Rubisco proton production can drive the elevation of CO_2 within condensates and carboxysomes

Authors:	**ORCID**	**ResearcherID**
Benedict M. Long [a,b,1]	0000-0002-4616-2967	B-3799-2008
Britta Förster [b]	0000-0002-5231-0954	O-7908-2015
Sacha B. Pulsford [b]	0000-0001-9179-9214	
G. Dean Price [a,b]	0000-0001-5906-4912	C-9505-2009
Murray R. Badger [a,b]	0000-0001-9428-2579	

Address:

[a] Realizing Increased Photosynthetic Efficiency (RIPE); and [b] Australian Research Council Centre of Excellence for Translational Photosynthesis

Plant Science Division
Research School of Biology
The Australian National University
134 Linnaeus Way
Acton ACT 2601
AUSTRALIA

Correspondence to: ben.long@anu.edu.au

ABSTRACT

Membraneless organelles containing the enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) are a common feature of organisms utilizing CO_2 concentrating mechanisms (CCMs) to enhance photosynthetic carbon acquisition. In cyanobacteria and proteobacteria, the Rubisco condensate is encapsulated in a proteinaceous shell, collectively termed a carboxysome, while some algae and hornworts have evolved Rubisco condensates known as pyrenoids. In both cases, CO_2 fixation is enhanced compared with the free enzyme. Previous mathematical models have attributed the improved function of carboxysomes to the generation of elevated CO_2 within the organelle via a co-localized carbonic anhydrase (CA), and inwardly diffusing HCO_3^- which has accumulated in the cytoplasm via dedicated transporters. Here we present a novel concept in which we consider the net of two protons produced in every Rubisco carboxylase reaction. We evaluate this in a reaction-diffusion, compartment model to investigate functional advantages these protons may provide Rubisco condensates and carboxysomes, prior to the evolution of HCO_3^- accumulation. Our model highlights that diffusional resistance to reaction species within a condensate allows Rubisco-derived protons to drive the conversion of HCO_3^- to CO_2 via co-localized CA, enhancing both condensate [CO_2] and Rubisco rate. Protonation of Rubisco substrate (RuBP) and product (PGA) plays an important role in modulating internal pH and CO_2 generation. Application of the model to putative evolutionary ancestors, prior to contemporary cellular HCO_3^- accumulation, revealed photosynthetic enhancements along a logical sequence of advancements, via Rubisco condensation, to fully-formed carboxysomes. Our model suggests that evolution of Rubisco condensation could be favored under low CO_2 and low light environments.

INTRODUCTION

Carbon dioxide (CO_2) fixation into the biosphere has been primarily dependent upon action of the enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) over geological timescales. Rubisco is distinguished by the competitive inhibition of its carboxylation activity by the alternative substrate molecular oxygen (O_2), leading to loss of CO_2 and metabolic energy via a photorespiratory pathway in most phototrophs (1). Almost certainly the most abundant enzyme on the planet (2), Rubisco's competing catalytic activities have required evolution of the enzyme, and/or its associated machinery, to maintain capture of sufficient carbon into organic molecules to drive life on Earth. In concert with geological weathering, the evolution of oxygenic photosynthesis ≈2.4 billion years ago has transformed the atmosphere from one rich in CO_2 and low in O_2 to one in which the relative abundances of these gases has overturned (3). Under these conditions, the Rubisco oxygenation reaction has increased, to the detriment of CO_2 capture. This catalytic paradox has led to different adaptive solutions to ensure effective rates of photosynthetic CO_2 fixation including; the evolution of the kinetic properties of the enzyme (4), increases in Rubisco abundance in the leaves of many terrestrial C_3 plants (5) and, the evolution of diverse and complex CO_2 concentrating mechanisms (CCMs) in many cyanobacteria, algae, and more recently hornworts, CAM, and C_4 plants (6, 7).

A defining characteristic of contemporary cyanobacteria is the encapsulation of their Rubisco enzymes within specialized, protein-encased micro-compartments called carboxysomes (8). These microbodies are central to the cyanobacterial CCM, in which cellular bicarbonate (HCO_3^-) is elevated by a combination of membrane-associated HCO_3^- pumps and CO_2-to-HCO_3^- converting complexes (9-11), to drive CO_2 production within the carboxysome by an internal carbonic anhydrase (CA; 12, 13). This process results in enhanced CO_2-fixation, with a concomitant decrease in oxygenation, and is a proposed evolutionary adaptation to a low CO_2 atmosphere (14, 15).

An analogous CCM operates in many algal and hornwort species which contain chloroplastic Rubisco condensates called pyrenoids (16, 17). Pyrenoids are liquid-liquid phase separated Rubisco aggregates which lack the protein shell of a carboxysome (18). These CCMs accumulate HCO_3^- and convert it to CO_2 within the pyrenoid to maximize CO_2-fixation. Common to cyanobacterial and algal systems is the presence of unique Rubisco-binding proteins, enabling condensation of Rubisco from the bulk cytoplasm (18-25). Condensation of proteins to form aggregates within the cell is an increasingly recognized as a means by which cellular processes can be segregated and organized, across a broad range of biological systems (26-28). The commonality of pyrenoid and carboxysome function (29) despite their disparate evolutionary histories (6), suggest a convergence of function driven by Rubisco condensation. In addition, dependency of functional CCMs on their pyrenoids or carboxysomes (30, 31) has led to the speculation that the evolution of Rubisco organization into membraneless organelles likely preceded systems which enabled elevated cellular HCO_3^- (14), raising the possibility that Rubisco condensation and encapsulation may have been the first steps in modern aquatic CCM evolution.

We consider here that, in a primordial model system without active HCO_3^- accumulation, co-condensation of Rubisco and CA enzymes is beneficial for the elevation of internal CO_2 because Rubisco carboxylation produces a net of two protons for every reaction turnover (*SI Appendix*, Fig. S1; 32, 33). These protons can be used within the condensate to convert HCO_3^- to CO_2, with pH lowered and CO_2 elevated as a result of restricted outward diffusion due to the high concentration of protein in the condensate and surrounding cell matrix acting as a barrier to diffusion. We propose that proton release within a primordial Rubisco condensate enabled the evolution of carboxysomes with enhanced carboxylation rates, prior to advancements which enabled cellular HCO_3^- accumulation.

RESULTS

The modelling of free Rubisco and Rubisco compartments

To demonstrate the feasibility of our proposal we initially consider a model of free Rubisco, a Rubisco condensate, and a carboxysome based on a set of compartmentalized reactions described in Fig. 1 and associated tables of parameters (Table 1, Table 2, *SI Appendix*, *SI Methods*). We present data for the tobacco Rubisco enzyme as an exemplar, noting that evaluations of other Rubisco enzymes in the model (Table 2) provide comparative outcomes. We assume a system with fixed external inorganic carbon (C_i; HCO_3^- and CO_2) supply in the absence of a functional CCM, simulating a primordial evolutionary state prior to the development of HCO_3^- accumulation in unicellular photosynthetic organisms (14).

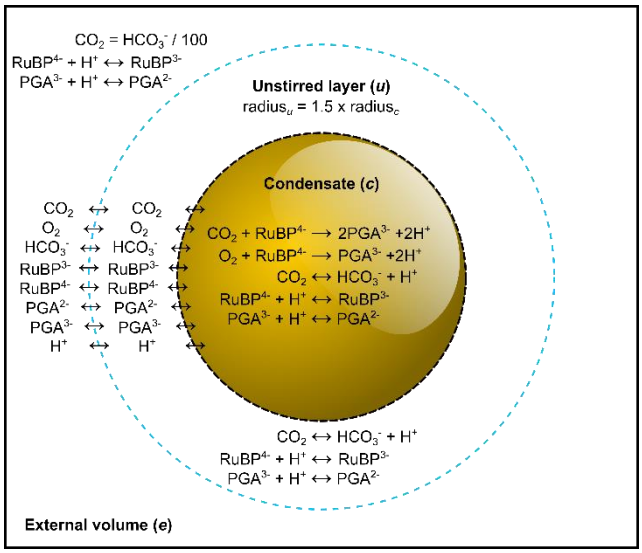

Fig. 1. Rubisco compartment model

A visual description of the compartment model used in this study. The model consists of three reaction compartments. The external compartment (*e*) is analogous to a static cellular cytoplasm in which we set the concentration of inorganic carbon (C_i) species (CO_2 and HCO_3^-), along with RuBP and PGA which can undergo reversible reactions with protons (H^+). Interconversion of C_i species in the unstirred (*u*) and condensate (*c*) compartments is catalyzed by carbonic anhydrase, whereas RuBP and PGA protonation/deprotonation is determined by the rate of conversion at physiological pH given *pKa* values of relevant functional groups (*SI Appendix*, Fig. S1). The central compartment of the model is a Rubisco condensate in which Rubisco carboxylation and oxygenation reactions occur, along with RuBP/PGA protonation and CA reactions. In modelling scenarios, we modify external CA by modulating its function in the unstirred layer. The diffusion of all reaction species between each compartment can be set in the model to simulate either a free Rubisco enzyme, a Rubisco condensate, or a carboxysome as described in Table 1. Model parameterization is described in detail in *SI Appendix, SI Methods*.

Our model consists of three nested compartments with a specialized Rubisco compartment (which can be described as a Rubisco condensate or carboxysome by modifying the compartment boundary permeabilities) at the center (Fig. 1). This compartment is surrounded by an unstirred boundary layer which we assume has diffusive resistance to substrate movement, and is bounded by an external compartment, at pH 8.0, which supplies reaction substrates. We contain Rubisco reactions within the central compartment but allow the protonation and deprotonation of reaction species (ribulose-1,5-bisphosphate [RuBP] and phosphoglycerate [PGA]) to occur in all compartments. We include the competing Rubisco substrates, O_2 and CO_2, and enable the latter to be interconverted with the more abundant HCO_3^- species through pH control and the interaction of carbonic anhydrase (CA), whose position in the model we manipulate. Specific details of the model and its parameterization are provided in Table 1, Methods, and *SI Appendix, SI Methods*.

Previous models consider the function of carboxysomes, for example, within cells capable of active accumulation of HCO_3^- in chemical disequilibrium with CO_2, and apply diffusional resistances to Rubisco reactants and products within a modelled reaction compartment (9, 34-42). We also apply diffusional resistances to all substrates in our model, but consider cytoplasmic CO_2 and HCO_3^- supply to be in chemical equilibrium, as would occur in the absence of an active CCM, in order to address any beneficial role of Rubisco compartmentation alone. The novel aspects of this model are; a chemical equilibrium of CO_2 and HCO_3^- in the compartment surrounding Rubisco, the inclusion of proton production by the Rubisco carboxylation and oxygenation reactions (and their equilibration across the carboxysome shell by diffusion) and, proton movement via protonated ribulose-1,5-bishposphate (RuBP) and phosphoglycerate (PGA) species. Application of the model to existing experimental data (*SI Appendix,* Fig. S2) provides a good estimation of the differential function of both the free Rubisco enzyme and carboxysomes isolated from the cyanobacterium *Cyanobium* (43), thus providing confidence in the model.

