nature plants

ARTICLES

https://doi.org/10.1038/s41477-022-01153-7

Check for updates

OPEN

Modelling the pyrenoid-based CO_2-concentrating mechanism provides insights into its operating principles and a roadmap for its engineering into crops

Chenyi Fei[1,2,6], Alexandra T. Wilson[1,3,6], Niall M. Mangan[4]✉, Ned S. Wingreen[1,2]✉ and Martin C. Jonikas[1,5]✉

Many eukaryotic photosynthetic organisms enhance their carbon uptake by supplying concentrated CO_2 to the CO_2-fixing enzyme Rubisco in an organelle called the pyrenoid. Ongoing efforts seek to engineer this pyrenoid-based CO_2-concentrating mechanism (PCCM) into crops to increase yields. Here we develop a computational model for a PCCM on the basis of the postulated mechanism in the green alga *Chlamydomonas reinhardtii*. Our model recapitulates all *Chlamydomonas* PCCM-deficient mutant phenotypes and yields general biophysical principles underlying the PCCM. We show that an effective and energetically efficient PCCM requires a physical barrier to reduce pyrenoid CO_2 leakage, as well as proper enzyme localization to reduce futile cycling between CO_2 and HCO_3^-. Importantly, our model demonstrates the feasibility of a purely passive CO_2 uptake strategy at air-level CO_2, while active HCO_3^- uptake proves advantageous at lower CO_2 levels. We propose a four-step engineering path to increase the rate of CO_2 fixation in the plant chloroplast up to threefold at a theoretical cost of only 1.3 ATP per CO_2 fixed, thereby offering a framework to guide the engineering of a PCCM into land plants.

The CO_2-fixing enzyme Rubisco mediates the entry of roughly 10^{14} kilograms of carbon into the biosphere each year[1–3]. However, in many plants Rubisco fixes CO_2 at less than one-third of its maximum rate under atmospheric levels of CO_2 (Supplementary Fig. 1)[4–6], which limits the growth of crops such as rice and wheat[7]. To overcome this limitation, many photosynthetic organisms, including C_4 plants[8,9], crassulacean acid metabolism (CAM) plants[10], algae[11,12] and cyanobacteria[13], enhance Rubisco's CO_2 fixation rate by supplying it with concentrated CO_2[14,15]. In algae, such a CO_2-concentrating mechanism occurs within a phase-separated organelle called the pyrenoid[16–19]. Pyrenoid-based CO_2-concentrating mechanisms (PCCMs) mediate approximately one-third of global CO_2 fixation[16].

While previous works have identified essential molecular components for the PCCM[16,20–29], key operating principles of this mechanism remain poorly understood due to a lack of quantitative and systematic analysis. At the same time, there is growing interest in engineering a PCCM into C_3 crops to improve yields and nitrogen- and water-use efficiency[30,31]. Key questions are: (1) What is the minimal set of components necessary to achieve a functional PCCM? (2) What is the energetic cost of operating a minimal PCCM?

To advance our understanding of the PCCM, we develop a reaction-diffusion model on the basis of the postulated mechanism in the green alga *Chlamydomonas reinhardtii* (*Chlamydomonas* hereafter; Fig. 1a)[31–33]: Briefly, external inorganic carbon (Ci: CO_2 and HCO_3^-) is transported across the plasma membrane by transporters LCI1 (Cre03.g162800) and HLA3 (Cre02.g097800)[23,24,34]. Cytosolic Ci becomes concentrated in the chloroplast stroma in the form of HCO_3^-, either via conversion of CO_2 to HCO_3^- by the putative stromal carbonic anhydrase LCIB/LCIC (Cre10.g452800/Cre06.g307500) complex (LCIB hereafter)[22,35,36] or via direct transport across the chloroplast membrane by the poorly characterized HCO_3^- transporter LCIA (Cre06.g309000)[24,37]. It is currently not known whether LCIA is a passive channel or a pump; therefore, in the model we first consider it as a passive channel (denoted by LCIA^C) and later consider it as an active pump (denoted by LCIA^P). Once in the stroma, HCO_3^- travels via the putative HCO_3^- channels BST1–3 (Cre16.g662600, Cre16.g663400 and Cre16.g663450)[25] into the thylakoid lumen, and diffuses via membrane tubules into the pyrenoid where the carbonic anhydrase CAH3 (Cre09.g415700)[38–40] converts HCO_3^- into CO_2. This CO_2 diffuses from the thylakoid tubule lumen into the pyrenoid matrix, where Rubisco catalyses fixation. Supplementary Table 1 summarizes the acronyms of key proteins in the *Chlamydomonas* PCCM.

We model the above enzymatic activities and Ci transport in a spherical chloroplast. We assume that carbonic anhydrases catalyse the bidirectional interconversion of CO_2 and HCO_3^-, producing a net flux in one direction where the two species are out of equilibrium. We consider three chloroplast compartments at constant pH values: a spherical pyrenoid matrix (pH 8, ref. [41]) in the centre, a surrounding stroma (pH 8, ref. [41,42]), and thylakoids (luminal pH 6, ref. [43]) traversing both the matrix and stroma (Fig. 1b and Supplementary Fig. 2). The flux balance of intracompartment reaction and diffusion and intercompartment exchange sets the steady-state concentration profiles of Ci species in all compartments (Methods). To account for the effect of Ci

[1]Department of Molecular Biology, Princeton University, Princeton, NJ, USA. [2]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA. [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. [4]Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, IL, USA. [5]Howard Hughes Medical Institute, Princeton University, Princeton, NJ, USA. [6]These authors contributed equally: Chenyi Fei, Alexandra T. Wilson. ✉e-mail: niall.mangan@northwestern.edu; wingreen@princeton.edu; mjonikas@princeton.edu

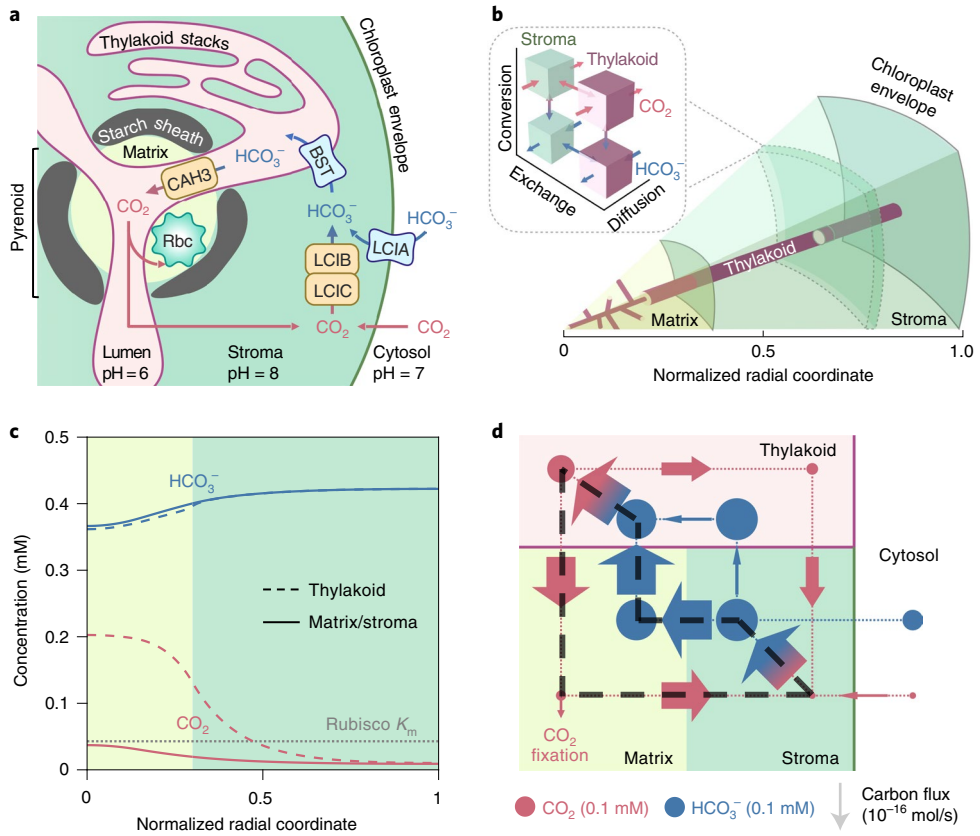

Fig. 1 | A multicompartment reaction-diffusion model describes the *Chlamydomonas* PCCM. a, Cartoon of a *Chlamydomonas* chloroplast with known PCCM components. HCO_3^- is transported across the chloroplast membrane by LCIA and across the thylakoid membranes by BST1–3 (referred to as BST henceforth for simplicity). In the acidic thylakoid lumen, a carbonic anhydrase CAH3 converts HCO_3^- into CO_2, which diffuses into the pyrenoid matrix where the CO_2-fixing enzyme Rubisco (Rbc) is localized. CO_2 leakage out of the matrix and the chloroplast can be impeded by potential diffusion barriers—a starch sheath and stacks of thylakoids—and by conversion to HCO_3^- by a CO_2-recapturing complex LCIB/LCIC (referred to as LCIB henceforth for simplicity) in the basic chloroplast stroma. **b**, A schematic of the modelled PCCM, which considers intracompartment diffusion and intercompartment exchange of CO_2 and HCO_3^-, as well as their interconversion, as indicated in the inset. Colour code as in **a**. The model is spherically symmetric and consists of a central pyrenoid matrix surrounded by a stroma. Thylakoids run through the matrix and stroma; their volume and surface area correspond to a reticulated network at the centre of the matrix extended by cylinders running radially outward. **c**, Concentration profiles of CO_2 and HCO_3^- in the thylakoid (dashed curves) and in the matrix/stroma (solid curves) for the baseline PCCM model that lacks LCIA activity and diffusion barriers. Dotted grey line indicates the effective Rubisco K_m for CO_2 (Methods). Colour code as in **a**. **d**, Net fluxes of inorganic carbon between the indicated compartments. The width of arrows is proportional to flux; the area of circles is proportional to the average molecular concentration in the corresponding regions. The black dashed loop denotes the major futile cycle of inorganic carbon in the chloroplast. Colour code as in **a**. For **c** and **d**, LCIA^C-mediated chloroplast membrane permeability to HCO_3^- $\kappa^{H^-}_{chlor} = 10^{-8}$ m s^{-1}, BST-mediated thylakoid membrane permeability to HCO_3^- $\kappa^{H^-}_{thy} = 10^{-2}$ m s^{-1}, LCIB rate $V_{LCIB} = 10^3$ s^{-1} and CAH3 rate $V_{CAH3} = 10^4$ s^{-1} (Methods). Other model parameters are estimated from experiments (Supplementary Table 2).

transport across the cell membrane, we simulate a broad range of surrounding cytosolic Ci pools from which the chloroplast can uptake Ci. We characterize the performance of the modelled PCCM with two metrics: (1) its efficacy, quantified by the computed CO_2 fixation flux normalized by the maximum possible flux through Rubisco; and (2) its efficiency, quantified by the ATP cost per CO_2 fixed (Methods).

Results

A baseline PCCM driven by intercompartmental pH differences. To identify the minimal components of a functional PCCM, we build a baseline model (Fig. 1c,d), with the carbonic anhydrase LCIB diffuse throughout the stroma, BST channels for HCO_3^- uniformly distributed across the thylakoid membranes, the carbonic anhydrase CAH3 localized to the thylakoid lumen within the pyrenoid, and Rubisco condensed within the pyrenoid matrix. This model lacks the HCO_3^- transporter LCIA and potential diffusion barriers to Ci. We first analyse modelled PCCM performance under air-level CO_2 (10 µM cytosolic); lower CO_2 conditions are discussed in later sections.

