The CO₂-fixing enzyme rubisco is responsible for almost all carbon fixation. This process frequently requires rubisco activase (Rca) machinery, which couples ATP hydrolysis to the removal of inhibitory sugar phosphates, including the rubisco substrate ribulose 1,5-bisphosphate (RuBP). Rubisco is sometimes compartmentalized in carboxysomes, bacterial microcompartments that enable a carbon dioxide concentrating mechanism (CCM). Characterized carboxysomal rubiscos, however, are not prone to inhibition, and often no activase machinery is associated with these enzymes. Here, we characterize two carboxysomal rubiscos of the form IA Clade that are associated with CbbQO-type Racas. These enzymes release RuBP at a much lower rate than the canonical carboxysomal rubisco from *Synechococcus* PCC6301. We found that CbbQO-type Racas encoded in carboxysome gene clusters can remove RuBP and the tight-binding transition state analog carboxyarabinitol 1,5-bisphosphate from cognate rubiscos. The *Acidithiobacillus ferrooxidans* genome encodes two form IA rubiscos associated with two sets of *cbbQ* and *cbbO* genes. We show that the two CbbQO activate systems display specificity for the rubisco enzyme encoded in the same gene cluster, and this property can be switched by substituting the C-terminal three residues of the large subunit. Our findings indicate that the kinetic and inhibitory properties of proteobacterial form IA rubiscos are diverse and predict that Racas may be necessary for some α-carboxysomal CCMs. These findings will have implications for efforts aiming to introduce biophysical CCMs into plants and other hosts for improvement of carbon fixation of crops.

The CO₂-fixing enzyme rubisco catalyzes the most quantitatively significant conversion of CO₂ gas to biomass, operating in the majority of autotrophic organisms. All genuine rubiscos possess an ancient and conserved active site architecture and a complex reaction mechanism (1). An inevitable consequence is a side reaction with oxygen (2), leading to the production of the metabolite 2-phosphoglycolate, which must be recycled via salvage pathways (“photorespiration”) to avoid the loss of carbon from the Calvin cycle (3). It is thought that the emergence of rubisco prior to the oxygenation of Earth’s atmosphere predisposed the enzyme to display a limited capacity to adapt to modern atmospheric conditions (4, 5). To compensate, diverse clades of rubisco-dependent autotrophs readily and convergently evolved a broad suite of carbon dioxide concentrating mechanisms (CCMs), many of which rely on spatial compartmentation of the carboxylase (6). A solution encountered in all cyanobacteria and many proteobacteria involves sequestration of cellular rubisco in protein-encapsulated microcompartments known as carboxysomes (7–9).

Rubisco requires the cofactors CO₂ and Mg²⁺ to bind at the active site prior to becoming catalytically competent (10). It has recently emerged that most if not all rubiscos are prone to inhibition, and often no activase machinery is associated with these enzymes. Here, we characterize two carboxysomal rubiscos of the form IA Clade that are associated with CbbQO-type Racas. These enzymes release RuBP at a much lower rate than the canonical carboxysomal rubisco from *Synechococcus* PCC6301. We found that CbbQO-type Racas encoded in carboxysome gene clusters can remove RuBP and the tight-binding transition state analog carboxyarabinitol 1,5-bisphosphate from cognate rubiscos. The *Acidithiobacillus ferrooxidans* genome encodes two form IA rubiscos associated with two sets of *cbbQ* and *cbbO* genes. We show that the two CbbQO activate systems display specificity for the rubisco enzyme encoded in the same gene cluster, and this property can be switched by substituting the C-terminal three residues of the large subunit. Our findings indicate that the kinetic and inhibitory properties of proteobacterial form IA rubiscos are diverse and predict that Racas may be necessary for some α-carboxysomal CCMs. These findings will have implications for efforts aiming to introduce biophysical CCMs into plants and other hosts for improvement of carbon fixation of crops.
they belong to the MoxR family of AAA+ proteins (27, 28) and rely on the CbbO adaptor protein for function. CbbO binds rubisco via its von Willebrand factor A domain, and a single CbbO subunit binds to the conserved concave surface of the CbbQ hexamer (29).

The cbbQ and cbbO genes found in the A. ferrooxidans carboxysomal cluster (Fig. 1) are therefore likely to encode an Rca that specifically remodels inhibited carboxysomal form IA \(^\text{C}\) rubiscos. This hypothesis was tested by Sutter et al. for Halothiobacillus neapolitanus; however, the reported activase assays used purified CbbQ complexes lacking the essential CbbO subunit. Analysis of fractured carboxysomes in that work showed that CbbQ associated with a shell enriched pellet fraction, suggesting an interaction of CbbQ with the carboxysome shell (30, 31).

