PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Atomic view of photosynthetic metabolite permeability pathways and confinement in synthetic carboxysome shells

Daipayan Sarkar[a], Christopher Maffeo[b], Markus Sutter[a,c,d], Aleksei Aksimentiev[b,e], Cheryl A. Kerfeld[a,c,d,f], and Josh V. Vermaas[a,f,1]

Affiliations are included on p. 7.

Edited by Jeffrey C. Cameron, University of Colorado Boulder, Boulder, CO; received February 15, 2024; accepted September 27, 2024 by Editorial Board Member Joseph J. Kieber

Carboxysomes are protein microcompartments found in cyanobacteria, whose shell encapsulates rubisco at the heart of carbon fixation in the Calvin cycle. Carboxysomes are thought to locally concentrate CO_2 in the shell interior to improve rubisco efficiency through selective metabolite permeability, creating a concentrated catalytic center. However, permeability coefficients have not previously been determined for these gases, or for Calvin-cycle intermediates such as bicarbonate (HCO_3^-), 3-phosphoglycerate, or ribulose-1,5-bisphosphate. Starting from a high-resolution cryogenic electron microscopy structure of a synthetic β-carboxysome shell, we perform unbiased all-atom molecular dynamics to track metabolite permeability across the shell. The synthetic carboxysome shell structure, lacking the bacterial microcompartment trimer proteins and encapsulation peptides, is found to have similar permeability coefficients for multiple metabolites, and is not selectively permeable to HCO_3^- relative to CO_2. To resolve how these comparable permeabilities can be reconciled with the clear role of the carboxysome in the CO_2-concentrating mechanism in cyanobacteria, complementary atomic-resolution Brownian Dynamics simulations estimate the mean first passage time for CO_2 assimilation in a crowded model carboxysome. Despite a relatively high CO_2 permeability of approximately 10^{-2} cm/s across the carboxysome shell, the shell proteins reflect enough CO_2 back toward rubisco that 2,650 CO_2 molecules can be fixed by rubisco for every 1 CO_2 molecule that escapes under typical conditions. The permeabilities determined from all-atom molecular simulation are key inputs into flux modeling, and the insight gained into carbon fixation can facilitate the engineering of carboxysomes and other bacterial microcompartments for multiple applications.

photosynthesis | permeability | bacterial microcompartments | carboxysomes | molecular simulation

Significance

Rubisco activity is essential for all life on earth, capturing atmospheric CO_2 and incorporating it into photosynthesis-driven metabolism. Cyanobacteria encapsulate rubisco into bacterial microcompartments called carboxysomes to elevate the local CO_2 concentration and enhance rubisco kinetics. The permeability across carboxysome shells is a key parameter to understand the CO_2 concentration mechanism in cyanobacteria, as it limits the concentration gradient between the inside and outside of the carboxysome. Through molecular simulation, we track the motion of photosynthetic metabolites across the synthetic carboxysome shell to determine a permeability coefficient, and further compare them to the rate of CO_2 leakage and turnover within model carboxysomes. This mechanistic insight is paramount to design and engineer such organelles for carbon fixation, bioenergy, and sustainability applications.

Carbon fixation in an oxygen-rich atmosphere is difficult. Rubisco, the key enzyme that adds a single CO_2 onto Ribulose 1,5-bisphosphate (RuBP) within the Calvin-cycle, also catalyzes RuBP reactivity with oxygen through photorespiration that results in a net loss of energy (1, 2). Thus, photosynthetic organisms such as cyanobacteria have evolved CO_2-concentration mechanisms (CCM) to grow in high oxygen conditions. The CCM tilts the local gas concentration near rubisco toward CO_2 rather than the O_2 abundant in the atmosphere (3). In cyanobacteria, a bacterial microcompartment (BMC) (4) known as the carboxysome forms the center of the CCM, encapsulating rubisco and carbonic anhydrase to create a locally high concentration of CO_2 that enables efficient carbon fixation (3, 5).

The carboxysome, and microcompartments generally, are bounded by protein components that oligomerize to form subassemblies with hexagonal or pentagonal symmetry that further assemble into complete shells with icosahedral symmetry (4, 7–13), similar in scale to some viruses. Since rubisco is encapsulated within the carboxysome (14), but enzymes that catalyze other Calvin-cycle reactions are not (Fig. 1), the carboxysome shell must simultaneously be intact to facilitate rubisco localization while also being permeable to the reactants and products for rubisco. Ideally, the carboxysome would concentrate CO_2 over O_2, reducing the nonproductive oxygenation reaction/photorespiration. Recent structural studies clearly demonstrate that carboxysomes remain intact, sometimes over several cycles of cell division (14, 15). However, the carboxysome shell permeability for individual metabolites has remained elusive (16–18) with multiple proposed models and assumptions (19–21). Missing quantitative permeation coefficients hinder efforts to better understand the limits to photosynthesis and develop strategies to further improve its efficiency or incorporate it into new contexts (22, 23).

The authors declare no competing interest.

This article is a PNAS Direct Submission. J.C.C. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: vermaasj@msu.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2402277121/-/DCSupplemental.

Published November 1, 2024.

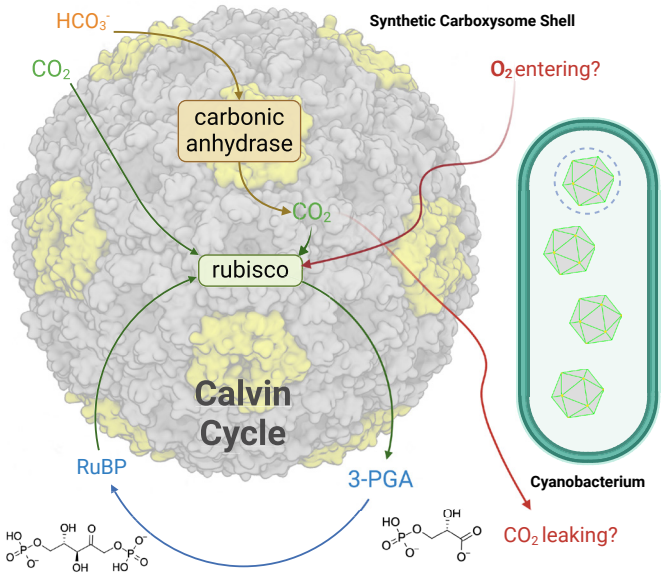

Fig. 1. Schematic of the cyanobacterial carboxysome and its function within the Calvin cycle. The carboxysome model used to represent the dashed cyanobacterial carboxysome is based on the synthetic carboxysome shell [PDB: 6OWG (6)]. In cryogenic electron microscopy (cryo-EM) atomic model, gray indicates CcmK hexamers, yellow CcmL pentamers; Ribulose 1,5-bisphosphate (RuBP) would enter the carboxysome, and bind to the rubisco enzyme. The RuBP can react either with CO_2 or O_2 that diffuses in from the cytosol, or with CO_2 that is generated by carbonic anhydrase (CA) activity. Rubisco converts RuBP into 3-phosphoglycerate (3-PGA), which must traverse the carboxysome shell again to continue the Calvin cycle.