Carboxysome and condensate proton permeability

An important assumption in our model is that there is some resistance to substrate movement across Rubisco compartment boundaries, including protons. The diffusion of protons across the carboxysome envelope has been considered previously (44). but within the context that pH stabilization is entirely dependent upon free diffusion through the shell, and in the absence of Rubisco activity which could lead to internally produced protons. In that study, pH-dependent fluorescent protein inside the carboxysome responded within millisecond time-scales to changes in the external pH, resulting in the conclusion that protons entered or exited the carboxysome freely. However, this result is also consistent with some level of diffusional resistance to protons, since considerable restriction to proton permeability can yield internal pH equilibration within even faster time-frames (*SI Appendix,* Fig. S3). Indeed, these previous findings have been shown to be consistent with a steady-state ΔpH across the carboxysome shell, where the relative rates of internal proton production and leakage across the shell can maintain an acidic interior (38). We therefore assume permeabilities to protons which are consistent with existing data, yet enable some restriction on proton movement. Molecular simulations suggest that pores

in the carboxysome shell favor negatively charged ions such as HCO_3^-, RuBP and PGA (45), and it is unlikely that the H_3O^+ ion will easily traverse the protein shell. For the diffusion of H_3O^+ in water, it is considered to have a higher diffusion rate than other ions in solution due to its participation in a proton wire system in collaboration with water (46).

Table 1. Typical initial values used in the COPASI biochemical compartment model in this study

Parameter		Model compartment		
		External (*e*)	Unstirred layer (*u*)	Condensate (*c*)
Rubisco sites (mol/m^3) [a]		absent	absent	10
Substrate permeabilities (m/s) [b]	Free Rubisco scenario			
	all substrates	n/a	1	1
	Condensate scenario			
	H^+	n/a	1×10^{-2}	1×10^{-2}
	CO_2, HCO_3^-, RuBP and PGA	n/a	1×10^{-4}	1×10^{-4}
	Carboxysome scenario			
	H^+	n/a	1×10^{-2}	1×10^{-4}
	CO_2, HCO_3^-, RuBP and PGA	n/a	1×10^{-4}	1×10^{-6}
Carbonic anhydrase catalysis factor [c]	Internal CA	constant ratio [d]	1	1×10^5
	Unstirred layer CA	constant ratio [d]	1×10^5	1
	Internal and unstirred layer CA	constant ratio [d]	1×10^5	1×10^5
Carbonic anhydrase rate constants (1/s) [e]	$CO_2 \rightarrow HCO_3^- + H^+$	n/a	0.05	0.05
	$HCO_3^- + H^+ \rightarrow CO_2$	n/a	100	100
Compartment volume (m^3) [f]		1	9.95×10^{-18}	4.19×10^{-18}
Species concentrations (mol/m^3)	HCO_3^-	0.01 – 25 [g]	[j]	[j]
	CO_2 (Rubisco substrate)	$0.001 \times [HCO_3^-]$	[j]	[j]
	O_2 (Rubisco substrate)	0.25 – 0.36 [h]	[j]	[j]
	H^+ (Rubisco product)	1×10^{-5} (pH 8.0) [i]	[j]	[j]
	$RuBP^{4-}$ (Rubisco substrate)	1×10^{-3} – 10 [g]	[j]	[j]
	$RuBP^{3-}$	[j]	[j]	[j]
	PGA^{3-} (Rubisco product)	[j]	[j]	[j]
	PGA^{2-}	[j]	[j]	[j]

[a] – Rubisco active site concentration is set at 10 mol/m^3 as an upper bound of likely concentrations which would allow for movement of holoenzymes within the compartment, and both small molecule and activation chaperone passage (see *SI Appendix, SI Methods*).
[b] – Permeabilities of the unstirred layer and condensate compartments to H^+, CO_2, HCO_3^-, RuBP and PGA are varied to simulate the free enzyme, condensate or carboxysome scenarios. A detailed description of the specific values utilized in the model is provided in *SI Appendix, SI Methods*.
[c] – Carbonic anhydrase (CA) catalysis factor > 1 indicates presence of CA in that compartment. Here we use the value of 1×10^5 as used by Reinhold, Kosloff and Kaplan (36).
[d] – A constant HCO_3^-:CO_2 ratio of 100:1 was used in the external compartment, assuming negligible effects of a single Rubisco compartment on the bulk external C_i species, regardless of CA level.
[e] – CA rate constants were kept the same in internal and unstirred layer compartments. Values of 0.05 and 100 (1/s) allow for the attainment of a HCO_3^-:CO_2 ratio of 100:1 (the approximate proportion, assuming the uncatalyzed interconversion of each species, at chemical equilibrium when pH is 8.0).
[f] – External compartment volume is 1 m^3 and the volumes of the unstirred and condensate are determined from the condensate radius, here set to 1×10^{-6} m for most scenarios, or to 5×10^{-8} when modelling small carboxysomes (*SI Appendix, SI Methods*).
[g] – The ranges of HCO_3^- and RuBP concentrations used here are typical ranges for these substrates in cyanobacterial and microalgal cells (41, 61, 62).
[h] – Concentrations of O_2 in water at 25 °C under current atmosphere (20% O_2 *v/v*) and atmospheres where CCMs may have evolved (30% *v/v*; 7, 48).
[i] – A pH of 8.0 in the external compartment approximates that of a typical cyanobacterial cell (38) and is fixed in modelled scenarios.
[j] – Values determined as output from the model.
n/a – not applicable

For modelling purposes here, we have assumed that negatively charged ions have a permeability of 10^{-6} m/s across the carboxysome shell while H_3O^+ has a higher value of 10^{-4} m/s (see *SI Appendix, SI Methods*). The

data in Fig. 2 show that proton permeability values greater than ≈10^{-3} m/s for the carboxysome shell leads to low carboxysome [CO_2] and lower Rubisco carboxylation turnover (under sub-saturating substrate supply).

Unlike the carboxysome, Rubisco condensate proton permeability does not appear to affect Rubisco carboxylation in our model under sub-saturating substrate conditions (Fig. 2). However, the condensate [CO_2] does appear to correlate with modelled changes in internal pH, suggesting a role for protons in determining condensate [CO_2] and carboxylation rates (Fig. S4). In the case of a condensate, $RuBP^{3-}$ is able to carry protons from outside to inside, and therefore provides protons required to convert HCO_3^- to CO_2. This can be observed within the model by varying RuBP permeability, with values above 10^{-6} m/s leading to increased compartment [CO_2], and enhanced carboxylation (Fig. 2). This value is consistent with application of the model to experimental data for carboxysomes (Fig. S2). RuBP permeabilities below 10^{-6} m/s also leads to rate-limiting concentrations of RuBP in all compartment types under sub-saturating substrate supply (Fig. 2). Variation of condensate permeabilities to protons and RuBP shows that fluxes of reaction species across the condensate boundary are permeability-dependent (Fig. S5).

Protons derived from Rubisco reactions influence both condensate and carboxysome function

The influence of Rubisco proton production on the response of carboxylation rate and Rubisco compartment [CO_2] to external substrate supply (RuBP and HCO_3^-) is considered in Fig. 3. For the purpose of demonstrating the role of protons, we assess model responses under sub-saturating substrate conditions where we observe greatest proton responses within the model (*SI Appendix,* Fig. S4), and account for differential changes in the 'apparent' $K_M RuBP$ arising from an assumed decreased permeability to this substrate in condensates and carboxysomes (Fig. 3A; *SI Appendix,* Fig. S6).

If we eliminate carboxylation-derived proton production within a Rubisco compartment in the model (zero protons; Fig. 3A) and assume diffusional limitations to proton movement, then proton-driven, CA-dependent CO_2 production within either a condensate or carboxysome becomes limited by the influx of protons from the external environment. With increasing compartment proton production per Rubisco carboxylation reaction (one and two protons; Fig. 3A), the lower pH (and increased [H^+] for CA-driven HCO_3^- dehydration) enhances CO_2 concentration within a condensate, and even more so within a carboxysome, increasing Rubisco CO_2 fixation rates. Levels of CO_2 are further enhanced if more protons are able to be produced per Rubisco turnover (e.g. three protons in Fig. 3A). In contrast to a condensate or carboxysome, the free Rubisco enzyme is unaffected by proton production due to the absence of a compartment with associated diffusional limitations.

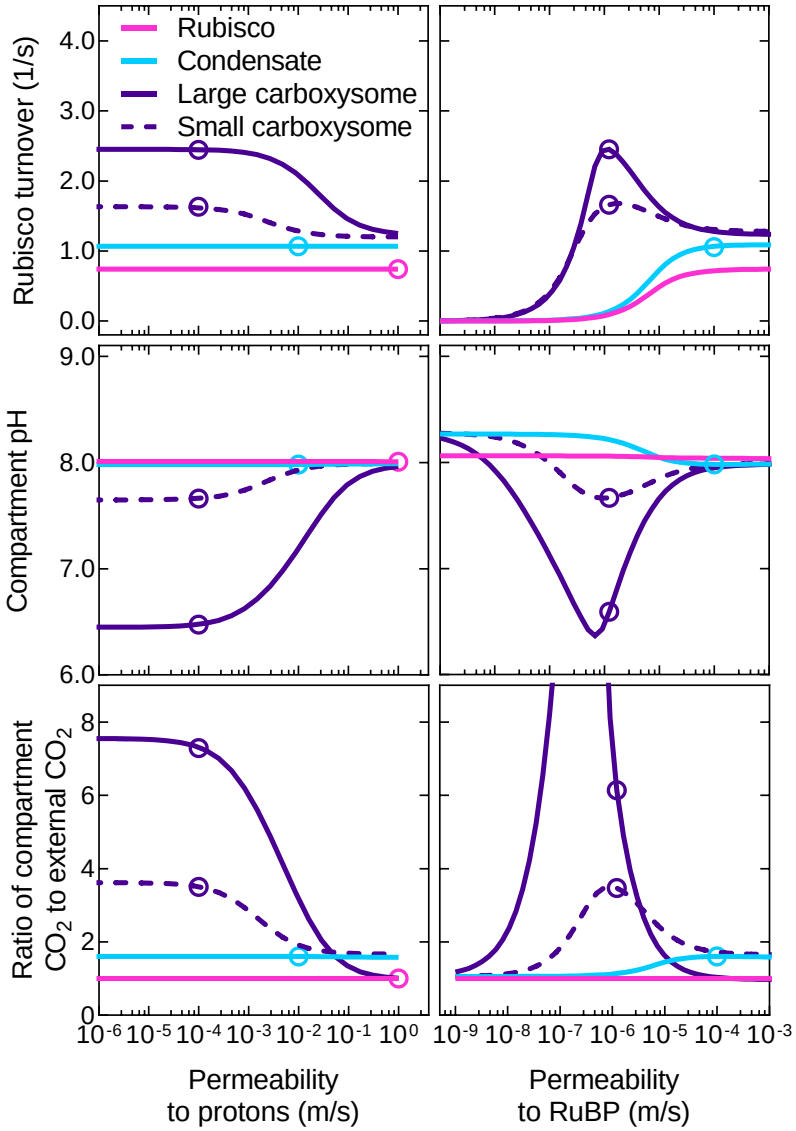

Fig. 2. Carboxysome and condensate function are dependent on proton and RuBP permeabilities