A caveat regarding the need for carboxysomal Rcas relates to the fact that previously examined carboxylases (belonging to the form IB clade) do not form stably inhibited complexes with the substrate RuBP (32–34). In addition, the frequent absence of candidate Rca genes in the genomes of cyanobacteria and proteobacteria that utilize carboxysomes suggest that Rca function is not strictly required for the carboxysomal CCM (35, 36). The existence of carboxysomal Rca systems will require ATP and ADP to pass the carboxysomal shell, which is widely believed to be selectively permeable and possibly even gated (37). With regard to form IB rubisco, it has very recently been shown that the Rca-like protein found in some cyanobacterial clades (38, 39) is a functional Rca in Nostoc sp. PCC7120 (40). Its mechanism involves an interaction with the rubisco large subunit N terminus, which is conserved in the higher plant system (41). The homolog encoded by Fremyella diplosiphon showed the expected carboxysomal localization (35).

Recently, there have been tremendous advances toward introducing carboxysomes and their associated rubiscos into higher plants (42, 43). In conjunction with a functional CCM, it is predicted that such a modification could enhance photosynthetic carbon assimilation of a C3 plant by 60% (44, 45). It has even been possible to transfer a powered and functional carboxysome CCM into Escherichia coli (46). In order to fully capitalize on these advances, it is imperative to resolve and define the role of Rcas in such compartments.

Here, using pure components, we functionally characterize the properties of rubiscos and define the Rca function of CbbQ complexes encoded in the carboxysomal clusters of A. ferrooxidans and H. neapolitanus, respectively. We found that both chaperones are indeed capable of removing both loose-binding RuBP and the exceptionally tight-binding transition-state analog carboxy-arabinitol 1,5-bisphosphate (CABP) from form IA \(^\text{C}\) rubisco active sites. The specificity of the two CbbQ complexes associated with the two form IA rubisco operons in A. ferrooxidans can be switched via amino acid substitutions in the RbcL C terminus. We propose that carboxysomal rubiscos exist on a continuum related to their kinetic properties and tendencies to bind inhibitors. Their position on the continuum determines the extent of activase dependency and permits activase loss in organisms possessing high-velocity and low-efficiency carboxylases.

Results
Bioinformatics analysis of form IA rubiscos and associated CbbQO sequences
We performed BLAST homology searches using the A. ferrooxidans IA \(^\text{Q}\) large subunit sequence (UniProt: B7JA24) to collect form IA CbbL sequences. The sequences were then classified according to their gene neighborhoods. Form IA \(^\text{C}\) sequences were followed by csoS2 and other carboxysome genes and contained cbbQ and cbbO in the operon. Form IA \(^\text{Q}\) genes were directly adjacent to cbbQ and did not encode carboxysome genes (Fig. 1). Among \(\alpha\)-carboxysome clusters encoding form IA \(^\text{C}\) rubisco and CbbQ and CbbO, there appear to be two arrangements. In one cluster type, cbbQ and cbbO genes are positioned directly downstream of the shell-encoding genes. In the other cluster type, cbbO and cbbQ are separated by genes encoding the dab complex (Fig. 1) (47).

We decided to introduce a new form IA \(^\text{C}\) division: these constituted form IA CbbL sequences encoded in carboxysomal clusters that did not contain cbbQ and cbbO genes and were designated as IAC, \(\text{Q}^-\). Form IA \(^\text{C}\) and IA \(^\text{Q}\) sequences are all proteobacterial, whereas the IAC, \(\text{Q}^-\) sequences are also found in \(\alpha\)-cyanobacteria and the endosymbiont of Paulinella chromatophora (48). A maximum likelihood tree indicated that

![Figure 1. Form I rubisco-associated operons in Acidithiobacillus ferrooxidans and Halothiobacillus neapolitanus.](image-url) The relative position and sequence identity between rubisco (cbbL and cbbS) and activase (cbbq and cbbO) encoding genes is indicated. The terminology used in this work for the different rubiscos and Rca complexes encoded in the operons is also shown. Rca, rubisco activase.
form IA\textsuperscript{C} and form IA\textsuperscript{Q} CbbL sequences do not form distinct clades, as observed previously (19) (Fig. S1A). In contrast, form IAC\textsubscript{Q} large subunits formed a well-supported separate clade. In addition, the multiple sequence alignment (Fig. S2) revealed that all the form IA\textsuperscript{C} and IA\textsuperscript{Q} CbbL sequences possessed a C-terminal motif (KLDXHK or KLXXXHK) consistent with the experimentally characterized CbbQO interaction (26). In contrast, all examined form IAC\textsuperscript{Q} sequences had diverse C termini, which generally terminated two residues earlier (KLDX) (Fig. S2).