Shell permeability coefficients are a key input to existing systems biology models that guide our understanding of cyanobacterial CCM function and engineering approaches (19, 20). Permeability is difficult to measure based on the rapid turnover within photosynthesis and the difficulty of probing the interior of the carboxysome, and so for simplicity, systems biology approaches have treated permeability as being equal to all metabolites, or an imposed parameter (19–21). Molecular simulation has revealed the free energy barriers to permeation for multiple photosynthetic metabolites and ions along the pore in the hexameric shell protein that makes up the majority of the carboxysome, determining that the energetic barriers were decidedly unequal, but stopped short of determining permeability coefficients (24, 25). Our approach is to conduct multimillion atom molecular dynamics (MD) simulation of a β-carboxysome synthetic shell (Movie S1) (PDB: 6OWG) (6) to explicitly track photosynthetic metabolites crossing the synthetic shell. While the current synthetic shell lacks the BMC timer protein CcmO, it was present in the operon used to express the shells (26), yet the protein was not present in the shell structure—which could be considered consistent with it being a minor component(27). Based on the rate of shell crossing events observed over 1.5 μs of simulation per metabolite and our known metabolite concentration, we can estimate the permeability coefficient directly from simulation (28). By tracking hundreds of transition events across the shell, we find that shell permeability coefficients for most metabolites, including CO_2 and O_2, are on the order of 10^{-2} cm/s.

Naturally, the high permeability for CO_2 measured from our simulations raises the question of how leaky the carboxysome is to CO_2? To address this question, we use Atomic Resolution Brownian Dynamics (ARBD) (29) to simulate CO_2 diffusion within a crowded carboxysome. Based on geometry considerations alone, we find that a CO_2 molecule is likely to find a rubisco active site prior to hitting the edge of a real carboxysome. When combining these results with estimates for how often CO_2 is reflected back into the carboxysome interior by shell proteins, we estimate that 2,650 CO_2 molecules can be fixed by rubisco for every 1 molecule that escapes the carboxysome. Thus, while our high CO_2 permeability estimates suggest that a CO_2 gradient cannot be maintained long term in the absence of enzymatic activity, replenishing CO_2 within the carboxysome from soluble bicarbonate via carbonic anhydrase activity is sufficient to keep the local CO_2 concentration around rubisco high within the carboxysome.

Results

Quantifying Metabolite Transition Events and Permeation Pathways Across Synthetic Carboxysome Shells. To quantify the permeability of small molecular compounds - CO_2, O_2, HCO_3^-, 3-phosphoglycerate (3-PGA), RuBP, we track permeation events for these metabolites across a synthetic carboxysome shell using classical MD simulations. Individual all-atom MD simulations are performed for each metabolite at high concentration (500 mM) to increase sampling of observed transition events for an aggregate of 1.5 μs per metabolite. To monitor for permeation events, we track the radial position for each metabolite molecule relative to the carboxysome center, and monitor trajectory traces for the radial position as a function of time (Fig. 2). If the trace for a specific small molecule can be found on both sides of the carboxysome shell, it must have crossed the shell, and is a permeation event.

The total permeation events for photosynthetic metabolites are aggregated in Fig. 3, with animations for individual metabolites through a single hexamer shown in Movies S2–S6. Gaseous compounds (CO_2 and O_2) have many more observed transitions in our simulations than the anionic compounds, suggesting that gases are more permeable across the synthetic carboxysome shell than other anionic metabolites. Transition events for the gases are also found to frequently track along the interface between the hexameric and pentameric proteins that compose

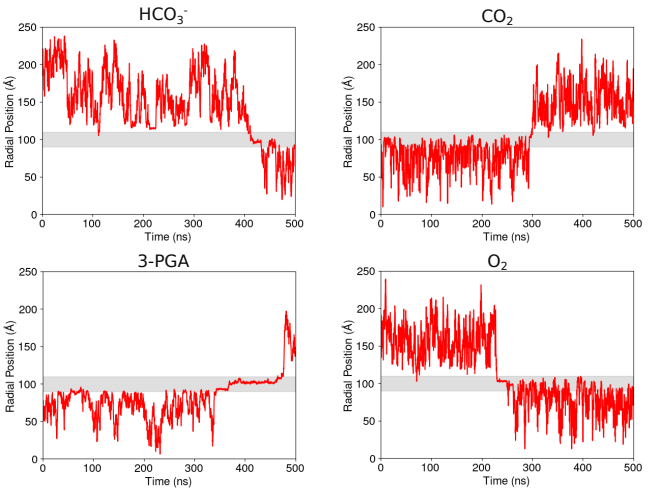

Fig. 2. Time series trace for example metabolite transitions across the carboxysome shell. The solid red line indicates the current radial position of a specific copy of the metabolite in our simulation. By following the position compared with the extent of the carboxysome shell proteins (gray region), we can readily count transition events where the metabolite crossed the carboxysome shell during our simulation.

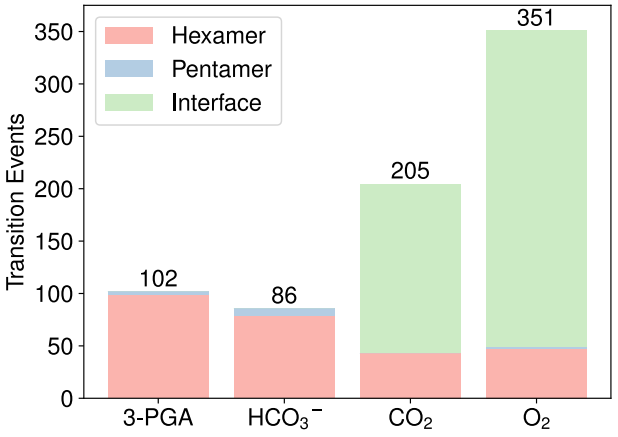

Fig. 3. Aggregated metabolite transition events, broken down by metabolite and pathway. Three pathways are possible for each metabolite, transitioning through the hexamer pore (red), the pentamer pore (blue), or the interface between hexagonal or pentagonal protein subunits.

the synthetic carboxysome shell model. By contrast, small anionic metabolites such as HCO_3^- and 3-PGA always transit through the central pores in the hexamers and pentamers. Considering the number of CcmK hexamers (30) and CcmL pentamers (12) assembled together to build the icosahdedral shell (Fig. 1), we would expect a 5:2 ratio of hexamer versus dimer transit events if there was no preference between the pores. However, the heavily skewed ratio observed in Fig. 3 highlights that the negatively charged compounds transit preferentially through the abundant hexameric pores that would dot the carboxysome surface. The transit through the hexameric pore for negatively charged compounds is facilitated by the positively charged KIGT motif, instead of the conserved KIGS motif (30) found in the central pore formed by CcmK hexamers (Fig. 4).

RuBP transits the fewest times within our simulations. We observe only six permeation events overall from 1.5 μs of MD simulation (Fig. 3), as many RuBP molecules get trapped in the hexamer central pore (Fig. 4*A*). One such transition of RuBP through a BMC-H hexamer protein is demonstrated in Movie S7. The strong charge–charge interactions between the RuBP and the KIGT motif increases RuBP residence time within the BMC-H hexamer pore. While HCO_3^- and 3-PGA also have substantial charge, RuBP uniquely has strong interactions on both ends of the molecule. As a result, RuBP frequently adopts conformations within the pore where the phosphates make strong contact with K36 (Fig. 4*B*), delaying RuBP transport across the hexamer protein. Our results suggest that RuBP has a substantially smaller permeability than other metabolites through the hexameric pore.