Rubisco carboxylation turnover (top panels), compartment pH (middle panels), and the ratio of Rubisco compartment CO_2 to external CO_2 (bottom panels) are dependent upon the permeability of the compartment to protons (left panels) and RuBP (right panels). Shown are modelled responses for free Rubisco (*pink lines*), a Rubisco condensate (*blue lines*), a large (1×10^{-6} m radius) carboxysome (*purple lines*), and a small (5×10^{-8} m radius) carboxysome (*purple dashed lines*) at sub-saturating substrate concentrations (1 mM HCO_3^- [0.01 mM CO_2] and either 35, 50, 87, or 1,300 µM RuBP for the free enzyme, condensate, small carboxysome and large carboxysome respectively; see Fig. 3 and *SI Appendix,* Fig. S4). *Open circles* represent the values obtained for typical permeabilities used in the model (Table 1). Data were generated using the COPASI (60) model run in parameter scan mode, achieving steady-state values over the range of proton and RuBP permeabilities indicated for the Rubisco compartment. For all cases CA activity was only present within the Rubisco compartment. Data presented are for the tobacco Rubisco with parameters listed in Table 2.

We also assessed the modelled responses of the free enzyme, its condensate, and carboxysomes to [RuBP] under sub-saturating HCO_3^- (1 mM; Fig. 3B). Again, proton production by Rubisco led to increases in condensate and carboxysome [CO_2], despite relatively high external HCO_3^- supply, resulting in corresponding increases in Rubisco carboxylation rate.

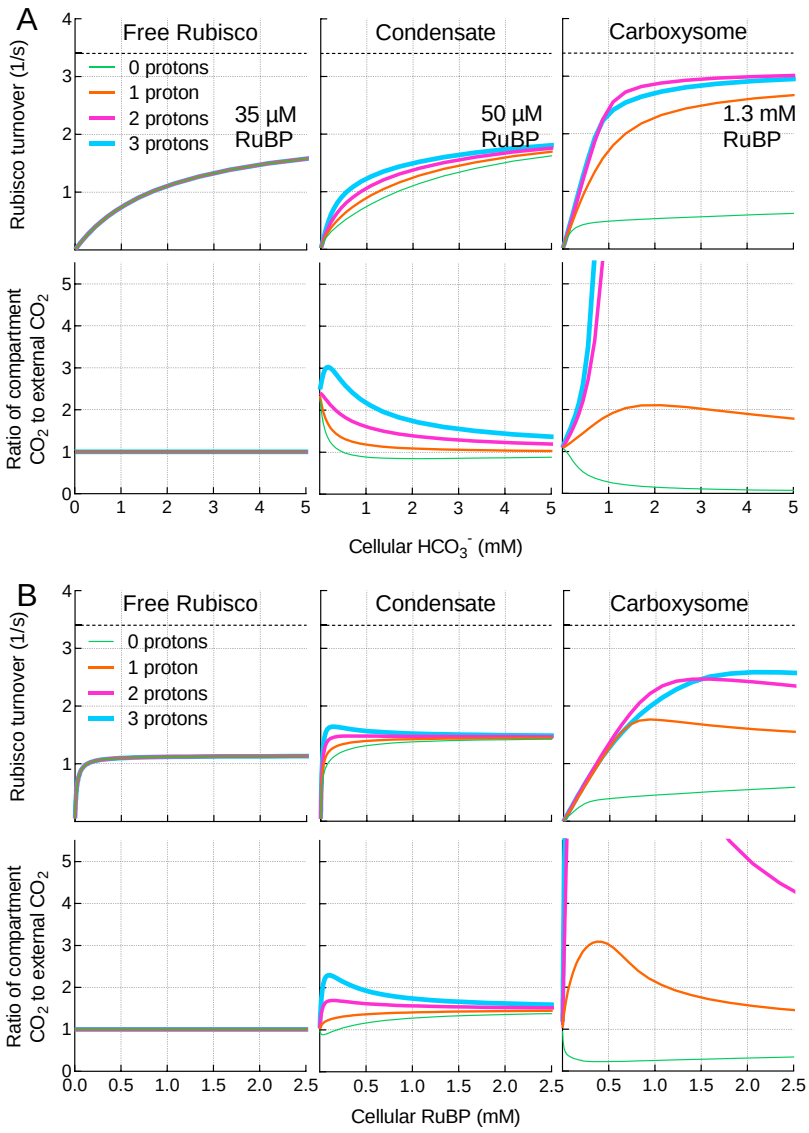

Fig. 3. Modelled responses of free Rubisco, condensate and carboxysome to external substrate supply

Modelled responses of free Rubisco, condensate and carboxysome to external HCO_3^- (A) and RuBP (B) supply. Carboxylation turnover rates (top panels) and the ratio of Rubisco compartment CO_2 to external CO_2 (lower panels) for the free enzyme, condensate and carboxysomes with zero, one, two or three protons being produced as products of the carboxylation reaction at sub-saturating RuBP (A) or HCO_3^- (B). Variation in permeabilities to RuBP between the free enzyme, condensate and carboxysome within the model result in increases in the apparent $K_M RuBP$ as permeabilities decline (*SI Appendix,* Fig. S6). For HCO_3^- responses, the RuBP concentration used in each scenario is indicated. For RuBP responses, the HCO_3^- supply set at 1 mM [0.01 mM CO_2]). External CO_2 is set to 1/100 × external HCO_3^-. Net proton production by Rubisco carboxylation is theoretically two protons (Fig. 1). In all cases modelled rates include carbonic anhydrase (CA) activity only in the Rubisco compartment. Rubisco turnover in a carboxysome producing three protons per carboxylation drops below that of the two-proton system due a proton-driven shift in the $RuBP^{3-}$:$RuBP^{4-}$ ratio, decreasing the concentration of the Rubisco substrate species, $RuBP^{4-}$. Maximum carboxylation turnover rate (k_{cat}^C; 3.4 [1/s]) of the tobacco Rubisco used in this example (Table 2), is indicated by the *dashed line* in the upper panels. The COPASI (60) model was run in parameter scan mode, achieving steady-state values over the range of HCO_3^- concentrations indicated. Proton number was manufactured by modifying the Rubisco carboxylation reaction stoichiometry to produce zero, one, two or three protons in the model.

In considering why these results are obtained, it is important to note that at pH 8.0, which we set for the bulk medium in our model, free [H^+] is only 10 nM. In the Rubisco compartment volume utilized in our model (≈5 x 10^{-18} m^3) this represents less than one proton, and means that the diffusional driving forces for proton exchange across the carboxysomal shell, for example, are 10^4–10^6 -fold smaller than those driving CO_2, HCO_3^-,

PGA and RuBP diffusion (which are in μM and mM ranges). Hence, inward H^+ diffusion will be rate-limiting depending on the permeability of the compartment to protons. Therefore, other proton sources must provide substrate for the CA/HCO_3^- dehydration reaction. The net outcome is that proton production by Rubisco carboxylation within a diffusion-limited compartment leads to decreased pH, elevated CO_2, and improvement in carboxylation turnover, compared with the free enzyme (*SI Appendix*, Fig. S7).

The model also shows an increase in carboxylation turnover resulting from the protons arising through oxygenation within a carboxysome, although this appears negligible in a Rubisco condensate (*SI Appendix*, Fig. S8). Notably, like the model of Mangan, Flamholz, Hood, Milo and Savage (38), we find that carboxysome function does not require specific diffusional limitation to O_2 influx in order to reduce oxygenation, due to competitive inhibition by the increase in CO_2.

Roles for RuBP, PGA and CO_2 in Rubisco compartment pH

When we run the full model, including net proton production by Rubisco, carboxysome pH becomes acidic at limiting substrate concentrations as a result of limited proton efflux (Fig. 4). As external substrate increases and internal CO_2 rises, the pH approaches that of the bulk external medium (which we set at pH 8.0 with the H^+ concentration). Here, increasing CO_2 efflux from the carboxysome is able to effectively dissipate protons from the Rubisco compartment, since there is a net loss of CO_2 that would otherwise be used for proton production by the CA hydration reaction ($CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$). This can be seen when we modify the rate of CO_2 efflux from the carboxysome by altering compartment CO_2 permeability. Slow CO_2 efflux leads to increased free proton concentrations within the compartment, and fast efflux enables a return to approximately external pH as substrate supply increases (Fig. 4).

When we consider reaction species' fluxes across the Rubisco compartment boundary, we find that both $RuBP^{3-}$ and PGA^{2-} can play a role in carrying protons out of both condensates and carboxysomes (Fig. 5, *SI Appendix*, Fig. S9), contributing to the stabilization of internal pH. In our model of carboxysomes, $RuBP^{3-}$ and PGA^{2-} efflux plays a significant role in pH balance at very low HCO_3^-. However, as external HCO_3^- rises, carboxysome CO_2 efflux replaces RuBP and PGA as the major proton carrier (Fig. 5) as described above. For a condensate, CO_2 efflux is the primary means of stabilizing internal pH within the model over a range of HCO_3^- concentrations (*SI Appendix*, Fig. S9).

Similar responses can be seen over a range of RuBP concentrations, where CO_2 efflux also plays a dominant role as a proton efflux carrier in the condensate, while $RuBP^{3-}$ and PGA^{2-} efflux are the major contributors to shuttling protons out of the carboxysome (Fig. 5, *SI Appendix*, Fig. S9). These results highlight that the inclusion of RuBP and PGA as proton carriers is essential in describing the functioning of the carboxysome, as they contribute to maintaining internal pH and, therefore, pH-sensitive Rubisco activity (47).

