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

2

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

4     [1]SynthSys & Institute of Molecular Plant Sciences, School of Biological Sciences, King's

5     Buildings, University of Edinburgh, EH9 3BF, UK

6     [2]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign,

7     1206 Gregory Drive, Urbana IL 61801.

8

9     **Summary**

10     Photosynthetic $CO_2$ fixation in plants is limited by the inefficiency of the $CO_2$-assimilating

11     enzyme Rubisco (D-ribulose-1,5-bisphosphate carboxylase/ oxygenase)[1-3]. In plants

12     possessing the C3 pathway, which includes most major staple crops, Rubisco is typically

13     evenly distributed throughout the chloroplast stroma. However, in almost all eukaryotic algae

14     Rubisco aggregates within a microcompartment known as the pyrenoid, in association with a

15     $CO_2$-concentrating mechanism that improves photosynthetic operating efficiency under

16     conditions of low inorganic carbon[4]. Recent work has shown that the pyrenoid matrix is a

17     phase-separated, liquid-like condensate[5]. In the alga *Chlamydomonas reinhardtii*,

18     condensation is mediated by two components: Rubisco and the linker protein EPYC1 (Essential

19     Pyrenoid Component 1)[6,7]. Here we show that expression of mature EPYC1 and a plant-algal

20     hybrid Rubisco leads to spontaneous condensation of Rubisco into a single phase-separated

21     compartment in Arabidopsis chloroplasts, with liquid-like properties similar to a pyrenoid

22     matrix. The condensate displaces the thylakoid membranes and is enriched in hybrid Rubisco

23     containing the algal Rubisco small subunit required for phase separation. Promisingly,

24     photosynthetic $CO_2$ fixation and growth is not impaired in stable transformants compared to

25     azygous segregants. These observations represent a significant initial step towards enhancing

26     photosynthesis in higher plants by introducing an algal $CO_2$-concentrating mechanism, which

27     is predicted to significantly increase the efficiency of photosynthetic $CO_2$ uptake[8,9].

28

29

30

31

32

33

34

**Main Text**

Rubisco 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]. In plants with the C3 photosynthetic pathway (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[10-12]. Improving the operating efficiency of Rubisco remains a key strategy to increase productivity in C3 crops[13]. This is an important goal given the steep trajectory of crop yield improvements required to meet the demands of our rising global population[14].

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 micro compartment within the chloroplast called the pyrenoid[4,15]. The CCM functions to enrich the pyrenoid with high concentrations of $CO_2$ to enhance the performance of Rubisco carboxylation and suppress oxygenation[16]. Introduction of such CCMs into C3 plants is predicted to lead to significant increases in $CO_2$-fixation efficiency and biomass yield[14,17,18]. 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[5]. Phase separation is facilitated by multivalent interactions between the intrinsically disordered Rubisco-binding protein EPYC1 and the small subunit of Rubisco (SSU)[7], of which residues on the α-helices of the native SSU isoforms are critical for binding to EPYC1[19,20]. 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[19]. 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[4,5,7,22].

We 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)[19]. 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

69    fused to a different version of GFP (turboGFP (tGFP) or enhanced GFP (eGFP)) to avoid

70    possible recombination events (Fig. 1a, Extended Data Fig. 1). The dual GFP construct

71    (EPYC1-dGFP) was transformed into WT plants or the Arabidopsis *1a3b* Rubisco mutant

72    complemented with an SSU from Chlamydomonas ($S2_{Cr}$)[21]. The resulting transgenic plants

73    (Ep) expressed both EPYC1::eGFP and EPYC1::tGFP, of which the latter was generally more

74    abundant at the protein level (Fig. 1b, Extended Data Fig. 2). Previously, immunoblots against

75    full length EPYC1 expressed in $S2_{Cr}$ or WT plants showed additional lower molecular weight

76    bands indicative of proteolytic degradation[19]. Here, degradation products of the mature EPYC1

77    were relatively reduced when expressed in $S2_{Cr}$ compared to WT plants.

78        The fluorescence signal for EPYC1-dGFP in WT plants was distributed evenly throughout

79    the chloroplast (Fig. 1c, Fig. 2a).  In contrast, EPYC1-dGFP in the hybrid $S2_{Cr}$ plants showed

80    only a single dense chloroplastic signal. Transmission electron microscopy confirmed the

81    presence of a single prominent condensed complex in the chloroplast stroma (Fig. 1d). The

82    condensates were spherical in shape and displaced native chlorophyll autofluorescence (Fig.

83    1e, f), indicating that the thylakoid membrane matrix was excluded from the condensate. In

84    protoplasts of leaf mesophyll cells, a condensate was visible in each chloroplast (Extended

85    Data Fig. 3a), and the average size of the condensates was related to the expression level of

86    EPYC1-dGFP (Fig. 1b, Extended Data Fig. 3b, c). We observed that the average diameter of

87    the condensates was 1.6±0.1 µm (n=126; 42 from three individual $S2_{Cr}$ transgenic lines), which

88    was comparable with that of the size range measured for the Chlamydomonas pyrenoid

89    (1.4±0.1 µm; n=55)[23] (Fig. 1g, Extended Data Fig. 3c). The estimated volume of the

90    condensates was 2.7±0.2 µm$^3$ (*ca*. 5% of the chloroplast volume). Variations in condensate

91    volume within individual $S2_{Cr}$ transgenic Ep lines were not correlated with chloroplast volume,

92    suggesting that regulation of condensate formation and size was largely independent of

93    chloroplast morphology, and likely limited by the influence of the surrounding network

94    stiffness (i.e. the stroma, which is densely packed with thylakoids) on the dynamics of droplet

95    ripening[24].

96        Condensates were also observed when EPYC1-dGFP was expressed in the Arabidopsis

97    *1a3b* Rubisco mutant  complemented with a native Arabidopsis SSU modified to contain the

98    two α-helices necessary for pyrenoid formation from the Chlamydomonas SSU

99    ($1A_{At}MOD$)[20,21] (Fig. 1c). However, condensates in the $1A_{At}MOD$ background were less

100    punctate (Fig. 2a), which is consistent with the lower affinity of the modified native SSU for

101    EPYC1 observed previously in yeast two-hybrid experiments[19]. Furthermore, we confirmed

102    that condensates were formed when either EPYC1::tGFP or EPYC1::eGFP expression

103    cassettes were transformed into S2$_{Cr}$ individually (Extended Data Fig. 4a). Expression of a full

104    length (i.e. non-truncated) variant of EPYC1-dGFP in Arabidopsis chloroplasts previously did

105    not result in phase separation[19], which was attributed to low levels of expression and an

106    incompatible stoichiometry between EPYC1 and Rubisco, and possible proteolytic

107    degradation. The results here indicate that the formation of a condensate is primarily due to the

108    expression of the mature variant of EPYC1 rather than the level of expression *per se*, and that

109    the required stoichiometry with Rubisco is achievable *in planta*. Furthermore, the observed

110    reduction in proteolytic degradation may be accounted for by sequestration of EPYC1 within

111    a phase-separated compartment (Fig. 1b), which would be less accessible to large protease

112    complexes[25].

113        Next we tested if the condensates exhibited internal mixing characteristics consistent with

114    the liquid-like behavior of pyrenoids (Fig. 2b, 2c). Fluorescence recovery after photobleaching

115    (FRAP) assays showed that condensates in live leaf cells had similar or increased rates of re-

116    homogenisation of EPYC1-dGFP (i.e. full recovery after 20-40 sec) compared to previous *in*

117    *vitro*[7] and *in alga*[5] reports. The more rapid interchange in condensates compared to

118    Chlamydomonas pyrenoids may be due to the relatively reduced availability of EPYC1 binding

119    sites on Rubisco in the S2$_{Cr}$ plant-algal hybrid Rubisco background[4]. In contrast, condensates

120    in leaf tissue chemically cross-linked with formaldehyde showed no recovery after

121    photobleaching, which is consistent with that observed for Chlamydomonas pyrenoids[5].

122        To test for the presence of Rubisco, the condensates were extracted from leaf tissue and

123    examined by immunoblot. The condensates could be pelleted by gentle centrifugation[6], and

124    isolated condensates were shown to be enriched in EPYC1-dGFP and Rubisco (Fig. 3a).

125    Protein samples were probed with either a polyclonal Rubisco antibody with a greater

126    specificity for higher plant LSU and SSUs, or an antibody raised against the RbcS2 SSU

127    isoform of Chlamydomonas (CrRbcS2). Comparison of the resulting immunoblot data showed

128    enrichment of the Chlamydomonas SSU in the pelleted condensate compared to native SSUs.

129    Subsequent Coomassie staining of denatured, gel-separated extracts revealed that nearly half

130    (49%) of Rubisco in the initial extract contained Chlamydomonas SSU, while 82% of Rubisco

131    in the pelleted condensate contained Chlamydomonas SSU (Fig. 3b). Pelleted condensates

132    coalesced into larger droplets when resuspended, as expected for the liquid-like behavior of

133    EPYC1-Rubisco interactions (Fig. 3c)[7]. Immunogold analysis of chloroplast TEM images from

134    S2$_{Cr}$ plants expressing EPYC1-dGFP showed approximately half (54%) of all Rubisco was

135    contained within the condensate (Fig. 3d, 3e, Extended Data Fig. 4b). Consistent with

136    Coomasie staining, 81% of Rubisco containing Chlamydomonas SSU was located in the

137    condensate. Thus, condensation of Rubisco is strongly associated with Rubisco complexes

138    bearing the Chlamydomonas SSU, which constituted approximately 50% of the Rubisco pool.

139    The latter is consistent with the expected expression levels of plant-algal hybrid Rubisco in

140    S2$_{Cr}$[21]. It is currently unclear if Rubisco can form a heterogenous L8S8 complex with different

141    SSU isoforms, or if only a single SSU isoform is favoured during assembly[26]. Thus, it remains

142    unclear whether the Rubisco pool within the condensate was comprised of a mixture of

143    homogeneous Rubisco complexes, or those containing both Arabidopsis and Chlamydomonas

144    SSUs.

145    We compared the growth performance of three separate T2 EPYC1-dGFP S2$_{Cr}$ transgenic

146    lines (Ep1-3) screened for the presence of condensates with their matching T2 azygous

147    segregant S2$_{Cr}$ lines where the EPYC1-dGFP transgene was absent (Az1-3). Plants were

148    cultivated under light levels typical for Arabidopsis growth (200 μmol photons m$^{-2}$ s$^{-1}$) or under

149    higher light levels (900 μmol photons m$^{-2}$ s$^{-1}$) where growth rates are more limited by Rubisco

150    activity (Fig. 4a, 4b, Extended Data Fig. 5)[27]. Regardless of the growth conditions, rosette

151    expansion rates or biomass accumulation were not distinguishable between transformants and

152    their segregant controls. Similarly, T2 EPYC1-dGFP WT plants (EpWT) showed no significant

153    differences compared to T2 segregant lines (AzWT). Due to the reduced Rubisco content in

154    the S2$_{Cr}$ background and differences in Rubisco catalytic characteristics, the growth

155    performance of S2$_{Cr}$ lines was slightly decreased compared to WT plants. The observed

156    differences in growth were in line with those reported previously for S2$_{Cr}$ and WT plants in the

157    absence of EPYC1[21].