Quantifying Permeability and Pathways for Photosynthetic Metabolites. By Fick's law, the net metabolite flux (J) is the product of the metabolite permeability (P) we are seeking to determine, the carboxysome shell surface area (A), and the concentration difference across the shell (ΔC) (28):

$$J = P \times A \times \Delta C \qquad [1]$$

The transition events observed in our simulation can be translated into a permeability coefficient, using the same methodology worked out for lipid bilayers (28). Briefly, the permeability is related to rate at which small molecules cross the partition ($r = \frac{\text{transition events}}{\text{time}}$), and the concentration of the metabolite in water (c_w) (28):

$$P = \frac{r}{2c_w} = \frac{\text{transition events}}{2c_w \text{time}} \qquad [2]$$

Taking the results from Fig. 3, we use Eq. **2** to tally the transition events, simulation time, and derived permeability coefficients for metabolite transport averaged across the full shell in Table 1.

The key finding is that the permeability coefficients for most small molecules are $\approx 10^{-2}$ cm/s (Table 1). This magnitude is consistent with optimal carboxysome permeability predictions from systems model (19), and two orders of magnitude higher than the optimal permeability predicted by a revised pH-aware CCM model (20). However, even the pH-aware model supports CCM activity over a range of permeabilities, including those on the order of 10^{-2} cm/s (20). The exceptions to this are the roughly 10 to 100 fold lower permeability predicted for RuBP, in line with assumed permeabilities between 10^{-3} cm/s and 10^{-4} cm/s in other models (31), and the higher permeability for water. Since water forms continuous channels through the CcmK pore (*SI Appendix*, Fig. S1), and protons are known to permeate across carboxysomes (32), a high permeability for water (Table 1) is natural, and suggests that there can be no osmotic gradient across the shell.

However, the high permeability for CO_2 and O_2 is a surprise, as prior simulations based on enhanced sampling methods predicted that there should be a barrier for CO_2 and O_2 to cross through the hexameric pore compared with larger photosynthetic metabolites (24, 25). Having hints from Fig. 3 that the pathways the different metabolites follow are distinct, we invert the observed probability distribution for the different metabolites along the radial and pore axis dimensions to create a free energy profile to compare with previous results (Fig. 5). As had been observed previously, 3-PGA and HCO_3^- have a free energy minimum of approximately −2 kcal/mol within the pore near the constriction site, where $Z \sim -10$ Å, and the radius is near zero. By contrast, CO_2 and O_2 are not found to sample this region of the free energy profile, and exhibit a free energy

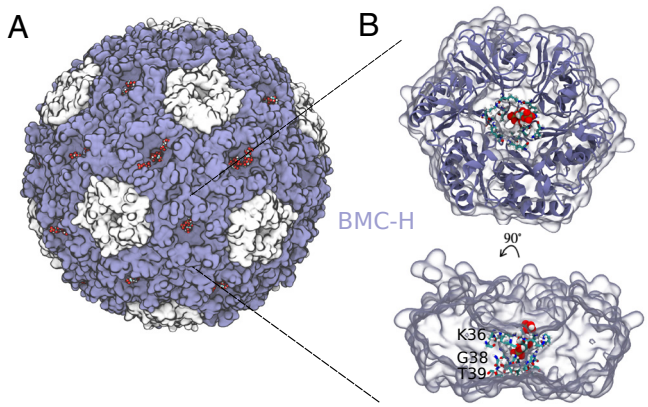

Fig. 4. RuBP routinely gets stuck within the pore of a CcmK hexamer. (*A*) Illustrates the different RuBP molecules (red oxygen, cyan carbon, white hydrogen) stuck at the pore of the hexamer protein (ice blue) after 500 ns of MD simulation. (*B*) illustrates the interaction between the negatively charged RuBP molecule and the positively charged pore lysine residues within the KIGT motif in the synthetic CcmK hexamer. Illustrated are the *Top*, where the BMC-H hexamer is illustrated in ribbon and by a transparent glass gray surface. Also illustrated is the lateral view (90°, where only the transparent surface is illustrated for clarity; indicating RuBP molecule trapped inside the BMC-H pore.

Table 1. Permeability coefficients of metabolites across synthetic β-carboxysome shell, using the formalism introduced in Eq. 2

Molecule	Transitions	Time (ns)	P (cm/s)
O_2	351	1,500	0.026 ± 0.008
CO_2	205	1,500	0.015 ± 0.003
HCO_3^-	86	1,500	0.0063 ± 0.0002
3-PGA	102	1,500	0.0074 ± 0.0009
RuBP	6	1,500	0.00044 ± 0.0002
H_2O	74,116	700	0.11 ± 0.002

For photosynthetic metabolites, the concentration in water was assumed to be 500 mM. For H_2O, whose transitions were measured from a shorter set of test simulations, we assumed a concentration of 55 M. Here, the mean permeability is calculated by taking the average of the transition events observed in three independent simulations, and the SE is computed from the SD over the 3 independent simulations.

maximum near the pore center. Since previous simulations were operating with an isolated hexamer, and restrained the sampling for small molecules to within a cylinder with a 15 Å radius, the maximum we observe is consistent with prior free energy profiles (24, 25). Based on the isolated hexamer simulation design, prior work did not sample the lower energy path for gas transit across the shell at the interface between hexameric tiles far from the central pore. The picture that emerges from Fig. 5 is that the primary pore within the hexamer is selective to larger metabolites, but that gases transit wherever there is a momentary pocket.

Assessing CO_2 Leakiness in a Crowded Carboxysome Context. The surprising corollary to the high CO_2 permeability in Table 1 is that once HCO_3^- is converted by carbonic anhydrase into CO_2 (Fig. 1), the CO_2 may leak from the carboxysome prior to CO_2 capture and fixation by rubisco. To estimate carboxysome leakiness for CO_2, we developed a model for a full carboxysome packed with rubisco (Fig. 6*A*) based on recent cryo-electron tomography (cryo-ET) experiments for marine cyanobacteria (33), α- (10, 11, 34) and β-carboxysomes (35). Our intent is to track the statistics of how often CO_2 will hit the carboxysome shell first or will instead come close to the rubisco active site. Since rubisco has multiple subunits, and over a hundred rubisco are densely packed within any real carboxysome, treating this system with atomistic MD as we did for the synthetic system is infeasible. Thus, we use ARBD methodology (29) to represent each of the 160 rubiscos as an interacting grid that can exert force on nearby CO_2 particles. The rubisco structure used in ARBD simulations should be in the *apo* or substrate-free form to allow CO_2 to access the well-conserved active site, and hence the PDB:7JN4 (36) structure from *Chlamydomonas reinhardtii* was used in place of similar models from other organisms (*SI Appendix*, Fig. S2). The *apo* rubisco model is then packed using inter-rubisco distance observed in cryo-electron tomograms obtained in marine cyanobacteria (33) to approximate the densely packed carboxysome interior.