Model output also emphasizes that compartment pH is highly dependent on the buffering capacity of RuBP and PGA. Not only does RuBP carry protons released in Rubisco reactions, both species also undergo

protonation and deprotonation at physiological pH. Modifying their *pKa* values *in silico* significantly alters Rubisco compartment pH (*SI Appendix*, Fig. S10).

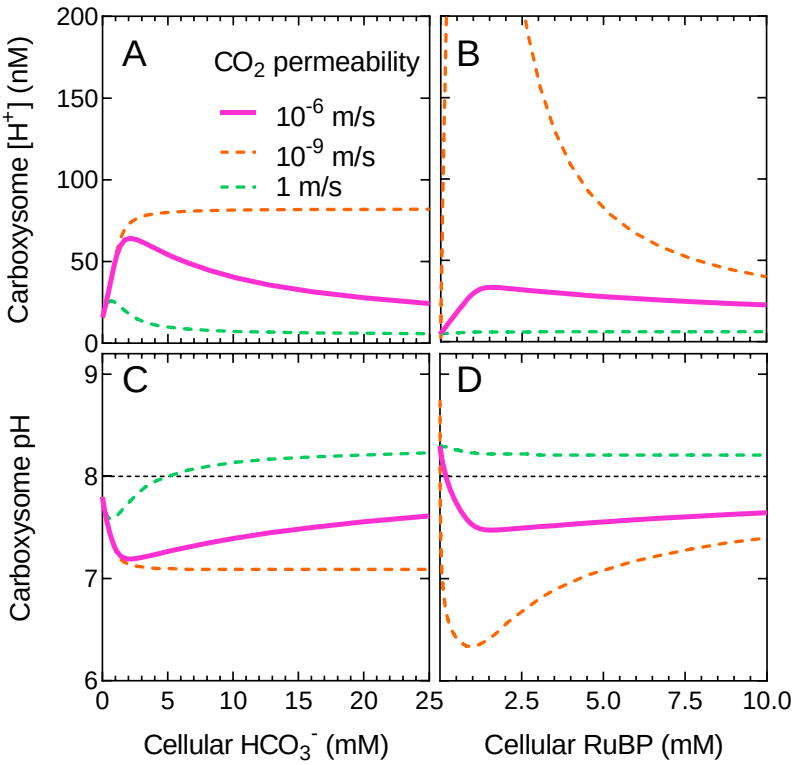

Fig. 4. Carboxysome proton concentration is modulated by CO_2 efflux

Carboxysome free proton concentration (A, B) and carboxysome pH (C, D) indicate that a functional carboxysome compartment undergoes net acidification at limiting HCO_3^- and RuBP supply (*SI Appendix,* Fig. S4). Plotted are proton concentration and pH over a range of [HCO_3^-] and [RuBP] for modelled carboxysomes with an internal carbonic anhydrase (CA), allowing for two protons to be produced per carboxylation reaction, and under typical modelled CO_2 permeability within the model (10^{-6} m/s; *solid pink lines*). If CO_2 efflux were rapid and unimpeded (CO_2 permeability 1 m/s; *dashed green lines*), pH rapidly returns to ≈8 (*black dashed line*, panels C and D) as external limiting substrate supply increases. Slow CO_2 efflux (CO_2 permeability 10^{-9} m/s; *dashed orange lines*) does not allow for dissipation of protons. Efflux of CO_2 from the carboxysome contributes to pH maintenance is it represents the loss of substrate for the CA hydration reaction ($CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$), which would otherwise lead to free proton release. Each dataset was modelled with an initial [RuBP] of 5 mM for HCO_3^- response curves, and 20 mM HCO_3^- in the case of RuBP response curves and CA activity is confined only to the carboxysome compartment. All other permeabilities under these conditions are set to 10^{-6} m/s as for a carboxysome (Table 1). The COPASI (60) model was run in parameter scan mode, achieving steady-state values over a range of substrate concentrations. Data presented are for the tobacco Rubisco (Table 2).

The need for CA

In previous models of carboxysome function there is a need for internal CA to accelerate the conversion of HCO_3^- to CO_2 to support high rates of Rubisco CO_2 fixation and CO_2 leakage out of the carboxysome. While those models consider functional CCMs with active cellular HCO_3^- accumulation, this need for CA is also true here, with internal interconversion needing acceleration to maximize Rubisco CO_2 fixation at saturating external HCO_3^- (*SI Appendix,* Fig. S11). Additionally, we find that CA inclusion within a condensate at high rates provides additional benefit to Rubisco carboxylation rates (*SI Appendix,* Fig. S11, Fig. S12).

It is also apparent that in the carboxysome CA function is dependent upon RuBP and HCO_3^- supply, emphasizing that provision of protons from the Rubisco reaction is essential for the production of CO_2 from

HCO_3^- via the CA enzyme (*SI Appendix*, Fig. S13). This dependency appears to be much less in a condensate due to our assumption of much higher permeabilities of $RuBP^{3-}$ and CO_2 to the condensate interior (*SI Appendix*, Fig. S11) and the low $K_M CO_2$ of the tobacco Rubisco modelled in this scenario.

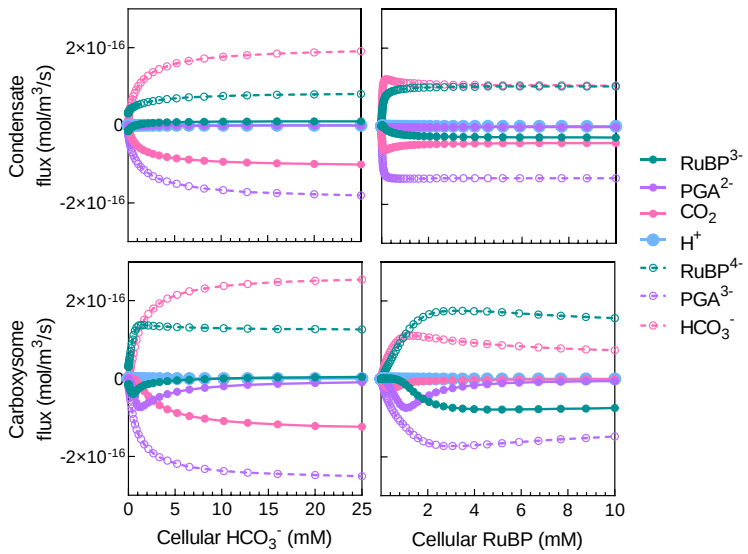

Fig. 5. Proton carriers help maintain compartment pH

Diffusional flux of chemical species across the condensate/carboxysome boundary over a range of HCO_3^- and RuBP concentrations in the model. Protons are carried by $RuBP^{3-}$, PGA^{2-} and CO_2 (as the substrate required for free proton release via the CA hydration reaction). Net free H^+ fluxes are extremely small and contributions to internal pH primarily arise through net fluxes of proton-carrier substrates (*solid circles*). The deprotonated $RuBP^{4-}$ and PGA^{3-} are the substrate and product of Rubisco carboxylation, respectively, within the model. Positive flux values indicate net influx into the compartment and negative values indicate net efflux. The COPASI (60) model was run in parameter scan mode, achieving steady-state at each substrate concentration. For HCO_3^- response curves, RuBP was set to 50 µM for a condensate and 1.3 mM for a carboxysome based on changes in apparent $K_{M}RuBP$ values arising from diffusional resistance (Fig. 3, *SI Appendix,* Fig. S6). External [HCO_3^-] was set to 1 mM for the generation of RuBP response curves. Data presented are for the tobacco Rubisco (Table 2) and model parameters for a condensate or a carboxysome are indicated in Table 1. These data are summarized in *SI Appendix,* Fig. S9.

Carboxysome evolution via Rubisco condensation

Our model shows that Rubisco co-condensed with CA gives improved function over its free enzyme (Fig. 3). Given that Rubisco condensation underpins carboxysome biogenesis (20, 25), we considered that the model may provide insights into carboxysome evolution via intermediate states utilizing Rubisco condensates, prior to HCO_3^- accumulation in the cell through inorganic carbon (C_i) transport.

To investigate this proposal, we analyzed the performance of potential evolutionary intermediates in a series of hypothesized pathways from free Rubisco to carboxysomes, in the absence of functional C_i accumulation (assuming this was a later evolutionary progression; 14). We first assumed a photorespiratory loss of ½ mole of CO_2 for every mole of O_2 fixed within the model (while carboxylation yields two molecules of PGA from one CO_2, oxygenation yields only one PGA and one CO_2 is lost via photorespiration), and by calculating average net carboxylation rates by each hypothesized evolutionary intermediate under different atmospheric conditions (see Methods). This allowed relative fitness comparisons of proposed evolutionary states under what might reasonably be considered low or high atmospheric CO_2, as may have been experienced

under atmospheres where CCMs arose (7, 48). We assume here that a greater average net carboxylation rate for a particular evolutionary intermediate (over a given range of HCO_3^-) would result in greater relative fitness and, therefore, evolutionary advantage. We apply this concept over HCO_3^- ranges within the model, rather than single-point comparisons, to provide a clearer view of Rubisco responses resulting from variations in the observed response curves and, to assess responses under different atmospheric CO_2 compositions (*SI Appendix*, Fig. S14).

We made these fitness calculations at both 20% and 30% (*v/v*) O_2, to simulate alternative atmospheres under which CCMs may have arisen (7, 48). We calculated the average net carboxylation rates for the defined HCO_3^- ranges (see Methods), and generated phenotype matrices to allow comparison of possible evolutionary states (Fig. 6; *SI Appendix, SI* datasets S1 and S2). In evolutionary state comparisons, those states which lead to larger average net carboxylation rates were considered to have greater fitness.

The relative fitness of Form I Rubisco enzymes (49) from a variety of sources were compared within the model, over all proposed evolutionary intermediates, to observe any differences resulting from varying Rubisco catalytic parameters (Table 2). A complete analysis of each Rubisco source and its performance at each proposed evolutionary step, under varied HCO_3^- and O_2 conditions is supplied in the *SI Appendix, SI* datasets S1 and S2.