158    Photosynthetic parameters derived from response curves of CO$_2$ assimilation rate to the

159    intercellular CO$_2$ concentration under saturating light were similar between respective Ep and

160    Az lines (Fig. 4c, 4d, Extended Data Table 1). The presence of condensates did not influence

161    the maximum achievable rates of Rubisco carboxylation ($V_{cmax}$). Notably, the CO$_2$ assimilation

162    rates at ambient concentrations of CO$_2$ for Ep and Az lines were comparable to WT lines when

163    normalized for Rubisco content ($A_{Rubisco}$) (Fig. 4d). This suggests that the modest reductions in

164    Rubisco turnover rate ($k_{cat}^c$) and specificity ($S_{C/O}$) for the plant-algal hybrid Rubisco in S2$_{Cr}$

165    compared to WT plants have only a mild impact on the efficiency of photosynthetic CO$_2$

166    assimilation and that the observed differences in growth rates are more associated with the

167    reduced levels of Rubisco in S2$_{Cr}$ plants[21]. Mesophyll conductance ($g_m$) levels were also

168    reduced in all S2$_{Cr}$ lines compared to WT plants, which is consistent with the impact of reduced

169    Rubisco content on $g_m$ observed in transplastomic tobacco[28]. Condensate formation in the

170    1A$_{At}$MOD background, where catalytic characteristics of the hybrid Rubisco are

171 indistinguishable from that of WT Rubisco, indicates that the SSU can be further engineered

172 to optimise phase separation, and Rubisco content and performance (Fig. 1c)[19]. Measurements

173 of the maximum electron transport rate ($J_{max}$) and the maximum potential quantum efficiency

174 of Photosystem II ($F_v/F_m$) were also indistinguishable between transformant and segregant

175 lines (Extended Data Table 1). Thus, the apparent displacement of the thylakoid membrane

176 matrix by the condensates (Fig. 1e) had no obvious impact on the efficiency of the light

177 reactions of photosynthesis.

178  Our findings show that EPYC1 and specific residues on the SSU are sufficient to aggregate

179 Rubisco into a single proto-pyrenoid condensate, which has no apparent negative impact on

180 plant growth. The overall photosynthetic performances of S2$_{Cr}$ transgenic lines appeared

181 unaffected by the condensate, which suggests that conditions inside higher plant chloroplasts

182 are highly compatible with the presence of pyrenoid-type bodies. Our data is arguably the key

183 step for the assembly of a pyrenoid-based CCM into plants that could increase crop yield

184 potentials by >60% (Fig. 4e)[8,9,29]. Previously described approaches for engineering the

185 cyanobacterial carboxysome-based CCM require engineering of the chloroplast-encoded

186 Rubisco large subunit, which is not generally feasible in major grain crops such as wheat and

187 rice[30], although expression of the large subunit from the nucleus of maize plants has been

188 recently demonstrated[31]. Furthermore, transgenic plants expressing cyanobacterial Rubisco

189 require high concentrations of $CO_2$ to grow. Here we demonstrate that condensation of Rubisco

190 is achievable through modification of the nuclear-encoded SSU, which is significantly more

191 amenable to genetic modification. Future work will focus on progressing towards a minimal

192 functional algal CCM by tethering the condensate to a thylakoid membrane for bicarbonate

193 delivery using appropriate transporters[32-34], as well as introduction of a lumenal carbonic

194 anhydrase and a stromal $CO_2$ salvage system[35-37].

195

196 **Methods**

197 **Plant material and growth conditions**

198 Arabidopsis (*Arabidopsis thaliana*, Col-0 background) seeds were sown on compost, stratified

199 for 3 d at 4°C and grown at 20°C, ambient $CO_2$ and 70% relative humidity under either 200 or

200 900 μmol photons $m^{-2}$ $s^{-1}$ supplied by cool white LED lights (Percival SE-41AR3cLED, CLF

201 PlantClimatics GmbH, Wertingen, Germany) in 12 h light, 12 h dark. For comparisons of

202 different genotypes, plants were grown from seeds of the same age and storage history,

203 harvested from plants grown in the same environmental conditions.

204

**Construct design and transformation**

The coding sequence of EPYC1 was codon optimised for expression in higher plants as in Atkinson et al.[19]. Mature EPYC1 was cloned directly into the level 0 acceptor vector pAGM1299 of the Plant MoClo system[38]. 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[39]. Homozygous transgenic and azygous lines were identified in the T2 generation using the pFAST-R selection cassette[40].

**Protein analyses**

Soluble protein was extracted from frozen leaf material of 21-d-old plants (sixth and seventh leaf) in protein extraction buffer (50mM HEPES-KOH pH 7.5 with 17.4% glycerol, 2% Triton X-100 and cOmplete Mini EDTA-free Protease Inhibitor Cocktail (Roche, Basel, Switzerland). Samples were heated at 70 °C for 15 min with 1 x Bolt LDS sample buffer (ThermoFisher Scientific, UK) and 200 mM 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[41], the SSU RbcS2 from Chlamydomonas (CrRbcS2) (raised to the C-terminal region of the SSU (KSARDWQPANKRSV) by Eurogentec, Southampton, UK) at 1:1,000 dilution, ACTIN (66009-1-Ig, Proteintech, UK) at 1:1000 dilution and EPYC1 at 1:2,000 dilution[6], 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).

**qRT-PCR analysis**

Total RNA was isolated from leaves of 21-d-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.[17], and run on a LightCycler 480 (Roche, Switzerland). Gene-specific primers were designed to amplify either eGFP (5'-CAGATTACGCCGTGAGCAAG-3' and 5'-GCTGAACTTGTGGCCGTTTA-3') or tGFP (5'-GCATCAGGGGTCTTGAAAGC-3' and 5'-TCCTTCAAAACGGTGGACCT-3') (IDT, Belgium). Amplification efficiency was determined with a calibration curve for each primer

239    pair. Two reference genes At4g26410 (RHIP1) and At1g13320 (PP2A) were used for

240    normalization[42]. The relative expression of eGFP and tGFP was calculated according to

241    Vandesompele et al.[43].

242

243    **Condensate extraction**

244    Soluble protein was extracted as before, then filtered through Miracloth (Merck Millipore,

245    Burlington, Massachusetts, USA), and centrifuged at 500 $g$ for 3 min at 4 °C, as in Mackinder

246    et al.[6]. The pellet was discarded, and the extract centrifuged again for 12 min. The resulting

247    pellet was washed once in protein extraction buffer, then re-suspended in a small volume of

248    buffer and centrifuged again for 5 min. Finally, the pellet was re-suspended in 25 µl of

249    extraction buffer and used in confocal analysis or SDS-PAGE electrophoresis.

250

251    **Growth analysis and photosynthetic measurements**

252    Rosette growth rates were quantified using an in-house imaging system[44]. Maximum quantum

253    yield of photosystem II (PSII) ($F_v/F_m$) was measured on 32-d-old plants using a Hansatech

254    Handy PEA continuous excitation chlorophyll fluorimeter (Hansatech Instruments Ltd, King's

255    Lynn, UK)[5]. Gas exchange and chlorophyll fluorescence were determined using a LI-COR LI-

256    6400 (LI-COR, Lincoln, Nebraska, USA) portable infra-red gas analyser with a 6400-40 leaf

257    chamber on either the sixth or seventh leaf of 35- to 45-d-old non-flowering rosettes grown in

258    large pots under 200 µmol photons $m^{-2}$ $s^{-1}$ to generate leaf area sufficient for gas exchange

259    measurements[46]. The response of $A$ to the intercellular $CO_2$ concentration ($C_i$) was measured

260    at various $CO_2$ concentrations (50, 100, 150, 200, 250, 300, 350, 400, 600, 800, 1000 and

261    1200 µmol $mol^{-1}$) under saturating light (1,500 µmol photons $m^{-2}$ $s^{-1}$). For all gas exchange

262    experiments, the flow rate was kept at 200 µmol $mol^{-1}$, leaf temperature was controlled at 25 °C

263    and chamber relative humidity was maintained at *ca*. 70%. Measurements were performed after

264    net assimilation and stomatal conductance had reached steady state. Gas exchange data were

265    corrected for $CO_2$ diffusion from the measuring chamber as in Bellasio et al.[47]. To estimate

266    $V_{cmax}$, $J_{max}$, $A_{Rubisco}$, the $CO_2$ compensation point (Γ) and mesophyll conductance to $CO_2$ ($g_m$)

267    the $A/C_i$ data were fitted to the $C_3$ photosynthesis model as in Ethier and Livingston[48] using the

268    catalytic parameters $K_c^{air}$ and affinity for $O_2$ ($K_o$) values for wild-type Arabidopsis Rubisco at

269    25 °C and the Rubisco content of WT and $S2_{Cr}$ lines[21].

270

271    **Confocal laser scanning and fluorescence recovery after photobleaching (FRAP)**

272 Leaves were imaged with a Leica TCS SP8 laser scanning confocal microscope (Leica

273 Microsystems, Milton Keynes, UK) as in Atkinson et al.[17]. Processing of images was done with

274 Leica LAS AF Lite software. Condensate and chloroplast dimensions were measured from

275 confocal images using Fiji (ImageJ, v1.52n)[49]. Condensate volume was calculated as a sphere,

276 while chloroplast volume was calculated as an ellipsoid, where depth was estimated as 25% of

277 the measured width. Chloroplast volumes varied between 24-102 $\mu m^3$, which was within the

278 expected size range and distribution for Arabidopsis chloroplasts[50]. Comparative

279 measurements of pyrenoid area were performed using Fiji on TEM cross-section images of WT

280 *C. reinhardtii* cells (cMJ030) as described in Itakura et al.[23]. FRAP was carried out on live leaf

281 tissue and tissue that had been fixed by infiltrating with 4% (v/v) formaldehyde for 90 min in

282 a vacuum chamber, then infiltrating three times for 5 min each with phosphate-buffered saline

283 (PBS) at pH 7.4. The SP8 microscope was used with a 40x water immersion objective and a

284 photomultiplier tube (PMT) detector. The 488 nm laser was set to 2% power for pre- and post-

285 bleach images, and 25% for the bleaching step. Pre-bleach images were captured at 220 ms

286 intervals (6 in total) and post-bleach images were captured at 400 ms intervals (at least 120).