CO_2 diffusing into the chloroplast is converted to HCO_3^- in the high-pH stroma where the equilibrium CO_2:HCO_3^- ratio is 1:80 (Methods). Since passive diffusion of HCO_3^- across the chloroplast envelope is very slow, this concentrated HCO_3^- becomes trapped in the stroma. The BST channels equilibrate HCO_3^- across the thylakoid membrane, so HCO_3^- also reaches a high concentration in the thylakoid lumen (Fig. 1c). The low pH in the thylakoid lumen favours a roughly equal equilibrium partition between CO_2 and HCO_3^-; however, HCO_3^- is not brought into equilibrium with CO_2 immediately upon entering the thylakoid outside the pyrenoid, since no carbonic anhydrase (CA) is present there. Instead, HCO_3^- diffuses within the thylakoid lumen towards the pyrenoid, where CAH3 localized within the pyrenoid radius rapidly converts HCO_3^- back to CO_2 (Fig. 1d). This CO_2 can diffuse across the thylakoid membrane into the pyrenoid matrix. This baseline model, driven solely by intercompartmental pH differences, achieves a pyrenoidal

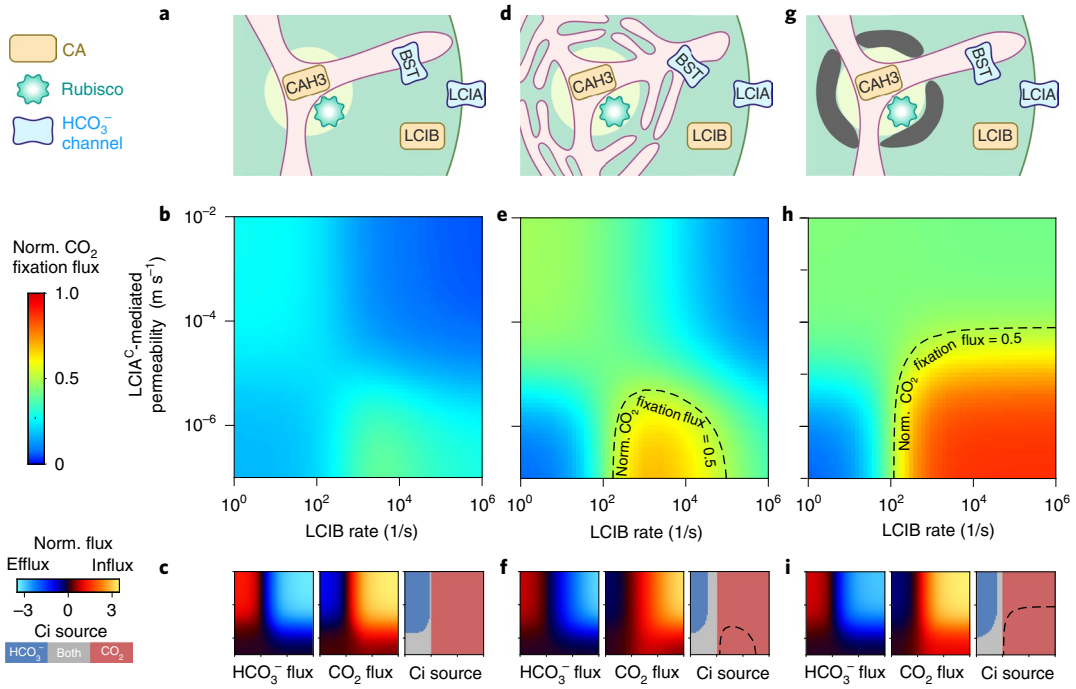

Fig. 2 | Barriers to CO_2 diffusion out of the pyrenoid matrix enable an effective PCCM driven only by intercompartmental pH differences. a–i, A model with no barrier to CO_2 diffusion out of the pyrenoid matrix (**a**–**c**) is compared to a model with thylakoid stacks slowing inorganic carbon diffusion in the stroma (**d**–**f**) and a model with an impermeable starch sheath (**g**–**i**) under air-level CO_2 (10 μM cytosolic). **a**,**d**,**g**, Schematics of the modelled chloroplast. **b**,**e**,**h**, Heatmaps of normalized CO_2 fixation flux, defined as the ratio of the total Rubisco carboxylation flux to its maximum if Rubisco were saturated, at varying LCIA^C-mediated chloroplast membrane permeabilities to HCO_3^- and varying LCIB rates. The BST-mediated thylakoid membrane permeability to HCO_3^- is the same as in Fig. 1c,d. For **e** and **h**, dashed black curves indicate a normalized CO_2 fixation flux of 0.5. **c**,**f**,**i**, Overall fluxes of HCO_3^- (left) and CO_2 (middle) into the chloroplast, normalized by the maximum CO_2 fixation flux if Rubisco were saturated, at varying LCIA^C-mediated chloroplast membrane permeabilities to HCO_3^- and varying LCIB rates. Negative values denote efflux out of the chloroplast. The inorganic carbon (Ci) species with a positive influx is defined as the Ci source (right). Axes are the same as in **b**, **e** and **h**.

CO_2 concentration approximately 2.5 times that found in a model with no PCCM.

The baseline PCCM suffers from pyrenoid CO_2 leakage. The substantial CO_2 leakage out of the matrix in the baseline model (Fig. 1d) is in part due to the relatively slow kinetics of Rubisco. During the average time required for a CO_2 molecule to be fixed by Rubisco in the pyrenoid, that CO_2 molecule can typically diffuse a distance larger than the pyrenoid radius (Supplementary Note I). Therefore, most of the CO_2 molecules entering the pyrenoid matrix will leave without being fixed by Rubisco (Supplementary Fig. 3). One might think that adding LCIA^C as a passive channel to enhance HCO_3^- diffusion into the chloroplast could overcome this deficit (Fig. 2a). However, even with the addition of LCIA^C to our baseline PCCM model, no combination of enzymatic activities and channel transport rates achieves an effective PCCM, that is, more than half-saturation of Rubisco with CO_2 (Fig. 2b and Supplementary Fig. 4). Thus, the pH-driven PCCM cannot operate effectively without a diffusion barrier.

Barriers to pyrenoidal CO_2 leakage enable a pH-driven PCCM. To operate a more effective PCCM, the cell must reduce CO_2 leakage from the pyrenoid matrix. A barrier to CO_2 diffusion has been regarded as essential for various CO_2-concentrating mechanisms[44–47]. Although the matrix is densely packed with Rubisco, our analysis suggests that the slowed diffusion of CO_2 in the pyrenoid matrix due to volume occupied by Rubisco can only account for a 10% decrease in CO_2 leakage (Supplementary Note VI.C). Thus, we consider alternative barriers in our model.

We speculate that thylakoid membrane sheets and the pyrenoid starch sheath could serve as effective barriers to decrease leakage of CO_2 from the matrix. Thylakoid membrane sheets could serve as effective barriers to CO_2 diffusion because molecules in the stroma must diffuse between and through the interdigitated membranes[45]. Indeed, our first-principle simulations suggest that the thylakoid stacks, modelled with realistic geometry[48], effectively slow the diffusion of Ci in the stroma (Supplementary Fig. 5). Evidence on the role of the starch sheath in the PCCM is limited and mixed. While early work suggested that a starchless *Chlamydomonas* mutant had normal PCCM performance in air[49], the phenotype was not compared to the appropriate parental strain. A more recent study found that a mutant (*sta2-1*) with a thinner starch sheath than wild-type strains displays decreased PCCM efficacy at very low CO_2[50]. On the basis of the latter work, we hypothesize that the starch sheath that surrounds the matrix may act as a barrier to CO_2 diffusion. Since the starch sheath consists of many lamellae of crystalline amylopectin[51–53], we model it as an essentially impermeable barrier equivalent to 10 lipid bilayers; in its presence, most CO_2 leakage out of the matrix occurs through the thylakoid tubules (Supplementary Fig. 6).

We next test whether the above two realistic diffusion barriers allow for an effective pH-driven PCCM. Adding either thylakoid stacks or a starch sheath to the baseline PCCM model above drastically reduces CO_2 leakage from the matrix to the stroma (Supplementary Fig. 7). The resulting PCCM is highly effective under air-level CO_2 (10 μM cytosolic) conditions: pyrenoidal CO_2 concentrations are raised above the effective half-saturation constant K_m of Rubisco (Methods) using only

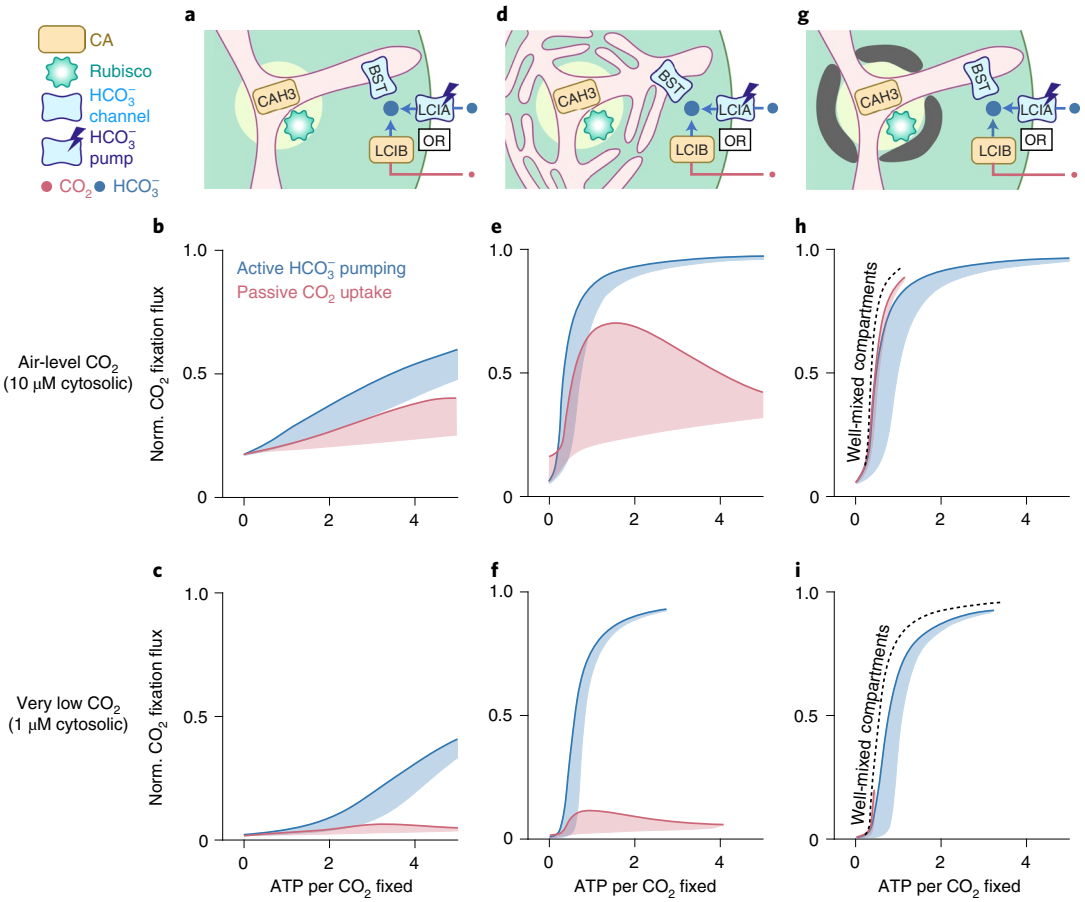

Fig. 3 | Feasible inorganic carbon uptake strategies for the chloroplast depend on the environmental level of CO_2. a–i, Results are shown for a model with no barrier to CO_2 diffusion out of the pyrenoid matrix (**a**–**c**), a model with thylakoid stacks serving as diffusion barriers (**d**–**f**) and a model with an impermeable starch sheath (**g**–**i**). **a**,**d**,**g**, Schematics of the modelled chloroplast employing LCIB for passive CO_2 uptake (red), or employing active $LCIA^P$-mediated HCO_3^- pumping across the chloroplast envelope and no LCIB activity (blue). PCCM performance under air-level CO_2 (10 μM cytosolic) (**b**,**e**,**h**) and under very low CO_2 (1 μM cytosolic) (**c**,**f**,**i**) are shown, as measured by normalized CO_2 fixation flux versus ATP spent per CO_2 fixed, for the two inorganic carbon uptake strategies in **a**, **d** and **g**. Solid curves indicate the minimum energy cost necessary to achieve a certain normalized CO_2 fixation flux. Shaded regions represent the range of possible performances found by varying HCO_3^- transport rates and LCIB rates. Colour code as in **a**. In **h** and **i**, dashed black curves indicate the optimal PCCM performance of a simplified model that assumes fast intracompartmental diffusion, fast HCO_3^- diffusion across the thylakoid membranes, and fast equilibrium between CO_2 and HCO_3^- catalysed by CAH3 in the thylakoid tubules inside the pyrenoid (Methods).

the intercompartmental pH differential and passive Ci uptake (Fig. 2e,h). PCCM performance with both barriers present closely resembles the impermeable starch sheath case (Supplementary Fig. 8); for simplicity, we omit such a combined model from further discussion.