Four 2 μs long ARBD simulations were performed to calculate the rate for a CO_2 molecule to bind to a rubisco active site (k_b) or to exit the shell (k_e), with the animation for a single simulation given in Movie S8. The atomic resolution offered by ARBD allows individual CO_2 molecules to be tracked relative to the well-conserved (*SI Appendix*, Fig. S2) rubisco active site (37), in our case K201, H294, and H327 shown to be involved in catalysis (38). In analyzing the ARBD trajectories, we define that a CO_2 molecule is bound once the molecule comes within 10 Å of the C_α atoms in this catalytic triad. The larger cutoff accounts for sticky spots on the van der Waals grid for rubisco once it has been discretized, the size of the residues, as well as conformational changes that might take place within rubisco that ARBD cannot capture. A CO_2 can also hit the bounding surface that is assumed to represent the carboxysome shell proteins. The ratio between these two events is related to the kinetic rate at which exiting and binding occur, as given by the simple Markov model derived in *SI Appendix*, section D.

From Fig. 6*E*, we see that the number of CO_2 molecules that bind to rubisco far exceed the number of CO_2 molecules that hit the carboxysome surface. However, the numbers are comparable, with the binding rate of CO_2 to rubisco only 2.3 ($\pm$ 0.05) times higher than the rate at which CO_2 hits the carboxysome surface. If this were the final number, appreciable CO_2 would leak out of the carboxysome during carbon fixation, and represents the scenario where the rubisco core was intact but the shell was absent. From Fig. 2 derived from our atomic simulations, we see that a typical CO_2 molecule is reflected by the shell many times before leaving the shell. Indeed, not shown in Fig. 2 are the many molecules that strike the carboxysome surface and never cross during our relatively short atomistic simulations. If we focus only on CO_2 molecules that are initially inside the synthetic carboxysome in our atomistic simulations, 72 will cross the carboxysome during simulation, compared with 82,797 recorded events where the CO_2 was at the carboxysome surface, and was reflected back at least 5 Å back into the interior. The rate of CO_2 leaving the carboxysome is therefore substantially slower than the rate of merely hitting the surface, dramatically increasing the ratio between carbons fixed by rubisco and carbons that leave the carboxysome. By our estimates, 2,650 CO_2 will be bound and fixed by rubisco for every 1 CO_2 molecule that manages to escape the carboxysome (*SI Appendix*, Eq. S5). Thus, while the carboxysome shell is permeable to CO_2, the effective diffusion length a CO_2 must travel before leaving the shell is amplified significantly by reflections off the shell surface, making it exceedingly likely a given molecule of CO_2 inside a carboxysome will find a primed rubisco.

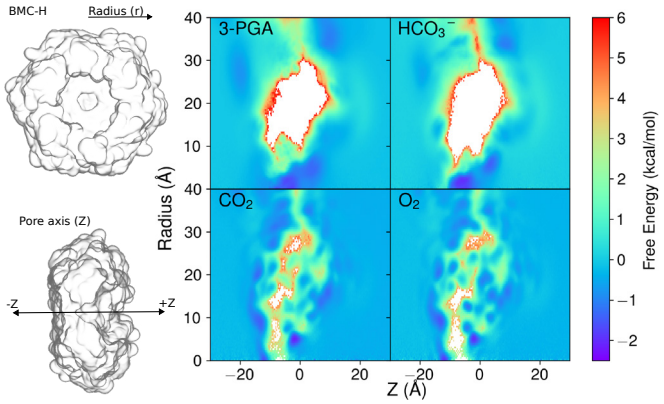

Fig. 5. Two-dimensional free energy surfaces for photosynthetic metabolite transport across the CcmK hexamer in the synthetic carboxysome shell. Note that the two dimensions here are the position along the radial distance away from the pore center (r, *Upper Right*) and the pore normal axis (Z, *Lower Right*). The free energies are directly computed by inverting the probability distribution for metabolite centers and are corrected for the larger surface area at increasing radius. In locations where no sampling has taken place, the probability distribution is not invertible, and so the free energy is represented as a white color.

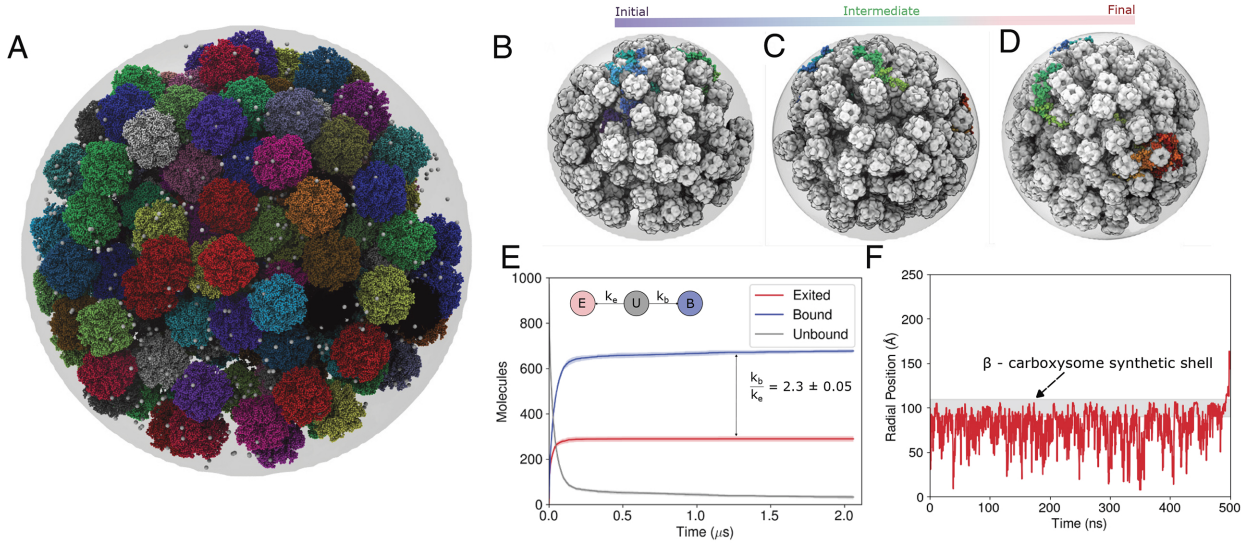

Fig. 6. (*A*) is the initial model for a crowded carboxysome interior, where each L8S8 (large and small subunits) rubisco *apo* enzyme is represented in a different color, and white spheres are the initial locations for CO_2 molecules. The bounding sphere (gray) represents the carboxysome shell. By tracking CO_2 molecules as they diffuse within the imposed bounding potential (*B*–*D*), where specific CO_2 molecules are followed as a color-changing timecourse, we can use these results to develop a simple Markov model (*SI Appendix*) to estimate the rates for exiting the carboxysome (k_e) or binding to rubisco (k_b). The time series for how many CO_2 molecules are unbound, hit the bounding sphere and attempt to exit, or bind to rubisco are reported in (*E*). (*F*) highlights a single trace for a CO_2 molecule within our trajectories, emphasizing that CO_2 may not transit across the carboxysome on the first attempt, and can be reflected multiple times off the carboxysome shell before exiting.

Discussion

The critical finding from this work is that carboxysome shells do not show large variation in permeability coefficient depending on the metabolite (20, 31, 39). Charged, uncharged, gaseous, or water-soluble, the permeability coefficients we determined through atomistic simulation across a synthetic carboxysome model are all within two orders of magnitude (Table 1). This is in stark contrast to lipid bilayers, where charge alone can vary the permeability by six orders of magnitude (40), and the dynamic range for passive permeability depending on the metabolite can be very large (28, 40). Thus rather than invoking selective permeability to generate a CCM that prefers CO_2 over O_2, to trap CO_2 internally, as has been hypothesized in prior literature (3, 16), we find that the biophysical principles behind CCM shell function are ingenious.

