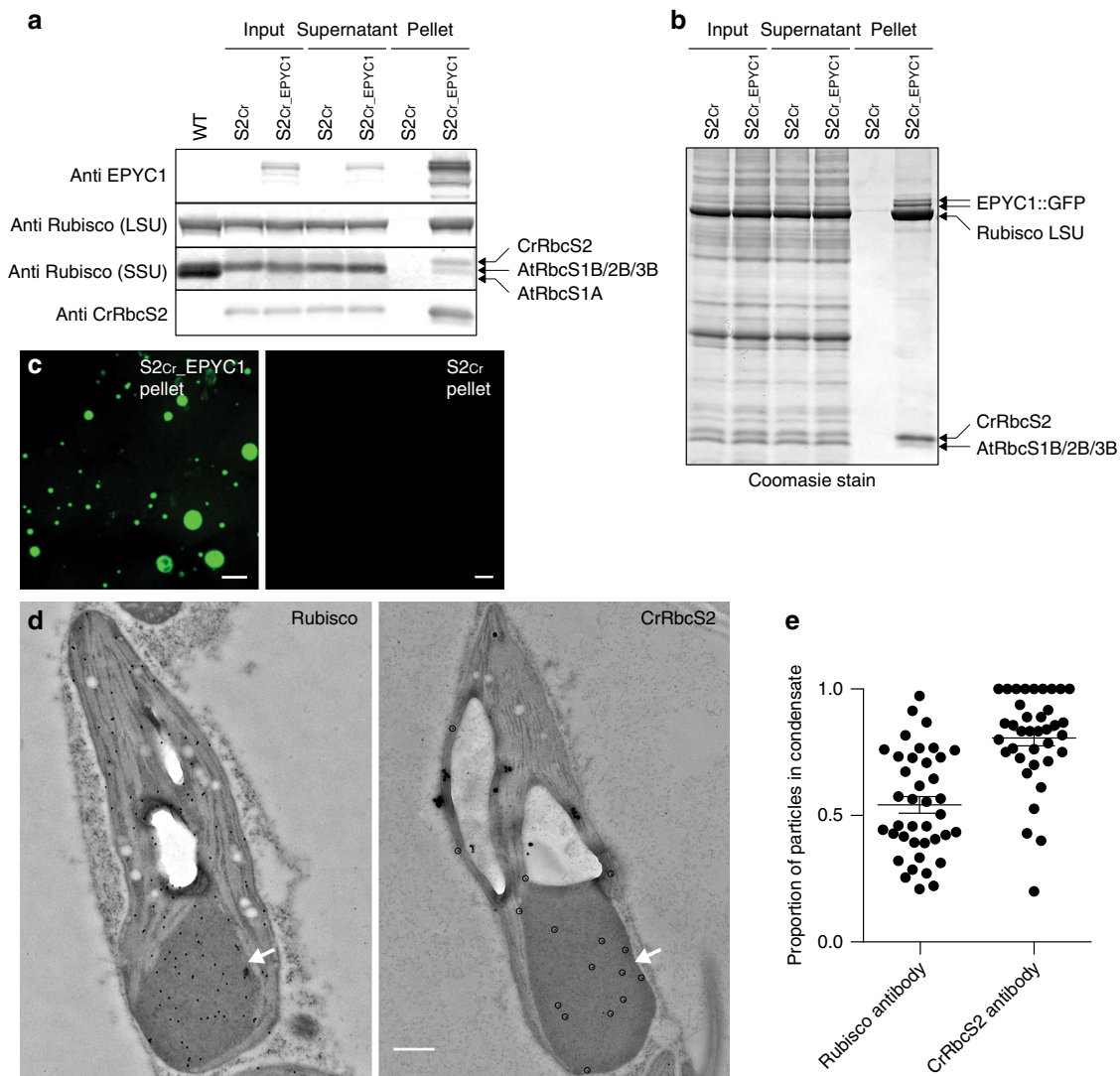
Fig. 3 Condensates contain EPYC1 and plant-algal hybrid Rubisco. a EPYC1 and Rubisco protein levels in whole leaf tissue (input), the supernatant following condensate extraction and centrifugation (supernatant) and the pellet (pellet) as assessed by immunoblot analyses with anti-EPYC1, anti-Rubisco (LSU and SSU shown) or anti-CrRbcS2 antibodies. Samples are shown for WT plants (WT), and S2$_{Cr}$ mutants not expressing EPYC1 (S2$_{Cr}$) and expressing EPYC1-dGFP (S2$_{Cr}$_EPYC1). The pellet is 40× more concentrated than the input and supernatant. Molecular weights: LSU, 55 kDa; RbcS1B, RbcS2B and RbcS3B, 14.8 kDa; AtRbcS1A, 14.7 kDa; CrRbcS2 15.5 kDa. The immunoblots shown were derived from the same experiment and gels/blots were processed in parallel. Immunoblots results were representative of four gels/blots for EPYC1, two for Rubisco, and one for CrRbcS2. **b** Coomassie-stained SDS-PAGE gel showing the composition of the pelleted condensate. Images are representative of three gels. **c** The condensates in the S2$_{Cr}$_EPYC1 pellet coalescence to form large liquid droplets. Scale bars = 25 µm. **d** Representative immunogold labelling of Rubisco in chloroplasts of an S2$_{Cr}$ transgenic line Ep3 expressing EPYC1-dGFP probed with polyclonal anti-Rubisco (left) or anti-CrRbcS2 (right) antibodies (dots are highlighted for the latter). The condensates are marked by a white arrowhead. Large white structures are starch granules. Scale bar = 0.5 µm for both images. **e** Proportion of gold nanoparticles inside the condensate compared to the chloroplast for each antibody. The mean ± SEM are shown for $n = 37$ individual chloroplast images. Source data are provided as a Source Data file.