The phylogenetic relationship between form IA rubisco-associated CbbS (Fig. S1B), CbbQ (Fig. S1C), and CbbO (Fig. S1D) amino acid sequences could not be used to predict whether they were associated with carboxysomal clusters (IA\textsuperscript{C} or IA\textsuperscript{Q} rubiscos). There is therefore no support for distinct ancestral form IA\textsuperscript{Q} or IA\textsuperscript{C} rubisco/activase systems. Instead form IA-associated CbbQO activase components appear to change their context over time, which would lead one to expect biochemical compatibility between non-native rubisco–Rca pairs.

**The inhibition and reaction kinetics of CbbQO-associated form IA rubiscos are diverse**

Existing data suggest that carboxysomal form I rubiscos do not form tight-binding complexes with RuBP in the apo state but readily release the inhibitor. This property correlates with high carboxylation velocities and Michaelis constants for CO\textsubscript{2}, but this generalization is exclusively derived from cyanobacterial form IB enzymes (32, 33, 40, 49, 50). It is not known whether carboxysomal form IA\textsuperscript{C} rubiscos follow the same trend. We therefore decided to compare kinetic parameters and the spontaneous release of RuBP in the form IA enzymes encoded in *A. ferrooxidans*, the carboxysomal form IA\textsuperscript{C} rubisco from *H. neapolitanus* and the carboxysomal form IB model enzyme from *Synechococcus* PCC6301. Following purification (Fig. 2A), the apoenzymes were preincubated with RuBP to form enzyme–RuBP (ER) (Fig. 2B) and subsequently added to a rubisco assay mixture. The activity of the ER complex (red traces) was then compared with that of the preactivated Enzyme–CO\textsubscript{2}–Mg\textsuperscript{2+} (ECM) holoenzyme (black traces) allowing the rate of RuBP release to be estimated by quantifying the formation of active ECM holoenzymes (32, 51) (Fig. 2B).

As reported earlier (26, 52, 53), the form IA\textsuperscript{Q} rubisco (AFL1S1) formed a stable binary complex with RuBP and had the slowest inhibitor release rate (Fig. 2C). The three carboxysomal enzymes displayed faster but highly variable reactivation rates (Fig. 2, D–G). The activity of fully activated Afl2S2 form IA\textsuperscript{C} enzyme was not linear over the time course, indicating a loss of functional active sites over the reaction time course (the “fallover” phenomenon commonly observed in higher plant rubiscos (32, 51)). Compared with the very loose RuBP-binding phenotype of the cyanobacterial form IB SynLS rubisco, all form IA enzymes assayed here had much slower RuBP release rates and would therefore be predicted to benefit from an associated activase system.

Kinetic characterization of the form IA enzymes indicated that the two *A. ferrooxidans* rubiscos had highly similar carboxylation velocities (~5.5 s\textsuperscript{-1}) and Michaelis constants (40–50 μM) for CO\textsubscript{2} (Table 1 and Fig. S3). The carboxysomal Afl2S2 form IA\textsuperscript{C} enzyme had a much reduced specificity factor of 32, compared with 56 for the form IA\textsuperscript{Q} rubisco. *H. neapolitanus* form IA\textsuperscript{C} (HnLS) rubisco kinetics were distinct, being faster and possessing a much higher Michaelis constant for CO\textsubscript{2}. Compared with the fast model carboxysomal form IB Synecochoccus enzyme SynLS (54–56), HnLS displayed an intermediate phenotype (Table 1).

The observed general trend of increased carboxylation velocities and reduced specificity factors for the carboxysomal enzymes is commonly observed for rubiscos that are supplied by CCMs. However, the three carboxysomal enzymes characterized here possessed a significant diversity in their kinetics. For the four enzymes evaluated here, there was a clear correlation between RuBP release rate and both the $K_{m}(CO_{2})$ and carboxylation turnover rate (Fig. 2, H and I). This observation is consistent with the hypothesis by Tcherkez et al. (57), which predicts that inhibitor binding should be enhanced in rubiscos that display high affinities for CO\textsubscript{2} and the associated low carboxylation velocities.