287 For photo-bleaching the laser was directed to a region with a diameter of 0.5-0.6 $\mu m$ on one

288 side of the EPYC1 aggregate. Recovery time was calculated by comparing GFP expression to

289 an unbleached region of the same size.

290

291 **Super-resolution image microscopy**

292 Super-resolution images were acquired using structured illumination microscopy. Samples

293 were prepared on high precision cover-glass (Zeiss, Jena, Germany). 3D SIM images were

294 acquired on a N-SIM (Nikon Instruments, UK) using a 100x 1.49NA lens and refractive index

295 matched immersion oil (Nikon Instruments). Samples were imaged using a Nikon Plan Apo

296 TIRF objective (NA 1.49, oil immersion) and an Andor DU-897X-5254 camera using a 488nm

297 laser line. Z-step size for z stacks was set to 0.120 $\mu m$ as required by manufacturer's software.

298 For each focal plane, 15 images (5 phases, 3 angles) were captured with the NIS-Elements

299 software. SIM image processing, reconstruction and analysis were carried out using the N-SIM

300 module of the NIS-Element Advanced Research software. Images were checked for artefacts

301 using the SIMcheck software (www.micron.ox.ac.uk/software/SIMCheck.php). Images were

302 reconstructed using NiS Elements software v4.6 (Nikon Instruments) from a z stack comprising

303 of no less than 1 $\mu m$ of optical sections. In all SIM image reconstructions the Wiener and

304 Apodization filter parameters were kept constant.

305

**Immunogold labelling and electron microscopy**

Leaf samples were taken from 21-d-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 hr 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 hr 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), incubated 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 antibody (Abcam, Cambridge, UK) in TBSTT was carried out for 1 hr 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 analysis of variance (ANOVA) to determine the significance of the difference between sample groups. When 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, Chicago, IL, USA).

**Data availability**

All relevant data and plant materials are available from the authors upon request. Raw data corresponding to the figures and results described in this manuscript are available online [*Edinburgh DataShare address to be added*]. Additional data reported in this paper are presented as Extended Data.

**References**

1       Andersson, I. & Backlund, A. Structure and function of Rubisco. *Plant Physiol Bioch* **46**, 275-291, doi:10.1016/j.plaphy.2008.01.001 (2008).

2       Salesse-Smith, C. E. *et al.* Overexpression of Rubisco subunits with RAF1 increases Rubisco content in maize. *Nat Plants* **4**, 802-810, doi:10.1038/s41477-018-0252-4 (2018).

3       Flamholz, A. I. *et al.* Revisiting Trade-offs between Rubisco Kinetic Parameters. *Biochemistry-Us* **58**, 3365-3376, doi:10.1021/acs.biochem.9b00237 (2019).

4       Mackinder, L. C. M. *et al.* A Spatial Interactome Reveals the Protein Organization of the Algal CO2-Concentrating Mechanism. *Cell* **171**, 133-147 e114, doi:10.1016/j.cell.2017.08.044 (2017).

5       Freeman Rosenzweig, E. S. *et al.* The Eukaryotic CO2-Concentrating Organelle Is Liquid-like and Exhibits Dynamic Reorganization. *Cell* **171**, 148-162 e119, doi:10.1016/j.cell.2017.08.008 (2017).

6       Mackinder, L. C. *et al.* A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 5958-5963, doi:10.1073/pnas.1522866113 (2016).

7       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. *Nature communications* **9**, 5076, doi:10.1038/s41467-018-07624-w (2018).

8       McGrath, J. M. & Long, S. P. Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. *Plant physiology* **164**, 2247-2261, doi:10.1104/pp.113.232611 (2014).

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

10      Galmes, J. *et al.* Expanding knowledge of the Rubisco kinetics variability in plant species: environmental and evolutionary trends. *Plant Cell and Environment* **37**, 1989-2001, doi:10.1111/pce.12335 (2014).

11      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, doi:10.1093/jxb/erw163 (2016).

12      Bathellier, C., Tcherkez, G., Lorimer, G. H. & Farquhar, G. D. Rubisco is not really so bad. *Plant Cell and Environment* **41**, 705-716, doi:10.1111/pce.13149 (2018).

13      Wilson, R. H., Whitney, S. M. in *Directed Enzyme Evolution: Advances and Applications* (ed M. Alcalde) Ch. 4, 101-126 (Springer International Publishing AG, 2017).

384   14   Long, S. P., Marshall-Colon, A. & Zhu, X. G. Meeting the global food demand of the
385        future by engineering crop photosynthesis and yield potential. *Cell* **161**, 56-66,
386        doi:10.1016/j.cell.2015.03.019 (2015).
387   15   Li, F. W. *et al.* Anthoceros genomes illuminate the origin of land plants and the unique
388        biology of hornworts. *Nat Plants* **6**, 259-272, doi:10.1038/s41477-020-0618-2 (2020).
389   16   Raven, J. A., Beardall, J. & Sanchez-Baracaldo, P. The possible evolution and future of
390        CO2-concentrating mechanisms. *Journal of Experimental Botany* **68**, 3701-3716,
391        doi:10.1093/jxb/erx110 (2017).
392   17   Atkinson, N. *et al.* Introducing an algal carbon-concentrating mechanism into higher
393        plants: location and incorporation of key components. *Plant biotechnology journal* **14**,
394        1302-1315, doi:10.1111/pbi.12497 (2016).
395   18   Rae, B. D. *et al.* Progress and challenges of engineering a biophysical CO2-
396        concentrating mechanism into higher plants. *J Exp Bot* **68**, 3717-3737,
397        doi:10.1093/jxb/erx133 (2017).
398   19   Atkinson, N. *et al.* The pyrenoidal linker protein EPYC1 phase separates with hybrid
399        Arabidopsis-Chlamydomonas Rubisco through interactions with the algal Rubisco
400        small subunit. *J Exp Bot* **70**, 5271-5285, doi:10.1093/jxb/erz275 (2019).
401   20   Meyer, M. T. *et al.* Rubisco small-subunit alpha-helices control pyrenoid formation in
402        Chlamydomonas. *Proceedings of the National Academy of Sciences of the United States*
403        *of America* **109**, 19474-19479, doi:10.1073/pnas.1210993109 (2012).
404   21   Atkinson, N. *et al.* Rubisco small subunits from the unicellular green alga
405        Chlamydomonas complement Rubisco-deficient mutants of Arabidopsis. *The New*
406        *phytologist* **214**, 655-667, doi:10.1111/nph.14414 (2017).
407   22   Wunder, T., Guo Oh, Z. & Mueller-Cajar, O. CO$_2$-fixing liquid droplets: Towards a
408        dissection of the microalgal pyrenoid. *Traffic* **20,** 380-389, doi: 10.1111/tra.12650
409        (2019).
410   23   Itakura, A. K. *et al.* A Rubisco-binding protein is required for normal pyrenoid number
411        and starch sheath morphology in Chlamydomonas reinhardtii. *Proceedings of the*
412        *National Academy of Sciences of the United States of America* **116**, 18445-18454,
413        doi:10.1073/pnas.1904587116 (2019).
414   24   Rosowski, K. A *et al.* Elastic ripening and inhibition of liquid-liquid phase separation.
415        *Nature Physics* **16,** 422-425 (2020).
416   25   van der Hoorn, R. A. L. & Rivas, S. Unravelling the mode of action of plant proteases.
417        *The New phytologist* **218**, 879-881, doi:10.1111/nph.15156 (2018).
418   26   Valegård, K., Hasse, D., Andersson, I. & Gunn, L. Structure of Rubisco from
419        Arabidopsis thaliana in complex with 2-carboxyarabinitol-1,5-bis-phosphate. *Acta*
420        *Crystallographica Section D-Structural Biology* **74**, 1-9 (2018).
421   27   Lauerer, M. *et al.* Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in
422        transgenic tobacco transformed with "antisense" rbcS. VI. Effect on photosynthesis in
423        plants grown at different irradiance. *Planta* **190**, 332-345, doi:10.1007/bf00196962
424        (1993).
425   28   Galmes, J., Perdomo, J. A., Flexas, J. & Whitney, S. M. Photosynthetic characterization
426        of Rubisco transplantomic lines reveals alterations on photochemistry and mesophyll
427        conductance. *Photosynthesis Research* **115**, 153-166, doi:10.1007/s11120-013-9848-8
428        (2013).
429   29   Price, G. D., Badger, M. R. & von Caemmerer, S. The prospect of using cyanobacterial
430        bicarbonate transporters to improve leaf photosynthesis in C3 crop plants. *Plant*
431        *physiology* **155**, 20-26, doi:10.1104/pp.110.164681 (2011).
432   30   Long, B. M. *et al.* Carboxysome encapsulation of the CO2-fixing enzyme Rubisco in
433        tobacco chloroplasts. *Nature communications* **9**, doi:Artn 3570

434        10.1038/S41467-018-06044-0 (2018).

435  31  Salesse-Smith, C., *et al*. Overexpression of Rubisco subunits with RAF1 increases
436      Rubisco content in maize. *Nature Plants* **4,** 802-810, doi: 10.1038/s41477-018-0252-4
437      (2018).

438  32  Engel, B. D. *et al.* Native architecture of the Chlamydomonas chloroplast revealed by
439      in situ cryo-electron tomography. *eLife* **4**, doi: e0488910.7554/eLife.04889 (2015).33
440      Yamano, T., Sato, E., Iguchi, H., Fukuda, Y. & Fukuzawa, H. Characterization of
441      cooperative bicarbonate uptake into chloroplast stroma in the green alga
442      Chlamydomonas reinhardtii. *Proceedings of the National Academy of Sciences of the*
443      *United States of America* **112**, 7315-7320, doi:10.1073/pnas.1501659112 (2015).

444  34  Mukherjee, A. *et al.* Thylakoid localized bestrophin-like proteins are essential for the
445      $CO_2$ concentrating mechanism of Chlamydomonas reinhardtii. *Proceedings of the*
446      *National Academy of Sciences of the United States of America* **116**, 16915-16920,
447      doi:10.1073/pnas.1909706116 (2019).