were cultivated under light levels typical for Arabidopsis growth (200 µmol photons m^{-2} s^{-1}) or under higher light levels (900 µmol photons m^{-2} s^{-1}) where growth rates are more limited by Rubisco activity (Figs. 4a, b and Supplementary Fig. 4)[27]. Regardless of the growth conditions, rosette expansion rates or biomass accumulation were not distinguishable between transformants and their segregant controls. Similarly, T2 EPYC1-dGFP WT plants (EpWT) showed no significant differences compared to T2 segregant lines (AzWT). Due to the reduced Rubisco content in the S2$_{Cr}$ background and differences in Rubisco catalytic characteristics, the growth performance of S2$_{Cr}$ lines was slightly decreased compared to WT plants. The

observed differences in growth were in line with those reported previously for S2$_{Cr}$ and WT plants in the absence of EPYC1[21].

Photosynthetic parameters derived from response curves of CO_2 assimilation rate to the intercellular CO_2 concentration under saturating light were similar between respective Ep and Az lines (Fig. 4c, d and Extended Data Table 1). The presence of condensates did not influence the maximum achievable rates of Rubisco carboxylation (V_{cmax}). Notably, the CO_2 assimilation rates at ambient concentrations of CO_2 for Ep and Az lines were comparable to WT lines when normalised for Rubisco content ($A_{Rubisco}$) (Fig. 4d). This suggests that the modest reductions in Rubisco turnover rate (k_{cat}^c) and specificity ($S_{C/O}$) for the plant-algal
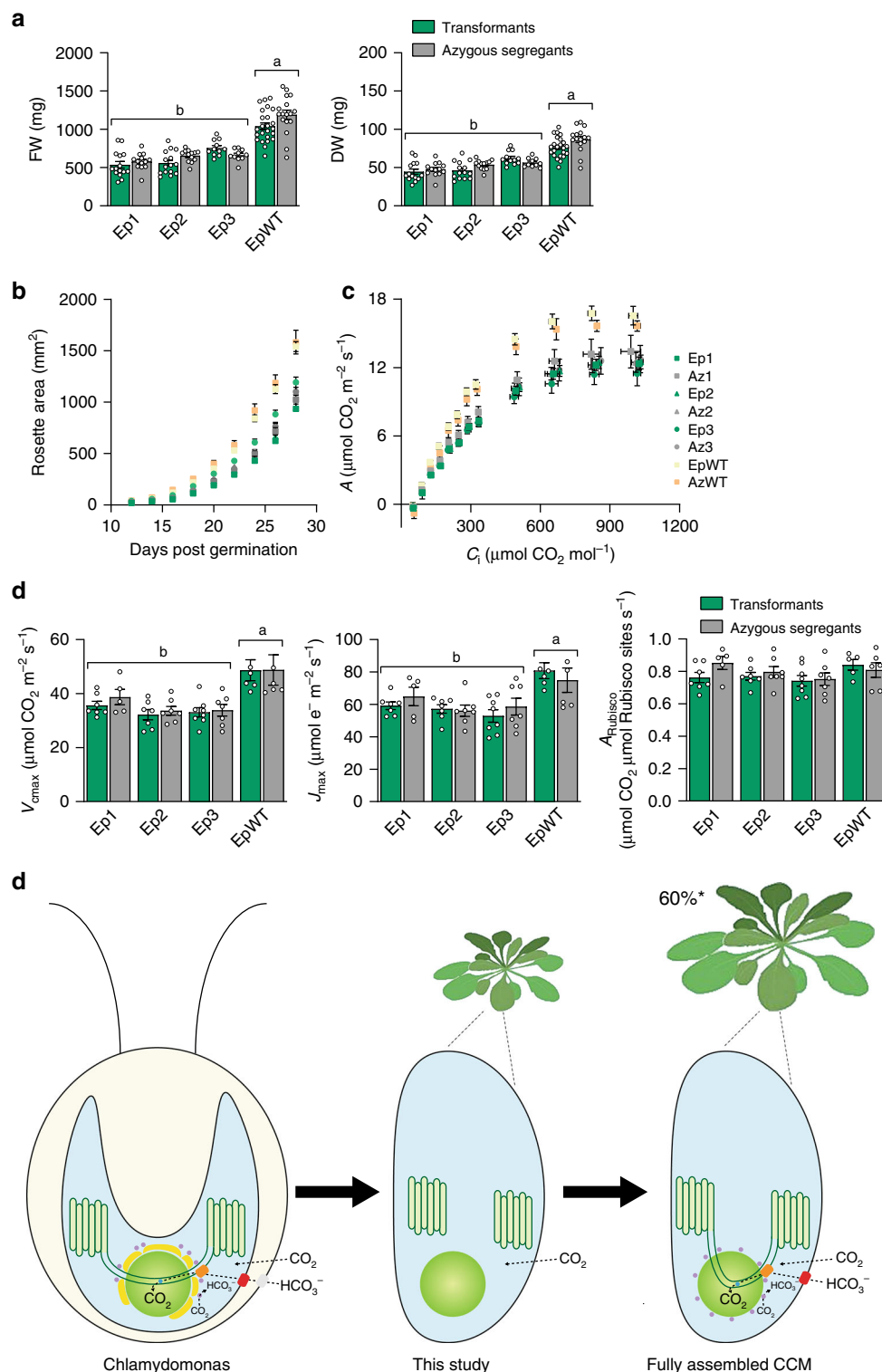
hybrid Rubisco in $S2_{Cr}$ compared to WT plants have only a mild impact on the efficiency of photosynthetic CO_2 assimilation and that the observed differences in growth rates are more associated with the reduced levels of Rubisco in $S2_{Cr}$ plants[21]. Mesophyll conductance (g_m) levels were also reduced in all $S2_{Cr}$ lines compared to WT plants, which is consistent with the impact of reduced Rubisco content on g_m observed in transplastomic tobacco[28]. Condensate formation in the $1A_{At}MOD$ background, where catalytic characteristics of the hybrid Rubisco are indistinguishable