In the cyanobacterial cytosol HCO_3^- will be the primary form of carbon that enters the carboxysome (20, 31, 39, 41–44). The concentration gradient that drives HCO_3^- flux into the carboxysome would be maintained by the carbonic anhydrase inside the carboxysome shell that locally generates CO_2 (42, 45–47). Most of the CO_2 that is generated by carbonic anhydrase will quickly find a rubisco active site on a microsecond timescale (Fig. 6), with kinetics effectively limited by how quickly generated CO_2 can diffuse to an active site. The diffusion is assisted by the multiple reflections from the shell surface that a CO_2 particle can experience before it finds a pathway to escape the carboxysome, creating an effective CCM even in a nominally leaky carboxysome shell.

The primary role of the carboxysome is to enhance the carboxylation over oxygenation reactions by rubisco, and hence the local concentration of O_2 inside the carboxysome is expected to be lower than the CO_2 concentration under most physiological conditions. Residual O_2 permeation across the carboxysome shell results in the limited photorespiration. The cytosolic O_2 concentration will be relatively low, as O_2 preferentially partitions to lipid membranes, and is generated during photosynthesis by photosystem II that is embedded within the thylakoid membrane (48). The carboxysome shell proteins are just as effective at reflecting O_2 entering from the cytosol as they are in retaining CO_2. The presence of the shell would act as a physical barrier to lower the O_2 availability near the rubisco active site compared to CO_2, thereby minimizing photorespiration. This is analogous to the reduction in O_2 availability within protein nanocages, where the proteins inhibit O_2 access to the nanocage interior (49). To be clear, as stated some O_2 will permeate across the carboxysome shell, and if it does so, we would anticipate that rubisco will push RuBP toward the photorespiratory pathway and 2-phosphoglycolate production. The ability for some O_2 to enter the carboxysome is consistent with previous experimental findings that intact photorespiration pathways are required for normal growth in cyanobacteria under standard atmospheric conditions (50).

Our results also highlight the engineering principles that dictate efficient carboxysome construction, which have been previously explored theoretically (51). The number of shell proteins needed to encapsulate a catalytic cluster of a specific radius grows as the radius squared, while the encapsulated enzyme count grows as the radius cubed. As protein synthesis costs energy, large carboxysomes should be favored to amortize the energetic cost of carboxysome biogenesis, by packing as many rubiscos as possible into a single carboxysome (51). However, since CO_2 will encounter a rubisco active site within a microsecond (Fig. 6), and the diffusion coefficient for CO_2 is 1,670 $nm^2/\mu s$ (52), the average pathlength a CO_2 can travel is only about 40 nm or so, with substantial variation due to stochastic motion. This means that there are diminishing returns if the distance between a carbonic anhydrase and a rubisco active site substantially exceeds that dimension, since distant rubiscos would not reliably be supplied with CO_2, and would effectively be a waste of resources to produce. Typical α-carboxysomes are measured to be about 100 to 150 nm in diameter (10, 11, 53, 54), which would seem to be sized appropriately if carbonic anhydrases were localized to the carboxysome periphery and CO_2 were to diffuse into

the interior. However, β-carboxysomes are much larger, with diameters up to 600 nm, suggesting that if carbonic anhydrases were only on the periphery, interior rubiscos would be starved of CO_2. Structural data from cryo-electron and super-resolution microscopy in α and β- carboxysomes respectively, suggests that CA enzymes inside the shell closely interact with rubisco, to form a supercomplex (47, 55). Experimentally in the case of β- carboxysomes the rubisco-CA localization is measured to a sub-shell layer of ~5 to 6 nm beneath the outer shell surface (55), consistent with a more uniform distribution of carbonic anhydrases. Integrating these details into the ARBD model (Fig. 6*A*) requires detailed knowledge for the number of CA enzymes inside the carboxysome shell and their distribution within an assembled microcompartment, which are areas of active research beyond our current scope.

RuBP Concentration Is High to Match Rubisco Stoichiometry. One of the big surprises in our results is how few permeation events we observe for RuBP compared with the other photosynthetic metabolites (Table 1). Since RuBP and CO_2 are used in a 1:1 stoichiometric ratio within the Calvin cycle, that implies that at steady state the fluxes for RuBP and HCO_3^- should be equal for continued catalysis. From the structure of Eq. **1**, the smaller permeability for RuBP relative to CO_2 or HCO_3^- implies that the concentration gradient needed to drive transport across the shell must be larger.

There are two reasonable ways to create a larger gradient for RuBP. If the external concentrations were similar, a smaller concentration difference for CO_2 would mean a high internal concentration of CO_2 equivalents inside the carboxysome. This seems unlikely, since carbon fixation accelerates when cyanobacteria are grown in a CO_2-rich environment (56, 57), and a high internal CO_2 concentration will inhibit carbonic anhydrase turnover. What we consider more plausible is that the cytosolic metabolite pool for RuBP is instead high and that the intracarboxysomal concentrations of both metabolites are similar in scale. In vivo measurements report millimolar concentration of RuBP in α and β-carboxysomes (58), which are consistent with estimated intracarboxysome concentrations for CO_2 in systems biology models (19, 20).

The missing piece of the RuBP puzzle is of course the other carboxysome shell components not featured in this study. The minimal model we simulate here (Fig. 1) contains two types of shell proteins, whereas all carboxysomes containing organisms encode at least one trimeric shell protein with hexagonal symmetry. Comparing available cryogenic electron microscopy (cryo-EM) structures for the BMC-trimer and hexamer shell proteins (4, 7, 59, 60), the stacked trimer subunit has a larger, gated pore that may facilitate RuBP transport. Future study will be required to assess the relative permeability between the two components, as the trimer is typically far less abundant on the carboxysome surface compared to the hexamer proteins.

pH Effects on the CCM. A compelling hypothesis in the wider carboxysome literature is that there is a pH gradient across the carboxysome shell (18) and that pH plays a role in carboxysome regulation, with lower intracarboxysomal pHs associated with higher metabolic activity (61). Lower pH supports robust carbonic anhydrase activity, which combines free protons (H^+) from solution with bicarbonate (HCO_3^-) to form CO_2 and H_2O. While it remains elusive and challenging to determine the role of pH in a carboxysome, systems biology predictions with a pH-aware model (20) shed light on the optimal interior pH for carboxysome function relative to the cyanobacterial cytosol pH (~8) (20, 31). In contrast to earlier models, where the optimal carboxysome permeability was between 10^{-2} to 10^{-3} cm/s (19), the pH-aware model has the optimal carboxysome permeability at 10^{-4} cm/s (20), however the revised model does support carbon fixation for CO_2 permeability of 10^{-2} cm/s. Next, comparing RuBP transport across the synthetic shell, from MD simulations, we derive the permeability to be 0.00044 cm/s (Table 1), which is similar to the optimal range reported by Long et al (31) for a carboxysome of radius 50 nm and compartment pH between 7 to 8.