Optimal passive Ci uptake uses cytosolic CO_2, not HCO_3^-. In addition to the requirement for a diffusion barrier, the efficacy of the pH-driven PCCM depends on the LCIB rate and the $LCIA^C$-mediated chloroplast membrane permeability to HCO_3^- (Fig. 2b,e,h). Depending on LCIB activity, our model suggests two distinct strategies to passively uptake Ci. If LCIB activity is low, CO_2 fixation flux increases with higher $LCIA^C$-mediated permeability to HCO_3^-, which facilitates the diffusion of cytosolic HCO_3^- into the stroma (Fig. 2c,f,i). In contrast, if LCIB activity is high, CO_2 fixation flux is maximized when $LCIA^C$-mediated permeability is low; in this case, a diffusive influx of CO_2 into the chloroplast is rapidly converted by LCIB into HCO_3^-, which becomes trapped and concentrated in the chloroplast. Under this scenario, permeability of the chloroplast membrane to HCO_3^- due to $LCIA^C$ is detrimental, since it allows HCO_3^- converted by LCIB in the stroma to diffuse back out to the cytosol (Fig. 2c,f,i).

Interestingly, the highest CO_2 fixation flux is achieved by passive CO_2 uptake mediated by the carbonic anhydrase activity of LCIB, not by passive HCO_3^- uptake via $LCIA^C$ channels (Fig. 2), even though HCO_3^- is more abundant than CO_2 in the cytosol. The key consideration is that the stroma (at pH 8) is more basic than the cytosol (at pH 7.1, ref. [54]), which allows LCIB to equilibrate passively acquired CO_2 with HCO_3^- to create an even higher HCO_3^- concentration in the stroma than in the cytosol.

The PCCM requires active Ci uptake under very low CO_2. While the passive CO_2 uptake strategy can power the pH-driven PCCM under air-level CO_2 (10 μM cytosolic), its Ci uptake rate is ultimately limited by the diffusion of CO_2 across the chloroplast envelope. Indeed, our simulations show that under very low CO_2 conditions (1 μM cytosolic)[55], a chloroplast using the passive CO_2 uptake strategy can only achieve at most 20% of its maximum CO_2 fixation flux, even in the presence of barriers to Ci diffusion (Fig. 3). Since passive HCO_3^- uptake cannot concentrate more Ci than passive CO_2 uptake (Fig. 2), we hypothesize that active Ci transport is required for an effective PCCM at very low CO_2. To test this idea, we consider a model employing active LCIA HCO_3^- pumps ($LCIA^P$) without LCIB activity (Fig. 3a,d,g). We find that, indeed, HCO_3^- pumping enables

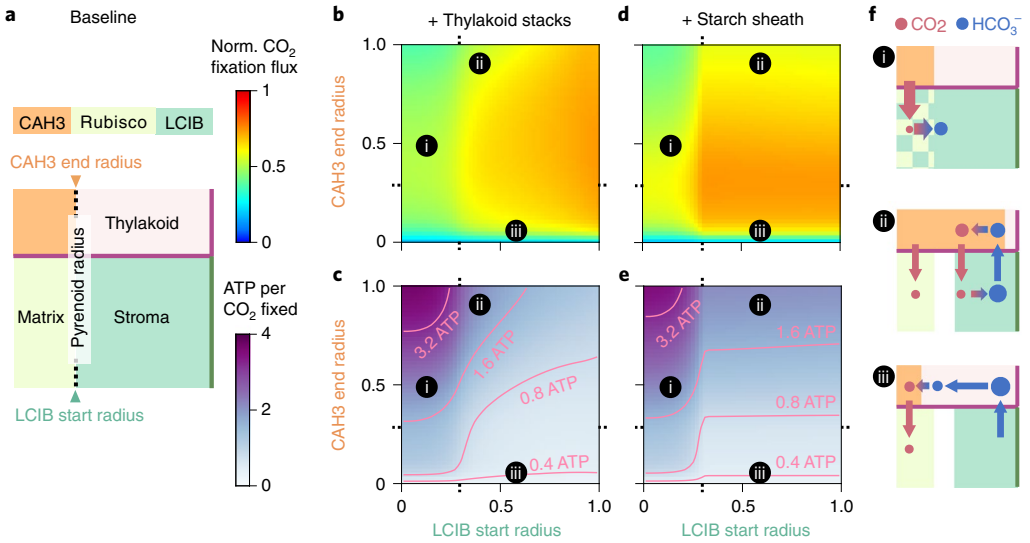

Fig. 4 | Proper localization of carbonic anhydrases enhances PCCM performance. a, Schematics of varying localization of carbonic anhydrases. The CAH3 domain starts in the centre of the intrapyrenoid tubules (radius $r = 0$) and the LCIB domain ends at the chloroplast envelope. Colour code as in Fig. 1d. Orange denotes region occupied by CAH3. **b**–**e**, CAH3 end radius and LCIB start radius are varied in a modelled chloroplast employing the passive CO_2 uptake strategy under air-level CO_2, with thylakoid stacks slowing inorganic carbon diffusion in the stroma (**b**,**c**) or with an impermeable starch sheath (**d**,**e**). Normalized CO_2 fixation flux (**b**,**d**) and ATP spent per CO_2 fixed (**c**,**e**) when the localizations of carbonic anhydrases are varied. **f**, Schematics of inorganic carbon fluxes for the localization patterns (i–iii) indicated in **b**–**e**. Colour code as in **a** and Fig. 1d. Dotted ticks in **b**–**e** denote pyrenoid radius as in **a**. Simulation parameters are the same as in Fig. 1c,d.

saturating CO_2 fixation flux under very low CO_2 conditions (Fig. 3 and Supplementary Fig. 12).

Both passive and active Ci uptake can have low energy cost. According to our model, both passive CO_2 uptake and active HCO_3^- pumping can support an effective PCCM under air-level CO_2. However, the latter directly consumes energy to achieve non-reversible transport. What is the total energy cost of a PCCM that employs active HCO_3^- uptake, and how does this cost compare to that of the passive CO_2 uptake strategy? To answer these questions, we used a nonequilibrium thermodynamics framework to compute the energy cost of different Ci uptake strategies (Supplementary Note II and Fig. 13)[56]. First, a PCCM without diffusion barriers is energetically expensive regardless of the Ci uptake strategies employed (Fig. 3a–c). Second, in the presence of diffusion barriers, we find that the passive CO_2 uptake strategy can achieve similar energy efficiency (~1 ATP cost per CO_2 fixed) to the active HCO_3^- uptake strategy (Fig. 3d–i). Thus, both strategies can achieve high PCCM performance at air-level CO_2; however, active HCO_3^- uptake is necessary to achieve high efficacy under lower CO_2.

The PCCM depends on cytosolic Ci and its chloroplast uptake. How does Ci transport across the cell's plasma membrane impact the feasible Ci uptake strategies at the chloroplast level? To explore this question in our chloroplast-scale model, we assess PCCM performance under a broad range of cytosolic CO_2 and HCO_3^- concentrations (Supplementary Fig. 15). Unsurprisingly, we find that the performance of a particular chloroplast Ci uptake strategy increases with the cytosolic level of its target Ci species. Thus, it is important to replenish cytosolic Ci species taken up by the chloroplast. Moreover, regardless of the makeup of the cytosolic Ci pool, a chloroplast lacking both passive CO_2 uptake and active HCO_3^- uptake fails to achieve high PCCM efficacy, unless the cytosolic CO_2 concentration is 100 µM or higher. Creating such a pool would presumably result in substantial CO_2 leakage across the plasma membrane and thus high energy cost. Therefore, effective mechanisms for Ci uptake from the external environment to the cytosol and from cytosol to the chloroplast are both essential for high PCCM performance.

Carbonic anhydrase localization alters modelled Ci fluxes. So far, we have only considered the carbonic anhydrase localization patterns that are thought to exist in *Chlamydomonas* under air-level CO_2[40,57]. To assess the benefits of such localization, we vary the localization of CAH3 and LCIB while maintaining the total number of molecules of each carbonic anhydrase (Fig. 4a). We find that ectopic carbonic anhydrase localization compromises PCCM performance. First, LCIB mislocalized to the basic pyrenoid matrix (pH 8) converts Rubisco's substrate CO_2 into HCO_3^-, and hence decreases CO_2 fixation (Fig. 4b–f, region i). Second, when CAH3 is distributed in the thylakoids outside the pyrenoid, CO_2 molecules produced by this CAH3 can diffuse directly into the stroma, making them less likely to be concentrated in the pyrenoid and thus decreasing the efficacy of the PCCM (Fig. 4b–f, region ii, and Supplementary Fig. 16). Moreover, CAH3 mislocalization outside the pyrenoid decreases PCCM efficiency as it leads to increased futile cycling of Ci between the stroma and thylakoid, increasing the energetic cost required to maintain the intercompartmental pH differences. Finally, concentrating CAH3 to a small region of thylakoid lumen in the centre of the pyrenoid increases the distance over which HCO_3^- needs to diffuse before it is converted to CO_2, thus lowering the CO_2 production flux by CAH3 (Fig. 4b–f, region iii). All these results hold true both at air-level CO_2 employing passive CO_2 uptake (Fig. 4) and at very low CO_2 employing active HCO_3^- uptake (Supplementary Fig. 17). Thus, our model shows that proper carbonic anhydrase localization is crucial to overall PCCM performance.

Effects of LCIB activity and localization at very low CO_2. When shifted from air levels to very low levels of CO_2 (~1 µM dissolved), *Chlamydomonas* relocalizes LCIB from diffuse throughout the stroma to localized around the pyrenoid periphery[57]. To better understand the value of LCIB localization to the pyrenoid periphery under very low CO_2, we vary both the end radius of stromal LCIB,

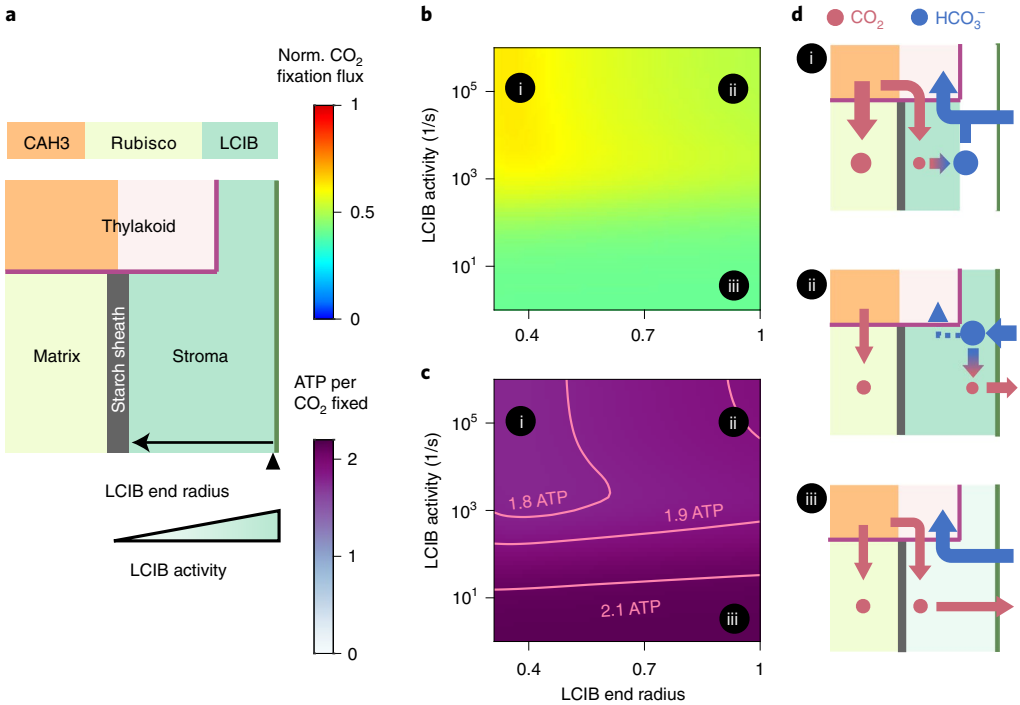

Fig. 5 | Localization of LCIB around the pyrenoid periphery reduces Ci leakage out of the chloroplast. a, Schematics of varying activity and end radius of LCIB in a modelled chloroplast employing an impermeable starch sheath and active HCO_3^- pumping across the chloroplast envelope under very low CO_2. Colour code as in Fig. 4a. The LCIB domain starts at the pyrenoid radius (0.3 on the *x* axis in **b** and **c**). **b**,**c**, Normalized CO_2 fixation flux (**b**) and ATP spent per CO_2 fixed (**c**) when the designated characteristics of LCIB are varied. **d**, Schematics of inorganic carbon fluxes for the LCIB states (i–iii) indicated in **b** and **c**. Colour code as in Fig. 4f. Simulation parameters as in Fig. 4. Active LCIA^P-mediated HCO_3^- pumping is described by the rate $\kappa^{H^-}_{chlor} = 10^{-4}$ m s^{-1} and the reversibility $\gamma = 10^{-4}$. To show a notable variation in normalized CO_2 fixation flux, a model with shortened thylakoid tubules is simulated (Methods). The qualitative results hold true independent of this specific choice.

which defines how far LCIB extends towards the chloroplast envelope, and the total number of LCIB molecules in a model employing a starch sheath barrier and active HCO_3^- uptake (Fig. 5a). Our analysis shows that it is energetically wasteful to allow concentrated CO_2 to leak out of the chloroplast (Supplementary Fig. 13). Consequently, LCIB relocalized near the starch sheath increases energy efficiency by recapturing CO_2 molecules that diffuse out of the matrix and trapping them as HCO_3^- in the chloroplast (Fig. 5b–c, region i). The energy cost is higher without any LCIB for CO_2 recapture (Fig. 5b–c, region iii), or with diffuse stromal LCIB, which allows incoming HCO_3^- to be converted into CO_2 near the chloroplast membrane at which point it can leak back to the cytosol (Fig. 5b–c, region ii, and Supplementary Fig. 19). Our model thus suggests that under very low CO_2 and in the presence of a strong CO_2 diffusion barrier around the pyrenoid, localizing LCIB at the pyrenoid periphery allows for efficient Ci recycling, therefore enhancing PCCM performance.