The small selection of kinetic data and inhibitory properties shown here and reported in the literature (42, 52, 53, 58–61) suggest that the form IA clade of rubiscos are likely to have diverse kinetics related to the relative ability of host organisms to acquire and concentrate CO\textsubscript{2}.

**The CbbQO complexes encoded in carboxysome clusters are functional Rcas**

Sutter et al. (30) demonstrated that the CbbQ protein encoded by the *H. neapolitanus* carboxysomal gene cluster physically associated with the α-carboxysome. Consistent with the requirement for association of the adaptor protein CbbO (26), in that study, the CbbQ hexamer was not able to activate inhibited carboxysomal rubisco. To clarify the function of the CbbQ and CbbO isoforms encoded by carboxysome gene clusters, we produced in *E. coli* and purified the form IA rubisco-associated CbbQO complexes encoded by *A. ferrooxidans* (IA\textsuperscript{Q},AfQ1O1 and IA\textsuperscript{C},AfQ3O3) and *H. neapolitanus* IA\textsuperscript{C} (HnQO) (Fig. 3A). The ATPase activity of the purified AfQ3O3 complex was similar to that of the AfQ1O1 complex. HnQO was much faster, showing an ATPase activity of ~8 min\textsuperscript{-1}/CbbQ protomer (Fig. 3B).

We then assayed carboxysomal AfQ3O3 for the ability to remove the transition state analog CABB from the *A. ferrooxidans* form IA\textsuperscript{C} (Afl2S2) holoenzyme (ECM), using conditions previously established for the *A. ferrooxidans* form IA\textsuperscript{Q} enzyme (26) (Fig. 3C). AfQ3O3 was almost nonfunctional, but increasing the concentration of the crowding agent PEG 3350 from 5% to 13% v/v resulted in an eightfold activity increase and thus robust Rca function (Fig. 3, D and F). The dependence of AfQ3O3 on high PEG concentrations was seen irrespective whether ECM component (ECMC) or ER (Fig. 3E) was used as inhibited rubisco, or whether the assays were
performed in the presence of 5 or 20 mM bicarbonate (Fig. 3F). AfQ3O3 is therefore the activase of the carboxysomal AfL2S2 (form IAC) rubisco, which is encoded in the same operon (Fig. 1A).

AfQ3O3 was able to remove CABP from its cognate form IAC rubisco HnLS even in the absence of PEG (Fig. 3G), although inclusion of the polymer also greatly enhanced Rca activity (Fig. 3G–I). Titrating both HnQO and PEG concentrations in this assay demonstrated a concentration dependence for both components and showed that Rca function was similar at 0.8 μM HnQO/13% PEG and 3.2 μM HnQO/5% PEG (concentration expressed as CbbQ protomer) (Fig. S4). This finding is consistent with the dominant volume exclusion contribution of macromolecular crowding. Here, the crowding agent occupies a substantial fraction of the volume, and this results in a smaller available volume for both rubisco and activase, leading to higher biochemical activities (62–64). The strong positive

**Table 1**

Kinetic parameters of the form I rubiscos

<table>
<thead>
<tr>
<th>Rubisco</th>
<th>( k_{\text{cat}} ) (s(^{-1}))</th>
<th>( K_c ) (μM)</th>
<th>( k_{\text{cat}}/K_c ) (s(^{-1}) μM(^{-1}))</th>
<th>( S_{\text{C/O}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>AfL1S1</td>
<td>5.2 ± 0.1</td>
<td>40 ± 3</td>
<td>130</td>
<td>56.4 ± 1.7</td>
</tr>
<tr>
<td>AfL2S2</td>
<td>5.5 ± 0.1</td>
<td>55 ± 3</td>
<td>100</td>
<td>32.2 ± 0.5</td>
</tr>
<tr>
<td>HnLS</td>
<td>8.0 ± 0.4</td>
<td>127 ± 16</td>
<td>63</td>
<td>43.6 ± 1.1</td>
</tr>
<tr>
<td>SynLS(^a)</td>
<td>11.7 ± 0.6</td>
<td>200 ± 4</td>
<td>59</td>
<td>42.7 ± 2.8</td>
</tr>
</tbody>
</table>

Michaelis–Menten plots used to derive the parameters are shown in Fig. S3.