from that of WT Rubisco, indicates that the SSU can be further engineered to optimise phase separation, and Rubisco content and performance (Fig. 1c)[18]. Measurements of the maximum electron transport rate (J_{max}) and the maximum potential quantum efficiency of Photosystem II (F_v/F_m) were also indistinguishable between transformant and segregant lines (Extended Data Table 1). Thus, the apparent displacement of the thylakoid membrane matrix by the condensates (Fig. 1e) had no obvious impact on the efficiency of the light reactions of photosynthesis.

Fig. 4 EPYC1-mediated condensation of Rubisco has no negative impact on growth and photosynthesis. a Fresh and dry weights of three T2 EPYC1-dGFP S2$_{Cr}$ transgenic lines (Ep1−3) and an EPYC1-dGFP WT transformant (EpWT) (both in green) with their respective azygous segregants (Az1−3 and AzWT) (in grey). Plants were measured after 32 days of growth under 200 µmol photons m^{-2} s^{-1} light. The mean ± SEM are shown for $n = 10-26$ individual plants for each line. **b** Rosette expansion of S2$_{Cr}$ and WT lines in (**a**). **c** Net CO_2 assimilation (A) based on intercellular [CO_2] (C_i) under saturating light (1500 µmol photons m^{-2} s^{-1}). Values show the mean ± SEM of measurements made on individual leaves from individual rosettes ($n = 5$ −8). **d** Variables derived from gas exchange data include maximum rate of Rubisco carboxylation (V_{cmax}), maximum electron transport rate (J_{max}), stomatal conductance (G_s), mesophyll conductance (G_m) and the net CO_2 assimilation rate at ambient concentrations of CO_2 normalised to Rubisco ($A_{Rubisco}$). Letters indicate significant difference ($p < 0.05$) of EpWT lines compared to Ep lines as determined by one-way ANOVA followed by Tukey's honestly significant difference (HSD) post-hoc tests (**a**, **d**). **e** Algal CCM components required for enhancing photosynthesis. Generating a pyrenoid-like condensate in a plant chloroplast provides a platform for introducing bicarbonate (HCO_3^-) channels/pumps at the chloroplast envelope (e.g. LCIA, shown in red)[34] and thylakoid membrane (e.g. BST1−3, shown in orange)[35], a lumenal carbonic anhydrase to convert HCO_3^- to CO_2 for release into the surrounding Rubisco condensate (CAH3, shown in blue)[36], mechanisms to capture CO_2 as HCO_3^- (LCIB and LCIC, shown in purple)[37,38] and traversing thylakoid membranes[33]. Current models suggest that introduction of a functional biophysical CCM into a C3 plant could lead to productivity gains of up to 60% [29−31]. Source data are provided as a Source Data file.