Intercompartmental pH differences are key to PCCM function. To determine the impact of thylakoid lumen and stromal pH on PCCM function, we vary the pH values of the two compartments (Fig. 6 and Supplementary Fig. 20). We find that regardless of Ci uptake strategy, the modelled PCCM achieves high efficacy only when the thylakoid lumen is much more acidic than the stroma (Fig. 6a,d). Indeed, carbonic anhydrase activity in a low-pH stroma (Fig. 6, region i) or in a high-pH intrapyrenoid tubule lumen (Fig. 6, region ii) would lead to low concentrations of HCO_3^- or CO_2, respectively, in those compartments; both would be detrimental to the PCCM. Interestingly, variation in pH differentially influences the energy efficiency of the PCCM employing passive CO_2 uptake (Fig. 6a–c) and the PCCM employing active HCO_3^- pumping (Fig. 6d–f). Specifically, only the latter shows a dramatically increased energy cost when the stroma has a relatively low pH; in this case, most HCO_3^- pumped into the stroma is converted to CO_2 and is subsequently lost to the cytosol (Fig. 6e,f, regions i and ii). Thus, our results suggest that high PCCM performance requires maintenance of a high-pH stroma and a low-pH thylakoid lumen.

The model recapitulates *Chlamydomonas* PCCM mutant phenotypes. We next explore whether our model can account for the phenotypes of known *Chlamydomonas* PCCM-deficient mutants. We select model parameters to best represent the effect of each mutation, assuming that the *Chlamydomonas* PCCM switches from passive CO_2 uptake under air-level CO_2 to active HCO_3^- uptake under very low CO_2 (Supplementary Figs. 23 and 24). Our simulation results show semi-quantitative agreement with experimental results for all published mutants (Supplementary Table 5) and provide mechanistic explanations for all recorded phenotypes. For example, our model captures that the *lcib* mutant fails to grow in air, presumably due to a defect in passive CO_2 uptake. This phenotype implies that *Chlamydomonas* does not pump HCO_3^- into the chloroplast under air-level CO_2 because a modelled *lcib* mutant employing HCO_3^- pumping has a PCCM effective enough to drive growth in air. Notably, the *lcib* mutant recovers growth under very low CO_2, which we attribute to the activation of an HCO_3^- uptake system under this condition[22,57,58]. Indeed, knockdown of the gene encoding the LCIA HCO_3^- transporters in the *lcib* mutant background results in a dramatic decrease in CO_2 fixation and growth under very low CO_2[57].

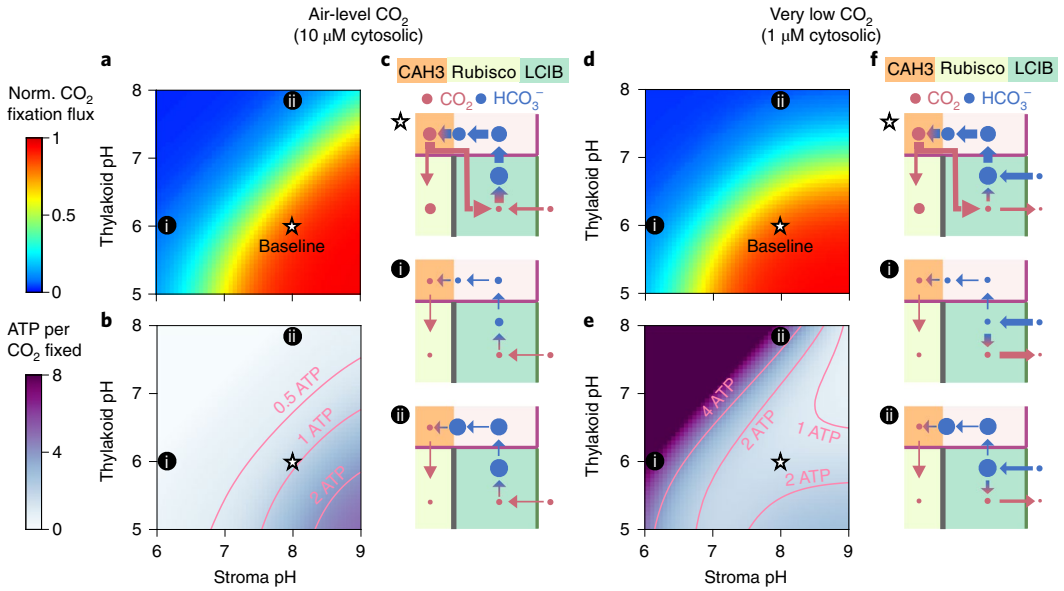

Fig. 6 | High PCCM performance requires low-pH thylakoids and a high-pH stroma. a–f, pH values of the thylakoid lumen and the stroma are varied in a modelled chloroplast with an impermeable starch sheath employing passive CO_2 uptake under air-level CO_2 (**a**–**c**) (10 μM cytosolic; parameters as in Fig. 4d,e) or active HCO_3^- pumping under very low CO_2 (**d**–**f**) (1 μM cytosolic, parameters as in Supplementary Fig. 17c,d). Normalized CO_2 fixation flux (**a**,**d**) and ATP spent per CO_2 fixed (**b**,**e**) as functions of the pH values in the two compartments are shown. **c**,**f**, Schematics of inorganic carbon pools and fluxes for the pH values indicated in **a**, **b**, **d** and **e**. White stars indicate the baseline pH values used in all other simulations.

More broadly, our model recapitulates phenotypes of *Chlamydomonas* mutants lacking the HCO_3^- transporter HLA3 or the CO_2 transporter LCI1 at the plasma membrane. Indeed, knockdown of the gene encoding HLA3 (simulated as a lower level of cytosolic HCO_3^-) leads to a dramatic decrease in PCCM efficacy under very low CO_2, presumably due to reduced HCO_3^- import into the cell and thus into the chloroplast[23,24]. In contrast, the *lci1* single mutant shows a moderate decrease in PCCM efficacy under air-level CO_2, presumably due to a reduced CO_2 influx into the cytosol and thus into the chloroplast, but no effect on the PCCM under very low CO_2, presumably due to the activation of an active HCO_3^- uptake system under this condition[34].

Finally, our model captures the phenotypes of *Chlamydomonas* starch mutants, which survive under both air-level and very low CO_2 conditions presumably because thylakoid stacks can effectively block CO_2 leakage from the pyrenoid in the absence of a starch sheath. The existence of non-starch diffusion barriers, such as the thylakoid stacks, may also help explain why some other pyrenoid-containing algae do not have a starch sheath[59].

Various thylakoid architectures can support PCCM function. The analysis of Ci fluxes in our model supports the long-held view that the thylakoid tubules traversing the pyrenoid in *Chlamydomonas* can deliver stromal HCO_3^- to the pyrenoid, where it can be converted to CO_2 by CAH3[32,60]. However, is a *Chlamydomonas*-like thylakoid architecture necessary to a functional PCCM? Certainly, eukaryotic algae display a variety of thylakoid morphologies, such as multiple non-connecting parallel thylakoid stacks passing through the pyrenoid, a single disc of thylakoids bisecting the pyrenoid matrix, or thylakoid sheets surrounding but not traversing the pyrenoid[61–64]. Our calculations show that different thylakoid morphologies could in principle support the functioning of an effective PCCM, as long as HCO_3^- can diffuse into the low-pH thylakoid lumen and the thylakoid carbonic anhydrase is localized to the pyrenoid-proximal lumen (Supplementary Fig. 25).

An effective PCCM needs Ci uptake, transport and trapping. Our model identifies a minimal PCCM configuration sufficient to effectively concentrate CO_2. Next, we ask: can alternative configurations of the same minimal elements achieve an effective PCCM? We restrict our focus to PCCMs employing passive Ci uptake strategies. We measured the efficacy and energy cost of 216 partial PCCM configurations in air, varying the presence and localization of Rubisco, thylakoid and stromal carbonic anhydrases, HCO_3^- channels on the thylakoid membranes and the chloroplast envelope, and diffusion barriers (Supplementary Fig. 26).

Our results summarize three central modules of an effective pH-driven PCCM (Fig. 7a): (i) a stromal carbonic anhydrase (LCIB) to convert passively acquired CO_2 into HCO_3^-, (ii) a thylakoid membrane HCO_3^- channel (BST) and a luminal carbonic anhydrase (CAH3) that together allow conversion of HCO_3^- to CO_2 near Rubisco, and (iii) a Rubisco condensate surrounded by diffusion barriers. We find that PCCM configurations lacking any one of these modules show a compromised ability to concentrate CO_2 (Fig. 7b). The *Chlamydomonas*-like PCCM configuration is the only configuration possessing all three modules; thus, this configuration is not only sufficient but also necessary to achieve an effective PCCM using the considered minimal elements.

Possible strategies for engineering a PCCM into land plants. Many land plants, including most crop plants, are thought to lack any form of CCM. Our analysis shows that a typical plant chloroplast configuration can only support ~30% of the maximum CO_2 fixation flux through Rubisco (Supplementary Table 6). Engineering a PCCM into crops has emerged as a promising strategy to increase yields through enhanced CO_2 fixation[30,31]. Despite early engineering advances including expressing individual PCCM components[65] and reconstituting a pyrenoid matrix in plants[66], the optimal order of engineering steps needed to establish an effective PCCM in a plant chloroplast remains unknown. Here we leverage our partial PCCM configurations to propose an engineering path that results in monotonic improvement of efficacy and avoids excessive energy costs.

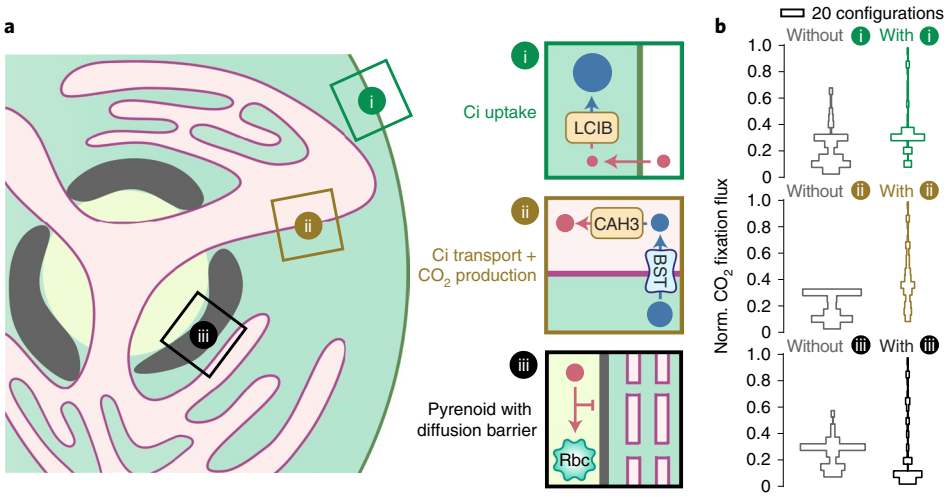

Fig. 7 | An effective PCCM is composed of three essential modules. a, Schematics of the three essential modules with designated functions (same style as in Fig. 1a). In *Chlamydomonas*, LCIB can be used for passive uptake of CO_2, which is then trapped in the stroma as HCO_3^- (module i); BST allows stromal HCO_3^- to diffuse into the thylakoid lumen where CAH3 converts HCO_3^- into CO_2 (module ii); and a starch sheath and thylakoid stacks could act as diffusion barriers to slow CO_2 escape out of the pyrenoid matrix (module iii). **b**, Histograms of normalized CO_2 fixation flux for CCM configurations without (left, grey) or with (right, coloured) the respective module. We tested 216 CCM configurations by varying the presence and/or localization of enzymes, HCO_3^- channels and diffusion barriers in the model (see Supplementary Fig. 26).