Table 2. Rubisco catalytic parameters used in competition modelling

Rubisco source	k_{cat}^C [1/s]	k_{cat}^O [1/s]	K_MCO_2 [µM]	K_MO_2 [µM]	K_MRuBP [µM]	Substrate specificity factor ($S_{C/O}$)	Carboxylation efficiency (k_{cat}^C/K_MCO_2) [1/s/µM]	References
Tobacco	3.40	1.14	10.7	295	18.0	82	0.318	(63)
Synechococcus	14.4	1.22[a]	172	585[a]	69.9	40	0.084	(61, 64)
Cyanobium	9.40	1.42[b]	169	1400[b]	40.0	55	0.056	(15, 43)
Ancestral F1A	4.77	1.42	113	2010	40.0[c]	60	0.042	(15, 43)
Ancestral F1B	4.72	0.50	120	641	69.9[c]	50	0.039	(15, 61)
Chlamydomonas	2.91	0.61	33.0	422	19.0[d]	61	0.088	(65, 66)

[a] – The k_{cat}^O and K_MO_2 for *Synechococcus* Rubisco are from Occhialini, Lin, Andralojc, Hanson and Parry (64).
[b] – The k_{cat}^O of the *Cyanobium* enzyme was estimated using the *Prochlorococcus* K_MO_2 (15), and the published values of k_{cat}^C, K_MCO_2 and $S_{C/O}$ for *Cyanobium* (43) using $k_{cat}^O = [(k_{cat}^C \times K_MO_2)/S_{C/O}]/K_MCO_2$.
[c] – K_MRuBP values for Ancestral F1A and F1B enzymes are those of *Cyanobium* and *Synechococcus* respectively, from the references highlighted.
[d] – The K_MRuBP value for *Chlamydomonas* is from Zhu and Spreitzer (65) while all other values are from Spreitzer, Peddi and Satagopan (66).

Regardless of the type of Rubisco used in this analysis, the same pattern of potential evolutionary augmentations was favored (Fig. 6, Fig. 7, and *SI Appendix, SI* datasets S1 and S2). In all cases, condensation of Rubisco in the absence of a CA enzyme, as an initial evolutionary step ('Rubisco - CA' to 'Condensate – int/ext CA'; Fig. 6), resulted in a decrease in fitness, emphasizing that the starting point for the evolution of Rubisco condensates and carboxysomes likely began with a cellular CA present ('Rubisco + CA'). Again, we emphasize here that our modelling assumes no active C_i accumulation as observed in modern aquatic CCMs where the occurrence of a cellular CA, outside a carboxysome for example, dissipates an accumulated HCO_3^- pool as membrane-permeable CO_2 and leads to a high-CO_2-requiring phenotype (42).

Percent change in average net carboxyla on rates - 20% O_2 0.01 - 1 mM cellular HCO_3^-	Next evolutionary state									
Initial state	Free Rubisco -CA	Free Rubisco +CA	Condensate +ext CA	Condensate +int/ext CA	Condensate +int CA	Carboxysome +int CA	Carboxysome +int/ext CA	Condensate -int/ext CA	Carboxysome -int/ext CA	Carboxysome +ext CA
Free Rubisco -CA	0	0	21	165	126	273	273	-135	-281	-276
Free Rubisco +CA	0	0	21	165	126	273	273	-135	-281	-276
Condensate +ext CA	-17	-17	0	119	87	208	208	-129	-250	-246
Condensate +int/ext CA	-62	-62	-54	0	-15	41	41	-113	-168	-166
Condensate +int CA	-56	-56	-47	17	0	65	65	-115	-180	-178
Carboxysome +int CA	-73	-73	-68	-29	-39	0	0	-109	-149	-147
Carboxysome +int/ext CA	-73	-73	-68	-29	-39	0	0	-109	-149	-147
Condensate -int/ext CA	389	389	449	865	753	1176	1177	0	-423	-408
Carboxysome -int/ext CA	155	155	167	246	225	306	306	81	0	3
Carboxysome +ext CA	157	157	169	251	228	312	312	80	-3	0

Fig. 6. Fitness matrix for proposed evolutionary steps from free Rubisco to contemporary carboxysomes, before the evolution of active C_i transport.

An example fitness matrix for the tobacco Rubisco enzyme showing percentage difference in average net Rubisco carboxylation rates (see Methods and *SI Appendix,* Fig. S14) over a 0.01 – 1 mM HCO_3^- in a 20% (*v/v*) O_2 atmosphere in systems lacking C_i transport and HCO_3^- accumulation. Within this table an initial evolutionary state (*left, rows*) can be compared with any potential next evolutionary state (*top, columns*). Values are percent changes in average net carboxylation rates between each state. Positive (*yellow*) values indicate an improvement in average net Rubisco carboxylation turnover when evolving from an initial state to the next evolutionary state. Negative (*blue*) values indicate a net detriment. As an example, evolution from free Rubisco with associated carbonic anhydrase (Free Rubisco + CA) to a Rubisco condensate with an internal CA (Condensate + int CA) shows a 126% improvement in average net carboxylation turnover. We assume that such an evolutionary adaptation would result in a competitive advantage over the initial state. Contrarily, a Rubisco condensate with both internal and external CA (Condensate +int/ext CA) shows a decrease in average net carboxylation turnover of 15% if the external CA was lost in an evolutionary adaptation (Condensate + int CA). Data presented here are for the tobacco Rubisco (Table 2), using the compartment model to simulate all potential evolutionary states (Table 1). External CA is modelled in the unstirred layer. The same pattern of potential evolutionary improvements is apparent regardless of the Rubisco source or carboxysome size used in the model, assuming sufficient RuBP supply (*SI Appendix, SI datasets* S1 and S2).

In Fig. 6 we provide an example evolutionary matrix for hypothesized progressions from a free Rubisco enzyme to a contemporary carboxysome. As a primary evolution, we speculated that condensation of Rubisco could have occurred either with or without co-condensation of CA ('Condensate + int CA' or 'Condensate + ext CA'), or an alternate evolution where some CA was co-condensed and some remained external to the condensate in the unstirred layer ('Condensate + int/ext CA'). All three possibilities gave rise to condensates with improved fitness over the free enzyme (Fig. 6, Fig. 7, and *SI Appendix, SI* datasets S1 and S2). While the greatest improvement was calculated for a condensate with both internal and external CA ('Condensate + int/ext CA'), this state showed a negative transition in evolving to a condensate with only internalized CA ('Condensate + int CA').

Following initial condensate formation, we propose that the next probable advancement would be the acquisition of bacterial micro-compartment proteins (8) to form a shell with enhanced diffusional resistance. Within the model, only condensates with internalized CA enzymes ('Condensate + int CA' and 'Condensate + int/ext CA') displayed an improved CO_2 fixation phenotype during a single step acquisition of a carboxysome shell (Fig. 6, Fig. 7 and *SI Appendix, SI* datasets S1 and S2). There was no difference in fitness phenotype between carboxysomes with internal CA ('Carboxysome + int CA') and those with CA both internal and

external in the unstirred layer ('Carboxysome +int/ext CA') in the model. Further improvements could be made by the addition of HCO_3^- transporters, the loss of unstirred layer CA, the acquisition of specific internal CAs, and evolution of Rubisco kinetic properties.

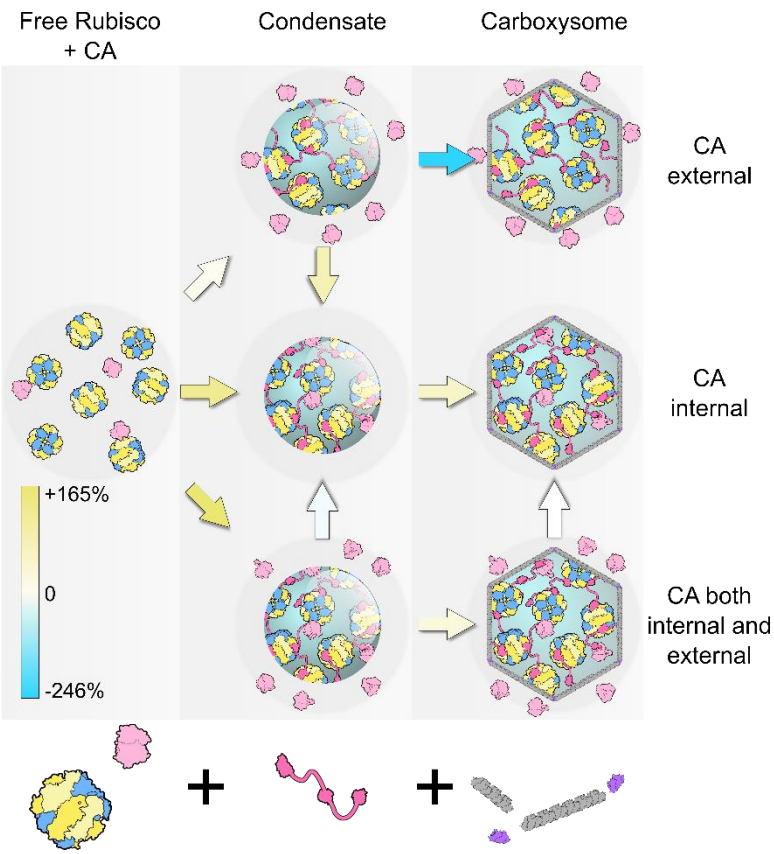

Fig. 7. Proposed evolution pathways to carboxysomes from free Rubisco via condensation, before the advent of C_i transport.

Model simulations propose condensation of Rubisco (*blue/yellow*) in the presence of a cellular carbonic anhydrase (CA) enzyme (*light pink*), here presented as three possible evolutionary alternatives with the CA external (in the unstirred layer), internal, or both external and internal of the condensate. More detailed analysis shows that evolution of a Rubisco condensate in the absence of a CA is not feasible (Fig. 6). Condensation is proposed to occur through the evolution of a condensing protein factor (here CcmM from β-carboxysomes, *bright pink*), and carboxysome formation via the acquisition of bacterial microcompartment shell proteins (*grey, purple*). Contemporary carboxysomes are represented by those containing only internal CA. Percent increase or decrease in average net carboxylation rates between each proposed evolutionary intermediate is indicated by the colored arrows, and the color-scale indicates the values presented in Fig. 6. Between proposed evolutionary stages, *yellow-shaded* arrows indicate an improvement in average net carboxylation rate and *blue-shaded* arrows a net decrease, suggesting a loss in competitive fitness. No net change is indicated by a *white arrow*. The same pattern of potential evolutionary improvements is apparent regardless of the Rubisco source or carboxysome size used in the model, assuming sufficient RuBP supply (*SI Appendix, SI datasets* S1 and S2). We assume the adaptation of increased cellular HCO_3^- followed as an evolutionary enhancement (14), hence the relative fitness of systems with CA external to the Rubisco compartment (here modelled in the unstirred layer), where in contemporary systems this is problematic (42).