\(^a\) Data from the study by Davidi et al. (56).

HnQO was able to remove CABP from its cognate form IAC rubisco HnLS even in the absence of PEG (Fig. 3G), although inclusion of the polymer also greatly enhanced Rca activity (Fig. 3, G–I). Titrating both HnQO and PEG concentrations in this assay demonstrated a concentration dependence for both components and showed that Rca function was similar at 0.8 μM HnQO/13% PEG and 3.2 μM HnQO/5% PEG (concentration expressed as CbbQ protomer) (Fig. S4). This finding is consistent with the dominant volume exclusion contribution of macromolecular crowding. Here, the crowding agent occupies a substantial fraction of the volume, and this results in a smaller available volume for both rubisco and activase, leading to higher biochemical activities (62–64).
effect of molecular crowding presumably relates to the carboxysomal microenvironment of AfL2S2 and HnLS. This property may be related to carboxysomal rubisco’s coacervation with the multivalent rubisco linker protein CsoS2 (65, 66). A strong positive effect of PEG is also observed when assaying green-type plant Rcas, which are found in the crowded chloroplast stroma (67, 68).

The two Rca complexes from A. ferrooxidans (AfQ3O3 and AfQ1O1) were also able to remove CABP from HnLS but displayed much lower activities in combination with a great dependence on high PEG concentrations (Fig. 3J).

These data establish that the gene products of cbbQ and cbbO encoded by carboxysomal gene clusters in both A. ferrooxidans and H. neapolitanus are capable of reactivating inhibited carboxysomal form IA\textsuperscript{−} rubisco complexes. The function of the A. ferrooxidans activase was greatly enhanced by high concentrations of the molecular crowding agent PEG.

The role of the RbCL C terminus in recognition by CbbQO

The RbCL C terminus is critical for CbbQO-mediated activation for both form I and form II rubiscos. The interaction is
mediated by a C-terminal HK/R motif, and modification of the C terminus (including single-residue truncations and extensions) eliminates the ability of rubisco to become activated by CbbQO (26) (Fig. 4A). Carboxysomal rubisco large subunits that do not encode cbbQ and cbbO in the operon (FormIA_C) are missing this motif and generally have a shorter C-terminal motif KLDXX. In contrast, when cbbQ and cbbO are present in the operon, the large subunit C-terminal motif is longer (KLDXXHK or KLDXXXR) (Figs. 4B and S2).

The form IA_C rubiscos from A. ferrooxidans and H. neapolitanus do not possess the conserved HK/R motif but instead encode an alternative “NR” C terminus (Fig. 4C). Two Afl2S2 rubisco C-terminal variants were produced and purified (Fig. 4D) and displayed similar carboxylation kinetics and spontaneous RuBP release under our assay conditions (Figs. 4E and S5). This permitted us to compare their activation kinetics using the two A. ferrooxidans CbbQO complexes AflQ1O1 and AflQ3O3.

Activation of the form IA_C Afl2S2 rubisco by either AflQ1O1 or AflQ3O3 was poor at 5% v/v PEG (Fig. 4, F and H). Substituting N472 with histidine or replacing the last three amino acids (VNR to AHK) resulted in a sixfold enhancement in AflQ1O1 activase function (Fig. 4, F and H). This indicates that the mutations significantly enhanced the low affinity of AflQ1O1 for the carboxysomal rubisco.

This dramatic increase in function was not observed when using the cognate AflQ3O3 assay at 13% v/v PEG (Fig. 4G). Whereas Afl2S2N472H permitted a 30% increase in AflQ3O3 function compared with the wildtype Afl2S2, Afl2S2 (AHK) showed a ~60% reduction in activation rate (Fig. 4, G and H). Substituting the last three residues of the C terminus therefore resulted in a specificity switch regarding the preference of the two A. ferrooxidans CbbQO complexes (Fig. 4H).

These results suggest that the as yet unidentified binding sites of A. ferrooxidans Q1O1 and Q3O3 to the RbcL C terminus are nonequivalent and can enable the activases to discriminate between substrate rubiscos. However, for both activases, a histidine residue at the penultimate position of RbcL is preferred to achieve high activase functionality in vitro. A crowded environment like the one found in a carboxysome may relax the requirement for a high rubisco activase activity.