Discussion

Our findings show that EPYC1 and specific residues on the SSU are sufficient to aggregate Rubisco into a single proto-pyrenoid condensate, which has no apparent negative impact on plant growth. The overall photosynthetic performances of S2$_{Cr}$ transgenic lines appeared unaffected by the condensate, which suggests that conditions inside higher plant chloroplasts are highly compatible with the presence of pyrenoid-type bodies. Our data are arguably the key step for the assembly of a pyrenoid-based CCM into plants that could increase crop yield potentials by up to 60% (Fig. 4e)[29−31]. Previously described approaches for engineering the cyanobacterial carboxysome-based CCM require engineering of the chloroplast-encoded Rubisco large subunit, which is not generally feasible in major grain crops such as wheat and rice[32], although expression of the large subunit from the nucleus of maize plants has been recently demonstrated[2]. Furthermore, transgenic plants expressing cyanobacterial Rubisco require high concentrations of CO_2 to grow. Here we demonstrate that condensation of Rubisco is achievable through modification of the nuclear-encoded SSU, which is significantly more amenable to genetic modification. Future work will focus on progressing towards a minimal functional algal CCM by tethering the condensate to a thylakoid membrane for bicarbonate delivery using appropriate transporters[33−35], as well as introduction of a lumenal carbonic anhydrase and a stromal CO_2 salvage system[36−38].

Methods

Plant material and growth conditions. Arabidopsis (*Arabidopsis thaliana*, Col-0 background) seeds were sown on compost, stratified for 3 days at 4 °C and grown at 20 °C, ambient CO_2 and 70% relative humidity under either 200 or 900 µmol photons m^{-2} s^{-1} supplied by cool white LED lights (Percival SE-41AR3cLED, CLF PlantClimatics GmbH, Wertingen, Germany) in 12 h light, 12 h dark. For comparisons of different genotypes, plants were grown from seeds of the same age and storage history, harvested from plants grown in the same environmental conditions.

Construct design and transformation. The coding sequence of EPYC1 was codon optimised for expression in higher plants as in Atkinson et al.[18]. Mature EPYC1 was cloned directly into the level 0 acceptor vector pAGM1299 of the Plant MoClo system[39]. To generate fusion proteins, gene expression constructs were assembled into binary level 2 acceptor vectors. The 35S cauliflower mosaic virus (CaMV) promoter and CsVMV (cassava vein mosaic virus) promoter were used to drive expression. Level 2 vectors were transformed into *Agrobacterium tumefaciens* (AGL1) for stable insertion in Arabidopsis plants by floral dipping[40]. Homozygous transgenic and azygous lines were identified in the T2 generation using the pFAST-R selection cassette[41].

Protein analyses. Soluble protein was extracted from frozen leaf material of 21-day-old plants (sixth and seventh leaf) in protein extraction buffer (50 mM HEPES-KOH pH 7.5 with 17.4% glycerol, 2% Triton X-100 and cOmplete Mini ethylenediaminetetraacetic acid (EDTA)-free Protease Inhibitor Cocktail (Roche, Basel, Switzerland). Samples were heated at 70 °C for 15 min with 1× Bolt LDS sample buffer (ThermoFisher Scientific, UK) and 200 mM dithiothreitol (DTT). Extracts

were centrifuged and the supernatants subjected to SDS-PAGE on a 12% (w/v) polyacrylamide gel and transferred to a nitrocellulose membrane. Membranes were probed with rabbit serum raised against wheat Rubisco at 1:10,000 dilution[42], the SSU RbcS2 from Chlamydomonas (CrRbcS2) (raised to the C-terminal region of the SSU (KSARDWQPANKRSV) by Eurogentec, Southampton, UK) at 1:1000 dilution, Actin (66009-1-Ig, Proteintech, UK) at 1:1000 dilution and EPYC1 at 1:2000 dilution[16], followed by IRDye 800CW goat anti-rabbit IgG (LI-COR Biotechnology, Cambridge, UK) at 1:10,000 dilution, and visualised using the Odyssey CLx imaging system (LI-COR Biotechnology), or by HRP-linked goat anti-rabbit IgG (Abcam) at 1:10,000 dilution, and visualised using Pierce ECL Western Blotting Substrate (Life Technologies).

qRT-PCR analysis. Total RNA was isolated from leaves of 21-day-old plants (as above) using the RNeasy plant mini kit (Qiagen, USA). Isolated RNA was treated with DNase (Qiagen, USA) and reverse transcribed with random primers using the GoScript Reverse Transcription kit (Promega, USA). Reverse transcription quantitative PCR (RT-qPCR) was carried out with Takyon No ROX SYBR mastermix (Eurogentec, Belgium) as in Atkinson et al.[12], and run on a LightCycler 480 (Roche, Switzerland). Gene-specific primers were designed to amplify either eGFP (5′-CAGATTACGCCGTGAGCAAG-3′ and 5′-GCTGAACTTGTGGCCGTTT A-3′) or tGFP (5′-GCATCAGGGGTCTTGAAAGC-3′ and 5′-TCCTTCAAAAC GGTGGACCT-3′) (IDT, Belgium). Amplification efficiency was determined with a calibration curve for each primer pair. Two reference genes At4g26410 (RHIP1) and At1g13320 (PP2A) were used for normalisation[43]. The relative expression of eGFP and tGFP was calculated according to Vandesompele et al.[44].