Low light may drive condensate and carboxysome evolution

Observing increased relative responses of compartment CO_2 to low RuBP concentrations in the model, where condensate pH is maximally decreased (*SI Appendix*, Fig. S4), we assessed the relative fitness of condensates and carboxysomes at sub-saturating RuBP (50 µM) over HCO_3^- ranges in the model. Low cellular RuBP can generally be attributed to light-limited RuBP regeneration via the Calvin Cycle in photoautotrophs (50). A concentration of 50 µM RuBP is approximately three-times the $K_M RuBP$ of the tobacco Rubisco used

here (Table 2) and supports ≈63% of the CO_2-saturated rate for a condensed Rubisco (*SI Appendix,* Fig. S6). At 50 µM RuBP we observe enhanced net carboxylation turnover in the condensate compared with the free enzyme, especially at low HCO_3^- (*SI Appendix,* Fig. S15). An additional benefit can be observed for very small carboxysomes since changes in the apparent $K_M RuBP$ are size-dependent (*SI Appendix,* Fig. S6 and Fig. S15). However, 50 µM RuBP is insufficient to support appreciable carboxylation in a modelled large carboxysome due to decreased substrate permeability (*SI Appendix,* Fig. S6 and Fig. S15).

DISCUSSION

The functional advantages of a condensate/carboxysome

The modelling of both Rubisco condensates and carboxysomes in this study demonstrates a number of factors which we predict play a key role in the function and evolution of these Rubisco "organelles". Firstly, the formation of a Rubisco condensate creates a localized environment in which HCO_3^- can be converted to CO_2 in the presence of CA. CO_2 can be elevated relative to the external environment by the creation of a viscous unstirred protein solution boundary layer. The presence of condensates or carboxysomes in prokaryotic cells or chloroplasts, where protein concentrations are high, would favor this (51, 52).

The condensation of Rubisco results in the co-localization of Rubisco-reaction protons. This enhances the potential to elevate CO_2 by driving both the conversion of HCO_3^- to CO_2 in the Rubisco compartment, and decreasing compartment pH under certain conditions (*SI Appendix,* Fig. S4). This role for protons is seen most clearly in carboxysomes where both CA and Rubisco activity are highly dependent on proton production by the Rubisco reaction (Fig. 3, and *SI Appendix,* Fig. S13). Rubisco condensates show a smaller enhancement by Rubisco proton production due to greater permeability to protonated RuBP and PGA, but under conditions of low RuBP and low HCO_3^-, condensate pH can be lowered and CO_2 elevated (Fig. 3 and *SI Appendix,* Fig. S4). In addition, the advantages of condensate formation are enhanced when the ratio of external HCO_3^- to CO_2 is increased (*SI Appendix,* Fig. S12) as would occur at higher cytoplasmic pH, and in the presence of active HCO_3^- accumulation and subsequent transfer of CA from the cytoplasm to the Rubisco compartment.

The modelling emphasizes that the exchange of protons between internal and external environments is probably independent of free proton or H_3O^+ diffusion. Instead, the exchange is dominated by the movement of protonated RuBP and PGA (with *pKa*'s around 6.6), and efflux of CO_2 which consumes a proton internally (Fig. 4 and Fig. 5; *SI Appendix,* Fig. S5).

The effects of Rubisco compartment pH and the relevance of sugar phosphate proton carriers

By accounting for the pK_a of physiologically relevant phosphate groups on RuBP and PGA (*SI Appendix,* Fig. S1), the model reveals that these species allow for sufficient ingress of protons into a Rubisco compartment to drive higher rates of carboxylation. This is because the concentration gradients for RuBP and PGA are

significantly high (in the mM range), compared with protons at pH 8.0 (in the nM range), enabling them to act as proton carriers in physiologically relevant concentrations. The sophisticated model of Mangan, Flamholz, Hood, Milo and Savage (38) also considers the net production of a proton within the carboxysome [i.e. consumption of one proton by the CA dehydration reaction and the generation of two protons by Rubisco carboxylation]. They calculate that a relatively acidic carboxysome is possible under steady-state conditions, depending on proton permeability. However, analysis of viral capsid shells which have some similarity to the carboxysome icosahedral shell, suggest proton transfer may need to be mediated by specific channels (53, 54). It is important to establish the real permeability of the carboxysome shell to the hydronium ion, and whether high rates of exchange are mediated by channels, proton wires or shuttles linked to protonated sugar phosphates suggested here.

Drivers of Rubisco compartment evolution

Our evolution analysis suggests that the specificity which Rubisco has for CO_2 over O_2 ($S_{C/O}$) is a likely key driver in determining fitness for the free Rubisco enzyme under low CO_2 atmospheres (*SI Appendix, SI Results*; Fig. S16). Indeed, the tobacco enzyme, having the greatest $S_{C/O}$ (Table 2) displayed the best performance under all low CO_2 scenarios (Fig. 8). Rubisco carboxylation efficiency (i.e. its carboxylation rate constant; $k_{cat}^{C}/K_M CO_2$) is also an important driver in determining relative fitness under low O_2 atmospheres, for both the free enzyme and condensates (Fig. S16 – Fig. S19). Furthermore, this analysis suggests that under increased atmospheric O_2 (as would have occured some 300 million years ago, when levels of O_2 rose and CO_2 fell; 14) there is selective pressure to reduce k_{cat}^{C} of the free enzyme (Fig. S16). This suggest that ancestral cyanobacterial Rubisco may have had kinetics similar to the tobacco Form IB enzyme, implying relatively high photorespiration rates in pre-CCM cyanobacteria. Notably, contemporary cyanobacteria appear to contain a full suite of photorespiratory genes (55), despite limited Rubisco oxygenase activity in modern carboxysomes (56).

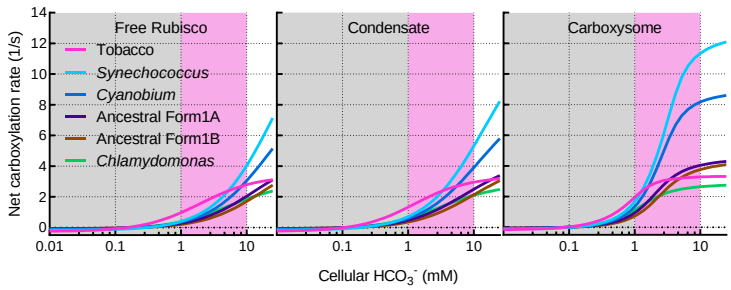

Fig. 8. Net Rubisco carboxylation turnover rates in competing enzymes

Net carboxylation turnover rates of competing Rubisco enzymes (Table 2) in the model as free enzymes, condensates or carboxysomes. Each state was modelled using the parameters outlined in Table 1 and as described in the Methods. Rates are depicted under low (0.01 – 1 mM HCO_3^-; *grey shaded area*) and high (1 – 10 mM HCO_3^-; *pink shaded area*) CO_2 environments. Data presented here are calculated under a 20% (*v/v*) O_2 atmosphere and assume no active accumulation of HCO_3^-.

Together, our evolution analysis suggests that a low CO_2 atmosphere may be a key driver in the initial formation of Rubisco condensates. At elevated CO_2, regardless of O_2 concentration, condensate fitness is unconstrained by Rubisco catalytic parameters (*SI Appendix,* Fig. S17 – Fig. S19). Large carboxysome formation likely provided compartment conditions which enabled the evolution of Rubisco enzymes with higher k_{cat}^{C}, but appear to be unconstrained by [O_2] in the atmosphere (*SI Appendix,* Fig. S20). The correlation between k_{cat}^{C} and fitness is apparent only in this scenario, since improving the maximum carboxylation turnover rate is the only means to improve net carboxylation rates when extremely high CO_2 concentrations around the enzyme can be achieved. Smaller carboxysomes, however, appear not to be driven by any Rubisco catalytic parameter at elevated CO_2 (*SI Appendix,* Fig. S21). Changes in atmospheric O_2 may have led to the selection of enzymes with better specificity, catalytic efficiency, and $K_M RuBP$ during intermediate stages of Rubisco condensate evolution (*SI Appendix,* Fig. S17 – Fig. S19).

The relative response of Rubisco condensates and small carboxysomes to low RuBP supply suggests that low light may also provide conditions conducive to condensate and carboxysome evolution. Low light generally leads to low RuBP (50). In the model such conditions lead to greater advantage of Rubisco condensates over large carboxysomes (*SI Appendix,* Fig. S15). This highlights that elaborations on simple Rubisco condensation can be afforded through facilitated substrate supply or potential diffusion barriers other than carboxysome formation. The association between pyrenoids and thylakoids, for example, suggests that luminal protons can possibly contribute towards the conversion of HCO_3^- to CO_2 (56), and allows the possibility that the pyrenoid function may be enhanced by the combined action of both Rubisco and thylakoid generated protons. This result highlights alternative evolution of Rubisco compartments such as pyrenoids which would provide enhancements exclusive of shell formation.

Limitations of the model

The model is designed as an idealized Rubisco compartment within a static environment. We use a single 'condensate' of 1 μm in radius for the demonstration of condensate function, approximating a large pyrenoid (57). Notably, both carboxysomes and pyrenoids range significantly in size (8, 58), and the effect of compartment size in the model is addressed in the *SI Appendix, SI Methods* and Fig. S22. It does not include an 'extracellular' compartment from which the cytoplasmic compartment can receive C_i via either diffusion or specific C_i pumps. We do not model the system as a functional CCM, holding a static equilibrium between C_i species in the system , rather than a disequilibrium which occurs in cyanobacterial cells (59). In the absence of active HCO_3^- pumps or a cytoplasmic CA, preferential diffusion of CO_2 through the cell membrane, coupled with its draw-down via carboxylation by Rubisco, would lead to relatively low cytoplasmic [C_i] with a HCO_3^-:CO_2 ratio possibly favoring CO_2 (*SI Appendix,* Fig. S12). However, notably there is no fitness benefit to Rubisco condensation in the absence of a CA (Fig. 6) and scenarios where the HCO_3^-:CO_2 ratio lean heavily toward CO_2 are less likely.

We do not apply a pH sensitivity to Rubisco catalysis in the model and omit unknown contributors to cellular buffering. This is both for simplicity and to highlight that without pH modulation in the model we would see dramatic changes in pH. Previous models (38) apply the pH-dependency of Rubisco as established by Badger (47), as would be applicable in this instance. We would expect to see a decrease in Rubisco activity in our model resulting from acidification or alkalization of the Rubisco compartment. However, we do not currently know how a condensate or carboxysome might modulate pH changes in reality, and how these might affect the enzymes within it. We also do not model CO_2 and Mg^{2+} dependencies upon Rubisco activation (47), which would be relevant considerations, especially at low CO_2 supply, from a physiological standpoint. These considerations may underpin mechanistic controls at low substrate supply within both compartment types.

Importantly, we note that the permeabilities for diffusion across the compartment interfaces have no physical measurements to support them. The values for carboxysomes are in line with previous models but it should be realized that these values have been derived from model analysis and not independent physical measurements. The values assumed for condensates, although seeming reasonable and supporting a functional model, have not been physically established independently.