**Discussion**

In this contribution, we conclusively demonstrate that CbbQO complexes encoded in proteobacterial carboxysomal operons are functional at removing inhibitory sugar phosphates from carboxysomal form IA_C rubisco active sites. The widespread evolutionary conservation of these genes, in conjunction with the functional green-type activase from Nostoc (40), therefore suggests that robust carboxysome CCM function relies on the ability to remove RuBP or other sugar phosphates from the rubisco active site in some bacteria. Our work does not inform on the nature of the physiological inhibitor(s), but these are likely to be reaction misfire products, such as xylulose 1,5-bisphosphate or possibly pentadiulose-1,5 bisphosphate (12, 13). Elucidation of this issue will necessitate a detailed biochemical study of rubiscos isolated from bacterial mutants with lesions in the associated cbbq/cbbO genes. RuBP is unlikely to be a relevant inhibitor because the concentration of Mg^{2+} and CO_{2} is expected to remain high in the carboxysome, which will disfavor decarbamylation and ER formation.

A. ferrooxidans possesses two form IA rubiscos, one form II enzyme and three CbbQO complexes (26, 56). Its form IA_C rubisco (Afl2S2) presents with a similar K_{m} (CO_{2}) but a reduced specificity factor compared with the form IA_Q enzyme (Afl1S1) (Table 1). The tendency of Afl2S2 to spontaneously release RuBP from the apoenzyme is less than that of the other carboxysomal rubiscos tested but more rapid than the form IA_Q rubisco Afl1S1 (Fig. 2). Our work thus shows that three carboxysomal rubiscos possess highly variable carboxylase kinetics and inhibitor binding. A well-characterized parallel regarding kinetics is the situation in pyrenoid-containing diatoms, where a wide variety of kinetic constants are also observed and interpreted to relate to relative CCM strength (69). A stronger CCM will permit the selection pressure on K_{m} (CO_{2}) and S_{c/o} to relax, and this will permit higher carboxylase velocities (57, 70). These properties may be related to the relative availability of metabolic energy and inorganic carbon for these organisms, which will determine the CO_{2} concentrations that can be achieved at the carboxysomal rubisco active site. It is possible that H. neapolitanus and cyanobacterial rubiscos experience higher CO_{2} concentrations because their metabolic modes (sulfur oxidation and oxygenic photosynthesis) generate copious amounts of energy. In contrast, A. ferrooxidans grows at low pH, which translates to an absence of the bicarbonate pool, and uses the energetically challenging iron-oxidizing metabolism (25). Its carboxysomal rubisco may thus be exposed to lower CO_{2} levels.

Our findings also show that in A. ferrooxidans, the CbbQO complexes are specific for the form IA rubisco encoded in their gene cluster, and that this specificity is dependent on the identity of the RbcL C terminus, which has been implicated in CbbQO function previously (26) (Fig. 4). This biochemical finding implies that in A. ferrooxidans, it will be important to package the correct (Afl2S2) rubisco in the carboxysome together with its AflQ3O3 activase. Recent discoveries elucidating the mechanisms of α-carboxysomal rubisco compartmentalization via the CsoS2 rubisco linker protein provide a straightforward experimental framework to explore these issues (65, 71).

The requirement for activase-mediated rubisco remodeling in carboxysome function has been controversial (30, 35), and in particular, the presence of cbbQ and cbbO in carboxysomal gene clusters has remained unexplained (45, 72).

In light of our findings, we propose that dependence of organisms bearing carboxysome CCMs on activase function will vary and relate to both phylogenetic history and to a continuum of kinetic and inhibitory properties of the associated carboxylases. These properties are sculpted by the
**Figure 4. The RbcL C terminus encodes Rca specificity.** A, model of CbbQO function. The large subunit C terminus (red) is required for interaction with the activase and likely interacts with the CbbO adaptor protein. B, Logo motifs of the C termini generated from the RbcL sequences used for the phylogenetic tree in Fig. S1. C, alignment of the form IA RbcL C termini used in this study. D, SDS-PAGE analysis of purified AFL2S2 variants. AHK, the VNR C terminus is substituted with AHK. E, carboxylase activity of fully activated (ECM) AFL2S2 wildtype and variant proteins assayed at 20 mM NaHCO3. F, mutation at the C terminus of AFL2S2 enhances activation of ECM complexes by Q101 (5% v/v PEG and 20 mM NaHCO3). G, the N472H mutation enhances AFL2S2(VNR) function, whereas the AHK substitution reduces it (13% v/v PEG and 20 mM NaHCO3). H, quantified Rca activities for AFL2S2 and its variants. I, scheme summarizing how mutations of the RbcL C tail influence activase function. Rca, rubisco activase.
prevailing CO₂ concentrations experienced by the active sites during evolution (59, 69). On one extreme of the spectrum, there are carboxysomal rubiscos, such as the A. ferrooxidans form IA² enzyme (AfL2S2), which do not readily release RuBP from the apoenzyme, and are only modestly faster than the cytoplasmic form IA⁰ enzymes (AfL1S1). Such carboxysomal enzymes will be predicted to strictly coexist with Rca-encoding genes. At the other extreme of the continuum, the majority of cyanobacteria do not encode Rca homologs. Consistently, the fast Synechococcus form IB enzyme rapidly dissociates from inhibitory RuBP and XuBP (32). However, its Kₘ (CO₂) is exceptionally high, and concordantly, the carboxylase efficiency Kₑ/Kₐ is poor.