448  35  Karlsson, J. *et al.* A novel alpha-type carbonic anhydrase associated with the thylakoid
449      membrane in Chlamydomonas reinhardtii is required for growth at ambient $CO_2$. *The*
450      *EMBO journal* **17**, 1208-1216, doi:10.1093/emboj/17.5.1208 (1998).

451  36  Yamano, T. *et al.* Light and low-$CO_2$-dependent LCIB-LCIC complex localization in
452      the chloroplast supports the carbon-concentrating mechanism in Chlamydomonas
453      reinhardtii. *Plant & cell physiology* **51**, 1453-1468, doi:10.1093/pcp/pcq105 (2010).

454  37  Toyokawa, C., Y. T., Fukuzawa H. Pyrenoid starch sheath is required for LCIB
455      localization and the $CO_2$-concentrating mechanism in green algae. *Plant physiology*,
456      doi:10.1104/pp.19.01587 (2020).

457  38  Engler, C. *et al.* A Golden Gate Modular Cloning Toolbox for Plants. *Acs Synth Biol* **3**,
458      839-843, doi:10.1021/sb4001504 (2014).

459  39  Clough, S. J. & Bent, A. F. Floral dip: a simplified method for Agrobacterium-mediated
460      transformation of Arabidopsis thaliana. *The Plant journal : for cell and molecular*
461      *biology* **16**, 735-743 (1998).

462  40  Shimada, T. L., Shimada, T. & Hara-Nishimura, I. A rapid and non-destructive
463      screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana.
464      *Plant Journal* **61**, 519-528, doi:10.1111/j.1365-313X.2009.04060.x (2010).

465  41  Howe, C. J. *et al.* Location and nucleotide sequence of the gene for the proton-
466      translocating subunit of wheat chloroplast ATP synthase. *Proceedings of the National*
467      *Academy of Sciences of the United States of America* **79**, 6903-6907 (1982).

468  42  Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K. & Scheible, W. R. Genome-
469      wide identification and testing of superior reference genes for transcript normalization
470      in Arabidopsis. *Plant Physiology* **139**, 5–17 (2005).

471  43  Vandesompele, J. *et al.* Accurate normalization of real-time quantitative RT-PCR data
472      by geometric averaging of multiple internal control genes. *Genome Biology* **3,** 7 (2002).

473  44  Dobrescu, A., Scorza, L. C. T., Tsaftaris, S. A. & McCormick, A. J. A "Do-It-Yourself"
474      phenotyping system: measuring growth and morphology throughout the diel cycle in
475      rosette shaped plants. *Plant methods* **13**, 95 (2017).

476  45  Maxwell, K. & Johnson, G. N. Chlorophyll fluorescence--a practical guide. *J Exp Bot*
477      **51**, 659-668 (2000).

478  46  Flexas, J. *et al.* Mesophyll conductance to $CO_2$ in Arabidopsis thaliana. *The New*
479      *phytologist* **175**, 501-511, doi:10.1111/j.1469-8137.2007.02111.x (2007).

480  47  Bellasio, C., Beerling, D. J. & Griffiths, H. An Excel tool for deriving key
481      photosynthetic parameters from combined gas exchange and chlorophyll fluorescence:
482      theory and practice. *Plant, cell & environment*, doi:10.1111/pce.12560 (2015).

483   48   Ethier, G. J. & Livingston, N. J. On the need to incorporate sensitivity to CO2 transfer
484        conductance into the Farquhar-von Caemmerer-Berry leaf photosynthesis model. *Plant,*
485        *Cell and Environment* **27**, 137-153, doi:10.1111/j.1365-3040.2004.01140.x (2004).
486   49   Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature*
487        *methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).
488   50   Crumpton-Taylor, M., Grandison, S., Png, K. M., Bushby, A. J. & Smith, A. M. Control
489        of starch granule numbers in Arabidopsis chloroplasts. *Plant physiology* **158**, 905-916,
490        doi:10.1104/pp.111.186957 (2012).
491   51   Diamos, A. G. & Mason, H. S. Chimeric 3' flanking regions strongly enhance gene
492        expression in plants. *Plant biotechnology journal* **16**, 1971-1982,
493        doi:10.1111/pbi.12931 (2018).
494   52   Sade, N. *et al.* Differential tissue-specific expression of NtAQP1 in Arabidopsis
495        thaliana reveals a role for this protein in stomatal and mesophyll conductance of CO2
496        under standard and salt-stress conditions. *Planta* **239**, 357-366, doi:10.1007/s00425-
497        013-1988-8 (2014).
498
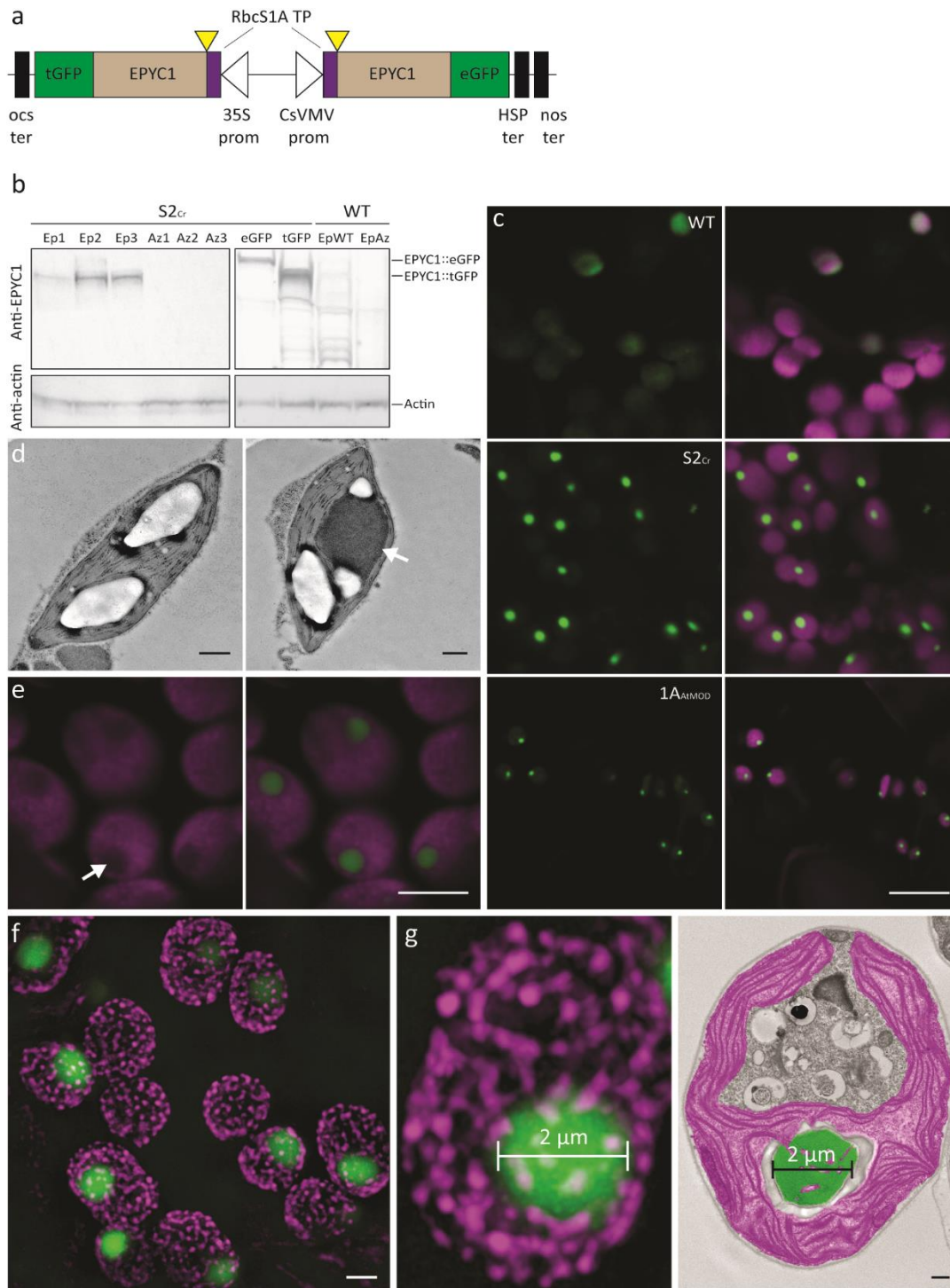
**Figure 1. Expression of EPYC1 in the Arabidopsis line S2$_{Cr}$ results in condensate formation. a**, Schematic representation of the dual GFP expression system (EPYC1-dGFP) for EPYC1 truncated at amino acid residue 27 (as indicated by yellow triangles) and fused at the N-terminus to the chloroplastic 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

508　used to increase expression[51]. **b**, EPYC1 protein levels in Arabidopsis plants as assessed by

509　immunoblot analysis with anti-EPYC1 antibodies. Shown are three T2 $S2_{Cr}$ transgenic plants

510　expressing EPYC1-dGFP (Ep1-3) and azygous segregants (Az1-3), and $S2_{Cr}$ plants

511　transformed with only EPYC1::tGFP (55.4 kDa) or EPYC1::eGFP (63.9 kDa). Also displayed

512　are a T2 EPYC1-dGFP WT transformant (EpWT) and azygous segregant (AzWT). Anti-actin

513　is shown as a loading control underneath. **c**, Expression of EPYC1-dGFP in WT, $S2_{Cr}$ and

514　$1A_{At}MOD$ backgrounds. Green and purple signals are GFP and chlorophyll autofluorescence,

515　respectively. Overlapping signals are white. Scale bars = 10 µm. **d**, TEM images of chloroplasts

516　from $S2_{Cr}$ plants with (right) and without (left) expression of EPYC1. A white arrowhead

517　indicates the dense dark grey area of the EPYC1 condensate. The large white structures are

518　starch granules. Scale bars = 0.5 µm. A representative chloroplast from a wild-type plant

519　expressing EPYC1-dGFP is shown for comparison in Extended Data Figure 3. **e**, Chlorophyll

520　autofluorescence is reduced at the site of EPYC1-dGFP accumulation (white arrow). Scale bar

521　= 5 µm. **f**, SIM microscopy showing EPYC1-dGFP condensates inside the chloroplast. The

522　magenta puncta show the position of grana stacks. Light magenta puncta indicate grana stacks

523　behind the condensate. Scale bar = 2 µm. **g**, Example comparison of condensate size (left, 2

524　µm) with that of a pyrenoid in Chlamydomonas (right, representative TEM image highlighting

525　the pyrenoid in green and chloroplast in purple). Scale bar for TEM image = 0.5 µm. Images