Condensate extraction. Soluble protein was extracted as before, then filtered through Miracloth (Merck Millipore, Burlington, Massachusetts, USA), and centrifuged at 500 × g for 3 min at 4 °C, as in Mackinder et al.[16]. The pellet was discarded, and the extract centrifuged again for 12 min. The resulting pellet was washed once in protein extraction buffer, then resuspended in a small volume of buffer and centrifuged again for 5 min. Finally, the pellet was resuspended in 25 µl of extraction buffer and used in confocal analysis or SDS-PAGE electrophoresis.

Growth analysis and photosynthetic measurements. Rosette growth rates were quantified using an in-house imaging system[45]. Maximum quantum yield of photosystem II (PSII) (F_v/F_m) was measured on 32-day-old plants using a Hansatech Handy PEA continuous excitation chlorophyll fluorimeter (Hansatech Instruments Ltd, King's Lynn, UK)[46]. Gas exchange and chlorophyll fluorescence were determined using a LI-COR LI-6400 (LI-COR, Lincoln, Nebraska, USA) portable infra-red gas analyser with a 6400-40 leaf chamber on either the sixth or seventh leaf of 35- to 45-day-old non-flowering rosettes grown in large pots under 200 µmol photons m^{-2} s^{-1} to generate leaf area sufficient for gas exchange measurements[47]. The response of A to the intercellular CO_2 concentration (C_i) was measured at various CO_2 concentrations (50, 100, 150, 200, 250, 300, 350, 400, 600, 800, 1000 and 1200 µmol mol^{-1}) under saturating light (1500 µmol photons m^{-2} s^{-1}). For all gas exchange experiments, the flow rate was kept at 200 µmol mol^{-1}, leaf temperature was controlled at 25 °C and chamber relative humidity was maintained at ca. 70%. Measurements were performed after net assimilation and stomatal conductance had reached steady state. Gas exchange data were corrected for CO_2 diffusion from the measuring chamber as in Bellasio et al.[48]. To estimate V_{cmax}, J_{max}, $A_{Rubisco}$, the CO_2 compensation point (Γ) and mesophyll conductance to CO_2 (g_m) the A/C_i data were fitted to the C3 photosynthesis model as in Ethier and Livingston[49,50] using the catalytic parameters K_c^{air} and affinity for O_2 (K_o) values for wild-type Arabidopsis Rubisco at 25 °C and the Rubisco content of WT and S2$_{Cr}$ lines[21].

Confocal laser scanning and fluorescence recovery after photobleaching.
Leaves were imaged with a Leica TCS SP8 laser scanning confocal microscope
(Leica Microsystems, Milton Keynes, UK) as in Atkinson et al.[12]. Processing of
images was done with Leica LAS AF Lite software. Condensate and chloroplast
dimensions were measured from confocal images using Fiji (ImageJ, v1.52n)[51].
Condensate volume was calculated as a sphere, while chloroplast volume was
calculated as an ellipsoid, where depth was estimated as 25% of the measured
width. Chloroplast volumes varied between 24 and 102 μm^3, which was within the
expected size range and distribution for Arabidopsis chloroplasts[52]. Comparative
measurements of pyrenoid area were performed using Fiji on TEM cross-section
images of WT *C. reinhardtii* cells (cMJ030) as described in Itakura et al.[23]. FRAP
was carried out on live leaf tissue and tissue that had been fixed by infiltrating
with 4% (v/v) formaldehyde for 90 min in a vacuum chamber, then infiltrating
three times for 5 min each with phosphate-buffered saline (PBS) at pH 7.4. The
SP8 microscope was used with a ×40 water immersion objective and a photo-
multiplier tube (PMT) detector. The 488 nm laser was set to 2% power for pre-
and post-bleach images, and 25% for the bleaching step. Pre-bleach images were
captured at 220 ms intervals (6 in total) and post-bleach images were captured at
400 ms intervals (at least 120). For photobleaching the laser was directed to a
region with a diameter of 0.5−0.6 μm on one side of the EPYC1 aggregate.
Recovery time was calculated by comparing GFP expression to an unbleached
region of the same size.