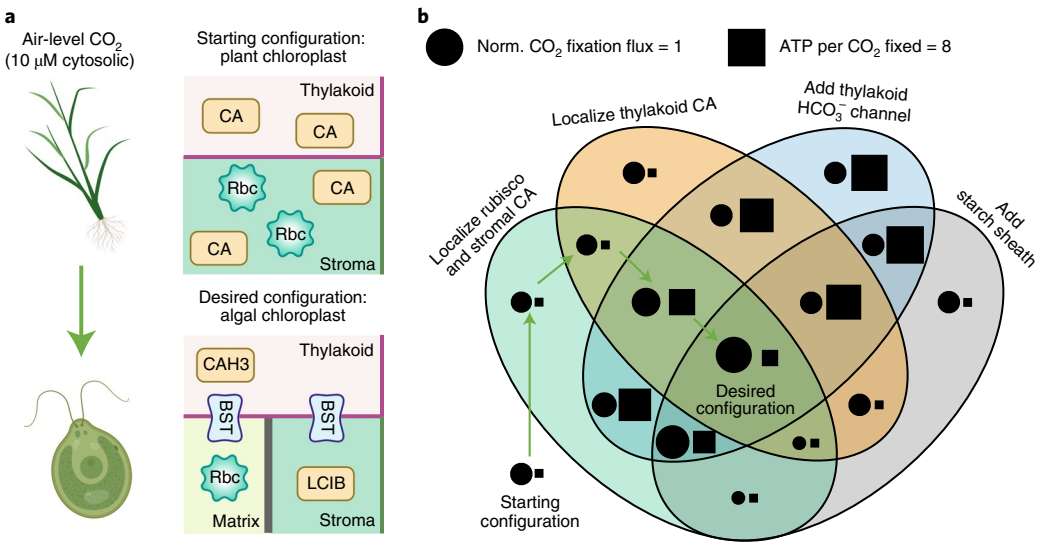

Fig. 8 | Proposed engineering path for installing a minimal PCCM into land plants. a, Top: schematics of the starting configuration representing a typical plant chloroplast that contains diffuse thylakoid carbonic anhydrase, diffuse stromal carbonic anhydrase, and diffuse Rubisco, and lacks HCO_3^- transporters and diffusion barriers. Bottom: the desired configuration representing a *Chlamydomonas* chloroplast that employs the passive CO_2 uptake strategy and a starch sheath (as in Fig. 2g). **b**, Venn diagram showing the normalized CO_2 fixation flux (circle, area in proportion to magnitude) and ATP spent per CO_2 fixed (square, area in proportion to magnitude) of various configurations after implementing the designated changes. Arrows denote the proposed sequential steps to transform the starting configuration into the desired configuration (see text). The starting configuration has a normalized CO_2 fixation flux of 0.31 and negligible ATP cost. All costs below 0.25 ATP per CO_2 fixed are represented by a square of the minimal size.

To the best of our knowledge, the plant chloroplast contains diffuse carbonic anhydrase and diffuse plant Rubisco in the stroma, and lacks HCO_3^- channels and diffusion barriers[67]. We note that plant Rubisco has a lower K_m for CO_2 than *Chlamydomonas* Rubisco; our engineering calculations account for this and employ values from plant Rubisco. Studies have also suggested that native plant carbonic anhydrases are diffuse in the thylakoid lumen[68], which we therefore assume in our modelled plant chloroplast configuration (Fig. 8, starting configuration). This configuration contains only one of the three essential modules for an effective PCCM (Fig. 7a), that is, the passive CO_2 uptake system.

After exploring all possible stepwise paths to install the remaining two modules to achieve the *Chlamydomonas*-like PCCM configuration (Fig. 8, desired configuration), we suggest the following path consisting of four minimal engineering steps (Fig. 8b, arrows). The first step is the localization of plant Rubisco to a

pyrenoid matrix, which we assume would inherently exclude the plant stromal carbonic anhydrase, as the tight packing of Rubisco in the matrix appears to exclude protein complexes greater than ~80 kDa[26,69]. The second step is the localization of the thylakoid carbonic anhydrase to thylakoids that border or traverse the matrix. These first two steps do not yield notable changes to either the efficacy or the efficiency of the PCCM. The next step is to introduce HCO_3^- channels to the thylakoid membranes, which increases the CO_2 fixation flux to ~175% of that of the starting configuration. This step also increases the cost of the PCCM to around 4 ATPs per CO_2 fixed. Such a high-cost step cannot be avoided, and all other possible paths with increasing efficacy at each step have more costly intermediate configurations (Fig. 8b and Supplementary Table 6). Importantly for engineering, the increased CO_2 fixation flux resulting from this step would provide evidence that the installed channels are functional. The final step of the suggested path is to add a starch sheath to block CO_2 leakage from the pyrenoid matrix, which triples the CO_2 fixation flux compared with the starting configuration and reduces the cost to only 1.3 ATPs per CO_2 fixed.

Selecting an alternative implementation order for the four minimal engineering steps leads to decreased performance of the PCCM in intermediate stages. For example, adding HCO_3^- channels on the thylakoid membranes before the stromal and thylakoid carbonic anhydrases are localized (Fig. 8b, blue oval) leads to futile cycling generated by overlapping carbonic anhydrases (Fig. 4, region ii). Additionally, adding a starch sheath before HCO_3^- channels are added to the thylakoids could decrease CO_2 fixation (Fig. 8b, grey oval); without channels, HCO_3^- cannot readily diffuse to the thylakoid carbonic anhydrase to produce CO_2, and the starch sheath impedes diffusion of CO_2 from the stroma to Rubisco. Thus, our suggested path avoids intermediate configurations with decreased efficacy or excessive energy cost.

Discussion

To better understand the composition and function of a minimal PCCM, we developed a multicompartment reaction-diffusion model on the basis of the *Chlamydomonas* PCCM. The model not only accounts for all published *Chlamydomonas* PCCM mutants, but also lays the quantitative and biophysical groundwork for understanding the operating principles of a minimal PCCM. Systematic analysis of the model suggests that keys to an effective and energetically efficient PCCM are barriers preventing CO_2 efflux from the pyrenoid matrix and carbonic anhydrase localizations preventing futile Ci fluxes. The model demonstrates the feasibility of passive CO_2 uptake at air-level CO_2, and shows that at lower external CO_2 levels, an effective PCCM requires active import of HCO_3^-. Both uptake strategies can function at a low energy cost.

While not explicitly considered in our model, protons are produced in Rubisco-catalysed CO_2-fixing reactions[5] and are consumed in CAH3-catalysed HCO_3^--to-CO_2 conversions. Protons must then be depleted in the pyrenoid matrix and replenished in the intrapyrenoid thylakoid lumen to maintain physiological pH values[41,43]. However, our flux-balance analysis shows that the concentrations of free protons are too low to account for the expected proton depletion/replenishment fluxes by free proton diffusion (Supplementary Note VI.D and Fig. 27). Thus, efficient transport of protons must employ alternative mechanisms. One possibility, suggested by recent modelling work[70], is that proton carriers such as RuBP and 3-PGA could be present at millimolar concentrations[71] and hence could enable sufficient flux to transport protons between compartments. Understanding the molecular mechanisms underlying proton transport will be an important topic for future studies.

Another class of CCM is the carboxysome-based CCM (CCCM) employed by cyanobacteria[13]. In the CCCM, HCO_3^- becomes concentrated in the cytosol via active transport[72] and diffuses into carboxysomes—compartments that are typically 100 to 400 nm in diameter, each composed of an icosahedral protein shell enclosing Rubisco[73]. The protein shell is thought to serve as a diffusion barrier, which is necessary for an effective CCCM[46,47]. Whereas the pyrenoid matrix does not appear to have a carbonic anhydrase, the carboxysome matrix contains a carbonic anhydrase that converts HCO_3^- to CO_2 to locally feed Rubisco. Recent studies suggest that protons produced during Rubisco's carboxylation could acidify the carboxysome, which in turn favours the carbonic anhydrase-catalysed production of CO_2[70]. One may ask: what are the benefits of operating a PCCM versus a CCCM? One possibility is that the PCCM uses more complex spatial organization to segregate Rubisco from the thylakoid lumen carbonic anhydrase, which allows the two enzymes to operate at pH values optimal for their respective catalytic functions. Thus, the PCCM may require a smaller Ci pool than the CCCM to produce sufficient CO_2 in the vicinity of Rubisco. Indeed, cyanobacteria appear to accumulate roughly 30 mM intracellular HCO_3^-[74,75], while *Chlamydomonas* creates an internal HCO_3^- pool of only 1 mM[76]. Future experimentation comparing the performance of the PCCM and the CCCM will advance our understanding of the two distinct mechanisms.

The PCCM has the potential to be transferred into crop plants to improve yields. Our model provides a framework to evaluate overall performance, considering both the efficacy and the energetic efficiency of the PCCM (Supplementary Fig. 28), and allows us to propose a favoured order of engineering steps. Moreover, we expect that our model will help engineers narrow down potential challenges by providing a minimal design for a functional PCCM. If the native plant carbonic anhydrases are inactive or absent, it might be favourable to express and localize other carbonic anhydrases with known activities. Additionally, a key step will be to test whether heterologously expressed *Chlamydomonas* BST channels function as HCO_3^- channels and to verify that they do not interfere with native ion channels in plants. We hope that our model provides practical information for engineers aiming to install a minimal PCCM into plants, and that it will serve as a useful quantitative tool to guide basic PCCM studies in the future.

Methods

Reaction-diffusion model. To better understand the operation of the PCCM, we developed a multicompartment reaction-diffusion model on the basis of the postulated mechanism in *Chlamydomonas*. The model takes into account the key PCCM enzymes and transporters and the relevant architecture of the *Chlamydomonas* chloroplast[48]. For simplicity, our model assumes spherical symmetry and considers a spherical chloroplast of radius R_{chlor} in an infinite cytosol. Thus, all model quantities can be expressed as functions of the radial distance r from the centre of the chloroplast (Fig. 1b). The modelled chloroplast consists of three compartments: a spherical pyrenoid matrix of radius R_{pyr} (pH 8) in the centre, surrounded by a stroma (pH 8), with thylakoids (luminal pH 6) traversing both the matrix and stroma (Fig. 1)[41–43]. At steady state, flux-balance equations set the spatially dependent concentrations of CO_2, HCO_3^-, and H_2CO_3 in their respective compartments (indicated by subscripts; see Supplementary Table 2 and Note I):

$$D^C\nabla^2_{thy}C_{thy} - j_{CAH3} - j_{sp} - j^C_{mem}f_s = 0 \quad (1a)$$

$$D^C\nabla^2_{pyr}C_{pyr} - j_{LCIB} - j_{sp} - j_{Rbc} + j^C_{mem}\frac{f_sf_v}{1-f_v} = 0 \quad (1b)$$

$$D^C_{str}\nabla^2_{str}C_{str} - j_{LCIB} - j_{sp} - j_{Rbc} + j^C_{mem}\frac{f_sf_v}{1-f_v} = 0 \quad (1c)$$

$$D^H\nabla^2_{thy}H_{thy} + j_{CAH3} + j_{sp} - j^H_{mem}f_s = 0 \quad (1d)$$

$$D^H\nabla^2_{pyr}H_{pyr} + j_{LCIB} + j_{sp} + j^H_{mem}\frac{f_sf_v}{1-f_v} = 0 \quad (1e)$$

$$D^H_{str}\nabla^2_{str}H_{str} + j_{LCIB} + j_{sp} + j^H_{mem}\frac{f_sf_v}{1-f_v} = 0. \quad (1f)$$

Here, C denotes the concentration of CO_2, and H denotes the combined concentration of HCO_3^- and H_2CO_3, which are assumed to be in fast equilibrium[77]. Thus, their respective concentrations are given by $H^- = \frac{\eta}{1+\eta}H$ for HCO_3^- and $H^0 = \frac{1}{1+\eta}H$ for H_2CO_3, where $\eta = 10^{\mathrm{pH}-\mathrm{pK}a_1}$ is a pH-dependent partition factor and $\mathrm{pK}a_1 = 3.4$ is the negative log of the first acid dissociation constant of H_2CO_3[78]. The first terms in equations (1a–1f) describe the diffusive fluxes of inorganic carbon (Ci) within compartments. D^C and D^H respectively denote the diffusion coefficients of CO_2, and HCO_3^- and H_2CO_3 combined, in aqueous solution. In a model with thylakoid stacks slowing Ci diffusion in the stroma, the effective diffusion coefficients $D_{\mathrm{str}}^{C/H}$ are obtained using a standard homogenization approach (see Supplementary Fig. 5 and Note I.G); $D_{\mathrm{str}}^{C/H} = D^{C/H}$ otherwise. The other flux terms (j_X) in equations (1a–1f) describe enzymatic reactions and intercompartment Ci transport, and the factors f_s and f_v describe the geometry of the thylakoids. Their expressions are provided in subsequent sections.