526　of EPYC1-dGFP condensates in the $S2_{Cr}$ background are from line Ep3.
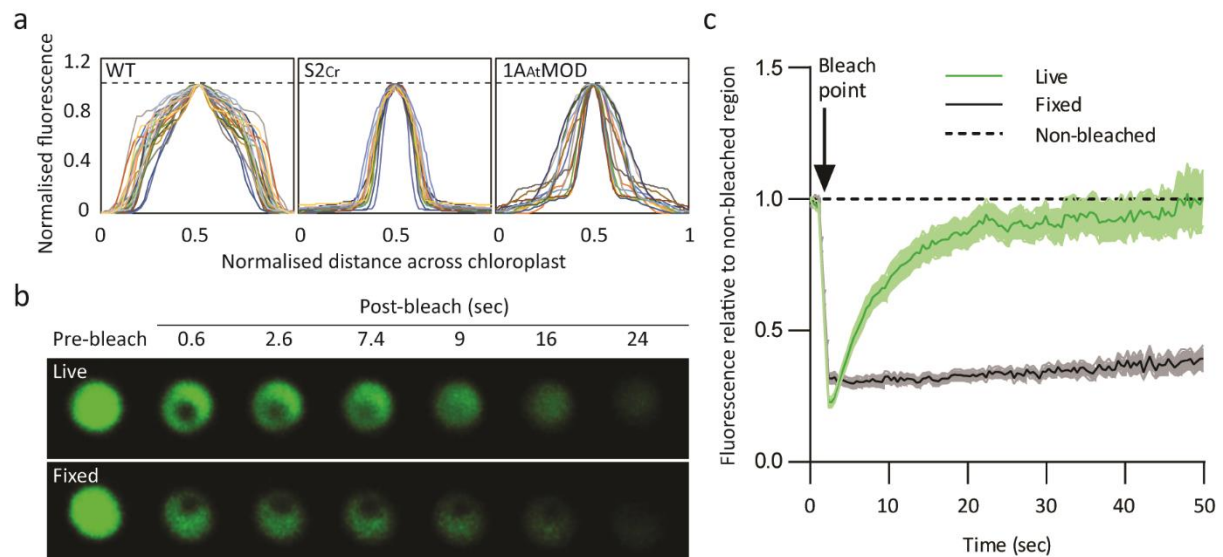
**Figure 2. *In planta* condensates behave like liquid-liquid phase separated micro compartments. a**, Fluorescence distribution plots of EPYC1-dGFP across the chloroplast. The intensity of the GFP fluorescence signal over the cross section of a chloroplast is shown in WT (n=28), $S2_{Cr}$ (line Ep3, 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. **c**, Fluorescence recovery of the bleached area in relation to the non-bleached area of condensates. The mean and SEM are shown for 13-16 chloroplasts.
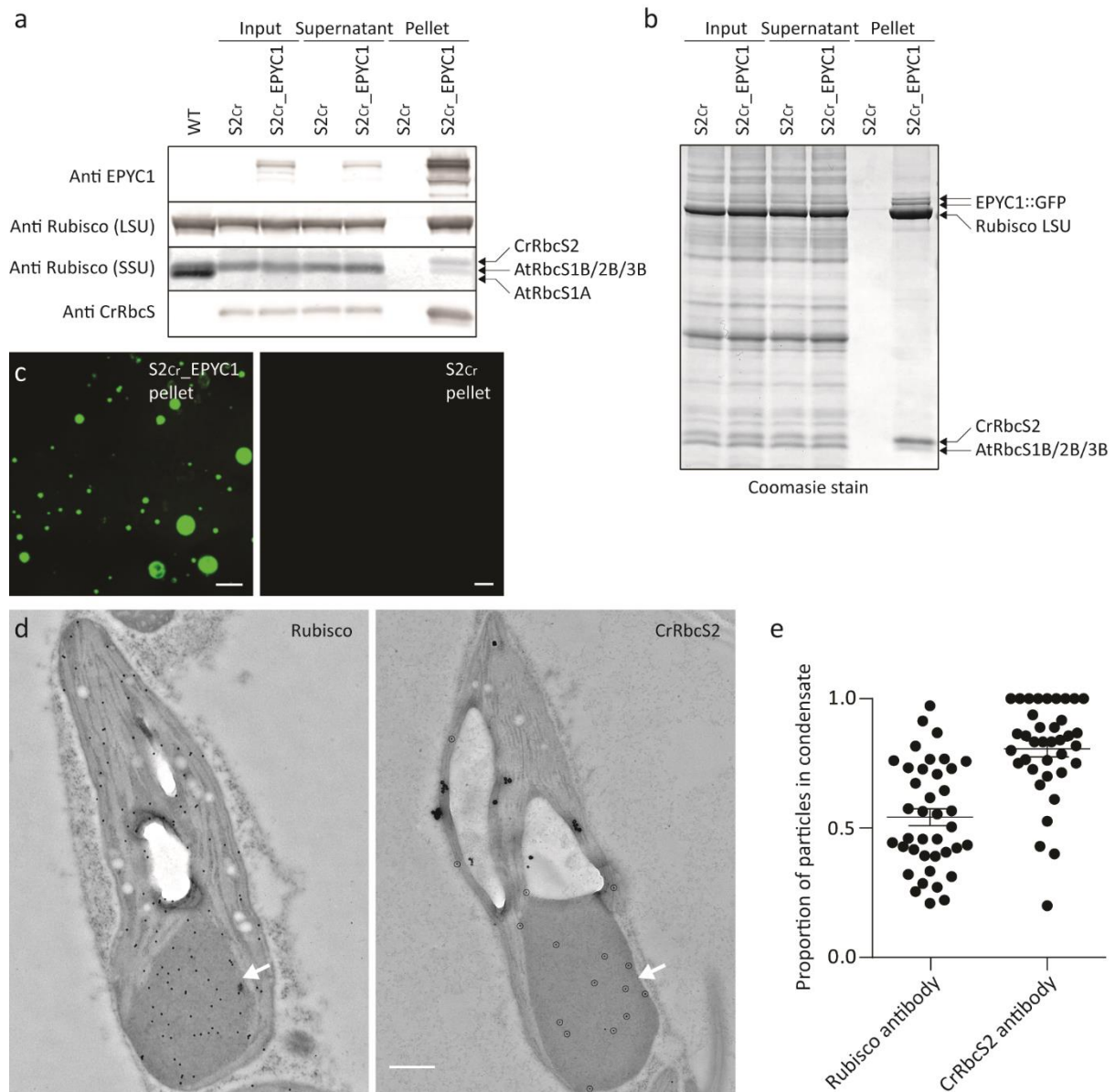
**Figure 3. Condensates contain EPYC1 and plant-algal hybrid Rubisco. a**, EPYC1 and Rubisco protein levels in whole leaf tissue (input), the supernatant following condensate extraction and centrifugation (supernatant) and the pellet (pellet) as assessed by immunoblot analyses with anti-EPYC1, anti-Rubisco (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 40x more concentrated than the input and supernatant. Molecular weights: LSU, 55 kD; RbcS1B, RbcS2B and RbcS3B, 14.8 kD; AtRbcS1A, 14.7kD; CrRbcS2 15.5 kD. **b**, Coomassie-stained SDS-PAGE gel showing the composition of the pelleted condensate. **c**, The condensates in the $S2_{Cr}$_EPYC1 pellet coalescence to form large liquid droplets. Scale bar = 50 µm. **d**, Representative immunogold

550  labelling of Rubisco in chloroplasts of an S2$_{Cr}$ transgenic line Ep3 expressing EPYC1-dGFP

551  probed with polyclonal anti-Rubisco (left) or anti-CrRbcS2 (right) antibodies (dots are

552  highlighted for the latter). The condensates are marked by a white arrowhead. Large white

553  structures are starch granules. Scale bar = 0.5 µm. **e**, Proportion of gold nanoparticles inside

554  the condensate compared to the chloroplast for each antibody. The mean and SEM are shown

555  for 37-39 chloroplasts. Scale bar = 0.5 µm.
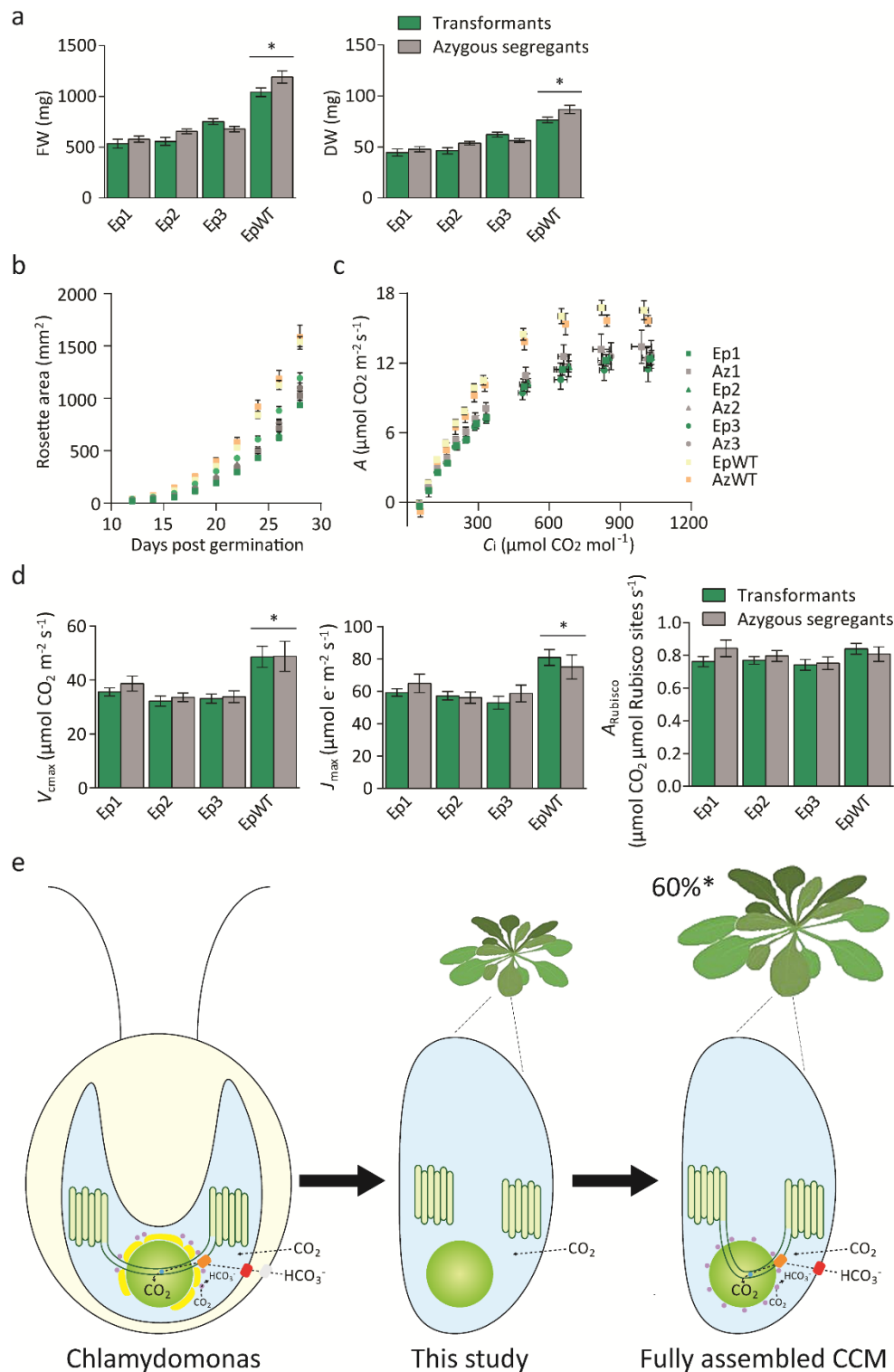
**Figure 4. EPYC1-mediated condensation of Rubisco has no negative impact on growth and photosynthesis. a**, Fresh and dry weights of three T2 EPYC1-dGFP $S2_{Cr}$ transgenic lines (Ep1-3) and an EPYC1-dGFP WT transformant (EpWT) (both in green) with their respective azygous segregants (Az1-3 and AzWT) (in grey). Plants were measured after 32 days of growth under 200 μmol photons $m^{-2}$ $s^{-1}$ light. The mean and SEM are shown for ≥12 individual plants for each line. **b**, Rosette expansion of $S2_{Cr}$ and WT lines in (a). **c**, Net $CO_2$ assimilation (*A*)