The enzyme from H. neapolitanus is intermediate to the extreme examples. Inhibitory compounds are released more readily, the enzyme is fast and possesses a high Kₘ (CO₂) but not as exaggerated as the cyanobacterial rubisco. In addition, experimental results suggest that deleting CbbQ or CbbO has a negligible effect on growth under energy-limited conditions (47). It is possible that a fitness defect would become apparent if growth under energy-limited conditions was explored (73). Transferring the H. neapolitanus carboxysomal genes and other CCM components into E. coli was successful, even if CbbQ was nonfunctional (46). It is possible that in this scenario, the fitness defect is too subtle to be detected in the context tested in the laboratory.

Although the outlined physiological data suggest that the H. neapolitanus CbbQO activase is not critical, here, we show that biochemically the system is functional. In addition, our bioinformatics analysis of CbbL sequences, we did not encounter IAC₂/Q branches inside the form IA²/Q cluster (Fig. S1). Evolutionary loss of the CbbQO activase system thus appears to be a rare event. Finally, an article published during revision of this article compared E. coli-produced H. neapolitanus carboxysomes. It was found that if HnQO was incorporated, the carboxysomes were more functional at fixing CO₂ in the presence of substoichiometric amounts of CABP (74).

Further systematic characterization of carboxysomal rubisco kinetics, their inhibitory properties, and activase dependencies will contribute to our appreciation of the carboxylase intrinsic tradeoffs (57, 75–77). It is likely that biotechnological exploitation of carboxysomes in plants and other organisms will benefit from taking into account a comprehensive understanding of these factors.

**Experimental procedures**

**Bioinformatics analysis**

BLAST searches were performed using the AfCbbL1 sequence (UniProt: B7JA24) to query the Integrated Microbial Genome database (78) and the National Center for Biotechnology Information. Classification into form IA⁰, IA², and IAC₂/Q was performed manually by inspecting the gene neighborhood of the selected genes. Maximum likelihood phylogenetic trees were constructed using MegaX (www.megasoftware.net) (79) and visualized using iTOL (https://itol.embl.de) (80). Positions containing gaps or missing data were eliminated. The sequences used are provided in Dataset S1. Multiple sequence alignments were drawn using ESPript (https://espirpt.ibcp.fr) (81). Gene neighborhoods were visualized using clinker (https://github.com/gamcil/clinker) (82).

**Molecular biology**

Rubisco- and activase-encoding genes were amplified from genomic DNA of H. neapolitanus ATCC23641 and A. ferrooxidans ATCC23270. The H. neapolitanus cbbL (Hneap_0922) and cbbS (Hneap_0921) genes including the intergenic region were cloned between NdeI and HindIII sites of pET30b to yield pET30bHnLS. The carboxysome CbbQO-encoding genes HncbbQ (Hneap_0905)/AfqO3 (AFE_1678) and HncbbO (Hneap_0910)/AfqO3 (AFE_1677) were cloned into pHue and pBad33 in the same context as pHueAfCbbQ1 and pBad33LubAfqO1 (26).

Site-directed mutagenesis was used to introduce mutations into the AfqO3 (AFE_1678). The corresponding primers and plasmids used in this study are listed in Tables S1 and S2. Protein-encoding sequences were verified using DNA sequencing.

**Proteins**

Proteobacterial rubiscos were produced in E. coli BL21(DE3) cells transformed with the appropriate plasmid and purified using anion exchange chromatography followed by size-exclusion chromatography as described for AfL1S1 (26). The SynLS purification protocol was described (55).