Super-resolution image microscopy. Super-resolution images were acquired
using structured illumination microscopy. Samples were prepared on high preci-
sion cover-glass (Zeiss, Jena, Germany). 3D SIM images were acquired on an N-
SIM (Nikon Instruments, UK) using a ×100 1.49NA lens and refractive index
matched immersion oil (Nikon Instruments). Samples were imaged using a Nikon
Plan Apo TIRF objective (NA 1.49, oil immersion) and an Andor DU-897X-5254
camera using a 488 nm laser line. Z-step size for z stacks was set to 0.120 μm as
required by the manufacturer's software. For each focal plane, 15 images (5 phases,
3 angles) were captured with the NIS-Elements software. SIM image processing,
reconstruction and analysis were carried out using the N-SIM module of the NIS-
Element Advanced Research software. Images were checked for artefacts using the
SIMcheck software (www.micron.ox.ac.uk/software/SIMCheck.php). Images were
reconstructed using NiS Elements software (v4.6, Nikon Instruments) from a z
stack comprising of no less than 1 μm of optical sections. In all SIM image
reconstructions, the Wiener and Apodization filter parameters were kept constant.

Immunogold labelling and electron microscopy. Leaf samples were taken from
21-day-old $S2_{Cr}$ plants and $S2_{Cr}$ transgenic lines expressing EPYC1-dGFP and fixed
with with 4% (v/v) paraformaldehyde, 0.5% (v/v) glutaraldehyde and 0.05 M
sodium cacodylate (pH 7.2). Leaf strips (1 mm wide) were vacuum infiltrated with
fixative three times for 15 min, then rotated overnight at 4 °C. Samples were rinsed
three times with PBS (pH 7.4) then dehydrated sequentially by vacuum infiltrating
with 50%, 70%, 80% and 90% ethanol (v/v) for 1 h each, then three times with
100%. Samples were infiltrated with increasing concentrations of LR White Resin
(30%, 50%, 70% [w/v]) mixed with ethanol for 1 h each, then 100% resin three
times. The resin was polymerised in capsules at 50 °C overnight. Sections (1-μm
thick) were cut on a Leica Ultracut ultramicrotome, stained with Toluidine Blue,
and viewed in a light microscope to select suitable areas for investigation. Ultrathin
sections (60-nm thick) were cut from selected areas and mounted onto plastic-
coated copper grids. Grids were blocked with 1% (w/v) BSA in TBSTT (Tris-
buffered saline with 0.05% [v/v] Triton X-100 and 0.05% [v/v] Tween 20), incu-
bated overnight with anti-Rubisco antibody in TBSTT at 1:250 dilution or anti-
CrRbcS2 antibody at 1:50 dilution, and washed twice each with TBSTT and water.
Incubation with 15 nm gold particle-conjugated goat anti-rabbit secondary anti-
body (Abcam, Cambridge, UK) in TBSTT was carried out for 1 h at 1:200 dilution
for Rubisco labelling or 1:10 for CrRbcS2 labelling, before washing as before. Grids
were stained in 2% (w/v) uranyl acetate then viewed in a JEOL JEM-1400 Plus
TEM (JEOL, Peabody, Massachusetts, USA). Images were collected on a GATAN
OneView camera (GATAN, Pleasanton, California, USA).

Statistical analyses. Results were subjected to one-way analysis of variance
(ANOVA) to determine the significance of the difference between sample groups.
When one-way ANOVA was performed, Tukey's honestly significant difference
(HSD) post-hoc tests were conducted to determine the differences between the
individual treatments (IBM SPSS Statistics, Ver. 26.0).

Reporting summary. Further information on experimental design is available in
the Nature Research Reporting Summary linked to this paper.

Data availability
The datasets and plant materials generated and analysed during the current study are
available from the corresponding author upon request. A reporting summary for this
article is available as a Supplementary Information file. Data supporting the findings of
this work are available within the paper and its Supplementary Information files. Raw
image data corresponding to Fig. 1, Fig. 3e and Supplementary Fig. 2c are available online

at the open access Edinburgh DataShare repository [https://doi.org/10.7488/ds/
2945]. Source data are provided with this paper.