564   based on intercellular [$CO_2$] ($C_i$) under saturating light (1500 µmol photons m$^{-2}$ s$^{-1}$). Values

565   show the mean and SEM of measurements made on individual leaves from ten or more

566   individual rosettes. **d**, Variables derived from gas exchange data include maximum rate of

567   Rubisco carboxylation ($V_{cmax}$), maximum electron transport rate ($J_{max}$), stomatal conductance

568   ($G_s$), mesophyll conductance ($G_m$) and the net $CO_2$ assimilation rate at ambient concentrations

569   of $CO_2$ normalized to Rubisco ($A_{Rubisco}$). Asterisks indicate significant difference ($P<0.05$) of

570   EpWT lines compared to Ep lines as determined by ANOVA. **e**, Algal CCM components

571   required for enhancing photosynthesis. Generating a pyrenoid-like condensate in a plant

572   chloroplast provides a platform for introducing bicarbonate ($HCO_3^-$) channels/pumps at the

573   chloroplast envelope (e.g. LCIA, shown in red)[33] and thylakoid membrane (e.g. BST1-3, shown

574   in orange)[35], a lumenal carbonic anhydrase to convert $HCO_3^-$ to $CO_2$ for release into the

575   surrounding Rubisco condensate (CAH3, shown in blue)[35], mechanisms to capture $CO_2$ as

576   $HCO_3^-$ (LCIB and LCIC, shown in purple)[36,37] and traversing thylakoid membranes[32]. Current

577   models suggest that introduction of a functional biophysical CCM into a C3 plant could lead

578   to productivity gains of up to 60%[8,9,29].

579 **Extended Data Table 1.** Photosynthetic parameters from gas exchange and fluorescence measurements for $S2_{Cr}$ transgenic lines of Arabidopsis.
580 The mean and SEM are shown for seven 35- to 45-d-old rosettes for gas exchange variables and twelve 32-d-old rosettes for $F_v/F_m$. $F_v/F_m$ is shown
581 for attached leaves dark-adapted for 45 min prior to fluorescence measurements. Letters above the SEM indicate significant difference (P <0.05)
582 as determined by ANOVA followed by Tukey's HSD tests. Values followed by the same letter are not statistically significantly different.
583 Abbreviations: $\Gamma$, $CO_2$ compensation point ($C_i$–A); $F_v/F_m$, maximum potential quantum efficiency of photosystem II; $g_s$, stomatal conductance to
584 water vapour; $g_m$, mesophyll conductance to $CO_2$ (i.e. conductance of $CO_2$ across the pathway from intercellular airspace to chloroplast stroma)
585 using the Ethier and Livingston method[48,52]; $J_{max}$, maximum electron transport rate; $V_{cmax}$, maximum rate of Rubisco carboxylation.

586

587

| | Ep1 | Az1 | Ep2 | Az2 | Ep3 | Az3 | EpWt | AzWt |
|---|---|---|---|---|---|---|---|---|
| $V_{cmax}$ (µmol $CO_2$ m$^{-2}$ s$^{-1}$) | 35.6±1.5$^a$ | 36.4±2.0$^a$ | 32.2±1.9$^a$ | 33.6±1.6$^a$ | 33.1±1.9$^a$ | 33.8±2.2$^a$ | 44.9±1.6$^b$ | 43.3±1.7$^b$ |
| $J_{max}$ (µmol e$^-$ m$^{-2}$ s$^{-1}$) | 59.2±2.3$^a$ | 61.9±6.3$^a$ | 57.2±2.6$^a$ | 56.1±3.5$^a$ | 52.9±4.4$^a$ | 58.6±5.2$^a$ | 76.4±2.4$^b$ | 74.9±7.5$^b$ |
| $\Gamma$ (µmol $CO_2$ m$^{-2}$ s$^{-1}$) | 63±8$^a$ | 53±5$^a$ | 52±6$^a$ | 54±7$^a$ | 53±7$^a$ | 56±8$^a$ | 51±7$^a$ | 64±12$^a$ |
| $g_s$ (mol $H_2O$ m$^{-2}$ s$^{-1}$) | 0.249±0.031$^a$ | 0.279±0.051$^a$ | 0.233±0.017$^a$ | 0.251±0.015$^a$ | 0.233±0.021$^a$ | 0.236±0.016$^a$ | 0.287±0.018$^a$ | 0.306±0.011$^a$ |
| $g_m$ (mol $CO_2$ m$^{-2}$ s$^{-1}$) | 0.034±0.001$^b$ | 0.035±0.003$^b$ | 0.032±0.002$^b$ | 0.033±0.002$^b$ | 0.034±0.003$^b$ | 0.032±0.002$^b$ | 0.045±0.002$^a$ | 0.046±0.003$^a$ |
| $F_v/F_m$ (ML) | 0.848±0.002$^a$ | 0.849±0.002$^a$ | 0.848±0.001$^a$ | 0.847±0.001$^a$ | 0.847±0.002$^a$ | 0.845±0.002$^a$ | 0.851±0.002$^a$ | 0.850±0.001$^a$ |
| $F_v/F_m$ (HL) | 0.852±0.002$^a$ | 0.845±0.002$^a$ | 0.850±0.001$^a$ | 0.855±0.004$^a$ | 0.846±0.002$^a$ | 0.849±0.001$^a$ | 0.850±0.003$^a$ | 0.852±0.002$^a$ |

588

589 **Extended Data Figure 1. Full sequence of the EPYC1 expression cassettes.** Sequence parts
590 are colour coded, and MoClo overhangs are shown in lower case.
591
592
593
594
595
596
597
598
599
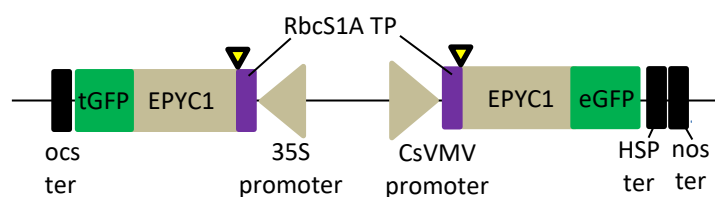