Rca complexes were produced in E. coli BL21(DE3) harboring the appropriate plasmids (Table S2) and purified using a combination of affinity chromatography, ion exchange chromatography, and size-exclusion chromatography. The purification of AfQ1O1 was described previously (26). AfQ3O3 and HnQO were obtained using the same strategy except that the ion exchange chromatography step was omitted for AfQ3O3. Protein concentrations were quantified using the Bradford assay using bovine serum albumin as the standard.

**Enzymatic assays**

All assays were performed at 25 °C. ATPase activity was measured using a coupled spectrophotometric assay (83). The CbbQO concentration used was 0.27 μM oligomer. RuBP was synthesized from ribose 5-phosphate and purified by anion exchange chromatography (84, 85). Rubisco and Rca assays were performed using the coupled spectrophotometric rubisco assay (26, 86) and quantified as described (87). ECM holoenzyme was obtained by incubating rubisco (10 μM active sites) in 20 mM Tris–HCl, pH 8.0, 50 mM NaCl, 50 mM NaHCO₃, and 10 mM MgCl₂ for 10 min at 25 °C. A fourfold molar excess of carboxypentitol bisphosphate (synthesized from RuBP following Ref. (88)) was added to ECM to form ECMC. Excess ligand was then removed by buffer exchange into 20 mM Tris–HCl, pH 8.0, and 50 mM NaCl using Micro-Bio-Spin 6 (BioRad) chromatography columns.
Inhibited ER complex (10 μM) was obtained by incubating apo-rubisco in 20 mM Tris–HCl (pH 8.0), 50 mM NaCl, and 4 mM EDTA for 10 min, followed by addition of RuBP to a final concentration of 0.4 mM for 30 min at 25 °C. Rca assays contained either 5 or 13% v/v PEG 3350 (Sigma–Aldrich).

Rubisco kinetics were determined using radiometric rubisco assays (89). 14CO2 fixation assays (0.5 ml of reaction volume) were performed at 25 °C in 7.7 ml septum-capped glass scintillation vials with assay buffer (100 mM N-(2-hydroxyethyl)piperazine-N′- (3-propanesulfonic acid)–NaOH, pH 8.0, 20 mM MgCl2, and 1 mM EDTA), 10 mg/ml carbonic anhydrase, and 1 mm RuBP. All the assay components were equilibrated with nitrogen gas prior to the addition of [1-3H]RuBP.

The reaction was stopped after 2 min using 200 μl of 50% (v/v) formic acid. The specific activity of 14CO2 was measured using the highest 14CO2 concentration and 5.2 nmol of RuBP. The reaction was stopped after 30 min, and the activity was ~220 CPM/nmol RuBP. Active site concentrations were determined using 14C-CABP binding assays (90). CO2/O2 specificity assays were performed as described (89, 91). The reacting were equilibrated in a gas mixture (999,300 ppm O2; 700 ppm CO2; Air Liquide Singapore Private Limited) prior to the addition of [1-3H]-RuBP.

Data availability
All data are contained within the article and supporting information.

Supporting information—This article contains supporting information (26, 55, 92).

Acknowledgments—We thank Avi Flamholz for the careful reading and insightful comments on this article.

Author contributions—Y.-C. C. T. and O. M.-C. conceptualization; Y.-C. C. T., I. L., Z. G., and D. L. investigation; Y.-C. C. T. writing–original draft; O. M.-C. writing–review & editing; O. M.-C. supervision; O. M.-C. funding acquisition.

Funding and additional information—This work was funded by Ministry of Education (MOE) of Singapore tier 2 grants (MOE2016-T2-2-088 and MOE-T2EP30120-0005) and the National Research Foundation (NRF) of Singapore grant (NRF2017-NRF-ISF002-2667) (to O. M.-C.).

Conflict of interest—The Mueller-Cajar laboratory receives funding from Flagship Labs 79, Inc. All the other authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: CABP, carboxy-arabinitol 1,5-bisphosphate; CCM, carbon dioxide concentrating mechanism; ECM, Enzyme-CO2-Mg2+; ECMC, ECM-CABP; ER, enzyme–RuBP; MOE, Ministry of Education; NRF, National Research Foundation; Rca, rubisco activase; RuBP, ribulose 1,5-bisphosphate.

References
Carboxysomal CbbQO-type rubisco activases are functional


Carboxysomal CbbQO-type rubisco activases are functional