References
1. Andersson, I. & Backlund, A. Structure and function of Rubisco. *Plant Physiol. Bioch.* **46**, 275–291 (2008).
2. Salesse-Smith, C. E. et al. Overexpression of Rubisco subunits with RAF1 increases Rubisco content in maize. *Nat. Plants* **4**, 802–810 (2018).
3. Flamholz, A. I. et al. Revisiting trade-offs between Rubisco kinetic parameters. *Biochem.-Us* **58**, 3365–3376 (2019).
4. Galmes, J. et al. Expanding knowledge of the Rubisco kinetics variability in plant species: environmental and evolutionary trends. *Plant Cell Environ.* **37**, 1989–2001 (2014).
5. Young, J. N. et al. Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *J. Exp. Bot.* **67**, 3445–3456 (2016).
6. Bathellier, C., Tcherkez, G., Lorimer, G. H. & Farquhar, G. D. Rubisco is not really so bad. *Plant Cell Environ.* **41**, 705–716 (2018).
7. Wilson, R. H. & Whitney, S. M. in *Directed Enzyme Evolution: Advances and Applications* (ed. Alcalde, M.) Ch. 4, 101−126 (Springer International Publishing AG, 2017).
8. Long, S. P., Marshall-Colon, A. & Zhu, X. G. Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* **161**, 56–66 (2015).
9. Mackinder, L. C. M. et al. A spatial interactome reveals the protein organization of the algal CO2-concentrating mechanism. *Cell* **171**, 133–147.e114 (2017).
10. Li, F. W. et al. Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. *Nat. Plants* **6**, 259–272 (2020).
11. Raven, J. A., Beardall, J. & Sanchez-Baracaldo, P. The possible evolution and future of CO2-concentrating mechanisms. *J. Exp. Bot.* **68**, 3701–3716 (2017).
12. Atkinson, N. et al. Introducing an algal carbon-concentrating mechanism into higher plants: location and incorporation of key components. *Plant Biotechnol. J.* **14**, 1302–1315 (2016).
13. Rae, B. D. et al. Progress and challenges of engineering a biophysical CO2-concentrating mechanism into higher plants. *J. Exp. Bot.* **68**, 3717–3737 (2017).
14. Freeman Rosenzweig, E. S. et al. The eukaryotic CO2-concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell* **171**, 148–162.e119 (2017).
15. Meyer, M. T. et al. Rubisco small-subunit alpha-helices control pyrenoid formation in Chlamydomonas. *Proc. Natl Acad. Sci. USA* **109**, 19474–19479 (2012).
16. Mackinder, L. C. et al. A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proc. Natl Acad. Sci. USA* **113**, 5958–5963 (2016).
17. Wunder, T., Cheng, S. L. H., Lai, S. K., Li, H. Y. & Mueller-Cajar, O. The phase separation underlying the pyrenoid-based microalgal Rubisco supercharger. *Nat. Commun.* **9**, 5076 (2018).
18. Atkinson, N. et al. The pyrenoidal linker protein EPYC1 phase separates with hybrid Arabidopsis-Chlamydomonas Rubisco through interactions with the algal Rubisco small subunit. *J. Exp. Bot.* **70**, 5271–5285 (2019).
19. He, S. et al. The structural basis of Rubisco phase separation in the pyrenoid. *Nat. Plants* https://doi.org/10.1101/2020.08.16.252809 (2020).
20. Meyer, M. T. et al. Assembly of the algal CO2-fixing organelle, the pyrenoid, is guided by a Rubisco-binding motif. *Sci. Adv.* https://doi.org/10.1126/sciadv.abd2408 (2020).
21. Atkinson, N. et al. Rubisco small subunits from the unicellular green alga Chlamydomonas complement Rubisco-deficient mutants of Arabidopsis. *New Phytol.* **214**, 655–667 (2017).
22. Wunder, T., Oh, Z. G. & Mueller-Cajar, O. CO2-fixing liquid droplets: towards a dissection of the microalgal pyrenoid. *Traffic* **20**, 380–389 (2019).
23. Itakura, A. K. et al. A Rubisco-binding protein is required for normal pyrenoid number and starch sheath morphology in Chlamydomonas reinhardtii. *Proc. Natl Acad. Sci. USA* **116**, 18445–18454 (2019).
24. Rosowski, K. A. et al. Elastic ripening and inhibition of liquid-liquid phase separation. *Nat. Phys.* **16**, 422–42 (2020).
25. van der Hoorn, R. A. L. & Rivas, S. Unravelling the mode of action of plant proteases. *New Phytol.* **218**, 879–881 (2018).
26. Valegard, K., Hasse, D., Andersson, I. & Gunn, L. H. Structure of Rubisco from Arabidopsis thaliana in complex with 2-carboxyarabinitol-1,5-bisphosphate. *Acta Crystallogr. D* **74**, 1–9 (2018).

27. Lauerer, M. et al. Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with "antisense" rbcS. VI. Effect on photosynthesis in plants grown at different irradiance. *Planta* **190**, 332–345 (1993).

28. Galmes, J., Perdomo, J. A., Flexas, J. & Whitney, S. M. Photosynthetic characterization of Rubisco transplantomic lines reveals alterations on photochemistry and mesophyll conductance. *Photosynthesis Res.* **115**, 153–166 (2013).

29. Price, G. D., Badger, M. R. & von Caemmerer, S. The prospect of using cyanobacterial bicarbonate transporters to improve leaf photosynthesis in C3 crop plants. *Plant Physiol.* **155**, 20–26 (2011).

30. McGrath, J. M. & Long, S. P. Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. *Plant Physiol.* **164**, 2247–2261 (2014).

31. Long, S. P., Burgess, S. & Causton, I. in *Sustaining Global Food Security: The Nexus of Science and Policy* (ed. Zeigler, R. S.) Ch. 9 (CSIRO Publishing, 2019).