600
601 ocs terminator
602 turbo GFP (Pontellina pluma)
603 mature EPYC1
604 Rbcs1A transit peptide
605 35S promoter
606 linking region
607 CSVMV promoter
608 eGFP (with N-terminal HA-tag)
609 HSP terminator
610 nos terminator
611
612 GTCCTGCTGAGCCTCGACATGTTGTCGCAAAATTCGCCCTGGACCCGCCCAACGATTTGTCGTCACTGTCAAGGT
613 TTGACCTGCACTTCATTTGGGGCCCACATACACCAAAAAAATGCTGCATAATTCTCGGGGCAGCAAGTCGGTTAC
614 CCGGCCGCCGTGCTGGACCGGGTTGAATGGTGCCCGTAACTTTCGGTAGAGCGGACGGCCAATACTCAACTTCAA
615 GGAATCTCACCCATGCGCGCCGGCGGGGAACCGGAGTTCCCTTCAGTGAGCGTTATTAGTTCGCCGCTCGGTGTG
616 TCGTAGATACTAGCCCCTGGGGCACTTTTGAAATTTGAATAAGATTTATGTAATCAGTCTTTTAGGTTTGACCGG
617 TTCTGCCGCTTTTTTTAAAATTGGATTTGTAATAATAAAACGCAATTGTTTGTTATTGTGGCGCTCTATCATAGA
618 TGTCGCTATAAACCTATTCAGCACAATATATTGTTTTCATTTTAATATTGTACATATAAGTAGTAGGGTACAATC
619 AGTAAATTGAACGGAGAATATTATTCATAAAAATACGATAGTAACGGGTGATATATTCATTAGAATGAACCGAAA
620 CCGGCGGTAAGGATCTGAGCTACACATGCTCAGGTTTTTTACAACGTGCACAACAGAATTGAAAGCAAATATCAT
621 GCGATCATAGGCTTCTCGCATATCTCATTAAAGCAGGACaagcTTATCATTCCTCACCAGCATCAGCATCAGGGG
622 TCTTGAAAGCATGTTGGTACTCAACGATTCCAAGCTCGGTGTTAGAGTGATCTTCCTCAACTCTTCTGAAAGCGA
623 ACATAGGTCCACCGTTTTGAAGGATAGAAGGGTGGATAGCAGACTTGAAGTGCATGTGAGAATCCACCACAGAAG
624 AGTAGTAACCACCATCTCTAAGTGAGAAGGTTCTGGTGAAAGATCCATCGAGATCGTTATCTCCCATAGGATGAA
625 GATGCTCAACAGTAGCGTTAGACCTGATGATCTTGTCGGTGAAGATAACAGAATCCTCAGGGAATCCAGTTCCCA
626 TAACCTGCACATCAACAAATTTTGGTCATATATTAGAAAAGTTATAAATTAAAATATACACACTTATAAACTACA
627 GAAAAGCAATTGCTATATACTACATTCTTTTATTTTGAAAAAAATATTTGAAATATTATATTACTACTAATTAAT
628 GATAATTATTATATATATATCAAAGGTAGAAGCAGAAACTTACCTTGAAATCTCCGATCACTCTTCCAGCCTCGT
629 ATCTGTAAGAGAAGCTAACGTGAAGAACACCACCATCCTCGTACTTCTCGATCCTAGTGTTGGTGTATCCACCGT
630 TGTTGATAGCATGAAGGAAAGGGTTCTCGTATCCAGATGGGTAAGTTCCGAAGTGGTAGAATCCGTATCCCATAA
631 CGTGAGAAAGAAGGTATGGAGAGAAGGTAAGAGCACCCTTGGTAGACTTCATCTTGTTAGTCATTCTTCCCTGCT
632 CAGGAGTTCCCTCACCACCTCCAACAAGCTCGAACTCAACACCGTTAAGGGTTCCAGTGATTCTACACTCGATTT
633 CCATAGCAGGAAGTCCAGACTCATCAGACTCAGATCCAGATCCTCTcgaaAGGCCCTTTCTCCAGTCTGCAGGGA
634 GAGCCGTTCTCTTAATTTCTGGCTTGCTTTTACCTGTCCAAGGATTAGTTCCGGCTTTGTCAGCGGAGGAACTAG
635 ATGACGAAGCGGGAGCGTCTCTCCAGGAAGCGGATGATGATCTAGCAGGTGCAGGGCACTGCTTGCAGGTGCGG
636 GACTGTTGGATCGGAGTGACTCCAACTCCTGTTTCCAATTACTTGGAAGGGCCGATCTACTAGGGGTGACTGCTT
637 TTTTGCTCGCCGATGAGCTACGAGCTGGGGCGGAAGAAGCAGGAGCTGCATCCCTCCAACTAGCTGAGGATGACC
638 TAGCTGGTGCTGGCGCGCTAGAAGCTGGTGCTGGACTGCTTGATCGAAGGGATTCAAGTTCCTGTTTCCAGTTAG
639 AAGGCAGAGCTGAACGAGACGGAGTGACGGCCTTCTTGCTGGAAGAACTCCTAGCTGGAGCGCTAGATGCAGGCG
640 CTGCGTCCCTCCATGAGGCAGAACTACTTCTGGCCGGAGCTGGCGCTGATGAAGCCGGTGCAGGGCTGGAGCTTC
641 TCAATGACTCCAATTCCTGCTTCCAATTTGAAGGCAAGGCACTTCTTGAAGGAGTAACGGCCTTCTTTGAAGCTG
642 AGCTACTTCTGGCAGGCGCTGATGATGCAGGAGCGGCGTCTCGCCACGATGCGCTGCTCGATCTCGCAGGAGCAG
643 GAGCAGAGGACGCCGCCGACGAGCTGCCGTTACCATTTCTGAGTGATTCAAGTTCTTGCCTCCAATTAGCTGGGA
644 GAACGCTTCTTGTTGGGCTCACTCTGTTCGTTGCACTACTAGCTCTTGAGGCTTGAACTGTAGCAGTAGATGACT
645 CTCTCCAAGATCCTCTTGCTGCacctCCGCAGTTAACTCTTCCGCCGTTGCTTGTGATGGAAGTAATGTCGTTGT
646 TAGCCTTGCGGGTGGCTGGGAAGGCAGCGGAGGACTTAAGTCCGTTGAAAGGAGCGACCATAGTGGCCTGAGCCG
647 GAGAGGCAACCATAGTAGCGGAAGAGAGCATAGAGGAAGCcattGTATCGATAATTGTAAATGTAATTGTAATGT

```
648   TGTTTGTTGTTTGTTGTTGTTGGTAATTGTTGTAAAAATGAGCTCTTATACTCGAGCGTGTCCTCTCCAAATGAA
649   ATGAACTTCCTTATATAGAGGAAGGGTCTTGCGAAGGATAGTGGGATTGTGCGTCATCCCTTACGTCAGTGGAGA
650   TGTCACATCAATCCACTTGCTTTGTAGACGTGGTTGGAACCTCTTCTTTTTCCACGATGCTCCTCGTGGGTGGGG
651   GTCCATCTTTGGGACCACTGTCGGCAGAGAGATCTTGAATGATAGCCTTTCCTTTATCGCAATGATGGCATTTGT
652   AGGAGCCACCTTCCTTTTCTACTGTCCTTTCGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTC
653   CCGAAATTATCCTTTGTTGAAAAGTCTCAATAGCCCTTTGATCTTCTGAGACTGTATCTTTGACATTTTTGGAGT
654   AGACCAGAGTGTCGTGCTCCACCATGTTGACcctcGCAAGAATTCAAGCTTggagCCAGAAGGTAATTATCCAAG
655   ATGTAGCATCAAGAATCCAATGTTTACGGGAAAAACTATGGAAGTATTATGTAAGCTCAGCAAGAAGCAGATCAA
656   TATGCGGCACATATGCAACCTATGTTCAAAAATGAAGAATGTACAGATACAAGATCCTATACTGCCAGAATACGA
657   AGAAGAATACGTAGAAATTGAAAAAGAAGAACCAGGCGAAGAAAAGAATCTTGATGACGTAAGCACTGACGACAA
658   CAATGAAAAGAAGAAGATAAGGTCGGTGATTGTGAAAGAGACATAGAGGACACATGTAAGGTGGAAAATGTAAGG
659   GCGGAAAGTAACCTTATCACAAAGGAATCTTATCCCCCACTACTTATCCTTTTATATTTTTCCGTGTCATTTTTG
660   CCCTTGAGTTTTCCTATATAAGGAACCAAGTTCGGCATTTGTGAAAACAAGAAAAAATTTGGTGTAAGCTATTTT
661   CTTTGAAGTACTGAGGATACAACTTCAGAGAAATTTGTAAGTTTGTaatgGCTTCCTCTATGCTCTCTTCCGCTA
662   CTATGGTTGCCTCTCCGGCTCAGGCCACTATGGTCGCTCCTTTCAACGGACTTAAGTCCTCCGCTGCCTTCCCAG
663   CCACCCGCAAGGCTAACAACGACATTACTTCCATCACAAGCAACGGCGGAAGAGTTAACTGCGGaggtGCAGCAA
664   GAGGATCTTGGAGAGAGTCATCTACTGCTACAGTTCAAGCCTCAAGAGCTAGTAGTGCAACGAACAGAGTGAGCC
665   CAACAAGAAGCGTTCTCCCAGCTAATTGGAGGCAAGAACTTGAATCACTCAGAAATGGTAACGGCAGCTCGTCGG
666   CGGCGTCCTCTGCTCCTGCTCCTGCGAGATCGAGCAGCGCATCGTGGCGAGACGCCGCTCCTGCATCATCAGCGC
667   CTGCCAGAAGTAGCTCAGCTTCAAAGAAGGCCGTTACTCCTTCAAGAAGTGCCTTGCCTTCAAATTGGAAGCAGG
668   AATTGGAGTCATTGAGAAGCTCCAGCCCTGCACCGGCTTCATCAGCGCCAGCTCCGGCCAGAAGTAGTTCTGCCT
669   CATGGAGGGACGCAGCGCCTGCATCTAGCGCTCCAGCTAGGAGTTCTTCCAGCAAGAAGGCCGTCACTCCGTCTC
670   GTTCAGCTCTGCCTTCTAACTGGAAACAGGAACTTGAATCCCTTCGATCAAGCAGTCCAGCACCAGCTTCTAGCG
671   CGCCAGCACCAGCTAGGTCATCCTCAGCTAGTTGGAGGGATGCAGCTCCTGCTTCTTCCGCCCCAGCTCGTAGCT
672   CATCGGCGAGCAAAAAAGCAGTCACCCCTAGTAGATCGGCCCTTCCAAGTAATTGGAAACAGGAGTTGGAGTCAC
673   TCCGATCCAACAGTCCCGCACCTGCAAGCAGTGCCCCTGCACCTGCTAGATCATCATCCGCTTCCTGGAGAGACG
674   CTCCCGCTTCGTCATCTAGTTCCTCCGCTGACAAAGCCGGAACTAATCCTTGGACAGGTAAAAGCAAGCCAGAAA
675   TTAAGAGAACGGCTCTCCCTGCAGACTGGAGAAAGGGCCTttcgTACCCATACGATGTTCCTGACTATGCGGGCT
676   ATCCCTATGACGTCCCGGACTATGCAGGATTGTATCCATATGACGTTCCAGATTACGCCACTAGAGCTGCTTACC
677   CATACGATGTTCCTGACTATGCGGGCTATCCCTATGACGTCCCGGACTATGCAGGATTGTATCCATATGACGTTC
678   CAGATTACGCCGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACG
679   TAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCA
680   TCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCA
681   GCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCA
682   CCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACC
683   GCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACA
684   GCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCG
685   AGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCG
686   ACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGG
687   AGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAgcttATATGAAGATGAAGATGA
688   AATATTTGGTGTGTCAAATAAAAAGCTTGTGTGCTTAAGTTTGTGTTTTTTTCTTGGCTTGTTGTGTTATGAATT
689   TGTGGCTTTTTCTAATATTAAATGAATGTAAGATCTCATTATAATGAATAAACAAATGTTTCTATAATCCATTGT
690   GAATGTTTTGTTGGATCTCTTCTGCAGCATATAACTACTGTATGTGCTATGGTATGGACTATGGAATATGATTAA
691   AGATAAGagatGTCAAGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCT
692   TGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATT
693   TATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGC
694   AAACTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCGA
```

695 **Extended Data Figure 2. Transcript abundances of tGFP and eGFP in EPYC1-dGFP**

696 **plant lines.** qRT-PCR was carried out on three T2 $S2_{Cr}$ transgenic plants expressing EPYC1-

697 dGFP (Ep1- Ep3, as in Fig. 1B) using gene-specific primers for tGFP and eGFP. Abundances

698 of tGFP and eGFP transcripts are shown relative to the highest tGFP expression level and

699 normalised to reference genes PP2A (At1g13320) and RHIP1 (At4g26410). Error bars show

700 the SEM for two biological replicates.