Where we think our MD may really inform the pH discussion is in noting that many water molecules can readily transit through the carboxysome shell (Table 1 and *SI Appendix*, Fig. S1). Hydrogen bond arrangements that could conduct a proton across the shell through a Grothuss-type mechanism postulated previously (31) are trivial to identify from our trajectories. Thus any pH gradient cannot be manufactured by active transport processes, as the gradient would be run down by these Grothuss proton movements, and lead to proton movements observed in prior experiments (32) and systems modeling predictions (31). Instead, based on the electrostatic surface analysis of the synthetic shell reveals a more negatively charged exterior relative to the interior (6), and is a lever evolution can turn to optimize the interior pH of a carboxysome by modifying the Nernst potential relative to the cytosol, as well as the metabolite permeability. Mutation is the clearest path forward to modify any potential pH gradient, as the carboxysome is too small for the number of free protons to be a useful indicator of pH. An approximately spherical 120 nm diameter solvated carboxysome will hold only 30.3 million water molecules,* thus even an empty shell will have fewer than 1 H_3O^+ in its interior at pH 7. Instead, the pH environment within the carboxysome is fixed by the protonation equilibria of the shell and its cargo, which in our fixed charge simulation models take the predominant protonation state at pH 7.

Rubisco Packing Considerations. In the in-silico carboxysome shell model we developed (Fig. 6) used PACKMOL (62) to randomly insert 160 copies of rubisco inside a spherical volume with a 95 nm approximate diameter. This would represent a carboxysome on the smaller side, chosen for computational efficiency, and as a worst-case scenario with abundant opportunities for CO_2 to hit the bounding surface. Assuming a spherical carboxysome, this works out to 2,800 nm^3 per rubisco. From two different high-resolution cryo-ET experiments, the authors report about 200 to 250 rubisco complexes are encapsulated in a lattice configuration inside an α - carboxysome of diameter ~120 nm (11), or 3600 nm^3 per rubisco. Similarly, analyzing tomograms obtained from rubisco assembly in cyanobium and Halo carboxysomes, the authors quantify about 244 ± 26 and 274 ± 72 respectively, in shells that are 100 to 120 nm wide (10, 53), ranging from 2,100 to 3,300 nm^3 per rubisco. The story is somewhat different from recent *Synechococcus* β - carboxysome tomograms (63), with approximately 1,800 nm^3 per rubisco, the tightest packing seen to this point in the literature.

While our rubisco packing density is reasonable given the experimental variation observed in tomograms, with our packing density in the middle of the range for an α-carboxysome. In the densely packed β-carboxysomes, we would anticipate that the balance between CO_2 egress and catalysis would shift even further toward catalysis from what we measure in our simulations, simply as CO_2 would be even more likely to hit an active site.

* $\frac{4\pi}{3}$ (60nm)^3 × $\frac{1kg}{L}$ × $\frac{mol}{18g}$ × 6.02310^{23} $\frac{molecules}{mol}$ ≈ 30300000 H_2O molecules.

We did not impose any structure on the packing within our model carboxysome that has recently emerged from tomography (10, 11, 53). In our view, ordered structure within the carboxysome would likely not substantially change the diffusion behavior for photosynthetic metabolites, especially in such densely crowded microcompartments, and that ordered conformations may arise from dynamics at longer timescales than our simulations can capture.

Current Limitations

As mentioned, the cryo-EM structure of the synthetic β-carboxysome shell (PDB: 6OWG) lacks the BMC trimer protein ccmO, encapsulation peptides (13, 64, 65), and lacks encapsulating enzymes (rubisco and CA), making the overall size of the synthetic shell smaller compared to that of a naturally occurring α- (11, 66) or β- (63) carboxysome. Lack of these proteins has an effect in the topology of the shell, which is more spherical instead of icosahedral. Such topological variations can result in different mechanical properties for the BMC shell, which may modulate the permeability. Continuum modeling by varying the bending rigidity ratios has predicted the surface patterns on a three component (BMC-H, BMC-P, and BMC-T) shell (67). However, atomic force microscopy experiments of viral capsids of size 30 nm and icosahedral symmetry have demonstrated that the Föppl-von Kármán thin-shell elasticity model is not sufficient to quantify the mechanical responses, and a coarse-grained molecular dynamics model was found to be better suited to predict the force response curves (68). The interactions between the protein subunits and the presence of the encapsulation peptides and inner cargo proteins can have an impact on the shell permeability—especially for gas molecules which indicate to transport through the protein–protein interfaces for the experimentally determined synthetic carboxysome structure (6OWG) (6).

Conclusion

The primary finding from our simulation results is that the carboxysome shells are not adept at excluding gas molecules from transiting the shell. If anything, gases permeate more readily at an equal concentration gradient than photosynthetic metabolites, particularly RuBP, which has the slowest permeability of the compounds tested. However, by concentrating rubisco and carbonic anhydrase, and substantially increasing the path length that gases must travel to leave or enter the shell, the carboxysome creates an efficient CCM from nominally leaky components. Indeed, the greater than 2,600:1 ratio between fixed and lost CO_2 in a typical carboxysome is an astounding feat of engineering by evolution.

Materials and Methods

Atomistic MD Simulation. The synthetic T = 4 carboxysome cryo-EM structure (6) was solvated in a water cube with side length 29 nm and prepared for simulation (Movie S1) with the CHARMM36m force field for proteins (69) using Visual Molecular Dynamics (VMD) (70). To track metabolite permeation, we add 500 mM concentrations of CO_2, O_2, HCO_3^-, 3-PGA, and RuBP to independent simulation systems (*SI Appendix*, Fig. S3). After a brief equilibration period utilizing NAMD 2.14 (71), each simulation was converted using TopoGromacs (72) to GROMACS (73) for production simulations, which were 3 replicates for 500 ns each. Particle tracking and analysis was performed with python-enabled VMD (70), extensively using numpy (74), scipy (75), and Matplotlib (76) libraries.

Atomic Resolution Brownian Dynamics simulations. Based on existing cryo-electron tomography literature, 160 copies of rubisco [PDB:7JN4 (36)] were packed into a 95 nm diameter sphere using PACKMOL (62), along with 1,000 noninteracting CO_2 particles. For atomic resolution Brownian dynamics, the diffusion parameters required for simulation were determined by HYDROPRO (77), as in prior work. Further details related to preparing grid potentials that drive the interaction between proteins and particles are included as *SI Appendix, Atomic Resolution Brownian Dynamics*. Analysis again used VMD and purpose-built python scripts.

Data, Materials, and Software Availability. Data (inputs and trajectory analysis in machine-readable form) have been deposited in Zenodo (10.5281/zenodo.10632916)(79). Some study data are available: Full trajectories are available upon request, which are not uploaded to Zenodo due to their size.

ACKNOWLEDGMENTS. D.S., M.S., C.A.K., and J.V.V. are supported by the U.S. Department of Energy, Office of Basic Energy Sciences under grant number DE-FG02-91ER20021. The software Atomic Resolution Brownian Dynamics (ARBD) development is supported by a NSF, Office of Advanced Cyberinfrastructure under award number #2311550 to C.M. and A.A. D.S. acknowledges the discussion with Dr. Abhishek Singharoy, Dr. Taras Pogorelov, and Dr. Nandan Haloi for initial discussion related to ARBD. D.S. also acknowledges the discussion with Dr. Leandro Martínez, regarding the use of PACKMOL to prepare the initial system for ARBD simulations. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231 using NERSC award BES-ERCAP0021030. This work used Anvil at Purdue University during the early user period through allocation BIO210061 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support (ACCESS) program (Boerner et al.). ACCESS is supported by NSF grants #2138259, #2138286, #2138307, #2137603, and #2138296. This work was supported in part through computational resources and services provided by the Institute for Cyber-Enabled Research at Michigan State University. Fig. 1 was composed with BioRender.com.