32. Long, B. M. et al. Carboxysome encapsulation of the CO2-fixing enzyme Rubisco in tobacco chloroplasts. *Nat. Commun.* https://doi.org/10.1038/S41467-018-06044-0 (2018).

33. Engel, B. D. et al. Native architecture of the Chlamydomonas chloroplast revealed by in situ cryo-electron tomography. *eLife* **4**, e04889 (2015).

34. Yamano, T., Sato, E., Iguchi, H., Fukuda, Y. & Fukuzawa, H. Characterization of cooperative bicarbonate uptake into chloroplast stroma in the green alga Chlamydomonas reinhardtii. *Proc. Natl Acad. Sci. USA* **112**, 7315–7320 (2015).

35. Mukherjee, A. et al. Thylakoid localized bestrophin-like proteins are essential for the CO2 concentrating mechanism of Chlamydomonas reinhardtii. *Proc. Natl Acad. Sci. USA* **116**, 16915–16920 (2019).

36. Karlsson, J. et al. A novel alpha-type carbonic anhydrase associated with the thylakoid membrane in Chlamydomonas reinhardtii is required for growth at ambient CO2. *EMBO J.* **17**, 1208–1216 (1998).

37. Yamano, T. et al. Light and low-CO2-dependent LCIB-LCIC complex localization in the chloroplast supports the carbon-concentrating mechanism in Chlamydomonas reinhardtii. *Plant Cell Physiol.* **51**, 1453–1468 (2010).

38. Toyokawa, C. Y. T. & Fukuzawa, H. Pyrenoid starch sheath is required for LCIB localization and the CO2-concentrating mechanism in green algae. *Plant Physiol.* https://doi.org/10.1104/pp.19.01587 (2020).

39. Engler, C. et al. A golden gate modular cloning toolbox for plants. *Acs Synth. Biol.* **3**, 839–843 (2014).

40. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.: Cell Mol. Biol.* **16**, 735–743 (1998).

41. Shimada, T. L., Shimada, T. & Hara-Nishimura, I. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. *Plant J.* **61**, 519–528 (2010).

42. Howe, C. J. et al. Location and nucleotide sequence of the gene for the proton-translocating subunit of wheat chloroplast ATP synthase. *Proc. Natl Acad. Sci. USA* **79**, 6903–6907 (1982).

43. Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K. & Scheible, W. R. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol.* **139**, 5–17 (2005).

44. Vandesompele, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* https://doi.org/10.1186/gb-2002-3-7-research0034 (2002).

45. Dobrescu, A., Scorza, L. C. T., Tsaftaris, S. A. & McCormick, A. J. A "Do-It-Yourself" phenotyping system: measuring growth and morphology throughout the diel cycle in rosette shaped plants. *Plant Methods* **13**, 95 (2017).

46. Maxwell, K. & Johnson, G. N. Chlorophyll fluorescence—a practical guide. *J. Exp. Bot.* **51**, 659–668 (2000).

47. Flexas, J. et al. Mesophyll conductance to CO2 in Arabidopsis thaliana. *New Phytol.* **175**, 501–511 (2007).

48. Bellasio, C., Beerling, D. J. & Griffiths, H. An Excel tool for deriving key photosynthetic parameters from combined gas exchange and chlorophyll fluorescence: theory and practice. *Plant Cell Environ.* https://doi.org/10.1111/pce.12560 (2015).

49. Ethier, G. J. & Livingston, N. J. On the need to incorporate sensitivity to CO2 transfer conductance into the Farquhar-von Caemmerer-Berry leaf photosynthesis model. *Plant Cell Environ.* **27**, 137–153 (2004).

50. Sade, N. et al. Differential tissue-specific expression of NtAQP1 in Arabidopsis thaliana reveals a role for this protein in stomatal and mesophyll conductance of CO2 under standard and salt-stress conditions. *Planta* **239**, 357–366 (2014).

51. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

52. Crumpton-Taylor, M., Grandison, S., Png, K. M., Bushby, A. J. & Smith, A. M. Control of starch granule numbers in Arabidopsis chloroplasts. *Plant Physiol.* **158**, 905–916 (2012).

53. Diamos, A. G. & Mason, H. S. Chimeric 3′ flanking regions strongly enhance gene expression in plants. *Plant Biotechnol. J.* **16**, 1971–1982 (2018).

Acknowledgements

Author contributions

Conceptualisation, N.A. and A.J.M.; methodology, N.A., K.X.C. and A.J.M.; formal analysis, N.A., Y.M., K.X.C. and A.J.M.; investigation, N.A., Y.M., K.X.C. and A.J.M.; writing-original draft, N.A. and A.J.M.; writing-review and editing, N.A., Y.M., K.X.C. and A.J.M.; visualisation, N.A., Y.M., K.X.C. and A.J.M.; supervision, A.J.M.; project administration, A.J.M.; funding acquisition, A.J.M.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-20132-0.

Correspondence and requests for materials should be addressed to A.J.M.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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

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