Conclusions

Accounting for proton production by Rubisco reactions, their utilization in the CA reaction, their transport via RuBP and PGA, and applying diffusional resistance to their movement, our model highlights a role for protons in Rubisco condensate and carboxysome function and evolution. Application of our model to proposed evolutionary intermediate states prior to contemporary carboxysomes provides a hypothetical series of advancements, and suggest that low CO_2 and low light environments may be key environmental drivers in the evolutionary formation of Rubisco condensates, while increases in atmospheric O_2 may have played a role in Rubisco catalytic parameter selection. Our modelled outcomes are achieved through assumption of diffusional resistances to reaction species, which align well with previous models but remain to be determined experimentally. Taken together, our analysis provides insights into the function of phase-separated condensates of proton-driven enzyme reactions.

METHODS

Mathematical modelling

Modelling of Rubisco compartment scenarios and data output were carried out using the biochemical network simulation program COPASI (copasi.org), described by Hoops, *et al.* (60). COPASI (v4.25, build 207) was used to simulate reaction time-courses achieving steady-state conditions in a three-compartment model where reaction species are linked in a biochemical network (Fig. 1). For standard modelling conditions catalytic parameters of the tobacco Rubisco were used, while those of *C. reinhardtii*, *C. marinum* PCC7001 and *S.*

elongatus PCC7942, and predicted ancestral Form IA and Form IB enzymes (15) were also used in evolutionary fitness analysis (Table 2). Michaelis-Menten rate equations were applied to Rubisco catalysis as dependent upon substrate and inhibitor concentrations within the Rubisco compartment (*SI Appendix, SI Methods*). O_2 was applied as an inhibitor of the carboxylation reaction and CO_2 an inhibitor of oxygenation. Greater detail of model parameterization is provided in the *SI Appendix, SI Methods*. The ordinary differential equations describing the model components can be found in the *SI Appendix*. Reactions, reaction species and model parameters can be found in *SI Appendix*, Table S1, Table S2, and Table S3 respectively.

Variation in compartment types (i.e. the free enzyme, a condensate, or a carboxysome) was simulated in the model by varying unstirred boundary and condensate permeabilities to all reaction species (Table 1, *SI Appendix, SI Methods*). The permeabilities to $RuBP^{3-}$ and $RuBP^{4-}$ were assumed to be the same, as were those for PGA^{2-} and PGA^{3-}, and a single permeability value applied to either RuBP or PGA species. Proton production by carboxylation and oxygenation reactions was varied by adjusting the proton stoichiometry for either reaction (*SI Appendix, SI Equations,* Table S1). Protonation and deprotonation of RuBP and PGA in each compartment was enabled by assigning rate constants equivalent to their *pKa* values (*SI Appendix,* Table S3).

The size of the external compartment was set to 1 m^3 within the model, and both the Rubisco compartment and unstirred boundary layer volumes determined by setting the Rubisco compartment radius. For standard modelling procedures we used a spherical Rubisco compartment radius of 1×10^{-6} m^3 and the unstirred boundary layer volume was determined as a simple multiplier of the Rubisco compartment radius. The Rubisco compartment radius used in our modelling generates a large condensate or carboxysome, akin to contemporary pyrenoids, however variation in the compartment size has little effect on the conclusions of the modelled outcomes since even small carboxysomes display higher Rubisco turnover rates that large condensates in the model (*SI Appendix,* Fig. S22). Extremely small carboxysomes have lower amplitude responses to proton and RuBP permeabilities (Fig. 2) and size may have led to favorable Rubisco kinetics during evolution to modern carboxysomes (*SI Appendix,* Fig. S21, *SI dataset* S2).

CO_2 concentration in the external compartment was set to 0.01 × external [HCO_3^-] (*SI Appendix, SI Equations*), assuming negligible effects of a single Rubisco compartment on the bulk external C_i species. Interconversion between CO_2 and HCO_3^- was allowed to proceed in the unstirred and Rubisco compartments with rate constants of 0.05 for the forward reaction ($CO_2 \rightarrow HCO_3^- + H^+$), and 100 for the back reaction ($HCO_3^- + H^+ \rightarrow CO_2$). CA contribution was enabled by applying a multiplying factor to each rate, such that a factor of 1 simulates the absence of CA. Typical CA multiplying factors for each type of modelled compartment are listed in Table 1. CA function external to a Rubisco compartment was modelled by modifying CA in the unstirred layer (Table 1).

O_2 concentration in the model was typically set at contemporary atmospheric levels 20% (*v/v*) by assigning a concentration of 0.25 mM, the water-soluble concentration at 25 °C. For simulations at 30% (*v/v*) O_2 (an estimated volumetric concentration in the atmosphere when it is proposed CCMs arose; 14), a concentration of 0.36 mM was used.

Steady-state reactions were initialized by setting the reactant concentrations in the external compartment and running the model to achieve steady-state. External pH was set at 8.0 using a compartment [H^+] of 1×10^{-5} mM. For saturating Rubisco substrate concentrations, initial [HCO_3^-] was set to 20 mM and [$RuBP^{4-}$] was set to 5 mM. Interconversion of $RuBP^{4-}$ and $RuBP^{3-}$ at *pKa* 6.7 and pH 8.0 results in ≈95% of all RuBP as the $RuBP^{4-}$ species.

Rubisco site concentrations were typically set to 10 mM. This value is similar to that calculated for α- and β-carboxysomes (61) although higher than that estimated for pyrenoids (18). It is nonetheless a reasonable upper limit for the purposes of examining system responses within the model.

Evolution analysis

Hypothesized free enzyme, condensate or carboxysome evolutionary intermediates (Fig. 6) were generated and analyzed within the model using the parameters in Table 1, applying the catalytic parameters of each Rubisco enzyme in Table 2. For each scenario the model was run over a range of [HCO_3^-] from 0.01 to 25 mM, at saturating RuBP (5 mM). Both carboxylation and oxygenation rates were output for each scenario and converted to turnover numbers by accounting for active site concentrations within the model. Net carboxylation turnover rates (1/s) were calculated by assuming a photosynthetic cost of ½ mole of CO_2 loss for each mole O_2 fixed (*SI Appendix*, Fig. S14). Modelling was carried out at both 20% and 30% O_2 (*v/v*; Table 1). Performance comparisons of each hypothesized evolutionary state were made for both low CO_2 (0.01 – 1 mM HCO_3^-) and high CO_2 (1 – 10 mM HCO_3^-) ranges by calculating the average net carboxylation rates for each scenario at each CO_2 range and O_2 concentration (*SI Appendix*, Fig. S14). Fitness comparisons were determined from the absolute differences in average net carboxylation rates between each modelled scenario (Fig. 6; *SI Appendix, SI datasets* S1 and S2). An additional dataset was calculated for the tobacco Rubisco enzyme at 20% (*v/v*) O_2 and 50 µM RuBP to determine compartment-type performance under conditions simulating low light.

Correlations between average net carboxylation rates (over specified modelled atmospheres) and Rubisco catalytic parameters (Table 2) for each proposed evolutionary state (Fig. 6; *SI Appendix*, Fig. S16 – Fig. S21) were calculated as Pearson correlation statistics, with *p* values calculated as two-tailed distribution of calculated *t*-statistics (*SI Appendix, SI Results, SI datasets* S1 and S2). In our results we highlight correlations where $p < 0.05$. For plotting purposes (*SI Appendix, SI Results*, Fig. S16 – Fig. S21) Rubisco catalytic parameters (Table 2) were normalized to the largest value in each parameter set and plotted against average net carboxylation rates.

AUTHOR CONTRIBUTIONS

M.R.B designed and developed the model. B.M.L. and B.F. carried out modelling experiments. B.M.L. and M.R.B generated the Figures. B.M.L. wrote the manuscript draft. B.M.L., B.F., S.B.P., G.D.P. and M.R.B. discussed, wrote and edited manuscript revisions.

ACKNOWLEDGEMENTS

This work is supported by a sub-award from the University of Illinois as part of the research project Realizing Increased Photosynthetic Efficiency (RIPE) that is funded by the Bill & Melinda Gates Foundation, Foundation for Food and Agriculture Research, and the U.K. Government's Department for International Development under grant number OPP1172157, and an award from The Australian Research Council, Centre of Excellence grant for Translational Photosynthesis (CE140100015) to M.R.B., G.D.P. We thank Angus Long for enabling rapid processing of COPASI data output, and Avi Flamholz, Luke Oltrogge and Matthew Mortimer for comments on the manuscript.

REFERENCES

1. M. Hagemann *et al.*, Evolution of photorespiration from cyanobacteria to land plants, considering protein phylogenies and acquisition of carbon concentrating mechanisms. *J. Exp. Bot.* **67**, 2963-2976 (2016).
2. Y. M. Bar-On, R. Milo, The global mass and average rate of rubisco. *Proc. Natl. Acad. Sci. USA* **116**, 4738-4743 (2019).
3. P. M. Shih, Photosynthesis and early Earth. *Curr. Biol.* **25**, R855-R859 (2015).
4. G. G. Tcherkez, G. D. Farquhar, T. J. Andrews, 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. J. Galmés *et al.*, Expanding knowledge of the Rubisco kinetics variability in plant species: environmental and evolutionary trends. *Plant, Cell Environ.* **37**, 1989-2001 (2014).
6. J. A. Raven, J. Beardall, P. Sanchez-Baracaldo, The possible evolution and future of CO_2-concentrating mechanisms. *J. Exp. Bot.* **68**, 3701-3716 (2017).
7. A. Flamholz, P. M. Shih, Cell biology of photosynthesis over geologic time. *Curr. Biol.* **30**, R490-R494 (2020).
8. B. D. Rae, B. M. Long, M. R. Badger, G. D. Price, 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).
9. L. Fridlyand, A. Kaplan, L. Reinhold, Quantitative evaluation of the role of a putative CO_2-scavenging entity in the cyanobacterial CO_2-concentrating mechanism. *BioSyst.* **37**, 229-238 (1996).
10. S.-I. Maeda, M. R. Badger, G. D. Price, Novel gene products associated with NdhD3/D4-containing NDH-1 complexes are involved in photosynthetic CO_2 hydration in the cyanobacterium, *Synechococcus* sp. PCC7942. *Mol. Microbiol.* **43**, 425-435 (2002).
11. G. D. Price, Inorganic carbon transporters of the cyanobacterial CO_2 concentrating mechanism. *Photosynthesis Res.* **109**, 47-57 (2011).
12. S. S. Cot, A. K. So, G. S. Espie, A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. *J. Bacteriol.* **190**, 936-945 (2008).
13. C. de Araujo *et al.*, Identification and characterization of a carboxysomal γ-carbonic anhydrase from the cyanobacterium *Nostoc* sp. PCC 7120. *Photosynthesis Res.* **121**, 135-150 (2014).
14. M. R. Badger, D. Hanson, G. D. Price, Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. *Funct. Plant Biol.* **29**, 161-173 (2002).
15. P. M. Shih *et al.*, Biochemical characterization of predicted Precambrian RuBisCO. *Nat. Commun.* **7**, 10382 (2016).
16. L. C. M. Mackinder, The *Chlamydomonas* CO_2-concentrating mechanism and its potential for engineering photosynthesis in plants. *New Phytol.* 10.1111/nph.14749, 10.1111/nph.14749 (2017).