Author affiliations: [a]MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI 48824; [b]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801; [c]Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; [d]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; [e]Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801; and [f]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824

Author contributions: D.S., M.S., C.A.K., and J.V.V. designed research; D.S., C.M., and J.V.V. performed research; C.M. and A.A. contributed new reagents/analytic tools; D.S., C.M., and J.V.V. analyzed data; and D.S., C.M., M.S., A.A., C.A.K., and J.V.V. wrote the paper.

1. T. D. Sharkey, The discovery of rubisco. *J. Exp. Bot.* **74**, 510–519 (2023).
2. F. A. Busch, Photorespiration in the context of Rubisco biochemistry, CO_2 diffusion and metabolism. *Plant J.* **101**, 919–939 (2020).
3. C. A. Huffine, R. Zhao, Y. J. Tang, J. C. Cameron, Role of carboxysomes in cyanobacterial CO_2 assimilation: CO_2 concentrating mechanisms and metabolon implications. *Environ. Microbiol.* **25**, 219–228 (2023).
4. M. Sutter, M. R. Melnicki, F. Schulz, T. Woyke, C. A. Kerfeld, A catalog of the diversity and ubiquity of bacterial microcompartments. *Nat. Commun.* **12**, 3809 (2021).
5. M. R. Badger, CO_2 concentrating mechanisms in cyanobacteria: Molecular components, their diversity and evolution. *J. Exp. Bot.* **54**, 609–622 (2003).
6. M. Sutter *et al.*, Structure of a synthetic β-Carboxysome shell. *Plant Physiol.* **181**, 1050–1058 (2019).
7. F. Cai *et al.*, The structure of CcmP, a tandem bacterial microcompartment domain protein from the β-Carboxysome, forms a subcompartment within a microcompartment. *J. Biol. Chem.* **288**, 16055–16063 (2013).
8. B. J. Greber, M. Sutter, C. A. Kerfeld, The plasticity of molecular interactions governs bacterial microcompartment shell assembly. *Structure* **27**, 749–763.e4 (2019).
9. G. Kalnins *et al.*, Encapsulation mechanisms and structural studies of GRM2 bacterial microcompartment particles. *Nat. Commun.* **11**, 388 (2020).