701



702
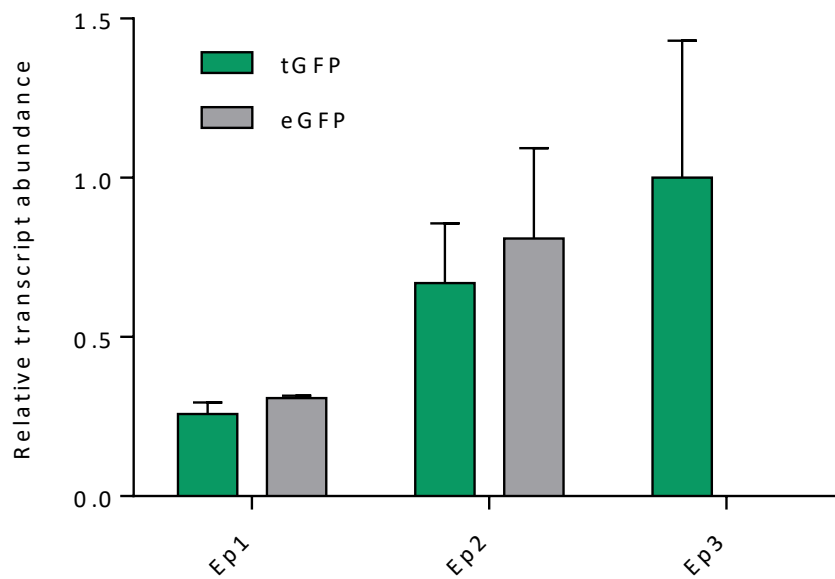
**Extended Data Figure 3. Characterisation of condensates in S2$_{Cr}$ plants. a**, Maximum projection of a z stack showing condensates in every chloroplast. Scale bar = 5 µm. **b**, The size of condensates is dependent on the expression level of EPYC1. Arabidopsis lines Ep1-3 with different expression levels of EPYC1-dGFP (see Figure. 1b) have different sizes of condensates. Scale bar = 10 µm. **c**, Data derived from confocal images of Chlamydomonas pyrenoids (n=55) and chloroplasts (n=42) from each of the three Ep transgenic lines (Ep1-3) showing the mean diameter and SEM of pyrenoids and condensates (left). The volume of the condensate is shown plotted against estimated chloroplast volume (centre) and the estimated proportion of chloroplast volume occupied by the condensate (right) (n=27 chloroplasts for each line).
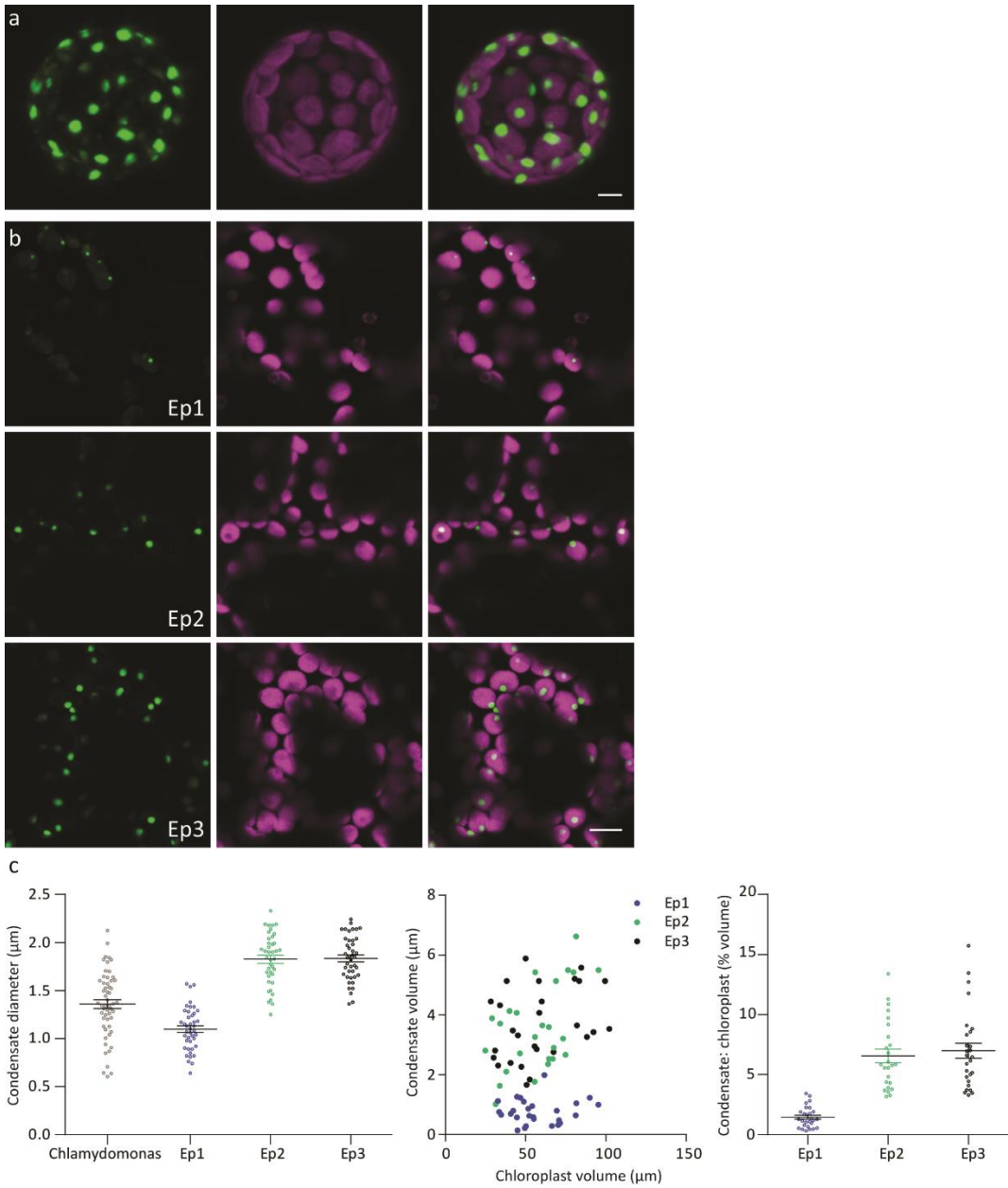
715  **Extended Data Figure 4. Condensate formation in S2$_{Cr}$ plants**. **a**, Condensation occurs in

716  S2$_{Cr}$ plants expressing a single EPYC1 expression cassette. Representative images are shown

717  for EPYC1 fused at the C-terminus to either tGFP (top) or eGFP (bottom). Scale bar = 10 µm.

718  **b**, Representative TEM image of a mesophyll cell cross-section showing chloroplasts with

719  EPYC1-dGFP condensates. Visible condensates are marked by a white arrowhead. The section

720  has been probed by immunogold labelling with anti-Rubisco antibodies.  Scale bar = 1 µm. c,

721  Representative TEM image of a chloroplast from a wild-type plant expressing EPYC1-dGFP.

722  No condensates were observed in this background. Scale bar = 0.5 µm.

723



724
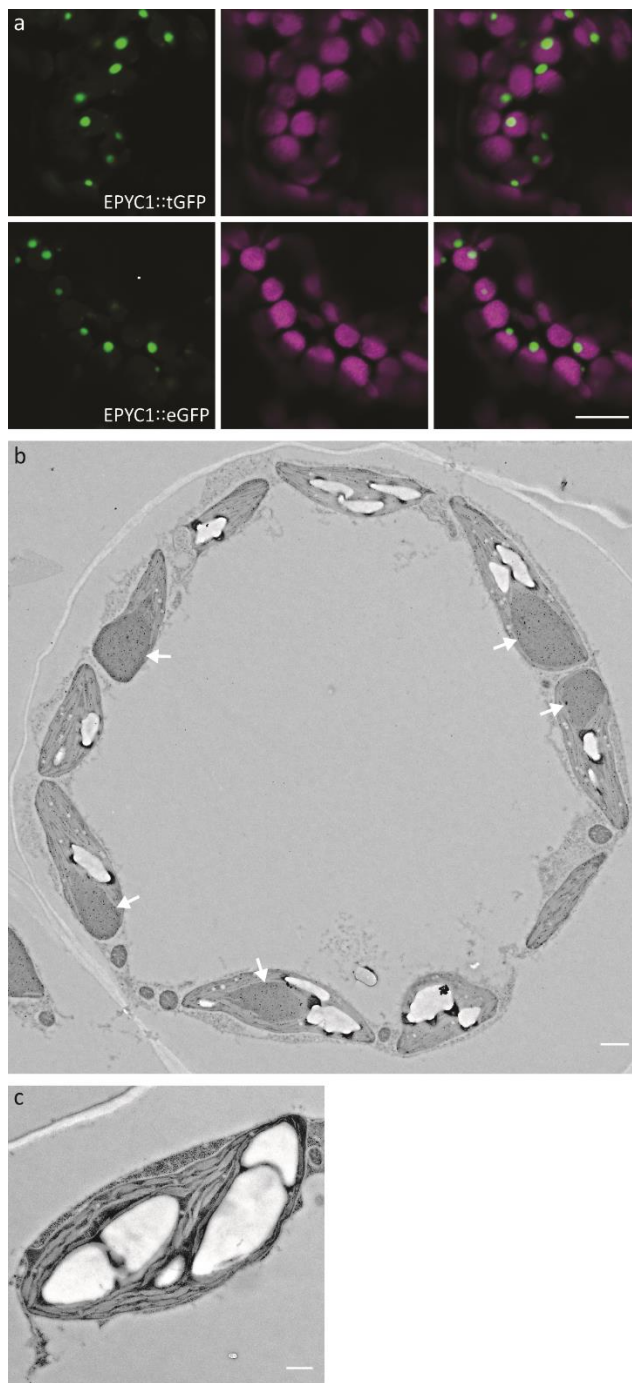
725    **Extended Data Figure 5. Growth of plant lines expressing EPYC1 under different light**

726    **levels. a,** Fresh (FW) and dry weight (DW) of three T2 EPYC1-dGFP S2$_{Cr}$ transgenic lines

727    (Ep1-3) and an EPYC1-dGFP WT transformant (EpWT) with their respective azygous

728    segregants (Az1-3 and AzWT) measured after 32 days of growth at 200 or 900 µmol photons

729    m$^{-2}$ s$^{-1}$ light. Asterisks indicate significant difference ($P<0.05$) as determined by ANOVA. **b,**

730    Rosette area expansion rates for individual transgenic lines and azygous segregants at two

731    different light levels. The mean and SEM are shown for ≥12 individual plants for each line.

732



733