17. J. C. Villarreal, S. S. Renner, Hornwort pyrenoids, carbon-concentrating structures, evolved and were lost at least five times during the last 100 million years. *Proc. Natl. Acad. Sci. USA* **109**, 18873-18878 (2012).
18. E. S. Freeman Rosenzweig *et al.*, The eukaryotic CO_2-concentrating organelle is liquid-like and exhibits dynamic reorganization. *Cell* **171**, 148-162 e119 (2017).
19. T. Wunder, Z. G. Oh, O. Mueller-Cajar, CO_2-fixing liquid droplets: Towards a dissection of the microalgal pyrenoid. *Traffic* **20**, 380-389 (2019).
20. H. Wang *et al.*, Rubisco condensate formation by CcmM in β-carboxysome biogenesis. *Nature* **566**, 131-135 (2019).
21. L. C. M. Mackinder *et al.*, A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proc. Natl. Acad. Sci. USA* **113**, 5958-5963 (2016).
22. T. Wunder, S. L. H. Cheng, S.-K. Lai, H.-Y. Li, O. Mueller-Cajar, The phase separation underlying the pyrenoid-based microalgal Rubisco supercharger. *Nat. Commun.* **9**, 5076 (2018).
23. T. Chaijarasphong *et al.*, Programmed ribosomal frameshifting mediates expression of the α-carboxysome. *J. Mol. Biol.* **428**, 153-164 (2016).
24. F. Cai *et al.*, Advances in understanding carboxysome assembly in *Prochlorococcus* and *Synechococcus* implicate CsoS2 as a critical component. *Life* **5**, 1141-1171 (2015).
25. L. M. Oltrogge *et al.*, Multivalent interactions between CsoS2 and Rubisco mediate α-carboxysome formation. *Nat. Struct. Mol. Biol.* 10.1038/s41594-020-0387-7 (2020).
26. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285 (2017).
27. A. A. Hyman, C. A. Weber, F. Jülicher, Liquid-liquid phase separation in biology. *Annu. Rev. Cell. Dev. Biol.* **30**, 39-58 (2014).
28. S. Alberti, Phase separation in biology. *Curr. Biol.* **27**, R1097-R1102 (2017).
29. B. D. Rae *et al.*, Progress and challenges of engineering a biophysical carbon dioxide-concentrating mechanism into higher plants. *J. Exp. Bot.* **68**, 3717–3737 (2017).
30. O. D. Caspari *et al.*, Pyrenoid loss in *Chlamydomonas reinhardtii* causes limitations in CO_2 supply, but not thylakoid operating efficiency. *J. Exp. Bot.* **68**, 3903-3913 (2017).
31. B. M. Long, M. R. Badger, S. M. Whitney, G. D. Price, Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. *J. Biol. Chem.* **282**, 29323-29335 (2007).
32. M. R. Harpel, F. C. Hartman, Facilitation of the terminal proton transfer reaction of ribulose 1,5-bisphosphate carboxylase/oxygenase by active-site Lys166. *Biochemistry* **35**, 13865-13870 (1996).
33. G. G. B. Tcherkez *et al.*, D_2O solvent isotope effects suggest uniform energy barriers in ribulose-1,5-bisphosphate carboxylase/oxygenase catalysis. *Biochemistry* **52**, 869-877 (2013).
34. L. Reinhold, M. Zviman, A. Kaplan, "Inorganic carbon fluxes and photosynthesis in cyanobacteria — a quantitative model" in Progress in Photosynthesis Research: Volume 4 Proceedings of the VIIth International Congress on Photosynthesis Providence, Rhode Island, USA, August 10–15, 1986, J. Biggins, Ed. (Springer Netherlands, Dordrecht, 1987), 10.1007/978-94-017-0519-6_62, pp. 289-296.
35. L. Reinhold, M. Zviman, A. Kaplan, A quantitative model for inorganic carbon fluxes and photosynthesis in cyanobacteria. *Plant Physiol. Biochem.* **27**, 945-954 (1989).
36. L. Reinhold, R. Kosloff, A. Kaplan, A model for inorganic carbon fluxes and photosynthesis in cyanobacterial carboxysomes. *Can. J. Bot.* **69**, 984-988 (1991).
37. N. Mangan, M. Brenner, Systems analysis of the CO_2 concentrating mechanism in cyanobacteria. *eLife* 10.7554/eLife.02043, e02043 (2014).
38. N. M. Mangan, A. Flamholz, R. D. Hood, R. Milo, D. F. Savage, pH determines the energetic efficiency of the cyanobacterial CO_2 concentrating mechanism. *Proc. Natl. Acad. Sci. USA* **113**, E5354–E5362 (2016).
39. M. R. Badger, G. D. Price, J. W. Yu, Selection and analysis of mutants of the CO_2-concentrating mechanism in cyanobacteria. *Can. J. Bot.* **69**, 974-983 (1991).
40. J. M. McGrath, S. P. Long, Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. *Plant Physiol.* **164**, 2247-2261 (2014).
41. G. D. Price, M. R. Badger, Isolation and characterization of high CO_2-requiring-mutants of the cyanobacterium *Synechococcus* PCC7942 : two phenotypes that accumulate inorganic carbon but are apparently unable to generate CO_2 within the carboxysome. *Plant Physiol.* **91**, 514-525 (1989).

42. G. D. Price, M. R. Badger, 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).
43. B. M. Long *et al.*, Carboxysome encapsulation of the CO_2-fixing enzyme Rubisco in tobacco chloroplasts. *Nat. Commun.* **9** (2018).
44. B. B. Menon, S. Heinhorst, J. M. Shively, G. C. Cannon, The carboxysome shell is permeable to protons. *J. Bacteriol.* **192**, 5881-5886 (2010).
45. P. Mahinthichaichan, D. M. Morris, Y. Wang, G. J. Jensen, E. Tajkhorshid, Selective permeability of carboxysome shell pores to anionic molecules. *J. Phys. Chem. B* **122**, 9110-9118 (2018).
46. A. Hassanali, F. Giberti, J. Cuny, T. D. Kühne, M. Parrinello, Proton transfer through the water gossamer. *Proc. Natl. Acad. Sci. USA* **110**, 13723-13728 (2013).
47. M. R. Badger, Kinetic properties of ribulose 1,5-bisphosphate carboxylase/oxygenase from Anabaena variabilis. *Arch. Biochem. Biophys.* **201**, 247-254 (1980).
48. M. R. Badger, G. D. Price, CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J. Exp. Bot.* **54**, 609-622 (2003).
49. F. R. Tabita, S. Satagopan, T. E. Hanson, N. E. Kreel, S. S. Scott, Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *J. Exp. Bot.* **59**, 1515-1524 (2007).
50. M. R. Badger, T. D. Sharkey, S. von Caemmerer, The relationship between steady-state gas exchange of bean leaves and the levels of carbon-reduction-cycle intermediates. *Planta* **160**, 305-313 (1984).
51. J. Spitzer, B. Poolman, How crowded is the prokaryotic cytoplasm? *FEBS Lett.* **587**, 2094-2098 (2013).
52. M. Tabaka, T. Kalwarczyk, J. Szymanski, S. Hou, R. Holyst, The effect of macromolecular crowding on mobility of biomolecules, association kinetics, and gene expression in living cells. *Frontiers in Physics* **2** (2014).
53. A. Akole, J. M. Warner, Model of influenza virus acidification. *PLoS One* **14**, e0214448 (2019).
54. J. F. Viso *et al.*, Multiscale modelization in a small virus: Mechanism of proton channeling and its role in triggering capsid disassembly. *PLoS Comput. Biol.* **14**, e1006082 (2018).
55. M. Eisenhut *et al.*, The plant-like C2 glycolate cycle and the bacterial-like glycerate pathway cooperate in phosphoglycolate metabolism in cyanobacteria. *Plant Physiol.* **142**, 333-342 (2006).
56. M. R. Badger *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).
57. R. M. L. McKay, S. P. Gibbs, K. C. Vaughn, RuBisCo activase is present in the pyrenoid of green algae. *Protoplasma* **162**, 38-45 (1991).
58. K. C. Vaughn, E. O. Campbell, J. Hasegawa, H. A. Owen, K. S. Renzaglia, The pyrenoid is the site of ribulose 1,5-bisphosphate carboxylase/oxygenase accumulation in the hornwort (Bryophyta: Anthocerotae) chloroplast. *Protoplasma* **156**, 117-129 (1990).
59. G. D. Price, D. Sültemeyer, B. Klughammer, M. Ludwig, M. R. Badger, The functioning of the CO_2 concentrating mechanism in several cyanobacterial strains: a review of general physiological characteristics, genes, proteins, and recent advances. *Can. J. Bot.* **76**, 973-1002 (1998).
60. S. Hoops *et al.*, COPASI—a COmplex PAthway SImulator. *Bioinformatics* **22**, 3067-3074 (2006).
61. L. Whitehead, B. M. Long, G. D. Price, M. R. Badger, Comparing the *in vivo* function of α-carboxysomes and β-carboxysomes in two model cyanobacteria. *Plant Physiol.* **165**, 398-411 (2014).
62. M. R. Badger, A. Kaplan, J. A. Berry, Internal inorganic carbon pool of *Chlamydomonas reinhardtii*. *Plant Physiol.* **66**, 407-413 (1980).
63. S. M. Whitney, P. Baldet, G. S. Hudson, T. J. Andrews, Form I Rubiscos from non-green algae are expressed abundantly but not assembled in tobacco chloroplasts. *Plant J.* **26**, 535-547 (2001).
64. A. Occhialini, M. T. Lin, P. J. Andralojc, M. R. Hanson, M. A. J. Parry, Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO_2. *Plant J.* **85**, 148-160 (2016).
65. G. Zhu, R. J. Spreitzer, Directed mutagenesis of chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase: Loop 6 substitutions complement for structural stability but decrease catalytic efficiency. *J. Biol. Chem.* **271**, 18494-18498 (1996).
66. R. J. Spreitzer, S. R. Peddi, S. Satagopan, Phylogenetic engineering at an interface between large and small subunits imparts land-plant kinetic properties to algal Rubisco. *Proc. Natl. Acad. Sci. USA* **102**, 17225-17230 (2005).