10. T. Ni *et al*., Structure and assembly of cargo Rubisco in two native α-carboxysomes. *Nat. Commun.* **13**, 4299 (2022).
11. L. A. Metskas *et al*., Rubisco forms a lattice inside alpha-carboxysomes. *Nat. Commun.* **13**, 4863 (2022).
12. T. Ni *et al*., Intrinsically disordered CsoS2 acts as a general molecular thread for α-carboxysome shell assembly. *Nat. Commun.* **14**, 5512 (2023).
13. L. M. Oltrogge, A. W. Chen, T. Chaijarasphong, J. B. Turnšek, D. F. Savage, α-Carboxysome size is controlled by the disordered scaffold protein CsoS2. *Biochemistry* **63**, 219–229 (2024).
14. J. C. Cameron, S. C. Wilson, S. L. Bernstein, C. A. Kerfeld, Biogenesis of a bacterial organelle: The carboxysome assembly pathway. *Cell* **155**, 1131–1140 (2013).
15. N. C. Hill, J. W. Tay, S. Altus, D. M. Bortz, J. C. Cameron, Life cycle of a cyanobacterial carboxysome. *Sci. Adv.* **6**, eaba1269 (2020).
16. C. A. Kerfeld, C. Aussignargues, J. Zarzycki, F. Cai, M. Sutter, Bacterial microcompartments. *Nat. Rev. Microbiol.* **16**, 277–290 (2018).
17. W. B. Carpenter *et al*., Ratiometric sensing of redox environments inside individual carboxysomes trapped in solution. *J. Phys. Chem. Lett.* **13**, 4455–4462 (2022).
18. J. Huang *et al*., Probing the internal pH and permeability of a carboxysome shell. *Biomacromolecules* **23**, 4339–4348 (2022).
19. N. M. Mangan, M. P. Brenner, Systems analysis of the CO_2 concentrating mechanism in cyanobacteria. *eLife* **3**, e02043 (2014).
20. 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. U.S.A.* **113**, E5354–E5362 (2016).
21. C. A. Huffine, L. C. Wheeler, B. Wing, J. C. Cameron, Computational modeling and evolutionary implications of biochemical reactions in bacterial microcompartments. *Curr. Opin. Microbiol.* **65**, 15–23 (2022).
22. J. H. Hennacy, M. C. Jonikas, Prospects for engineering biophysical CO_2 concentrating mechanisms into land plants to enhance yields. *Annu. Rev. Plant Biol.* **71**, 461–485 (2020).
23. L. N. Liu, Advances in the bacterial organelles for CO_2 fixation. *Trend. Microbiol.* **30**, 567–580 (2022).
24. 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).
25. M. Faulkner *et al*., Molecular simulations unravel the molecular principles that mediate selective permeability of carboxysome shell protein. *Sci. Rep.* **10**, 17501 (2020).
26. F. Cai, S. L. Bernstein, S. C. Wilson, C. A. Kerfeld, Production and characterization of synthetic carboxysome shells with incorporated luminal proteins. *Plant Physiol.* **170**, 1868–1877 (2016).
27. Y. Sun, A. J. M. Wollman, F. Huang, M. C. Leake, L. N. Liu, Single-organelle quantification reveals stoichiometric and structural variability of carboxysomes dependent on the environment. *Plant Cell* **31**, 1648–1664 (2019).
28. R. M. Venable, A. Krämer, R. W. Pastor, Molecular dynamics simulations of membrane permeability. *Chem. Rev.* **119**, 5954–5997 (2019).
29. J. Comer, A. Aksimentiev, Predicting the DNA sequence dependence of nanopore ion current using atomic-resolution brownian dynamics. *J. Phys. Chem. C* **116**, 3376–3393 (2012).
30. C. A. Kerfeld *et al*., Protein structures forming the shell of primitive bacterial organelles. *Science* **309**, 936–938 (2005).
31. B. M. Long, B. Förster, S. B. Pulsford, G. D. Price, M. R. Badger, Rubisco proton production can drive the elevation of CO_2 within condensates and carboxysomes. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2014406118 (2021).
32. B. B. Menon, S. Heinhorst, J. M. Shively, G. C. Cannon, The carboxysome shell is permeable to protons. *J. Bacteriol.* **192**, 5881–5886 (2010).
33. W. Dai *et al*., Visualizing individual RuBisCO and its assembly into carboxysomes in marine cyanobacteria by cryo-electron tomography. *J. Mol. Biol.* **430**, 4156–4167 (2018).
34. L. Evans, M. K. Cameron, P. Tiwary, Computing committors via Mahalanobis diffusion maps with enhanced sampling data. *J. Chem. Phys.* **157**, 214107 (2022).
35. H. Wang *et al*., Rubisco condensate formation by CcmM in β-carboxysome biogenesis. *Nature* **566**, 131–135 (2019).
36. S. He *et al*., The structural basis of Rubisco phase separation in the pyrenoid. *Nat. Plants* **6**, 1480–1490 (2020).
37. S. Poudel *et al*., Biophysical analysis of the structural evolution of substrate specificity in RuBisCO. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30451–30457 (2020).
38. O. A. Douglas-Gallardo, J. A. Murillo-López, J. Oller, A. J. Mulholland, E. Vöhringer-Martinez, Carbon dioxide fixation in RuBisCO is protonation-state-dependent and irreversible. *ACS Catalysis* **12**, 9418–9429 (2022).
39. L. Reinhold, R. Kosloff, A. Kaplan, A model for inorganic carbon fluxes and photosynthesis in cyanobacterial carboxysomes. *Can. J. Bot.* **69**, 984–988 (1991).
40. J. V. Vermaas *et al*., Passive membrane transport of lignin-related compounds. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 23117–23123 (2019).
41. 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).
42. R. L. Clark, J. C. Cameron, T. W. Root, B. F. Pfleger, Insights into the industrial growth of cyanobacteria from a model of the carbon-concentrating mechanism. *AIChE J.* **60**, 1269–1277 (2014).
43. R. Burnap, M. Hagemann, A. Kaplan, Regulation of CO_2 concentrating mechanism in cyanobacteria. *Life* **5**, 348–371 (2015).
44. B. M. Long, B. D. Rae, V. Rolland, B. Förster, G. D. Price, Cyanobacterial CO2-concentrating mechanism components: Function and prospects for plant metabolic engineering. *Curr. Opin. Biol.* **31**, 1–8 (2016).
45. S. S. W. Cot, A. K. C. So, G. S. Espie, A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. *J. Bacteriol.* **190**, 936–945 (2008).
46. G. C. Cannon, S. Heinhorst, C. A. Kerfeld, Carboxysomal carbonic anhydrases: Structure and role in microbial CO_2 fixation. *Biochim. Biophys. Acta Proteins Proteomics* **1804**, 382–392 (2010).
47. C. Blikstad *et al*., Identification of a carbonic anhydrase–Rubisco complex within the alpha-carboxysome. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2308600120 (2023).
48. A. Ghysels *et al*., Permeability of membranes in the liquid ordered and liquid disordered phases. *Nat. Commun.* **10**, 5616 (2019).
49. R. Gao *et al*., A prototype protein nanocage minimized from carboxysomes with gated oxygen permeability. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2104964119 (2022).
50. M. Eisenhut *et al*., The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17199–17204 (2008).
51. F. Hinzpeter, U. Gerland, F. Tostevin, Optimal compartmentalization strategies for metabolic microcompartments. *Biophys. J.* **112**, 767–779 (2017).
52. W. M. Haynes, *CRC Handbook of Chemistry and Physics* (CRC Press, 2014).
53. S. L. Evans *et al*., Single-particle cryo-EM analysis of the shell architecture and internal organization of an intact α-carboxysome. *Structure* **31**, 677–688.e4 (2023).
54. Y. Sun *et al*., Decoding the absolute stoichiometric composition and structural plasticity of *alpha*-carboxysomes. *mBio* **13**, e0362921 (2022).
55. M. J. Niederhuber, T. J. Lambert, C. Yapp, P. A. Silver, J. K. Polka, Superresolution microscopy of the β-carboxysome reveals a homogeneous matrix. *MBoC* **28**, 2734–2745 (2017).
56. X. Ji, J. M. H. Verspagen, D. B. Van de Waal, B. Rost, J. Huisman, Phenotypic plasticity of carbon fixation stimulates cyanobacterial blooms at elevated CO_2. *Sci. Adv.* **6**, eaax2926 (2020).
57. S. J. Hurley, B. A. Wing, C. E. Jasper, N. C. Hill, J. C. Cameron, Carbon isotope evidence for the global physiology of Proterozoic cyanobacteria. *Sci. Adv.* **7**, eabc8998 (2021).
58. 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).
59. A. M. Larsson, D. Hasse, K. Valegård, I. Andersson, Crystal structures of β-carboxysome shell protein CcmP: Ligand binding correlates with the closed or open central pore. *J. Exp. Bot.* **68**, 3857–3867 (2017).
60. M. G. Klein *et al*., Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. *J. Mol. Biol.* **392**, 319–333 (2009).
61. J. S. MacCready, A. G. Vecchiarelli, Positioning the model bacterial organelle, the carboxysome. *mBio* **12**, 10.1128/mbio.02519-19 (2021).
62. L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **30**, 2157–2164 (2009).
63. W. W. Kong *et al*., Cryo-electron tomography reveals the packaging pattern of RuBisCOs in Synechococcus β-carboxysome. *Structure* **32**, S0969212624001849 (2024).
64. L. M. Oltrogge *et al*., Multivalent interactions between CsoS2 and Rubisco mediate α-carboxysome formation. *Nat. Struct. Mol. Biol.* **27**, 281–287 (2020).
65. J. B. Turnšek, L. M. Oltrogge, D. F. Savage, Conserved and repetitive motifs in an intrinsically disordered protein drive α-carboxysome assembly. *J. Biol. Chem.* **300**, 107532 (2024).
66. T. Li *et al*., Uncovering the roles of the scaffolding protein CsoS2 in mediating the assembly and shape of the *alpha*-carboxysome shell. *mBio* **29**, e0135824 (2024).
67. S. Li, D. A. Matoz-Fernandez, M. Olvera De La Cruz, Effect of mechanical properties on multicomponent shell patterning. *ACS Nano* **15**, 14804–14812 (2021).
68. W. H. Roos *et al*., Squeezing protein shells: How continuum elastic models, molecular dynamics simulations, and experiments coalesce at the nanoscale. *Biophys. J.* **99**, 1175–1181 (2010).
69. J. Huang *et al*., CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
70. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
71. J. C. Phillips *et al*., Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **153**, 044130 (2020).
72. J. V. Vermaas, D. J. Hardy, J. E. Stone, E. Tajkhorshid, A. KohlMeyer, TopoGromacs: Automated topology conversion from CHARMM to GROMACS within VMD. *J. Chem. Inf. Model.* **56**, 1112–1116 (2016).
73. M. J. Abraham *et al*., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
74. C. R. Harris *et al*., Array programming with NumPy. *Nature* **585**, 357–362 (2020).
75. P. Virtanen *et al*., SciPy 1.0.: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
76. J. D. Hunter, Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
77. A. Ortega, D. Amorós, J. García de la Torre, Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101**, 892–898 (2011).
78. T. J. Boerner *et al*., "Advancing innovation: NSF's advanced cyberinfrastructure coordination ecosystem: services & support" in *Practice and Experience in Advanced Research Computing, PEARC '23* (Association for Computing Machinery, New York, NY, USA, 2023), pp. 173–176.
79. D. Sarkar *et al*., Input data for "Atomic View of Photosynthetic Metabolite Permeability Pathways and Confinement in Cyanobacterial Carboxysomes". Zenodo. https://zenodo.org/records/10632916. Deposited 8 February 2024.