The boundary conditions at $r = R_{\mathrm{pyr}}$ are determined by the diffusive flux of Ci across the starch sheath at the matrix–stroma interface, that is,

$$-D^C\partial_r C_{\mathrm{pyr}} = -D_{\mathrm{str}}^C\partial_r C_{\mathrm{str}} = \kappa_{\mathrm{starch}}\left(C_{\mathrm{pyr}} - C_{\mathrm{str}}\right) \tag{2a}$$

$$-D^H\partial_r H_{\mathrm{pyr}} = -D_{\mathrm{str}}^H\partial_r H_{\mathrm{str}} = \kappa_{\mathrm{starch}}(H_{\mathrm{pyr}} - H_{\mathrm{str}}), \tag{2b}$$

where ∂_r denotes derivative with respect to r, and the starch sheath is assumed to have the same permeability κ_{starch} for all Ci species. $\kappa_{\mathrm{starch}} \to \infty$ when there is no starch sheath and Ci can diffuse freely out of the matrix. $\kappa_{\mathrm{starch}} = 0$ describes an impermeable starch sheath (see Supplementary Note I.F). Similarly, Ci transport flux across the chloroplast envelope yields the boundary conditions at $r = R_{\mathrm{chlor}}$, that is,

$$D_{\mathrm{str}}^{\mathrm{C}}\partial_r C_{\mathrm{str}} = \kappa^{\mathrm{C}}\left(C_{\mathrm{cyt}} - C_{\mathrm{str}}\right) \tag{3a}$$

$$D_{\mathrm{str}}^{\mathrm{H}}\partial_r H_{\mathrm{str}} = \kappa^{H^0}\left(H_{\mathrm{cyt}}^0 - H_{\mathrm{str}}^0\right) + \kappa^{H^-}\left(H_{\mathrm{cyt}}^- - H_{\mathrm{str}}^-\right) + \kappa_{\mathrm{chlor}}^{H^-}\left(H_{\mathrm{cyt}}^- - \gamma H_{\mathrm{str}}^-\right), \tag{3b}$$

where $\kappa_{\mathrm{chlor}}^{H^-}$ and γ denote the rate and reversibility of inward HCO_3^- transport from the cytosol, representing the action of the uncharacterized chloroplast envelope HCO_3^- transporter LCIA[24,37]; $\gamma = 1$ corresponds to a passive bidirectional channel and $\gamma < 1$ corresponds to an active pump. The external CO_2 conditions are specified by cytosolic CO_2 concentration C_{cyt}. We set $C_{\mathrm{cyt}} = 10\,\mu\mathrm{M}$ for air-level CO_2 conditions, and $C_{\mathrm{cyt}} = 1\,\mu\mathrm{M}$ for very low CO_2 conditions. Unless otherwise specified, all cytosolic Ci species are assumed to be in equilibrium at pH 7.1[54].

Thylakoid geometry. The thylakoid geometry has been characterized by cryo-electron tomography in *Chlamydomonas*[48]. In our model, we account for this geometry by varying the local volume fraction f_v and surface-to-volume ratio f_s of the thylakoids. These fractions describe a tubule meshwork at the centre of the pyrenoid ($r \le R_{\mathrm{mesh}}$), extended radially by N_{tub} cylindrical tubules, each of radius a_{tub} (see Supplementary Note I.C), that is,

$$f_v = \begin{cases} (N_{\mathrm{tub}}a_{\mathrm{tub}}^2)/(4R_{\mathrm{mesh}}^2) & \text{for } r \le R_{\mathrm{mesh}} \\ (N_{\mathrm{tub}}a_{\mathrm{tub}}^2)/(4r^2) & \text{for } r > R_{\mathrm{mesh}} \end{cases}, \text{and } f_s = 2/a_{\mathrm{tub}}. \tag{4}$$

In the baseline model, the thylakoid tubules are assumed to extend to the chloroplast envelope, that is, the outer radius of tubules $R_{\mathrm{tub}} = R_{\mathrm{chlor}}$. In a model with shorter tubules, we choose $R_{\mathrm{tub}} = 0.4\,R_{\mathrm{chlor}}$, and set $f_v = 0$ and $f_s = 0$ for $r > R_{\mathrm{tub}}$. Thus, the Laplace–Beltrami operators in equation (1) are given by $\nabla_{\mathrm{thy}}^2 = r^{-2}f_v^{-1}\partial_r f_v r^2\partial_r$ for the thylakoid tubules, and by $\nabla_{\mathrm{pyr}}^2 = \nabla_{\mathrm{str}}^2 = r^{-2}(1-f_v)^{-1}\partial_r(1-f_v)r^2\partial_r$ for the matrix and stroma.

Enzyme kinetics. The model considers three key *Chlamydomonas* PCCM enzymes, that is, the carbonic anhydrases (CAs) CAH3 and LCIB and the CO_2-fixing enzyme Rubisco. The interconversion between CO_2 and HCO_3^- is catalysed by both CAs and follows reversible Michaelis-Menten kinetics[79]. The rate of CA-mediated CO_2-to-HCO_3^- conversion is given by

$$j_{\mathrm{CA}}(C, H^-) = \frac{(V_{\mathrm{max,CA}}^C/K_{\mathrm{m}}^C)(C - K^{\mathrm{eq}}H^-)}{1 + C/K_{\mathrm{m}}^C + H^-/K_{\mathrm{m}}^{H^-}}\mathcal{L}_{\mathrm{CA}}, \tag{5}$$

where $V_{\mathrm{max,CA}}^C$ denotes the maximum rate of CA, K_{m}^C and $K_{\mathrm{m}}^{H^-}$ respectively denote the half-saturation concentrations for CO_2 and HCO_3^-, and $V_{\mathrm{max,CA}}^C/K_{\mathrm{m}}^C$ denotes the first-order rate constant which we refer to as the 'rate' of the CA (Fig. 2). Finally, $K^{\mathrm{eq}} = 10^{\mathrm{pK}_{\mathrm{eff}}-\mathrm{pH}}$ denotes the equilibrium ratio of CO_2 to HCO_3^-, where the effective pKa is given by $\mathrm{pK}_{\mathrm{eff}} = 6.1$[80,81]. The localization function $\mathcal{L}_{\mathrm{CA}}$ is equal to one for r where CA is present and zero elsewhere. The uncatalysed spontaneous rate of CO_2-to-HCO_3^- conversion, with a first-order rate constant k_{sp}^C, is given by $j_{\mathrm{sp}} = k_{\mathrm{sp}}^C(C - K^{\mathrm{eq}}H^-)$[82]. Note that negative values of j_{CA} and j_{sp} denote fluxes of CO_2-to-HCO_3^- conversion.

The rate of CO_2 fixation catalysed by Rubisco is calculated from

$$j_{\mathrm{Rbc}}(C) = V_{\mathrm{max,Rbc}}^C\frac{C}{K_{\mathrm{m}}^{\mathrm{eff}} + C}\mathcal{L}_{\mathrm{Rbc}}. \tag{6}$$

Here, $V_{\mathrm{max,Rbc}}^C$ denotes the maximum rate, and the effective K_{m} (Rubisco K_{m} in Fig. 1) is given by $K_{\mathrm{m}}^{\mathrm{eff}} = K_{\mathrm{m,Rbc}}^C(1 + O/K_{\mathrm{m,Rbc}}^O)$ to account for competitive inhibition by O_2[83,84], where O denotes the concentration of O_2, and $K_{\mathrm{m,Rbc}}^C$ and $K_{\mathrm{m,Rbc}}^O$ denote the half-saturation substrate concentrations for CO_2 and O_2, respectively. $\mathcal{L}_{\mathrm{Rbc}}$ is equal to one where Rubisco is localized, and zero elsewhere.

In our baseline model, we assume that CAH3 is localized in the thylakoid tubules traversing the pyrenoid[40], LCIB is distributed diffusely in the stroma[57] and Rubisco is localized in the pyrenoid matrix[16]. To explore the effect of enzyme localization, we vary the start and end radii of the enzymes while maintaining a constant number of molecules (Figs. 4 and 5, and Supplementary Note III).

Transport of Ci across thylakoid membranes. The flux of CO_2 diffusing across the thylakoid membrane from the thylakoid lumen to the matrix or stroma is given by

$$j_{\mathrm{mem}}^C = \begin{cases} \kappa^C(C_{\mathrm{thy}} - C_{\mathrm{pyr}}) & \text{for } r \le R_{\mathrm{pyr}} \\ \kappa^C(C_{\mathrm{thy}} - C_{\mathrm{str}}) & \text{for } r > R_{\mathrm{pyr}} \end{cases}, \tag{7}$$

where κ^C denotes the permeability of thylakoid membranes to CO_2. Similarly, the cross-membrane diffusive flux of HCO_3^- and H_2CO_3, j_{mem}^H, is given by

$$j_{\mathrm{mem}}^H = \begin{cases} (\kappa^{H^-} + \kappa_{\mathrm{thy}}^{H^-})(H_{\mathrm{thy}}^- - H_{\mathrm{pyr}}^-) + \kappa^{H^0}(H_{\mathrm{thy}}^0 - H_{\mathrm{pyr}}^0) & \text{for } r \le R_{\mathrm{pyr}} \\ (\kappa^{H^-} + \kappa_{\mathrm{thy}}^{H^-})(H_{\mathrm{thy}}^- - H_{\mathrm{str}}^-) + \kappa^{H^0}(H_{\mathrm{thy}}^0 - H_{\mathrm{str}}^0) & \text{for } r > R_{\mathrm{pyr}} \end{cases}, \tag{8}$$

where κ^{H^-} and κ^{H^0} respectively denote the baseline membrane permeability to HCO_3^- and H_2CO_3, and $\kappa_{\mathrm{thy}}^{H^-}$ denotes the additional permeability of thylakoid membranes to HCO_3^- due to bestrophin-like channels[25]. Note that the final terms of equations (1a) and (1a–1c) differ by a factor of $\frac{f_v}{1-f_v}$ because the cross-membrane fluxes have a larger impact on the concentrations in the thylakoid compartment, which has a smaller volume fraction.

Choice of parameters and numerical simulations. The model parameters were estimated from experiment (see Supplementary Table 2 and references therein), except for the rates of LCIB and CAH3 and the kinetic parameters of the HCO_3^- transporters, which are not known. We performed a systematic scan for these unknown parameters within a range of reasonable values (Fig. 2 and Supplementary Fig. 4). The numerical solutions of equation (1) were obtained by performing simulations using a finite element method. Partial differential equations were converted to their equivalent weak forms, computationally discretized by first-order elements[85] and implemented in the open-source computing platform FEniCS[86]. A parameter sensitivity analysis was performed to verify the robustness of the model results (Supplementary Fig. 30). A convergence study was performed to ensure sufficient spatial discretization (Supplementary Fig. 31).

Energetic cost of the CCM. We computed the energetic cost using the framework of nonequilibrium thermodynamics[56] (see Supplementary Note II.B for details). In brief, the free-energy cost of any nonequilibrium process (reaction, diffusion, or transport) is given by $(j_+ - j_-)\ln(j_+/j_-)$ (in units of thermal energy RT), where j_+ and j_- denote the forward and backward flux, respectively. Summing the energetic cost of nonequilibrium processes described in equation (1), we show that the total energy required to operate the PCCM can be approximated (in units of RT) by

$$\dot{W}_{\mathrm{PCCM}} \approx J_{\mathrm{str}}^{C\to H^-}\ln\frac{K_{\mathrm{thy}}^{\mathrm{eq}}}{K_{\mathrm{str}}^{\mathrm{eq}}} + J_{\mathrm{chlor}}^C\ln\frac{\gamma^{-1}K_{\mathrm{thy}}^{\mathrm{eq}}}{K_{\mathrm{str}}^{\mathrm{eq}}} + J_{\mathrm{Rbc}}\ln\frac{\gamma^{-1}K_{\mathrm{thy}}^{\mathrm{eq}}}{K_{\mathrm{str}}^{\mathrm{eq}}},$$

Here, $J_{\mathrm{str}}^{C\to H^-} = -\int_0^{R_{\mathrm{chlor}}} 4\pi r^2(1-f_v)(j_{\mathrm{LCIB}} + j_{\mathrm{sp}})dr$ integrates the flux of LCIB-mediated and spontaneous conversion from CO_2 to HCO_3^- in the stroma, with $4\pi r^2(1-f_v)dr$ being the geometric factor. $J_{\mathrm{chlor}}^C = 4\pi R_{\mathrm{chlor}}^2\kappa^C(C_{\mathrm{str}}|_{r=R_{\mathrm{chlor}}} - C_{\mathrm{cyt}})$ denotes the flux of CO_2 diffusing from the stroma back out into the cytosol. $J_{\mathrm{Rbc}} = \int_0^{R_{\mathrm{chlor}}} 4\pi r^2(1-f_v)j_{\mathrm{Rbc}}dr$ integrates the flux of CO_2 fixation by Rubisco. The $\ln\gamma^{-1}$ and $\ln(K_{\mathrm{thy}}^{\mathrm{eq}}/K_{\mathrm{str}}^{\mathrm{eq}})$ terms denote the free-energy cost of pumping HCO_3^- across the chloroplast envelope and pumping protons across the thylakoid membranes, respectively. Using ATP hydrolysis energy $|\Delta G_{ATP}| = 51.5\,RT$[87], we compute the equivalent ATP spent per CO_2 fixed as $\dot{W}_{\mathrm{PCCM}}/J_{\mathrm{Rbc}}/|\Delta G_{\mathrm{ATP}}|$.

Well-mixed compartment model. To better understand the biophysical limit of the PCCM, we consider a well-mixed compartment simplification of the full model. Specifically, we assume that (i) the diffusion of Ci is fast in the matrix and stroma, and therefore the concentrations of CO_2 and HCO_3^- are constant across radii in each of the two compartments, taking values denoted by C_{pyr}, C_{str}, H_{pyr}^- and H_{str}^-; (ii) HCO_3^- transport across the thylakoid membranes is fast, and thus the thylakoid tubule concentration of HCO_3^- inside the pyrenoid is equal to H_{pyr}^-,

while the thylakoid tubule concentration outside the pyrenoid is equal to H_{str}^{-}; (iii) HCO_3^- and CO_2 are in equilibrium (catalysed by CAH3) in the thylakoid tubules inside the pyrenoid, and thus the CO_2 concentration therein is given by $C_{thy} = K_{thy}^{eq} H_{pyr}^{-}$; and (iv) the concentration of CO_2 in the thylakoid tubules approaches C_{str} toward the chloroplast envelope. Thus, the flux-balance conditions are described by a set of algebraic equations of 4 variables, C_{pyr}, C_{thy}, C_{str} and H_{str}^{-} (see Supplementary Notes IV and V). The algebraic equations are solved using the Python-based computing library SciPy (version 1.5.0)[88]. The energetic cost of the well-mixed compartment model is computed similarly as above.

Engineering paths. We are interested in how adding and removing individual components affects the overall functioning of the PCCM. We thus measured the efficacy and energy efficiency of 216 PCCM configurations, modulating the presence and localization of enzymes, HCO_3^- channels and diffusion barriers. Each configuration was simulated using the reaction-diffusion model above, with the appropriate parameters for that strategy (Supplementary Fig. 26).

To find all possible engineering paths between these configurations, we considered a graph on which each possible configuration is a node. Nodes were considered to be connected by an undirected edge if they were separated by one engineering step. Thus, by taking steps on the graph, we searched all possible engineering paths, given a start node with poor PCCM performance and a target node with good performance. A single engineering step could be the addition or removal of an enzyme, a channel, or a diffusion barrier, as well as the localization of a single enzyme. The exception is the localization of Rubisco, which we assumed can exclude LCIB from the matrix as it forms a phase-separated condensate[26]. We did not consider strategies employing both a starch sheath and thylakoid stacks as diffusion barriers. We used a custom depth-first search algorithm in MATLAB (R2020a) to identify all shortest engineering paths between a start and a target node.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data generated or analysed during this study are included in this Article and the supplementary tables. The raw datasets have been deposited in the Zenodo repository at https://doi.org/10.5281/zenodo.6406849.

Code availability

Custom simulation codes are available on GitHub at https://github.com/f-chenyi/Chlamydomonas-CCM.

Received: 4 August 2021; Accepted: 11 April 2022;
Published online: 19 May 2022

References

1. Phillips, R. & Milo, R. A feeling for the numbers in biology. *Proc. Natl Acad. Sci. USA* **106**, 21465–21471 (2009).
2. Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**, 237–240 (1998).
3. Bar-On, Y. M. & Milo, R. The global mass and average rate of Rubisco. *Proc. Natl Acad. Sci. USA* **116**, 4738–4743 (2019).
4. Tcherkez, G. G. B., Farquhar, G. D. & Andrews, T. J. Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proc. Natl Acad. Sci. USA* **103**, 7246–7251 (2006).
5. Andersson, I. Catalysis and regulation in Rubisco. *J. Exp. Bot.* **59**, 1555–1568 (2008).
6. Bauwe, H., Hagemann, M. & Fernie, A. R. Photorespiration: players, partners and origin. *Trends Plant Sci.* **15**, 330–336 (2010).
7. Walker, B. J., VanLoocke, A., Bernacchi, C. J. & Ort, D. R. The costs of photorespiration to food production now and in the future. *Annu. Rev. Plant Biol.* **67**, 107–129 (2016).
8. Berry, J. A. & Farquhar, G. D. The CO_2 concentrating function of C_4 photosynthesis. A biochemical model. In *Proc. 4th International Congress on Photosynthesis* (eds Hall, D. O. et al.) 119–131 (The Biochemical Society, 1978).
9. Leegood, R. C. C_4 photosynthesis: principles of CO_2 concentration and prospects for its introduction into C_3 plants. *J. Exp. Bot.* **53**, 581–590 (2002).
10. Osmond, C. B. Crassulacean acid metabolism: a curiosity in context. *Annu. Rev. Plant Physiol.* **29**, 379–414 (1978).
11. Giordano, M., Beardall, J. & Raven, J. A. CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annu. Rev. Plant Biol.* **56**, 99–131 (2005).
12. Reinfelder, J. R. Carbon concentrating mechanisms in eukaryotic marine phytoplankton. *Annu. Rev. Mar. Sci.* **3**, 291–315 (2011).
13. Badger, M. R. & Price, G. D. CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J. Exp. Bot.* **54**, 609–622 (2003).
14. Badger, M. R. et al. The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae. *Can. J. Bot.* **76**, 1052–1071 (1998).
15. Badger, M. R. & Andrews, T. J. in *Progress in Photosynthesis Research* Vol. 3 (ed. Biggins, J.) 601–609 (Springer, 1987).
16. Mackinder, L. C. M. et al. A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proc. Natl Acad. Sci. USA* **113**, 5958–5963 (2016).
17. Meyer, M. T., Whittaker, C. & Griffiths, H. The algal pyrenoid: key unanswered questions. *J. Exp. Bot.* **68**, 3739–3749 (2017).
18. Freeman Rosenzweig, E. S. et al. The eukaryotic CO_2-concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell* **171**, 148–162.e19 (2017).
19. He, S. et al. The structural basis of Rubisco phase separation in the pyrenoid. *Nat. Plants* **6**, 1480–1490 (2020).
20. Spalding, M. H., Spreitzer, R. J. & Ogren, W. L. Carbonic anhydrase-deficient mutant of *Chlamydomonas reinhardii* requires elevated carbon dioxide concentration for photoautotrophic growth. *Plant Physiol.* **73**, 268–272 (1983).
21. Spalding, M. H., Spreitzer, R. J. & Ogren, W. L. Reduced inorganic carbon transport in a CO_2-requiring mutant of *Chlamydomonas reinhardii*. *Plant Physiol.* **73**, 273–276 (1983).
22. Wang, Y. & Spalding, M. H. An inorganic carbon transport system responsible for acclimation specific to air levels of CO_2 in *Chlamydomonas reinhardtii*. *Proc. Natl Acad. Sci. USA* **103**, 10110–10115 (2006).
23. Duanmu, D., Miller, A. R., Horken, K. M., Weeks, D. P. & Spalding, M. H. Knockdown of limiting-CO_2–induced gene *HLA3* decreases HCO_3^- transport and photosynthetic Ci affinity in *Chlamydomonas reinhardtii*. *Proc. Natl Acad. Sci. USA* **106**, 5990–5995 (2009).
24. 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).
25. Mukherjee, A. et al. Thylakoid localized bestrophin-like proteins are essential for the CO_2 concentrating mechanism of *Chlamydomonas reinhardtii*. *Proc. Natl Acad. Sci. USA* **116**, 16915–16920 (2019).
26. Mackinder, L. C. M. et al. A spatial interactome reveals the protein organization of the algal CO_2-concentrating mechanism. *Cell* **171**, 133–147.e14 (2017).
27. 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).
28. Wang, L. et al. Chloroplast-mediated regulation of CO_2-concentrating mechanism by Ca^{2+}-binding protein CAS in the green alga *Chlamydomonas reinhardtii*. *Proc. Natl Acad. Sci. USA* **113**, 12586–12591 (2016).
29. Fukuzawa, H. et al. *Ccm1*, a regulatory gene controlling the induction of a carbon-concentrating mechanism in *Chlamydomonas reinhardtii* by sensing CO_2 availability. *Proc. Natl Acad. Sci. USA* **98**, 5347–5352 (2001).
30. Mackinder, L. C. M. The *Chlamydomonas* CO_2-concentrating mechanism and its potential for engineering photosynthesis in plants. *New Phytol.* **217**, 54–61 (2018).
31. Hennacy, J. H. & Jonikas, M. C. Prospects for engineering biophysical CO_2-concentrating mechanisms into land plants to enhance yields. *Annu. Rev. Plant Biol.* **71**, 461–485 (2020).
32. Raven, J. A. CO_2-concentrating mechanisms: a direct role for thylakoid lumen acidification? *Plant Cell Environ.* **20**, 147–154 (1997).
33. Wang, Y., Stessman, D. J. & Spalding, M. H. The CO_2 concentrating mechanism and photosynthetic carbon assimilation in limiting CO_2: how *Chlamydomonas* works against the gradient. *Plant J.* **82**, 429–448 (2015).
34. Kono, A. & Spalding, M. H. LCI1, a *Chlamydomonas reinhardtii* plasma membrane protein, functions in active CO_2 uptake under low CO_2. *Plant J.* **102**, 1127–1141 (2020).
35. Yamano, T. et al. Light and low-CO_2-dependent LCIB–LCIC complex localization in the chloroplast supports the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant Cell Physiol.* **51**, 1453–1468 (2010).
36. Jin, S. et al. Structural insights into the LCIB protein family reveals a new group of β-carbonic anhydrases. *Proc. Natl Acad. Sci. USA* **113**, 14716–14721 (2016).
37. Miura, K. et al. Expression profiling-based identification of CO_2-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant Physiol.* **135**, 1595–1607 (2004).
38. Karlsson, J. et al. A novel α-type carbonic anhydrase associated with the thylakoid membrane in *Chlamydomonas reinhardtii* is required for growth at ambient CO_2. *EMBO J.* **17**, 1208–1216 (1998).
39. Hanson, D. T., Franklin, L. A., Samuelsson, G. & Badger, M. R. The *Chlamydomonas reinhardtii cia3* mutant lacking a thylakoid lumen-localized carbonic anhydrase is limited by CO_2 supply to Rubisco and not photosystem II function in vivo. *Plant Physiol.* **132**, 2267–2275 (2003).

40. Blanco-Rivero, A., Shutova, T., Román, M. J., Villarejo, A. & Martinez, F. Phosphorylation controls the localization and activation of the lumenal carbonic anhydrase in *Chlamydomonas reinhardtii*. *PLoS ONE* **7**, e49063 (2012).
41. Freeman Rosenzweig, E. S. *Dynamics and Liquid-like Behavior of the Pyrenoid of the Green Alga Chlamydomonas reinhardtii*. PhD thesis, Stanford University (2017).
42. Heldt, H. W., Werdan, K., Milovancev, M. & Geller, G. Alkalization of the chloroplast stroma caused by light-dependent proton flux into the thylakoid space. *Biochim. Biophys. Acta* **314**, 224–241 (1973).
43. Kramer, D. M., Sacksteder, C. A. & Cruz, J. A. How acidic is the lumen? *Photosynth. Res.* **60**, 151–163 (1999).
44. Farquhar, G. D. On the nature of carbon isotope discrimination in C_4 species. *Aust. J. Plant Physiol.* **10**, 205–226 (1983).
45. Fridlyand, L. E. Models of CO_2 concentrating mechanisms in microalgae taking into account cell and chloroplast structure. *Biosystems* **44**, 41–57 (1997).
46. Reinhold, L., Kosloff, R. & Kaplan, A. A model for inorganic carbon fluxes and photosynthesis in cyanobacterial carboxysomes. *Can. J. Bot.* **69**, 984–988 (1991).
47. Mangan, N. M. & Brenner, M. P. Systems analysis of the CO_2 concentrating mechanism in cyanobacteria. *eLife* **3**, e02043 (2014).
48. Engel, B. D. et al. Native architecture of the *Chlamydomonas* chloroplast revealed by in situ cryo-electron tomography. *eLife* **4**, e04889 (2015).
49. Villarejo, A., Martinez, F., Plumed, M., del, P. & Ramazanov, Z. The induction of the CO_2 concentrating mechanism in a starch-less mutant of *Chlamydomonas reinhardtii*. *Physiol. Plant.* **98**, 798–802 (1996).
50. Toyokawa, C., Yamano, T. & Fukuzawa, H. Pyrenoid starch sheath is required for LCIB localization and the CO_2-concentrating mechanism in green algae. *Plant Physiol.* **182**, 1883–1893 (2020).
51. Imberty, A., Buléon, A., Tran, V. & Péerez, S. Recent advances in knowledge of starch structure. *Starch* **43**, 375–384 (1991).
52. Buleon, A. et al. Starches from A to C - *Chlamydomonas reinhardtii* as a model microbial system to investigate the biosynthesis of the plant amylopectin crystal. *Plant Physiol.* **115**, 949–957 (1997).
53. Zeeman, S. C., Kossmann, J. & Smith, A. M. Starch: its metabolism, evolution, and biotechnological modification in plants. *Annu. Rev. Plant Biol.* **61**, 209–234 (2010).
54. Braun, F.-J. & Hegemann, P. Direct measurement of cytosolic calcium and pH in living *Chlamydomonas reinhardtii* cells. *Eur. J. Cell Biol.* **78**, 199–208 (1999).
55. Vance, P. & Spalding, M. H. Growth, photosynthesis, and gene expression in *Chlamydomonas* over a range of CO_2 concentrations and CO_2/O_2 ratios: CO_2 regulates multiple acclimation states. *Can. J. Bot.* **83**, 796–809 (2005).
56. Beard, D. A. & Qian, H. *Chemical Biophysics: Quantitative Analysis of Cellular Systems*. (Cambridge Univ. Press, 2008).
57. Wang, Y. & Spalding, M. H. Acclimation to very low CO_2: contribution of limiting CO_2 inducible proteins, LCIB and LCIA, to inorganic carbon uptake in *Chlamydomonas reinhardtii*. *Plant Physiol.* **166**, 2040–2050 (2014).
58. Duanmu, D., Wang, Y. & Spalding, M. H. Thylakoid lumen carbonic anhydrase (*CAH3*) mutation suppresses air-dier phenotype of *LCIB* mutant in *Chlamydomonas reinhardtii*. *Plant Physiol.* **149**, 929–937 (2009).
59. Meyer, M. T., Goudet, M. M. M. & Griffiths, H. in *Photosynthesis in Algae: Biochemical and Physiological Mechanisms* (eds Larkum, A. W. D. et al.) 179–203 (Springer, 2020).
60. Pronina, N. A. & Semenenko, V. E. Role of the pyrenoid in concentration, generation and fixation of CO_2 in the chloroplast of microalgae. *Sov. Plant Physiol.* **39**, 470–476 (1992).
61. Ford, T. W. A comparative ultrastructural study of *Cyanidium caldarium* and the unicellular red alga *Rhodosorus marinus*. *Ann. Bot.* **53**, 285–294 (1984).
62. Gibbs, S. P. The ultrastructure of the pyrenoids of algae, exclusive of the green algae. *J. Ultrastruct. Res.* **7**, 247–261 (1962).
63. Kusel-Fetzmann, E. & Weidinger, M. Ultrastructure of five *Euglena* species positioned in the subdivision Serpentes. *Protoplasma* **233**, 209–222 (2008).
64. Dodge, J. D. *The Fine Structure of Algal Cells* (Academic Press, 1973).
65. 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).
66. Atkinson, N., Mao, Y., Chan, K. X. & McCormick, A. J. Condensation of Rubisco into a proto-pyrenoid in higher plant chloroplasts. *Nat. Commun.* **11**, 6303 (2020).
67. Poschenrieder, C. et al. Transport and use of bicarbonate in plants: current knowledge and challenges ahead. *Int. J. Mol. Sci.* **19**, 1352 (2018).
68. Ignatova, L., Rudenko, N., Zhurikova, E., Borisova-Mubarakshina, M. & Ivanov, B. Carbonic anhydrases in photosynthesizing cells of C3 higher plants. *Metabolites* **9**, 73 (2019).
69. DiMario, R. J., Clayton, H., Mukherjee, A., Ludwig, M. & Moroney, J. V. Plant carbonic anhydrases: structures, locations, evolution, and physiological roles. *Mol. Plant* **10**, 30–46 (2017).
70. Long, B. M., Förster, B., Pulsford, S. B., Price, G. D. & Badger, M. R. Rubisco proton production drives the elevation of CO_2 within condensates and carboxysomes. *Proc. Natl Acad. Sci. USA* **118**, e2014406118 (2021).
71. Küken, A. et al. Effects of microcompartmentation on flux distribution and metabolic pools in *Chlamydomonas reinhardtii* chloroplasts. *eLife* **7**, e37960 (2018).
72. Price, G. D. & Badger, M. R. Expression of human carbonic anhydrase in the cyanobacterium *Synechococcus PCC7942* creates a high CO_2-requiring phenotype: evidence for a central role for carboxysomes in the CO_2 concentrating mechanism. *Plant Physiol.* **91**, 505–513 (1989).
73. Rae, B. D., Long, B. M., Badger, M. R. & Price, G. D. Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria. *Microbiol. Mol. Biol. Rev.* **77**, 357–379 (2013).
74. Sültemeyer, D., Price, G. D., Yu, J.-W. & Badger, M. R. Characterisation of carbon dioxide and bicarbonate transport during steady-state photosynthesis in the marine cyanobacterium *Synechococcus* strain PCC7002. *Planta* **197**, 597–607 (1995).
75. Woodger, F. J., Badger, M. R. & Price, G. D. Sensing of inorganic carbon limitation in *Synechococcus* PCC7942 is correlated with the size of the internal inorganic carbon pool and involves oxygen. *Plant Physiol.* **139**, 1959–1969 (2005).
76. Badger, M. R., Kaplan, A. & Berry, J. A. Internal inorganic carbon pool of *Chlamydomonas reinhardtii*: evidence for a carbon dioxide-concentrating mechanism. *Plant Physiol.* **66**, 407–413 (1980).
77. Gibbons, B. H. & Edsall, J. T. Rate of hydration of carbon dioxide and dehydration of carbonic acid at 25 degrees. *J. Biol. Chem.* **238**, 3502–3507 (1963).
78. Adamczyk, K., Prémont-Schwarz, M., Pines, D., Pines, E. & Nibbering, E. T. J. Real-time observation of carbonic acid formation in aqueous solution. *Science* **326**, 1690–1694 (2009).
79. Lindskog, S. & Coleman, J. E. The catalytic mechanism of carbonic anhydrase. *Proc. Natl Acad. Sci. USA* **70**, 2505–2508 (1973).
80. Millero, F. J. Thermodynamics of the carbon dioxide system in the oceans. *Geochim. Cosmochim. Acta* **59**, 661–677 (1995).
81. Mangan, N. M., Flamholz, A., Hood, R. D., Milo, R. & Savage, D. F. pH determines the energetic efficiency of the cyanobacterial CO_2 concentrating mechanism. *Proc. Natl Acad. Sci. USA* **113**, E5354–E5362 (2016).
82. Kern, D. M. The hydration of carbon dioxide. *J. Chem. Educ.* **37**, 14 (1960).
83. von Caemmerer, S., Evans, J. R., Hudson, G. S. & Andrews, T. J. The kinetics of ribulose-1,5-bisphosphate carboxylase/oxygenase in vivo inferred from measurements of photosynthesis in leaves of transgenic tobacco. *Planta* **195**, 88–97 (1994).
84. Buchanan, B. B., Gruissem, W. & Jones, R. L. *Biochemistry and Molecular Biology of Plants* (John Wiley & Sons, 2015).
85. Langtangen, H. P. & Mardal, K.-A. *Introduction to Numerical Methods for Variational Problems* (Springer Nature, 2019).
86. Alnæs, M. et al. The FEniCS project version 1.5. *Arch. Numer. Softw.* **3**, 9–23 (2015).
87. Nelson, D. L., Lehninger, A. L. & Cox, M. M. *Lehninger Principles of Biochemistry* (Macmillan, 2008).
88. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

Acknowledgements

We thank members of the Jonikas and Wingreen groups for insightful discussions. This work was supported by the National Institutes of Health through grant 5R01GM140032-02 (N.S.W. and M.C.J.); the National Science Foundation through grant MCB-1935444 (M.C.J.) and through the Centre for the Physics of Biological Function PHY-1734030 (N.S.W.); and the Simons Foundation and Howard Hughes Medical Institute grant 55108535 (M.C.J.). M.C.J. is a Howard Hughes Medical Institute Investigator. Schematics for a subset of figures were created with BioRender.com.

Author contributions

C.F., A.T.W., N.M.M., N.S.W. and M.C.J. designed research; C.F., A.T.W., N.M.M. and N.S.W. performed modelling; C.F. and A.T.W. performed simulation; C.F., A.T.W., N.M.M., N.S.W. and M.C.J. analysed data; and C.F., A.T.W., N.M.M., N.S.W. and M.C.J. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41477-022-01153-7.

Correspondence and requests for materials should be addressed to Niall M. Mangan, Ned S. Wingreen or Martin C. Jonikas.

Peer review information *Nature Plants* thanks Yusuke Matsuda and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Martin Jonikas

Last updated by author(s): Apr 5, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☒	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Full details of custom simulation code are published and publicly available at https://github.com/f-chenyi/Chlamydomonas-CCM. An open-source computing platform FEniCS (v.2019.1.0) was used to perform simulations. An open-source Python library SciPy (version 1.5.0) was used to solve algebraic equations.
Data analysis	MATLAB (R2020a) was used for data analysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data generated or analysed during this study are included in this article, the extended data, and the supplementary tables. The raw datasets have been deposited in the Zenodo repository, https://doi.org/10.5281/zenodo.6406849.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	This does not apply, as this study does not have any experimental groups.
Data exclusions	No data was excluded.
Replication	This does not apply, as this study does not have any experimental groups. The findings are reproducible using the simulation code provided in the Code Availability section.
Randomization	This does not apply, as this study does not have any experimental groups.
Blinding	This does not apply, as this study does not have any experimental groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging