Molecular Plant
Research article

CellPress
Partner Journal

Unique biogenesis and kinetics of hornwort Rubiscos revealed by synthetic biology systems

Zhen Guo Oh[1,*], Tanner Ashton Robison[1,2], Dan Hong Loh[1], Warren Shou Leong Ang[2], Jediael Zheng Ying Ng[3], Fay-Wei Li[1,2,*] and Laura Helen Gunn[1,4,*]

[1]School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA

[2]Boyce Thompson Institute, Ithaca, NY 14853, USA

[3]Max Planck Institute for Terrestrial Microbiology, 35043 Marburg, Germany

[4]Department of Cell and Molecular Biology, Uppsala University, 751 24 Uppsala, Sweden

***Correspondence: Zhen Guo Oh (zgo2@cornell.edu), Fay-Wei Li (fl329@cornell.edu), Laura Helen Gunn (lhg42@cornell.edu)**

https://doi.org/10.1016/j.molp.2024.10.013

ABSTRACT

Hornworts are the only land plants that employ a pyrenoid to optimize Rubisco's CO_2 fixation, yet hornwort Rubisco remains poorly characterized. Here we assembled the hornwort *Anthoceros agrestis* Rubisco (*Aa*Rubisco) using the *Arabidopsis thaliana* SynBio expression system and observed the formation of stalled intermediates, prompting us to develop a new SynBio system with *A. agrestis* cognate chaperones. We successfully assembled *Aa*Rubisco and Rubisco from three other hornwort species. Unlike *A. thaliana* Rubisco, *Aa*Rubisco assembly is not dependent on RbcX or Raf2. Kinetic characterization reveals that hornwort Rubiscos exhibit a range of catalytic rates (3–10 s^{-1}), but with similar affinity (~30 μM) and specificity (~70) for CO_2. These results suggest that hornwort Rubiscos do not comply with the long-held canonical catalytic trade-off observed in other land plants, providing experimental support that Rubisco kinetics may be phylogenetically constrained. Unexpectedly, we observed a 50% increase in *Aa*Rubisco catalytic rates when RbcX was removed from our SynBio system, without any reduction in specificity. Structural biology, biochemistry, and proteomic analysis suggest that subtle differences in Rubisco large-subunit interactions, when RbcX is absent during biogenesis, increases the accessibility of active sites and catalytic turnover rate. Collectively, this study uncovered a previously unknown Rubisco kinetic parameter space and provides a SynBio chassis to expand the survey of other Rubisco kinetics. Our discoveries will contribute to developing new approaches for engineering Rubisco with superior kinetics.

Key words: hornworts, CO_2 fixation, Rubisco, synthetic biology, structural biology, assembly factors

Oh Z.G., Robison T.A., Loh D.H., Ang W.S.L., Ng J.Z.Y., Li F.-W., and Gunn L.H. (2024). Unique biogenesis and kinetics of hornwort Rubiscos revealed by synthetic biology systems. Mol. Plant. **17**, 1833–1849.

INTRODUCTION

In most ecosystems, land plants play an indispensable role as producers of sugar molecules by capturing carbon dioxide (CO_2) through photosynthesis. This CO_2 fixation process is initiated by the enzyme, ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). Despite Rubisco's central role in global primary productivity, it has two major limitations: slow catalytic activity and poor specificity for CO_2. A catalytic trade-off, an inverse relationship between specificity and catalytic rate, is reported to be exhibited by Rubisco from most C_3 plants (Bouvier et al., 2021; Tcherkez and Farquhar, 2021). To improve Rubisco's CO_2-fixing efficiency, some photosynthetic organisms have evolved biophysical CO_2-concentrating mechanisms (CCMs) (Vaughn et al., 1990). One such CCM involves a proteinaceous and membraneless organelle known as the pyrenoid, housing most of the cell's Rubisco (Freeman Rosenzweig et al., 2017). Together with inorganic carbon pumps and carbonic anhydrase, high concentrations of CO_2 are supplied to these Rubisco hubs, allowing Rubisco to function efficiently (Karlsson et al., 1998; Yamano et al., 2010; Wunder et al., 2019). With the vision of introducing pyrenoids into crops to enhance yield (Adler et al., 2022), extensive effort has been placed to characterize the biophysical CCM of the green alga *Chlamydomonas reinhardtii* (Mackinder et al., 2016; Freeman Rosenzweig et al., 2017; Itakura et al., 2019; Hennacy and Jonikas, 2020; Fei et al., 2022). Although computational modeling has predicted the most efficient

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

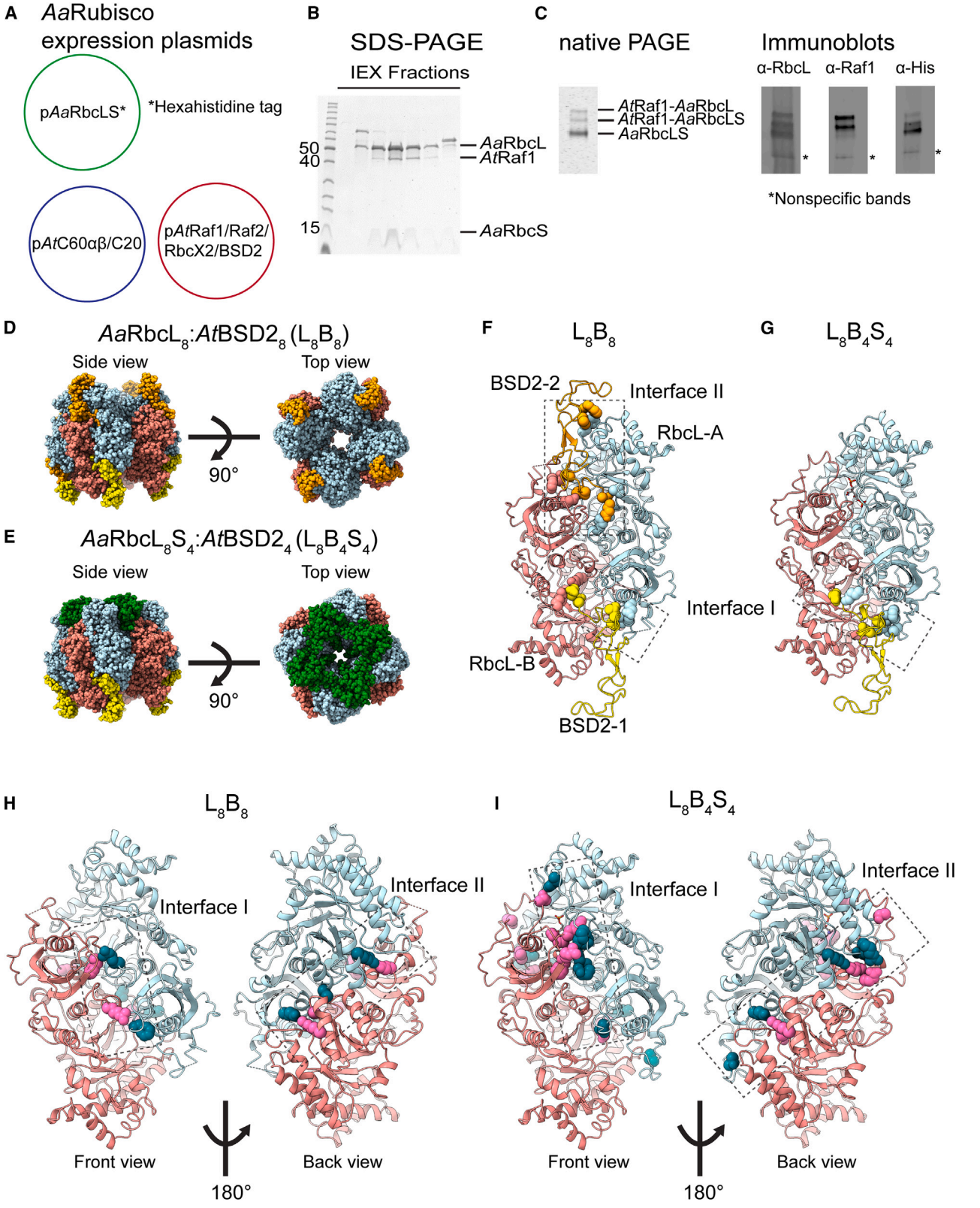


Figure 1. *Aa*Rubisco:*At*Chaperone complexes.

(A) Simple schematic of expressing *Aa*RbcL and C-terminal hexahistidine-tagged RbcS using the *At*Rubisco SynBio system.

(B) Coomassie-stained SDS–PAGE analysis of the produced *Aa*Rubisco intermediates purified by anion-exchange chromatography (IEX).

(legend continued on next page)

pathway to implement *C. reinhardtii* CCM into plant chloroplasts (Fei et al., 2022), experimental validation is still under way (Adler et al., 2022).

Hornworts, one of the three phyla of bryophytes, are the only land plants with pyrenoid-based CCMs (Smith and Griffiths, 1996). Compared to *Chlamydomonas*, hornworts share a more recent common ancestor with angiosperms, from which all economically significant crop plants are descended (Li et al., 2017). This suggests that hornwort CCMs might have higher compatibility with crop chloroplasts and could simplify crop-engineering efforts. With the exception of localizing Rubisco and Rubisco activase to hornwort pyrenoids (Vaughn et al., 1990; Lafferty et al., 2024), relatively limited effort has been put into delineating hornwort pyrenoid constituents. Biochemical/kinetic characterization of hornwort Rubisco is scarce, with catalytic turnover rates being reported for only three species (Hanson et al., 2002). A major challenge limiting a broader survey of hornwort Rubisco kinetics, in terms of both number of species and kinetic parameters (e.g., substrate affinity and substrate specificity), is purifying sufficient Rubisco from the source organism, in part because hornwort Rubisco constitutes only about 3%–4% of soluble protein (Hanson et al., 2002), as opposed to 30%–50% of soluble protein in vascular C_3 plant leaves (Phillips and Milo, 2009). Turning to alternative sources to heterologously express plant Rubisco was not possible until recently (Aigner et al., 2017).

Rubisco is a hexadecamer comprising eight large (RbcL) and eight small (RbcS) subunits, L_8S_8, where four RbcL dimers form an octamer core capped at both ends by four RbcS. Heterologous expression of *Arabidopsis thaliana* Rubisco in *Escherichia coli* requires co-expressing RbcL and RbcS, *A. thaliana* chaperonins (Cpn60αβ, Cpn20), and Rubisco assembly factors (Raf1, Raf2, RbcX2, BSD2). The Cpn60αβ/Cpn20 chaperonin heterocomplex encapsulates nascent RbcL chains to provide a favorable local environment for folding. Raf1, Raf2, RbcX2, and BSD2 stabilize folded RbcL intermediates until RbcS are positioned to form the complete L_8S_8 (Aigner et al., 2017). In addition to the *A. thaliana* synthetic biology (SynBio) system, two other systems exist for expressing *Nicotiana tabacum* Rubisco with either RbcX2 (Lin et al., 2020) or RbcX1 (Buck et al., 2023). Interestingly, interchangeability of *A. thaliana* and *N. tabacum* chaperonins and assembly factors permits assembly of either Rubisco to varying extents (Aigner et al., 2017; Lin et al., 2020).

In this study, we adapted these SynBio approaches to characterize Rubisco found in the pyrenoid of the model hornwort, *Anthoceros agrestis*. We found that *A. agrestis* Rubisco (*Aa*Rubisco) exhibits chaperone dependencies different to those of *A. thaliana*, with *Aa*Rubisco assembly not requiring Raf2 or RbcX, and that RbcX's involvement in Rubisco biogenesis can alter Rubisco catalytic rates. Our *A. agrestis* assembly chassis permits the assembly of Rubiscos from three other hornwort species: *Anthoceros fusiformis*, *Phymatoceros phymatodes*, and *Notothylas orbicularis*. We resolved four *Aa*Rubisco structures. Kinetic measurements indicate that carboxylation rates (k_{cat}^{c}) of hornwort Rubiscos are decoupled from their affinity (K_c) and specificity ($S_{c/o}$) for CO_2. This divergence from the canonical catalytic trade-off demonstrates that, indeed, Rubiscos are likely phylogenetically constrained (Bouvier et al., 2021). Further, *Aa*Rubisco assembled without RbcXs led to an increased accessibility of active sites and catalytic turnover. This new SynBio expression tool and Rubisco characterizations offer a pathway toward alternative approaches for enhancing carbon-assimilation efficiency by introducing the evolutionarily unique hornwort Rubisco and/or its pyrenoids into C_3 chloroplasts.

RESULTS

Non-cognate assembly factors result in *A. agrestis* Rubisco intermediate complexes

Attempts to purify *Aa*Rubisco from *A. agrestis* thallus yielded *Aa*Rubisco associated with higher-molecular-weight complexes and co-eluted with proteins containing chlorophyll pigments (Supplemental Figure 1). These heterogeneous and impure *Aa*Rubisco populations are unsuitable for downstream biochemical, kinetic, or structural work. It was surprising that *Aa*Rubisco proved so difficult to purify compared to algal Rubiscos that are housed inside pyrenoids (Wunder et al., 2018; Oh et al., 2023). However, given that these pyrenoids are found in distinct lineages and have independently evolved, this could potentially reflect differences in the pyrenoid matrix. For example, unlike *Chlamydomonas*, the *A. agrestis* pyrenoid is constitutively expressed and could be more tightly associated with its Rubisco. Further, Rubisco constitutes only about 3%–4% of soluble protein in hornworts (Hanson et al., 2002). Thus, we sought to heterologously express *Aa*Rubisco in *E. coli* BL21(DE3) cells. To establish whether an existing expression system would be sufficient to produce *Aa*Rubisco, we first utilized the established *A. thaliana* Rubisco SynBio system (Aigner et al., 2017; Wilson et al., 2019; Ng et al., 2020) by co-expressing *A. thaliana* chaperonins (p*At*C60αβ/C20) and assembly factors (p*At*Raf1/Raf2/RbcX2/BSD2) with *Aa*RbcL and C-terminal hexahistidine-tagged *Aa*RbcS (p*Aa*RbcLS) (Figure 1A and Supplemental Figure 2A).

Purified Rubisco (*Aa*RbcLS) co-eluted with *At*Raf1 on SDS–PAGE (Figure 1B and Supplemental Figure 2B) and migrated as three distinct bands on native PAGE, identified as *Aa*RbcL–*At*Raf1, *Aa*RbcLS–*At*Raf1, and *Aa*RbcLS (Figure 1C and Supplemental Figure 2C). These *Aa*Rubisco–*At*Raf1 complexes, representing assembly intermediates, are generally thought to exist transiently. More stable complexes have been obtained by using non-cognate Rubisco:chaperones, for example by

(C) Coomassie-stained native PAGE and western blot of *Aa*Rubisco intermediates using antibodies raised against Rubisco, Raf1, and the hexahistidine tag.

(D and E) Side view and top view of *Aa*RbcL_8:*At*BSD2_8 **(D)** and *Aa*RbcL_8S_4:*At*BSD2_4 **(E).**

(F and G) Polar contacts between BSD2 and RbcL of L_8B_8 **(F)** and $L_8B_4S_4$ **(G)**. BSD2 is colored orange and yellow, while RbcL is colored salmon and light blue. Amino acids involved in interface contacts are displayed as spheres.

(H and I) Polar contacts at RbcL monomer–monomer interfaces in L_8B_8 **(H)** and $L_8B_4S_4$ **(I)**. Interfaces I and II show interactions in the front view and back view, respectively of these monomer–monomers. Amino acids involved in these interactions are displayed as spheres, colored violet and jade.

expressing cyanobacterial RbcL with plant assembly factors, e.g., *Thermosynechococcus elongatus* BP1 Rubisco–*At*BSD2 (PDB: 6EKC) (Aigner et al., 2017) and *Synechococcus* sp. PCC6301 Rubisco–*At*Raf1–*At*BSD2 (PDB: 8ILB, 8ILM) (Wang et al., 2023). While cognate Rubisco:chaperone intermediate complexes were also observed in an *N. tabacum* Rubisco SynBio system, they have not been structurally characterized (Buck et al., 2023).

Our stable Rubisco:chaperone complexes present a unique opportunity to elucidate the Rubisco assembly process. Using cryo-electron microscopy (cryo-EM), we resolved the complexes into two distinct density maps and solved these structures, each representing different stages of Rubisco assembly (Figure 1D and 1E; Supplemental Figure 3). The first structure consists of an octameric *Aa*RbcL core in a symmetrical complex with eight units of *At*BSD2 (L_8B_8; 2.9 Å resolution), generally resembling PDB: 6EKC (Aigner et al., 2017). In the second structure, the octameric *Aa*RbcL core contains four units of *At*BSD2 on one end, while the opposite end is capped with four units of *Aa*RbcS ($L_8B_4S_4$; 2.9 Å resolution). We did not observe density corresponding to *Aa*Rubisco–*At*Raf1 (Supplemental Figure 3B), possibly because these interactions are unstable and were lost during cryo-EM preparation. Indeed, a previous study required the addition of chemical crosslinkers to stabilize cyanobacterial Rubisco–*At*Raf1 complexes (Wang et al., 2023).

PDBsum analysis (Laskowski et al., 1997) revealed several key interaction changes when four RbcS displaces four BSD2 and the complex transitions from L_8B_8 to $L_8B_4S_4$. First, when BSD2-2 interactions at interface II are interrupted by RbcS displacement, interactions between RbcL-A–BSD2-1 and RbcL-B–BSD2-1 change at interface I (Figure 1F and 1G; Supplemental Table 1). In particular, the C terminus of BSD2-1 loses several polar contacts with RbcL-B and forms new interactions with RbcL-A and RbcL-B. Displacement of BSD2-2 also results in RbcL-A and RbcL-B being oriented to form multiple new hydrogen bonds and salt bridges at the intra-dimer interface, interfaces I and II (Figure 1H and 1I; Supplemental Table 2). Notably, we identified interactions between RbcL and BSD2 not observed in cyanobacterial Rubisco–*At*BSD2 structures (Aigner et al., 2017; Wang et al., 2023) (PDB: 6EKC, 8ILB) (Supplemental Table 3). We resolved the usually flexible RbcL N terminus, 60s loop, and C terminus and observed movement of loop 6 toward the C terminus, covering the active site, once RbcS was bound. The substrate analog 2-carboxyarabinitol-1,5-bisphosphate (CABP) and Mg^{2+} ion were also observed bound to the active site on the side of the RbcL dimer with RbcS assembled (Supplemental Figure 4A and 4B). Indeed, carboxylase activity could be detected with activated enzymes (Supplemental Figure 2D). Lastly, RbcL residue Trp-411, which previously interacted with BSD2-2 (Supplemental Table 1), forms non-bonded interactions with two RbcS (Supplemental Figure 4C and Supplemental Table 4).

Overall, we identified multiple structural changes upon transition from L_8B_8 to $L_8B_4S_4$. L_8S_8 Rubisco holoenzymes were not identified in our cryo-EM analysis (Supplemental Figure 3B). This finding suggests incompatibilities between the non-cognate *A. thaliana* Rubisco assembly factors and *Aa*Rubisco subunits. For instance, amino acid differences between *A. thaliana* and *A. agrestis* RbcL, RbcS, and BSD2 (Supplemental Figure 2E–2G) could influence their tripartite protein–protein interaction. When *At*Rubisco is expressed with *A. thaliana* assembly factors, *At*RbcS could effectively displace *At*BSD2 that is bound to *At*RbcL to form L_8S_8 complexes (Aigner et al., 2017). However, when non-cognate *A. thaliana* assembly factors are used to assemble *Aa*Rubisco, *At*BSD2 is tightly bound to *Aa*RbcL in assembly intermediates (L_8B_8, $L_8B_4S_4$) and/or *Aa*RbcS could not effectively displace *At*BSD2 to form L_8S_8. Therefore, we reason that cognate *Aa*Rubisco assembly factors are essential toward successful L_8S_8 *Aa*Rubisco assembly.

*Aa*RbcX1 and *Aa*RbcX2 enhance heterologous expression of *A. agrestis* Rubisco

To assemble L_8S_8 *A. agrestis* Rubisco in *E. coli*, we selected *A. agrestis* chaperonins and assembly factors that were homologous to proteins in the *A. thaliana* SynBio system (Aigner et al., 2017). Only a single gene copy of each protein was found, due to a lack of whole-genome duplication in hornworts (Li et al., 2020). We developed a three-plasmid system emulating previous Rubisco SynBio system design (Figure 2A) (Aigner et al., 2017). p*Aa*RbcLS was co-expressed in *E. coli* BL21(DE3) cells with a plasmid encoding the *A. agrestis* chloroplast chaperonin protein machinery (p*Aa*C60$\alpha\beta$/C20) and a plasmid encoding *A. agrestis* Rubisco assembly factors (p*Aa*Raf1/Raf2/RbcX2/Bsd2/RbcX1).

Existing *A. thaliana* and *N. tabacum* Rubisco SynBio expression systems produce functional L_8S_8 Rubisco in the presence of only one of the RbcX isoforms, RbcX2 or RbcX1 (Aigner et al., 2017; Lin et al., 2020; Buck et al., 2023) (Supplemental Figure 6A and 6B). Both isoforms are also conserved throughout land plants (Supplemental Figure 6C). However, we found that this was not the case for *Aa*Rubisco assembly in *E. coli*. When *Aa*RbcX2, but not *Aa*RbcX1, was included during expression, two peaks were observed on the size-exclusion chromatography (SEC) elution profile (Figure 2B). Native PAGE analysis indicates that the first peak represents higher-molecular-weight complexes, likely where Rubisco is non-productively bound to chaperonins, and the second peak contains assembled L_8S_8 Rubisco holoenzymes (Figure 2C). Conversely, when only *Aa*RbcX1 (and not *Aa*RbcX2) was included (Supplemental Figure 6), we observed a marked improvement in Rubisco assembly (Figure 2B and 2C), with fewer higher-molecular-weight complexes (Figure 2B and 2C). Co-expressing both *Aa*RbcX1 and *Aa*RbcX2 in our expression system resulted in a more monodisperse chromatographic L_8S_8 Rubisco peak corresponding to the abundantly assembled *Aa*Rubisco complexes observed by Native PAGE analyses (Figure 2B and 2C). Thus, by co-expressing both RbcXs (along with the full complement of other chaperonins and chaperones), we were able to recombinantly assemble and purify homogeneous *Aa*Rubisco (Figure 2D). Expressing both *At*RbcX1 and *At*RbcX2 in the *A. thaliana* Rubisco SynBio system also results in the formation of more *At*Rubisco complexes (Supplemental Figure 7).

Prokaryotic and eukaryotic RbcX has been demonstrated to function by binding the C terminus of RbcL (Saschenbrecker et al., 2007; Liu et al., 2010; Bracher et al., 2015). We hypothesized that in our SynBio expression systems, the additional expression of RbcX1 functions to provide additional stability for Rubisco intermediates, thus facilitating additional holoenzyme assembly. Therefore, we examined whether we could observe such *Aa*Rubisco intermediates in an expression

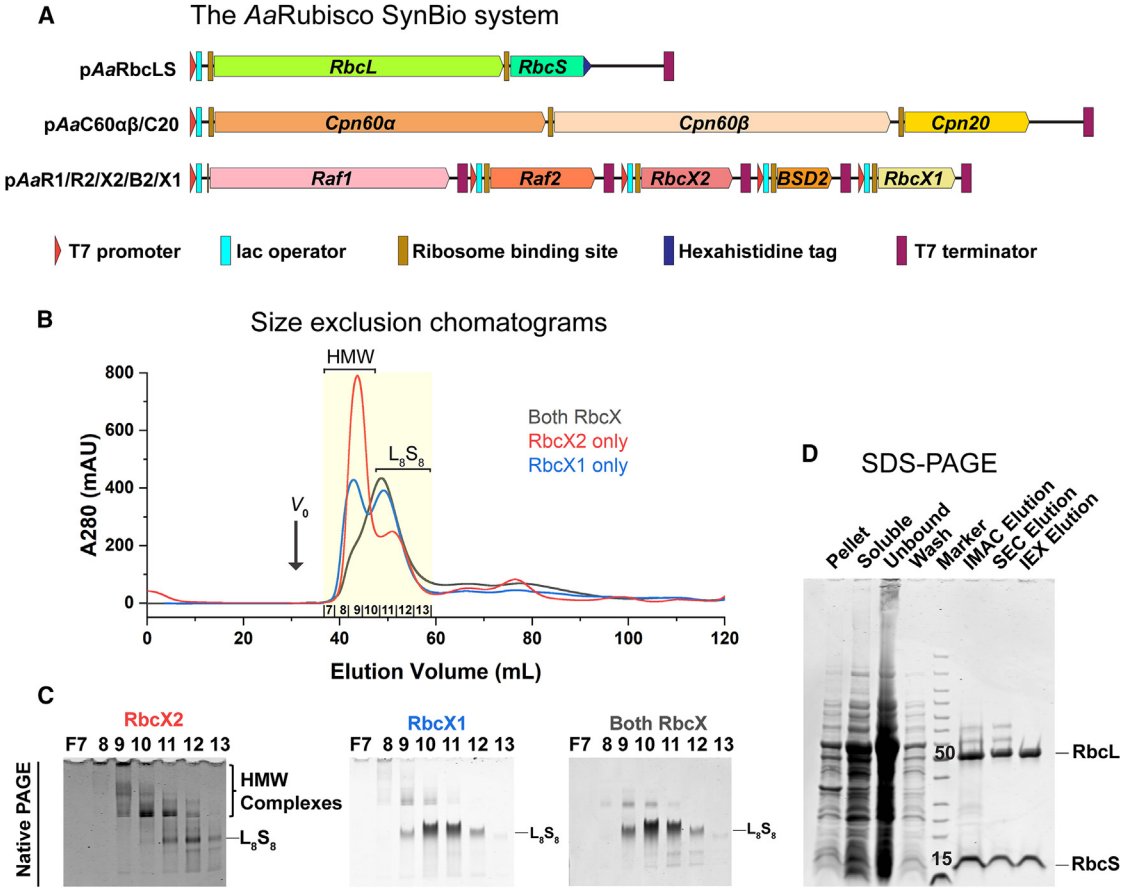


Figure 2. Purifying Rubisco from the *Aa*Rubisco SynBio system.

(A) Schematic of the three plasmids that constitute the *Aa*Rubisco SynBio system. p*Aa*RbcLS is used to express the *Aa*RbcL and C-terminal hexahistidine-tagged *Aa*RbcS, p*Aa*C60αβ/C20 to express chloroplast chaperonins 60αβ and chaperonin 20, and p*Aa*R1R2X2B2X1 for expressing Rubisco assembly factors Raf1, Raf2, RbcX2, BSD2, and RbcX1.

(B) Size-exclusion chromatography elution profile of *Aa*Rubisco assembled with chloroplast chaperonins, Raf1, Raf2, and BSD2, and either RbcX2, RbcX1, or both RbcX2 and RbcX1. Peaks corresponding to higher-molecular-weight complexes (HMW) and L_8S_8 Rubisco are indicated.

(C) Coomassie-stained native PAGE analysis of purified fractions. Yellow box on the chromatogram demarcates purified fractions from **(B)**.

(D) Coomassie-stained SDS–PAGE analysis of purified recombinant *Aa*Rubisco. Fractions analyzed are pellet, soluble, immobilized metal affinity chromatography (IMAC) unbound, wash and elution fractions, size-exclusion chromatography (SEC), and anion-exchange (IEX) chromatography elutions.

system that lacks RbcS by co-expressing p*Aa*C60αβ/C20 and p*Aa*Raf1/Raf2/RbcX2/Bsd2/RbcX1 with an N-terminal hexahistidine-tagged RbcL, p*Aa*NhisRbcL (Supplemental Figure 8A). We observed a single peak on the SEC elution profile that comprises mostly Cpn60 and trace amounts of DnaK, *Aa*RbcL and *Aa*Raf1 from the SDS–PAGE analysis. Correspondingly, native PAGE analysis showed Cpn60 complexes and a smear, suggesting that these Rubisco intermediates are heterogeneous and could have different chaperones bound to them (Supplemental Figure 8B and 8C). Future studies involving mutagenesis of RbcX and *in vitro* reconstitution with other assembly chaperones would be required to carefully evaluate the function of RbcX in Rubisco assembly.

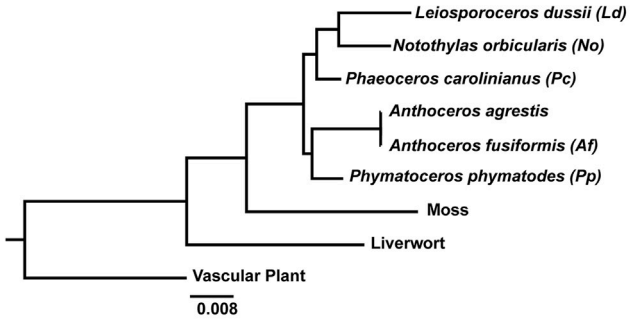
Other hornwort Rubiscos can be assembled using our *A. agrestis* expression chassis

We sought to evaluate whether our *Aa*Rubisco assembly chassis could assemble other hornwort Rubiscos by expressing different hornwort RbcL and RbcS with p*Aa*C60αβ/C20 and p*Aa*R1R2X2B2X1. Rubiscos from the hornwort species *A. fusiformis*, *N. orbicularis*, *Phaeoceros carolinianus*, *P. phymatodes*, and *Leiosporoceros dussii* were selected for testing, which together spans over 300 million years of evolution (Bechteler et al., 2023). RbcL from these hornwort species have high sequence identity compared to *A. agrestis*, with the exception of *A. fusiformis*, which is 100% identical (Figure 3A and 3B; Supplemental Figure 9A and 9B). The most abundant RbcS-encoding isoform was selected from transcript data for each species. All Rubiscos tested were able to be folded and assembled by the *A. agrestis* assembly chassis to varying degrees. Poor yields were obtained for *P. carolinianus* and *L. dussii*, while *A. fusiformis*, *N. orbicularis*, and *P. phymatodes* could be assembled relatively well (Figure 3C and Supplemental Figure 10). In particular, *A. fusiformis* Rubisco assembly was the best among the hornwort Rubiscos tested, in agreement with previous studies indicating that assembly factors Raf1, RbcX, and BSD2 mainly interact with RbcL (Liu et al., 2010; Hauser et al., 2015; Aigner et al.,

A Rubisco sequence identity (%)

Species	*Pc*	*No*	*Pp*	*Af*	*Ld*
RbcL	98.1	97.7	98.3	100	97.3
RbcS	76.8	72.6	74.2	87.2	71.8

B Rubisco Large subunit phylogenetic tree



C Native PAGE

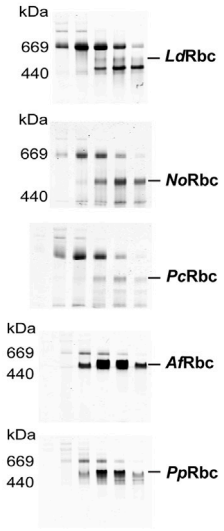


Figure 3. Assembly of other hornwort Rubiscos with *A. agrestis* SynBio system.
(A) Sequence identity of tested hornwort Rubiscos to *Aa*RbcL and *Aa*RbcS.
(B) Phylogenetic relationship of RbcL from hornworts (*Phaeoceros carolinianus*, *Notothylas orbicularis*, *Phymatoceros phymatodes*, *Anthoceros agrestis*, *Anthoceros fusiformis*, and *Leiosporoceros dussii*), liverwort (*Marchantia polymorpha*), moss (*Physcomitrella patens*), and vascular plant (*A. thaliana*). Sequence alignments were performed with AliView (Larsson, 2014) and uploaded to IQtree webserver (Nguyen et al., 2015). The output tree was visualized with FigTree (Rambaut, 2018).
(C) Coomassie-stained native PAGE analysis of size-exclusion chromatography purified hornwort Rubiscos assembled with the *A. agrestis* SynBio system from Figure 2.

2017) to assist in assembly during biogenesis. For the cases of *P. carolinianus* and *L. dussii*, it is likely that cognate Rubisco assembly factors and chaperonins are required for efficient and successful assembly.

A. agrestis Rubisco assembly is dependent on Raf1 and BSD2, but not Raf2

Heterologous *At*Rubisco assembly is dependent on the co-expression of assembly factors Raf1, Raf2, and BSD2 (Aigner et al., 2017). Raf2 null-mutant *Zea mays* and *A. thaliana* plants also have reduced Rubisco content (Feiz et al., 2014; Fristedt et al., 2018). To determine whether *Aa*Rubisco is similarly dependent on its cognate assembly factors, we evaluated *Aa*Rubisco assembly in the absence of each of these assembly factors. Omitting either Raf1 or Bsd2 from our expression system did not result in L_8S_8 Rubisco assembly (Figure 4A and 4B; Supplemental Figure 11), indicating that *Aa*Rubisco is dependent on both Raf1 and Bsd2 for assembly. However, when Raf2 was absent from our expression system, Rubisco holoenzyme assembly was observed. Thus, unlike *Arabidopsis*, *Aa*Rubisco is not dependent on Raf2. This finding suggests that hornworts, and possibly bryophytes (non-vascular plants), could have evolved a differential dependence on Raf2.

We previously demonstrated that both RbcX1 and RbcX2 are required for successful Rubisco holoenzyme assembly (Figure 2B and 2C). However, *Aa*Rubisco could also assemble in the absence of both RbcX isoforms (Figure 4 and Supplemental Figure 11). While a previous study reported that *At*Rubisco assembly was diminished ~50% when *At*RbcX2 was not expressed (Aigner et al., 2017), we did not observe such effects in *A. agrestis* (Supplememtal Figures 7 and 12). Overall, our findings suggest that the proper assembly of *Aa*Rubisco in *E. coli* requires either both or neither of RbcX to be present. We further characterized the minimal SynBio system by expressing p*Aa*RbcLS with p*Aa*C60αβ/C20 and p*Aa*Raf1/BSD2 (i.e., without RbcX1, RbcX2, or Raf2). Although we could observe the formation of Rubisco L_8S_8 complexes, Rubisco assembly intermediates were also present (Supplemental Figure 13). These data suggest that Raf2 and/or RbcX equilibrates the transition of Rubisco assembly intermediates to L_8S_8.

Hornwort Rubiscos have higher K_c and lower $S_{c/o}$ values compared with those of vascular plants

Hornwort Rubisco purified from our *Aa*Rubisco SynBio *E. coli* expression system was subjected to comprehensive kinetic analyses using heterologously expressed *At*Rubisco as a control (Supplemental Table 5 and Supplemental Figure 14A). Our measured *At*Rubisco k_{cat} (2.6 s^{-1}) was similar to previous records (Aigner et al., 2017), while its Michaelis–Menten constant for carboxylation, K_c was higher (23.1 μM) compared to other studies (9.9–10.7 μM) (Galmés et al., 2014; Atkinson et al., 2017). Generally, C_3 land plant Rubiscos exhibit lower K_c (15 ± 6 μM) and higher $S_{c/o}$ (97 ± 10) values compared to organisms with a CCM, as Rubisco is localized in the stromal environment (Flamholz et al., 2019).

Hornwort Rubiscos possess similar K_c (29.6–36.7 μM) and $S_{c/o}$ (64.5–70.5) values; however, different k_{cat}^c values were measured for *P. phymatodes* (3.0 s^{-1}), *A. fusiformis* (5.1 s^{-1}), and *A. agrestis* (6.4 s^{-1}) (Figure 5A and 5B; Supplemental Figure 14A and 14B; Supplemental Table 5). The higher k_{cat}^c values of these hornwort Rubiscos are thus accompanied by higher K_c and lower $S_{c/o}$ values. Although pyrenoids have not been observed in *A. fusiformis* and *P. phymatodes*, their lower $S_{c/o}$ values suggest they may cope through expression of carbonic anhydrase or employ an alternative CO_2-concentrating strategy. Alternatively, $S_{c/o}$ may be relatively invariable in hornwort Rubiscos regardless of CO_2 concentrations at active sites, as our data suggest that, unlike C_3 plant Rubiscos, hornwort Rubiscos completely break the canonical trade-off between $S_{c/o}$ and k_{cat}^c (Figure 5A and 5B).

Interestingly, *Aa*Rubisco assembled without *Aa*RbcX exhibited similar K_c and $S_{c/o}$ values but possessed a much higher k_{cat}^c (10.5 s^{-1}; 50% increase), further suggesting that hornwort Rubisco catalytic parameters are not as tightly coupled as in Rubiscos from

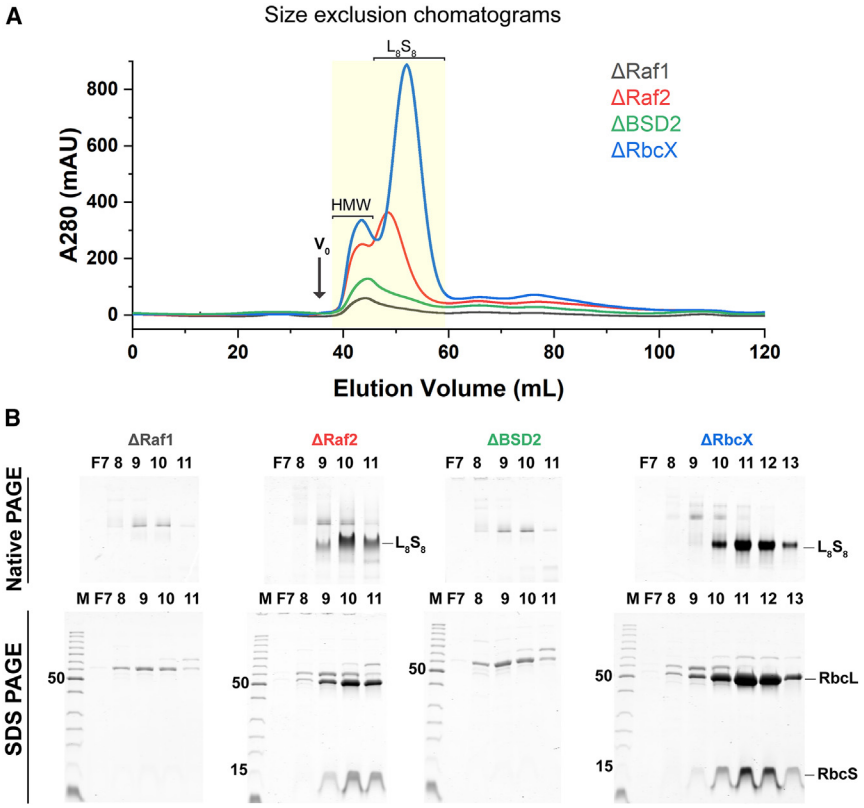


Figure 4. Assembly factor dependence of *Aa*Rubisco assembly

*Aa*Rubisco dependency on assembly factors assessed by removing respective assembly-factor-encoding genes from p*Aa*R1R2X2B2X1; Raf1 (ΔRaf1), Raf2 (ΔRaf2), BSD2 (ΔBSD2), and RbcX1 and RbcX2 (ΔRbcX).

(A) Size-exclusion chromatography assessment of Rubisco assembly, measured by absorbance at 280 nm.

(B) Coomassie-stained SDS–PAGE and native PAGE analysis of purified fractions. Yellow box on the chromatogram demarcates these purified fractions.

vascular plants. In contrast, *A. thaliana* Rubisco assembled with or without RbcXs had similar carboxylation rates (Supplemental Figure 14C). We also found that *Aa*Rubiscos assembled in the absence of RbcXs (such as ΔRbcXs and the minimal SynBio system; Figure 4B and Supplemental Figure 13) exhibit similar turnover rates. Contrastingly, *Aa*Rubiscos assembled with RbcXs (such as with RbcX1, both RbcXs, or ΔRaf2; Figures 2 and 4) exhibit lower turnover rates (Supplemental Figure 14D). This discovery prompted us to further investigate the role of RbcX in *Aa*Rubisco assembly.

Structural comparisons of *Aa*Rubiscos assembled with and without RbcX

We obtained high-resolution cryo-EM density maps of *Aa*Rubisco assembled with or without RbcX, resolved to 3.0 Å and 2.9 Å, respectively (Supplemental Figures 15 and 16). These two *Aa*Rubisco structures were similar to other form IB Rubisco structures, with an average root-mean-square deviation (RMSD) of 1.366 Å for all Cα atoms of the RbcL compared to either *A. thaliana* (PDB: 5IU0) (Valegård et al., 2018) or *C. reinhardtii* Rubisco structures (PDB: 1GK8) (Taylor et al., 2001). RbcL within the two different *Aa*Rubisco structures are highly similar, with an RMSD of 0.342 Å between all Cα atoms. CABP and an Mg^{2+} ion were also observed in the active sites of both Rubiscos (Supplemental Figures 15E and 16E). There were no obvious differences observed between these structures in terms of overall dimensions or secondary structure.

However, deeper analyses using PDBsum (Laskowski et al., 1997) revealed stark structural differences. Namely, differences in polar contacts were observed between RbcL monomers within L_2 dimers (Figure 6A–6C and Supplemental Table 6). More extensive interactions are observed between RbcL monomers at the front of the dimer in *Aa*Rubisco assembled in the presence of RbcX (interface I, Figure 6A and interface IV, Figure 6B). Differences in amino acid interactions lining the back of the dimers were also observed in the presence or absence of RbcX (interfaces II, III, V, Figure 6A and 6B; Supplemental Figure 17; Supplemental Table 6). Thus the higher k_{cat}^c *Aa*Rubisco assembled without RbcX has a looser RbcL monomer–monomer interface, consistent with the L_2 stapling action of RbcX during Rubisco biogenesis (Liu et al., 2010; Hauser et al., 2015; Aigner et al., 2017), and this differentially orients Rubisco dimers, resulting in a less accessible active site.

To investigate the accessibility of active sites in *Aa*Rubiscos, we characterized the rate of ribulose-1,5-bisphosphate (RuBP) release from RuBP-inhibited Rubisco complexes and the affinity for RuBP. While both *Aa*Rubiscos have a similar rate of inhibitor release, k_{obs} ($2.7 \pm 1.5 \times 10^{-3}$ and $2.7 \pm 8.6 \times 10^{-3}$ s^{-1}, respectively), Rubisco assembled with RbcX reached a faster final velocity, v_f (0.44 ± 0.10 s^{-1} compared to 0.21 ± 0.15 s^{-1}) (Supplemental Figure 18A). *Aa*Rubisco assembled without RbcX has a 4-fold lower K_M (28 ± 6 μM compared to 125 ± 15 μM), i.e., exhibiting higher affinity for RuBP (Supplemental Figure 18B–18E). Although chaperones were not observed bound to RbcL in cryo-EM structures (Supplemental Figures 15B and 16B), these observations suggest that a remnant amount of chaperones could still be bound to *Aa*Rubiscos. Indeed, RbcX was detected via mass spectrometry in *Aa*Rubisco assembled with RbcX (Supplemental Data 1). Thus, our evidence suggests that the dynamic binding of RbcX influences *Aa*Rubisco active-site accessibility, consistent with a non-competitive binding model of enzyme kinetics that affects maximal turnover rates.

DISCUSSION

In this study, we characterized hornwort Rubisco and its assembly processes and compared these to C_3 plants. Despite Rubisco evolving slowly (Bouvier et al., 2024), hornworts and vascular

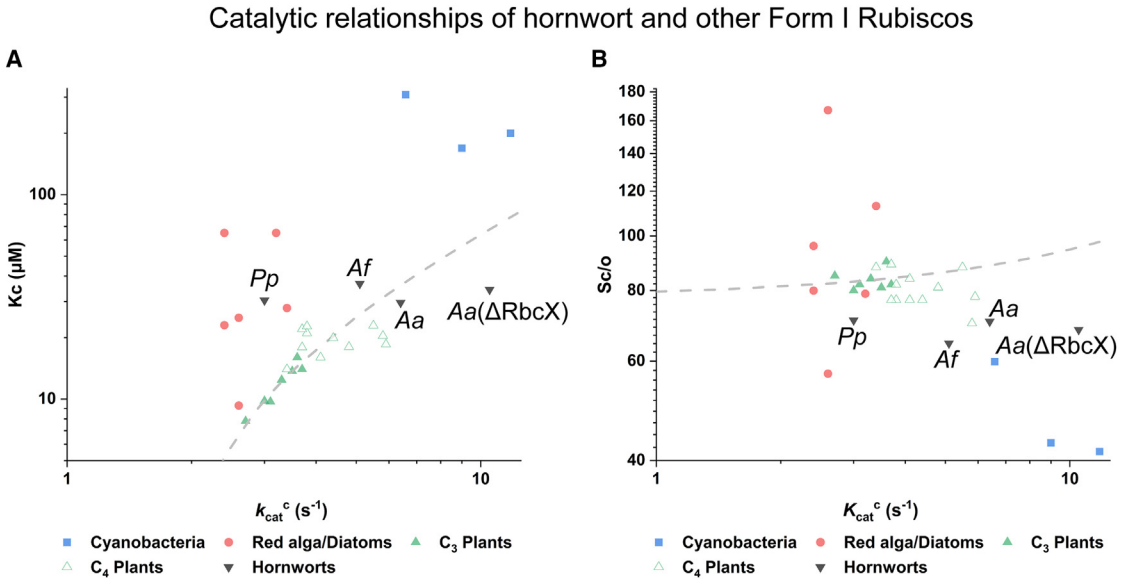


Figure 5. Kinetic characterization of hornwort Rubiscos.

(A and B) Catalytic relationships of hornwort Rubiscos and other form I Rubiscos. Contrasting Michaelis constant for CO_2 (K_c) **(A)**, specificity for CO_2 opposed to O_2 ($S_{c/o}$) **(B),** and the carboxylation speed (k_{cat}^{c}). Hornwort data points (black symbols) were compared with cyanobacteria (blue symbols), red algae/diatoms (red symbols), C_3 plants (green symbols), and C_4 plants (symbols with green outlines) data points (Sharwood, 2017). Gray dashed lines represent the trend lines generated with C_3 and C_4 plant data points, from which hornwort Rubiscos significantly deviate.

plants diverged from each other ~450 million years ago, and thus it is not unexpected that hornwort Rubiscos can deviate in their biogenesis requirements and kinetic properties. Indeed, other surveys of Rubiscos from divergent lineages (e.g., red algae; Young et al., 2016) and across time (e.g., ancestrally reconstructed Solanaceae Rubiscos; Lin et al., 2022) have revealed kinetic surprises that deviate from the "rules" written by vascular plant Rubiscos. Accessing the wealth of Rubisco evolution and kinetic data from diverse organisms has been, in part, hindered by the technical limitation of obtaining sufficient quantities and purity of Rubisco from the source material (Supplemental Figure 1). Our *A. agrestis* Rubisco SynBio system is a vital tool to circumvent these limitations for bryophyte Rubiscos.

Cognate Rubisco chaperones are important

Successful assembly of *N. tabacum* Rubisco through combining *A. thaliana* and *N. tabacum* chaperonins and chaperones demonstrates that the full cognate suite of chaperones is not always required (Lin et al., 2020). However, expressing *A. thaliana* chaperonins and chaperones was insufficient for *Aa*Rubisco assembly. Successful *Aa*Rubisco assembly was only achieved through expression of cognate *A. agrestis* chaperonins and chaperones. Our system also achieved varying success in assembling five other hornwort Rubiscos despite high RbcL conservation (Figure 3). Notably, *A. fusiformis* Rubisco, whose RbcL and RbcS sequences are 100% and 87.2% identical to *A. agrestis*, respectively, was assembled well by our system (Figure 3), supporting previous reports that Rubisco assembly factors mainly interact with RbcL and not RbcS (Liu et al., 2010; Hauser et al., 2015; Aigner et al., 2017), thus demonstrating that RbcL sequence variability limits the modularity of a universal hornwort Rubisco SynBio system and reiterating the need for cognate chaperones and chaperonins.

A new stable Rubisco assembly intermediate

By expressing *Aa*Rubisco with the *A. thaliana* Rubisco SynBio system (Aigner et al., 2017; Wilson et al., 2019; Ng et al., 2020), we observed eukaryotic Rubisco:chaperone bound intermediate complexes and could structurally resolve two of them, L_8B_8 and $L_8B_4S_4$. Increased stability through polar contacts between subunits (RbcL/RbcS/BSD2) is observed when L_8B_8 transitions to $L_8B_4S_4$. Notably, the four RbcS in this structure are positioned on one end of the intermediate. It is thus possible that addition of RbcS to L_8B_8 occurs sequentially on each end of the L_8 core, rather than simultaneously on both ends. While $RbcL_8Raf1_nBSD2_n$ (Wang et al., 2023) has been observed, the precise function of Raf2 remains unknown, nor is it known whether ternary complexes such as $RbcL_8RbcX_n$ $BSD2_n$ or cyanobacterial $RbcL_8RbcX_nRaf1_n$ (Li et al., 2022) complexes exist during eukaryotic Rubisco assembly. Future structural characterization of other eukaryotic Rubisco: chaperone intermediates would provide a more comprehensive view of the Rubisco biogenesis pathway. Nonetheless, we suggest that $L_8B_4S_4$, the first reported structure with this stoichiometry, is the penultimate assembly complex before the addition of the last four RbcS on the other end for the formation of the eukaryotic L_8S_8 holoenzyme (Figure 7A). $L_8B_4S_4$ populations have been inferred as being present in leaf chloroplasts and may thus be a physiologically relevant intermediate (Conlan et al., 2019).

Divergent biogenesis requirements

*Aa*Rubisco and *At*Rubisco exhibit similar, but different, chaperone requirements in Rubisco SynBio systems. In both

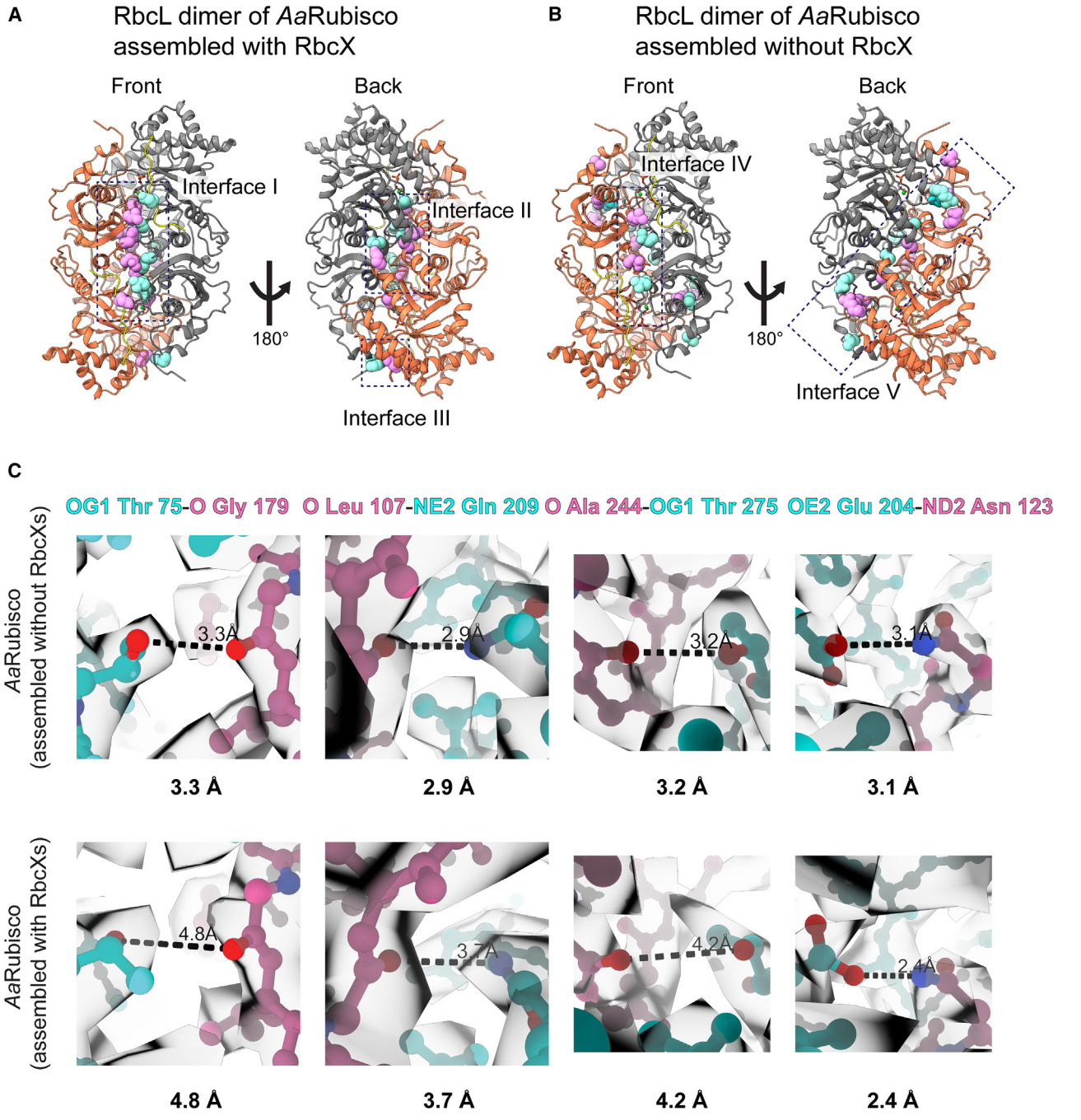


Figure 6. RbcX positions *Aa*Rubisco dimers closer.

(A and B) Dimer pairs of *Aa*Rubisco assembled with **(A)** and without **(B)** RbcX. Interfaces where there are differences in polar contacts between *Aa*Rubiscos are indicated. RbcLs are colored orange and gray. Amino acids involved in these interactions are shown as light-pink and teal spheres. Conserved RbcX-binding sites on the C terminus of RbcL monomers are shown in yellow.

(C) Density map and measured distances of protein–protein interactions between amino acids of RbcL monomers of both *Aa*Rubiscos.

systems, Raf1 and BSD2 serve essential roles in heterologous Rubisco biogenesis. We present the first evidence that expressing both RbcX1 and RbcX2 enhances both *Aa*Rubisco and *At*Rubisco assembly (Figure 2B and Supplemental Figure 7), and that absence of both RbcX1 and RbcX2 does not diminish yields (Figure 4B and Supplemental Figure 12). Contrastingly, while *At*Raf2 is necessary for *At*Rubisco assembly (Aigner et al., 2017), *Aa*Rubisco complexes can assemble without *Aa*Raf2 (Figure 4B). This suggests that non-vascular plants and vascular plants have diverged in their dependence on Raf2 for Rubisco biogenesis. Future ancestral reconstruction of Rubiscos and chaperones could provide insight into whether Raf2 reliance was gained by vascular plants or lost by hornworts.

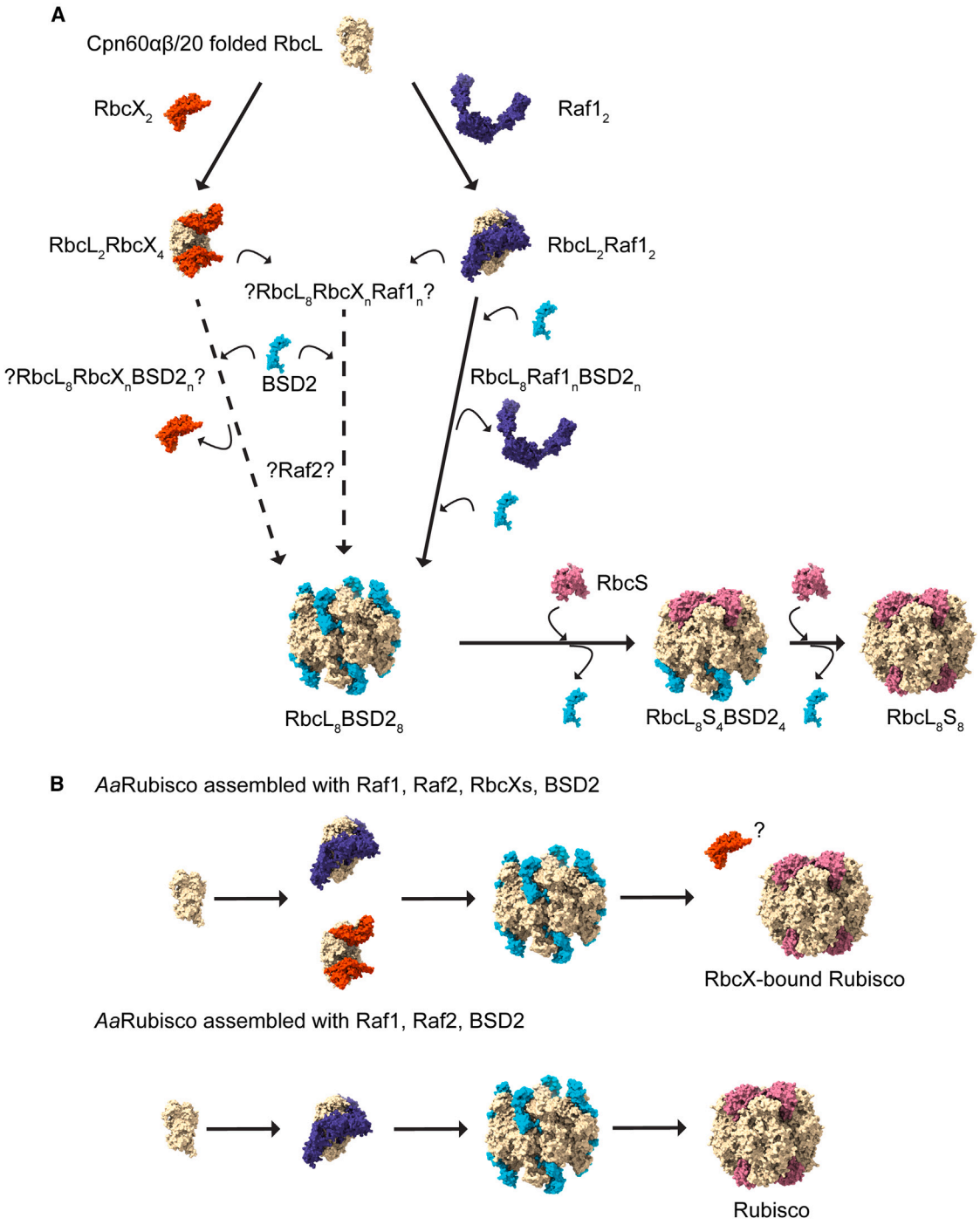


Figure 7. Proposed models of eukaryotic Rubisco assembly.

(A) Translated Rubisco large subunits are folded by Cpn60αβ/Cpn20. Next, Rubisco large subunit dimers are stabilized by Raf1 or RbcX dimers. Subsequently, binding of BSD2 displaces Raf1 or RbcX. Formation of $RbcL_8$:$Raf1_n$:$BSD2_n$ complexes (Wang et al., 2023) has been reported, while $RbcL_8$:$RbcX_n$:$BSD2_n$ and $RbcL_8$:$Raf1_n$:$RbcX_n$ complexes (Li et al., 2022) have yet to be identified in eukaryotic Rubisco assembly. RbcX and Raf1 could also act in cooperation. After the formation of $RbcL_8$:$BSD2_8$, RbcS displaces BSD2 sequentially on one of the ends, resulting in the formation of $RbcL_8S_4$:$BSD2_4$. Addition of the final RbcS displaces the remaining BSD2, forming the $RbcL_8S_8$ holoenzyme. The role of Raf2 remains unknown, and it possibly acts downstream of RbcX and Raf1.

(B) Simple diagram of *Aa*Rubisco assembly in the *Aa*Rubisco SynBio system. Presence and absence of RbcXs in the SynBio system affects the catalytic rate of assembled Rubisco. RbcL:Raf1 and RbcL:RbcX models were accessed from PDB: 3RG6 and 6KKM. Other models used were generated in this study. Question marks indicate complexes as inferred from the literature.

Divergent structure–function presents evolutionary and engineering opportunities

Different kinetics were measured for *Aa*Rubisco, but not *At*Rubisco, when expressed with or without RbcX, suggesting that this phenomenon might be unique to hornwort Rubiscos. Structural differences may cause these differences in Rubisco activity, as *Aa*Rubisco monomers are spaced slightly differently when assembled in the presence or absence of RbcX, with notable differences in the extent of protein–protein interactions at the RbcL monomer–monomer interface (Figures 5 and 6). These structural differences are observed after all chaperones have been displaced (Supplemental Figures 15 and 16), indicating that RbcX action on *Aa*Rubisco staples monomers together, altering the mature L_8S_8 structure. Differences in these interactions at the RbcL monomer–monomer interface could decrease holoenzyme stability and account for increased k_{cat}^c, consistent with previously described Rubisco activity–stability trade-offs (Studer et al., 2014). Different inhibitor-release kinetics also suggest that this different structure affects active-site accessibility (Figure 6).

We did not observe dramatic variation between these two structures, as even the active sites look similar (Supplemental Figures 15E and 16E). This is consistent with previously characterized Rubiscos that exhibit different kinetics but do not exhibit structural differences (Karkehabadi et al., 2005). Because assembled *At*Rubisco activity is not affected in the absence of *At*RbcX2 and residues involved in these monomer interfaces are conserved (Supplemental Figure 17), it is also possible that the concerted effort of *Aa*Raf1, *Aa*Raf2, and *Aa*BSD2 results in faster Rubisco activity or that this finding was a SynBio artifact that does not occur naturally. Alternatively, it is also possible that *A. agrestis* modulates its Rubisco kinetics through RbcX expression in response to environmental and/or developmental cues. To decipher this, future experiments are required to evaluate photosynthetic efficiency in *A. agrestis* *Aa*RbcX1 and *Aa*RbcX2 knockout mutants.

Rubisco's evolutionary potential is imposed by its ancestral lineage

Rubisco's turnover and affinity for CO_2 have often been classified to be at odds, whereby the improvement of one kinetic property comes at the cost of the other. However, recent studies propose a different perspective, namely that Rubisco instead is phylogenetically constrained and, hence, Rubisco's limited evolutionary capacity is imposed by its ancestry. In our study, we provide experimental data that suggest that this is indeed so. Our kinetic data show that hornwort Rubiscos exhibit higher k_{cat}^c, but higher K_c and lower $S_{c/o}$, values compared to Rubiscos from C_3 plants. This trend was observed even among hornworts that lack pyrenoids, suggesting that *P. phymatodes* and *A. fusiformis* possess other non-pyrenoid-based biophysical CCMs. Strikingly, hornwort Rubiscos break the canonical catalytic trade-off observed for vascular plants, with different hornwort Rubiscos exhibiting variable k_{cat}^c while K_c and $S_{c/o}$ remains more or less constant (Figure 5). Changing the chaperone complement not only affects assembly efficiency but also Rubisco's catalytic efficiency; *Aa*Rubiscos assembled without RbcX have much higher k_{cat}^c but similar K_c and $S_{c/o}$, further supporting that hornwort Rubiscos break the trade-off. Accordingly, these measured hornwort Rubisco kinetics do not fit with the "mechanistic coupling" or "selection within limits" model of Rubisco catalytic constraints (Flamholz et al., 2019). However, hornwort Rubisco kinetics could still be mechanistically constrained, consistent with the second Flamholz model, whereby differences in k_{cat}^c could be attributed to variation in RuBP enolization, although full oxygenation kinetics would need to be measured and derived to confirm this. This apparent kinetic uncoupling in hornwort Rubiscos warrants further survey of Rubiscos from diverse plant lineages, including other hornworts and bryophytes, which could deepen our mechanistic understanding of Rubisco kinetic divergence and potentially aid plant-engineering efforts.

Rubiscos that break the catalytic trade-off are of great interest for engineering efforts to enhance photosynthetic carbon assimilation in agricultural crop plants. For instance, eukaryotic form ID red-type Rubiscos that have high $S_{c/o}$ values have the potential to enhance C_3 plant growth (Young et al., 2016). However, efforts to functionally express these in C_3 plant chloroplasts were unsuccessful because the assembly factors for eukaryotic red-type Rubiscos are not known (Whitney et al., 2001; Lin and Hanson, 2018). Given that we have identified the required chaperonin/chaperone suite for functional expression of *Aa*Rubisco, endogenous proteins (Rubisco subunits, chaperonins, and assembly factors) could be removed before introducing our SynBio system, expressing this faster hornwort Rubisco into agricultural plants. *A. agrestis* and *A. fusiformis* Rubiscos exhibit distinct catalytic properties, despite 100% RbcL sequence conservation (Figure 5), which is consistent with previous reports as to the kinetic influence of RbcS (for review see Mao et al., 2023). Given the apparent lack of significant deviation in hornwort Rubisco $S_{c/o}$, future studies could evaluate the kinetics of hybrid Rubiscos comprising *Aa*RbcL and other bryophyte RbcS, circumventing the requirement to custom build a SynBio system for each species and permitting a wider kinetic survey. Alternatively, our findings suggest a possible engineering solution toward the evolutionary constraints faced by Rubisco through manipulation of native expressed Rubisco assembly factors that could either increase yield or improve kinetics.

RbcX: An engineering opportunity?

*Aa*Rubisco assembly was improved through the substitution of *Aa*RbcX2 with *Aa*RbcX1 or the expression of both *Aa*RbcXs. However, even with these optimizations, *Aa*Rubisco absorbance yield at 280 nm was ~400 mAU, which is half that of the *A. thaliana* Rubisco SynBio system, ~800 mAU (Figure 2B and Supplemental Figure 7A). The additional expression of *At*RbcX1 in the *A. thaliana* Rubisco SynBio system similarly doubled the yield of Rubisco complexes assembled, ~1600 mAU (Supplemental Figure 7B). Contrastingly, high yields of *Aa*Rubisco, ~800 mAU, could only be achieved in the absence of RbcXs (Figure 4B). The reasons for these different assembly efficiencies are not clear, but these data suggest potential synergistic RbcX1 and RbcX2 action whereby they could act in sequence or cooperatively. While experimental homodimeric structures of RbcX1 and RbcX2 exist from *A. thaliana* (PDB: 4GR2, 4GR6), it is not known whether RbcX1 and RbcX2 could form heterodimers *in planta* promoting different Rubisco assembly efficiency. In *A. thaliana*, RbcX2 is

transcribed at a stable level while RbcX1 expression increases under specific stress conditions (Kolesiński et al., 2011). Currently, no information is available on the transcriptomic regulation of *A. agrestis* RbcX1 and RbcX2 under relevant stress or developmental conditions. Further evaluation is needed to determine whether land plants evolved to switch their Rubisco assembly dependency between RbcX1, RbcX2, or both isoforms.

Differential *Aa*Rubisco assembly and activity in the presence/absence of RbcX raises the possibility of future chaperone engineering to modulate Rubisco kinetics. Directed evolution of molecular chaperones, by expressing the SynBio system in a Rubisco-dependent *E. coli* strain (Wilson et al., 2018), could provide a viable high-throughput route to achieving this. Domain swapping between *A. agrestis* and *A. thaliana* RbcX isoforms could pinpoint which of the myriad differences in amino acid residues and/or domains (Supplemental Figure 6) could be responsible for differential stapling action. Together, with experimental structures of *Aa*Rubisco in complex with RbcX, these experiments could provide a deeper mechanistic understanding of how *Aa*RbcX binding differentially orients RbcL monomers.

Overall, efforts from this study offer new opportunities to boost photosynthetic carbon assimilation in crops. Multiple attempts have been made to introduce a faster Rubisco (Occhialini et al., 2016; Gunn et al., 2020; Chen et al., 2023) in C_3 plants, with the goal of nitrogen re-assimilation for growth and fruit production, *Aa*Rubisco shows great potential to achieve this. In C_4 plants, increasing Rubisco abundance through overexpression of Raf1 or Raf2 (Salesse-Smith et al., 2018; Eshenour et al., 2024) has already led to improved photosynthesis and growth. Targeting RbcX presents yet another engineering approach to further increase Rubisco levels and enhance overall C_4 plant performance.

The *Aa*Rubisco SynBio system: A powerful tool for future research

Successful *in vitro* reconstruction of algal pyrenoids revealed detailed mechanistic properties of biophysical CCMs (Wunder et al., 2018; Oh et al., 2023). We envisage that the *A. agrestis* Rubisco SynBio system will be an essential tool for characterizing the land plant pyrenoid *in vitro*, especially for developing a mechanistic understanding of how hornwort Rubisco interacts with other pyrenoid components. Further, this SynBio tool can be used as a base system to address questions about Rubisco's catalytic evolution. Our system can fold other hornwort Rubiscos to varying degrees (Figure 3), and optimal expression will likely require replacing chaperone-encoding genes in our system with the cognate sequence for the Rubisco of interest. Our SynBio system also represents a basic chassis that can be adapted to assemble other bryophyte Rubiscos. This will permit a wider kinetic survey of hornwort and other bryophyte Rubisco kinetics, allowing more data points for examination of any potential catalytic constraints and trade-offs (or lack thereof). While *Aa*Rubisco exhibits the highest k_{cat}^{c} but lowest K_c and equivalent $S_{c/o}$ compared to the other hornwort Rubiscos characterized in this study, there is nothing to suggest that *Aa*Rubisco should be the most catalytically interesting bryophyte Rubisco: it is simply the model hornwort species. There may very well be other naturally evolved bryophyte Rubiscos that possess even more promising Rubiscos from both molecular evolution and crop-engineering angles.

Limitations of the study

We were unable to make direct comparisons between recombinant *Aa*Rubisco from our hornwort SynBio Rubisco expression system and bona fide *Aa*Rubisco (from *A. agrestis* thallus) due to technical difficulties. These included a significant culturing challenge to amass sufficient hornwort thallus, *Aa*Rubisco insolubility, and co-elution with chlorophyll-binding proteins. Further, the heterogeneity of *Aa*Rubisco populations from thallus due to differentially expressed *A. agrestis* Rubisco small subunit isoforms could influence kinetics.

From enzymatic assays, we observed a significant difference in *Aa*Rubisco kinetics based on the presence or absence of RbcX during assembly. This prompted further investigation to elucidate the structure–function relationship of these Rubiscos. The resolved density maps at 2.9 Å and 3.0 Å did not show large structural variation between them. Within each dataset, three-dimensional (3D) classifications only produced minor alternative classes that could not be refined to a high resolution (<3.5 Å) to build a good model for comparison against the main class. These data suggest that the main density map from each dataset was derived from the most stable Rubiscos and thus represents the majority population of the enzymes that were assayed. It also suggests that the differences between the two Rubisco types are small, which could only be discerned using intra-enzyme interaction analyses such as PDBsum.

Despite the above-mentioned evidence, we acknowledge the risks of drawing too strong conclusions about intermolecular forces from comparisons of circa 3.0-Å structures. This is because that the observed kinetics may not be conferred only by these structural variations, but other parameters such as the additional functions of hornwort RbcX (which were not explored in this study) remain possible. For now, our finding seems to be unique to hornworts, leaving the question open regarding the evolution of the role of RbcXs in Rubisco assembly across land plants.

METHODS

Plasmids

Gene fragments (*rbcL*, *RbcS*, *Raf1*,*Raf2*, *RbcX1*, *RbcX2*, and *BSD2*) were synthesized by Twist Bioscience. Fragments were subsequently amplified with KOD One PCR Master Mix (Toyobo) for cloning purposes. Chloroplast transit peptides of all genes, except *rbcL*, were removed based on homologs of *A. thaliana* genes. All plasmids of the recombinantly expressed proteins are listed in Supplemental Table 8.

Plasmid p*Aa*RbcLS contains genes, *rbcL* and *RbcS*, with a C-terminal 6×His. Both genes are under the control of a single T7 promoter and T7 terminator, with individual ribosome-binding sites for *rbcL* and *RbcS*. This plasmid was constructed with In-Fusion Cloning (ClonExpress Ultra One Step Cloning Kit C115; Vazyme) of amplified gene fragments of *rbcL*, *RbcS*, and the backbone plasmid of pRSFDuet-1 (Novogen), which contains a cassette encoding kanamycin resistance. All other hornwort Rubisco plasmids (pNhis*Aa*RbcL, p*Af*RbcLS, p*Pc*RbcLS, p*No*RbcLS, p*Pp*RbcLS, and p*Ld*RbcLS) were synthesized from GenScript

with protein expression elements similar to those described above for p*Aa*RbcLS.

Plasmid p*Aa*Cpn60αβCpn20 was synthesized from GenScript in the pET11a vector, which contains a marker for ampicillin resistance. Genes are under the control of a single T7 promoter and T7 terminator, with individual ribosome-binding sites for *Cpn60α*, *Cpn60β*, and *Cpn20*.

Plasmids p*Aa*Raf1Raf2RbcX2BSD2 and p*Aa*Raf1Raf2RbcX1BSD2 contain *Raf1*, *Raf2*, *RbcX1*, or *RbcX2* and *BSD2* genes, with each gene regulated and flanked by respective T7 promoters, ribosome-binding sites, and T7 terminators. These plasmids were constructed using the Golden Gate Cloning method with the type II restriction enzyme *Bsa*I with amplified gene fragments of *Raf1*, *Raf2*, *RbcX2*, or *RbcX1*, *BSD2*, and pCDF backbone, which contains a gene encoding streptomycin resistance. Plasmids p*Aa*Raf1Raf2RbcX2BSD2RbcX1, p*Aa*Raf2RbcX2BSD2RbcX1, p*Aa*Raf1RbcX2BSD2RbcX1, p*Aa*Raf1Raf2BSD2, p*Aa*Raf1Raf2RbcX2RbcX1, and p*Aa*Raf1BSD2 were synthesized from GenScript where genes are under the control of respective T7 promoters, ribosome-binding sites, and T7 terminators.

Plasmids p*At*Cpn60αβCpn20 and p*At*Raf1Raf2RbcX2BSD2 were a gift from Manajit Hayer-Hartl (Aigner et al., 2017). Plasmid p*At*RbcLS was a gift from Oliver Mueller Cajar (Ng et al., 2020).

Plasmids p*At*Raf1Raf2RbcX2BSD2RbcX1 and p*At*Raf1Raf2BSD2 were synthesized from GenScript. Design of these plasmids was kept similar to p*At*Raf1Raf2RbcX2BSD2, where all genes are regulated and flanked by respective T7 promoters, ribosome-binding sites, and T7 terminators except for *Raf1*, which shares the T7 terminator with *Raf2*.

Cultures

Luria–Bertani (LB) broth, pH 7.5 (tryptone 10 g/l, NaCl 5 g/l, yeast extract 5 g/l) was used to culture *E. coli* in both liquid cultures and solid agar plates (Bacto Agar 15 g/l). For molecular cloning and protein expression purposes, NEB 5α (NEB, USA) and BL21(DE3) (NEB) *E. coli* strains were used respectively. Antibiotics of ampicillin (200 μg/ml), kanamycin (30 μg/ml), and streptomycin (15 μg/ml) were added as required.

A. agrestis cultures

A. agrestis liquid cultures were grown in gametophyte growth medium with no sucrose added (Gunadi et al., 2022) in 250-ml flasks in a Percival culture chamber (CU-41L4). Growth conditions include a temperature of 22°C, 12:12-h day/night cycle, and light intensity of 73 μmol/m^2/s. Cool white light was emitted by F25T8 TL841 Alto fluorescent bulbs (Philips, USA). Cultures were also agitated on a Thermo MaxQ 2000 Digital Orbital Shaker at 130 rpm. Cultures were homogenized using a T18 digital ULTRA TURRAX homogenizer (IKA, Staufen, Germany) for approximately 10 s at 4000 rpm and subcultured with fresh medium approximately every 14 days to ensure optimal growth.

Lysis optimization from *Anthoceros* thallus

Two grams of *A. agrestis* or *A. fusiformis* tissue were harvested, the biomass was frozen with liquid N_2 and crushed to fine powder with a mortar and pestle. Ten milliliters of lysis buffer (20 mM Tris, 50 mM NaCl, 12.5% glycerol, 50 mM β-mercaptoethanol, 2% [v/v] Triton-X) supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF) and one protease inhibitor tablet (Roche) was added to resuspend the fine powder. Subsequent tests also include buffer screening through the addition of 500 mM L-arginine or L-glutamate to lysis buffers. The suspension was subjected to either ultrasonication (Branson Sonifier 250) at 30% amplitude for 2 min (10 s on/2 min off) or loaded on to a French pressure cell press (American Instrument Company, USA) at 1000 PSI. Polyvinylpolypyrrolidone 2% (PVPP) was added to lysates, and the mixture was sieved through four layers of Miracloth pre-wetted with the respective lysis buffer. The filtrate was centrifuged (21 000 *g* for 20 min at 4°C) to separate pellet and supernatant fractions before SDS–PAGE analysis.

Rubisco purification from *A. agrestis* thallus

One hundred grams of *A. agrestis* tissue was harvested from liquid cultures. The biomass was frozen with liquid N_2 and crushed to fine powder with a mortar and pestle. One hundred milliliters of *A. agrestis* lysis buffer (20 mM CAPS [pH 11.0], 50 mM NaCl, 12.5% glycerol) supplemented with 1 mM PMSF and two protease inhibitor tablets was added to resuspend the fine powder. The suspension was then subjected to lysis with a French pressure cell press at 1000 PSI. 2% PVPP was added to the French-pressed lysate, and the mixture was sieved through four layers of Miracloth pre-wetted with *A. agrestis* lysis buffer. The filtrate was centrifuged at 26 892 *g* with a Sorvall SS-34 rotor in the Sorvall RC-5B Refrigerated Superspeed Centrifuge at 4°C for 30 min. The resulting supernatant was then concentrated to 5 ml with a 30 kDa Amicon Ultra-15 Centrifugal Filter Unit (Merck) at 3220 *g* and 4°C.

The concentrated soluble lysate was loaded onto the SEC column, Superdex HiLoad 16/600 200 pg (Cytiva) equilibrated with buffer A (20 mM Tris–HCl [pH 8], 50 mM NaCl, 5% glycerol, 1 mM EDTA). Proteins were eluted in 5-ml fractions and analyzed with SDS–PAGE and native PAGE using 8%–16% and 7.5% Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad), respectively. Fractions containing Rubisco were flash frozen in liquid N_2 and stored at −80°C.

Antibodies

Commercial polyclonal rabbit antibodies raised against Rubisco (Agrisera, AS03-037A) were used in immunoblots to evaluate lysis optimization steps. Other Rubisco large subunit immunoblotting was performed with a polyclonal antibody raised in rabbits against the *A. agrestis* Rubisco large subunit peptide EVWKEIKFVFETIDTL and affinity purified (Life Technologies, USA). Rabbit antibodies raised against recombinant maize Raf1 were a gift from the David Stern Lab (Feiz et al., 2012). Monoclonal mouse antibodies raised against hexahistidine moieties were purchased from Millipore Sigma (SAB2702218).

Recombinant Rubisco purification

Recombinant Rubisco assemblies were produced and purified from *E. coli* BL21(DE3) cells harboring the relevant plasmids. In brief, 1 l of LB was inoculated with 40 ml of overnight grown starter cultures and incubated at 37°C/200 rpm until $OD_{600} \sim 0.6$. Isopropyl β-D-1-thiogalactopyranoside (1 mM) was added to induce protein production at 25°C/180 rpm overnight, and cells were harvested. Two liters of cell biomass was resuspended in 20 ml of lysis buffer (20 mM Tris–HCl [pH 8.0], 150 mM NaCl, 5% glycerol, 10 mM imidazole). Forty microliters of 0.5 M PMSF as a protease inhibitor, 40 μl of 1 M $MgCl_2$, and 4 μl of benzonase nuclease (Millipore) to degrade nucleic acids were added to the suspension. Cells were then lysed using a French pressure cell press at 1000 PSI. Cell lysate was centrifuged at 26 892 *g* with a Sorvall SS-34 rotor in the Sorvall RC-5B Refrigerated Superspeed Centrifuge at 4°C for 30 min.

Subsequently, soluble protein was subjected to a three-step purification strategy: (1) immobilized metal affinity chromatography (IMAC) to capture complexes containing His-tagged RbcS; (2) SEC to resolve $RbcL_8S_8$ from assembly intermediates; and (3) anion-exchange (IEX) chromatography to further purify these protein complexes. SEC and IEX chromatography were performed with columns connected to the NGC Quest 10 Plus Chromatography System (Bio-Rad).

Six milliliters of nickel–nitriloacetic acid (Ni–NTA) agarose (Qiagen) was washed with MilliQ water and resuspended in lysis buffer. Ni–NTA agarose was then added to the soluble cell lysate and incubated on a rotator (4°C, 45 min), to allow His-tagged RbcS to bind to the agarose. Thereafter the suspension was placed in an Econo-Column chromatography

column (2.5 × 10 cm, Bio-Rad), and unbound material was collected. To remove non-specific binding proteins, 50 ml of wash buffer (20 mM Tris–HCl [pH 8.0], 150 mM NaCl, 5% glycerol, 100 mM imidazole) was applied and collected as the wash fraction. Lastly, proteins were eluted by applying 25 ml of elution buffer (20 mM Tris–HCl [pH 8.0], 150 mM NaCl, 5% glycerol, 250 mM imidazole) and collected in 2-ml fractions. Fractions containing Rubisco were pooled and concentrated to 5 ml by spinning in a 30 kDa Amicon Ultra-15 Centrifugal Filter Unit (Merck) at 3220 *g* and 4°C.

Concentrated IMAC protein fractions were loaded onto the SEC column equilibrated with Buffer A. Proteins were eluted in 3-ml fractions. Fractions containing Rubisco were pooled and concentrated to 5 ml as described above.

Concentrated SEC protein fractions were loaded onto the IEX column CaptoQ HiRes 10/100 (Cytiva) column equilibrated with buffer A. Proteins were eluted in 3-ml fractions by applying buffer B (20 mM Tris [pH 8], 1 M NaCl, 5% glycerol, 1 mM EDTA) over an NaCl concentration gradient to 0.5 M NaCl. Fractions with pure Rubisco holoenzyme were concentrated, subjected to ultra-diafiltration with buffer A, aliquoted, flash frozen in liquid N_2, and stored at −80°C.

Sample preparation for cryo-EM

Rubisco samples were activated with 10 mM $MgCl_2$ and 40 mM $NaHCO_3$ for 15 min. Subsequently, transition-stated analog, CABP, was added to a final concentration of 0.4 mM and incubated for 15 min. CABP-bound Rubisco samples were then subjected to buffer exchange with Micro Bio-Spin P-6 Gel columns (Bio-Rad) equilibrated with buffer A (20 mM Tris [pH 8.0], 50 mM NaCl). Eluted proteins were diluted to 0.5–1 mg/ml.

Cryo-EM single-particle analysis and image processing

400-mesh Quantifoil 1.2/1.3 gold grids with 2 nm carbon support (Electron Microscopy Sciences) were glow discharged for 45 s. The sample chamber of the Vitrobot Mark IV (FEI) was maintained at 4°C with 100% relative humidity. Four microliters of sample was applied onto the grids and blotted (blot force of 0 for 4 s). The grids were then plunge frozen into liquid ethane. Grids were loaded onto a Talos Arctica 200 kV (Thermo Fisher) transmission electron microscope equipped with a K3 detection camera (Gatan). Grids containing appropriate sample quality and ice thickness were subjected to full automatic data acquisition using the software EPU (Thermo Scientific).

Movie parameters for each dataset are as follows: 1477 super-resolution movies of *Aa*Rubisco:*At*Chaperone intermediates were captured at 79 000× magnification corresponding to a pixel size of 0.545 Å; 1527 super-resolution movies of *Aa*Rubisco complexes were captured at 63 000× magnification corresponding to a pixel size of 0.695 Å; and 2004 movies of *Aa*Rubisco (assembled without RbcX) complexes were captured at 63 000× magnification corresponding to a pixel size of 1.39 Å. For all datasets, dose rate was set to a total of 65 electrons collected over 50 movie frames whereby each frame received a dose of 1.3 electrons per pixel. Spherical aberration constant of the objective lens was 2.7 mm with an objective aperture of 100 μm. All datasets were imaged at a defocus range between −0.5 and −2.5 μm. Movies were imported into Cryosparc v.4 (Punjani et al., 2017). Motion correction was conducted using the Patch Motion Correction function, and contrast transfer function (CTF) was corrected using the Patch CTF Estimation function. Particles were first extracted with either 2× or 4× binning and subjected to several rounds of reference-free 2D and 3D classification to remove bad particles. The good particles were then re-extracted without binning and 3D refined (homogeneous refinement) to produce the final electron density maps. Local resolution of the maps was estimated in Cryosparc.

Model building and fitting

The model of Rubisco from *C. reinhardtii* (Taylor et al., 2001) (PDB: 1GK8) was docked into the density map using the "fit in map" function in UCSF Chimera (Pettersen et al., 2004). The amino acid sequence was then changed to *A. agrestis* using CHAINSAW (Stein, 2008) and was used as rigid body docking in COOT (Emsley et al., 2010) to produce a working model of *A. agrestis* Rubisco. The model of BSD2 was acquired from PDB: 8ILB and 8ILM (Wang et al., 2023) and docked into the L8B8 and L8B4S4 density maps. Real space refinement was conducted using Phenix (Adams et al., 2010), and the models were then fine-tuned in COOT by manually going through the large and small subunits of Rubisco and BSD2. An additional iteration of real space refinement in Phenix and manual adjustment in COOT was conducted to produce the final models. Figures in this paper were produced using UCSF ChimeraX (Pettersen et al., 2021).

Spectrophotometric Rubisco assay

Rubisco activity measurements were conducted using a coupled-enzyme assay as described by Kubien et al. (2011), with omission of Triose-P isomerase/glycerol-3-phosphate dehydrogenase in the enzyme reaction mix. This results in two NADH oxidation events per RuBP carboxylated. Turnover was assessed by depletion of NADH absorbance at 340 nm at 25°C using a Cary 60 spectrophotometer and PCB-1500 circulating water bath (Agilent). Purified protein fractions were first activated by adding 50 mM $MgCl_2$ and 0.4 M $NaHCO_3$ to protein samples in a ratio of 2:1:7 and incubated at room temperature for 1 h.

Assays were performed in a 100-μl reaction mixture, where protein samples were added to an enzyme master mix comprising assay buffer (100 mM tricine [pH 8.0], 10 mM $MgCl_2$) with creatine phosphokinase (25 units/ml), glyceraldehyde-3-phosphate dehydrogenase (25 units/ml), 3-phosphoglycerate kinase (25 units/ml), 20 mM $NaHCO_3$, 0.5 mM NADH, 2 mM ATP, and 10 mM creatine phosphate. To initiate the reaction, RuBP was added to a final concentration of 1 mM. A_{340} was measured for 3 min, from which molar concentration of NADH depleted can be calculated with its extinction coefficient (0.0062 μM^{-1} cm^{-1}). Concentration of 3-phosphoglyceric acid (3-PGA) produced was derived from concentration of NADH depleted. 0.5 μM *At*Rubisco active sites were analyzed as positive controls, while omission of Rubisco or RuBP was analyzed as negative controls.

To determine K_M for RuBP for *Aa*Rubiscos, 0.4–2.5 μM activated Rubisco was added to the enzyme master mix as described above, with varying RuBP concentrations, 3.125–1000 μM. Turnover rates for each reaction were determined by dividing concentration of 3-PGA against time and concentration of Rubisco. Data were plotted with OriginPro 2023, and the Michaelis–Menten model-fitting function was used to determine K_M of RuBP.

To estimate the rate of inhibitor release from inhibited *Aa*Rubiscos, we prepared RuBP-inhibited enzymes by incubating 8 μM Rubisco with 4 mM EDTA for 10 min before adding a final concentration of 1 mM RuBP and further incubating for 50 min. 0.4 μM inhibited enzymes were added to the enzyme master mix as described above with 1 mM RuBP and assayed for 10 min. Product accumulation plotted against time was fitted with OriginPro 2023 using a non-linear fitting function (Pearce and Andrews, 2003; Pearce, 2006)

$$[\text{product}] = v_f \cdot t + ((v_i - v_f)(1 - \exp(-k_{obs} \cdot t)) / k_{obs}$$

to estimate initial velocity v_i, final velocity v_f, and k_{obs}, the first-order rate constant.

Radiometric kinetic assays

$^{14}CO_2$ fixation assays were performed as described by Ng et al. (2020), with 0.5-ml reaction volume at 25°C in 7.7 ml of septum-capped glass scintillation vials with assay buffer (100 mM 4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid [EPPS]–NaOH [pH 8.0], 15 mM $MgCl_2$, 1 mM EDTA), 10 μg/ml carbonic anhydrase, and 1 mM RuBP. All assay

components were equilibrated with nitrogen gas prior to addition of $^{14}CO_2$ concentration varying from 0.14 to 13.25 mM $NaH^{14}CO_3$ (corresponding to 7–162 μM $^{14}CO_2$). 5 μM purified Rubisco active sites were first activated by incubation in an assay buffer containing 10 mM $NaH^{14}CO_3$. Assays were initiated by addition of 20 μl of activated Rubisco and quenched after 1 min using 200 μl of 50% (v/v) formic acid. Specific activity of $^{14}CO_2$ was determined from complete turnover of 5.24 nmol RuBP under saturating $^{14}CO_2$ concentrations, ranging from 1350 to 2000 CPM/nmol RuBP. All reactions were dried using a heat block at 110°C before resuspension in 750 μl of water and 1 ml of Ultima Gold LSC cocktail for quantification by scintillation. Rubisco active sites were quantified using [^{14}C]CABP. In brief, [^{14}C]CABP-bound Rubisco was separated from free ligands by SEC (Econo-Column chromatography column, 0.7 × 30 cm, Bio-Rad) packed with 10 ml of Sephadex G-50 gel filtration medium (Cytiva) in column buffer (20 mM EPPS–NaOH [pH 8.0], 75 mM NaCl) and quantified by scintillation counting. Assay data were processed and plotted with OriginPro 2023, and the Michaelis–Menten model-fitting function was used to determine k_{cat}^c and *Kc*.

Specificity assay

CO_2:O_2 specificity ($S_{c/o}$) assays were carried out at 25°C as described by Kane et al. (1994). In brief, purified Rubiscos were assayed in 20 ml of septum-capped glass scintillation vials containing 1 ml of 30 mM triethanolamine–HCl (pH 8.3), 10 mM $MgSO_4$, and 10 μg/ml carbonic anhydrase. Reactions were equilibrated in a defined gas mixture (995 291 M O_2; 709 M CO_2) prior to the addition of [1-^{3}H]RuBP for assay initiation. After 1 h, reaction products were dephosphorylated using alkaline phosphatase (10 U/reaction) and separated on an Aminex HPX-87H column (Bio-Rad). Eluted peaks corresponding to radiolabeled glycerate and glycolate were quantified by scintillation for calculation of $S_{c/o}$.

Mass spectrometry

Trypsin-digested purified *Aa*Rubisco assembled with and without *Aa*RbcXs was analyzed with nano-liquid chromatography–tandem mass spectrometry using an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific) at the Boyce Thompson Institute Mass Spectrometry Center. Sample preparation and analysis were performed similarly to a previously described method (Zhang et al., 2024). Protein identification and quantification were performed with Proteome Discoverer 2.5 (Thermo Fisher Scientific) with reference to *E. coli* DE3 proteome (Uniprot Proteome: UP000002032) and expressed recombinant proteins in the *Aa*Rubisco SynBio system (*Aa*Rubisco subunits; *Aa*Chaperonins and *Aa*Rubisco assembly factors). Mass spectrometry data were sorted to identify fold differences between *Aa*Rubisco samples.

DATA AND CODE AVAILABILITY

Electron density maps have been deposited in the Electron Microscopy Data Bank with the accession codes EMD-45607 for L_8B_8, EMD-45608 for $L_8B_4S_4$, EMD-45641 for L_8S_8, and EMD-45605 for L_8S_8 assembled without RbcX. Molecular models have been deposited in the Protein Data Bank with associated accession numbers PDB: 9CI1 for L_8B_8, 9CI2 for $L_8B_4S_4$, 9CK5 for L_8S_8, and 9CHZ for L_8S_8 assembled without RbcX. The mass spectrometry proteomic data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD052489.

FUNDING

This research was supported by National Science Foundation grant no. MCB-2213840 to L.H.G. and MCB-2213841 to F.-W.L.

ACKNOWLEDGMENTS

We thank Manajit Hayer-Hartl for the gift of plasmids p*At*C60αβ/C20 and p*At*Raf1/Raf2/RbcX2/BSD2, Tsai Yi-Chin and Oliver Mueller-Cajar for the gift of RuBP and p*At*RbcLS, and Kathryn Eshenour, Amber Hotto, and David Stern for the gift of CABP and Raf1 antibodies. We acknowledge the Boyce Thompson Institute Mass Spectrometry Center for mass spectrometry services, Venkatesh Thirumalaikumar and Aleksandra Skirycz for performing the mass spectrometry, and Bingsen Zhang for analyzing the data. We acknowledge the Cornell Center for Materials Research facility and Mariena Silvestry-Ramos for providing technical expertise and maintenance of electron microscopes. We thank Christopher Fromme for access to computing resources for processing initial cryo-EM datasets. Authors declared no competing interests.

AUTHOR CONTRIBUTIONS

Z.G.O., F.-W.L., and L.H.G. conceived the project. Z.G.O planned and performed most experiments, designed the plasmids for the *Aa*Rubisco SynBio system, and performed protein purification, enzyme kinetic measurements, model building of *Aa*Rubiscos, and PDBsum analysis. T.A.R. cultivated hornwort liquid cultures and retrieved and analyzed hornwort Rubisco subunit, chaperonin, and chaperone sequences. D.H.L. assisted in protein expression and purification. W.S.L.A. performed cryo-EM sample preparation, data collection, and processing and assisted Z.G.O. in model building. J.Z.Y.N. purified protein and conducted specificity assay experiments. F.-W.L. and L.H.G. supervised the study. Z.G.O wrote the manuscript with contributions and comments from all authors.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Molecular Plant Online*.

Received: July 3, 2024
Revised: October 14, 2024
Accepted: October 31, 2024
Published: November 2, 2024

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. **66**:213–221.

Adler, L., Díaz-Ramos, A., Mao, Y., Pukacz, K.R., Fei, C., and McCormick, A.J. (2022). New horizons for building pyrenoid-based CO_2-concentrating mechanisms in plants to improve yields. Plant Physiol. **190**:1609–1627.

Aigner, H., Wilson, R.H., Bracher, A., Calisse, L., Bhat, J.Y., Hartl, F.U., and Hayer-Hartl, M. (2017). Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD_2. Science **358**:1272–1278.

Atkinson, N., Leitão, N., Orr, D.J., Meyer, M.T., Carmo-Silva, E., Griffiths, H., Smith, A.M., and McCormick, A.J. (2017). Rubisco small subunits from the unicellular green alga Chlamydomonas complement Rubisco-deficient mutants of *Arabidopsis*. New Phytol. **214**:655–667.

Bechteler, J., Peñaloza-Bojacá, G., Bell, D., Gordon Burleigh, J., McDaniel, S.F., Christine Davis, E., Sessa, E.B., Bippus, A., Christine Cargill, D., Chantanoarrapint, S., et al. (2023). Comprehensive phylogenomic time tree of bryophytes reveals deep relationships and uncovers gene incongruences in the last 500 million years of diversification. Am. J. Bot. **110**:e16249.

Bouvier, J.W., Emms, D.M., Rhodes, T., Bolton, J.S., Brasnett, A., Eddershaw, A., Nielsen, J.R., Unitt, A., Whitney, S.M., and Kelly, S. (2021). Rubisco Adaptation Is More Limited by Phylogenetic Constraint Than by Catalytic Trade-off. Mol. Biol. Evol. **38**:2880–2896.

Bouvier, J.W., Emms, D.M., and Kelly, S. (2024). Rubisco is evolving for improved catalytic efficiency and CO_2 assimilation in plants. Proc. Natl. Acad. Sci. USA **121**:e2321050121.

Bracher, A., Hauser, T., Liu, C., Hartl, F.U., and Hayer-Hartl, M. (2015). Structural Analysis of the Rubisco-Assembly Chaperone RbcX-II from Chlamydomonas reinhardtii. PLoS One **10**:e0135448.

Buck, S., Rhodes, T., Gionfriddo, M., Skinner, T., Yuan, D., Birch, R., Kapralov, M.V., and Whitney, S.M. (2023). Escherichia coli expressing chloroplast chaperones as a proxy to test heterologous Rubisco production in leaves. J. Exp. Bot. **74**:664–676.

Chen, T., Riaz, S., Davey, P., Zhao, Z., Sun, Y., Dykes, G.F., Zhou, F., Hartwell, J., Lawson, T., Nixon, P.J., et al. (2023). Producing fast and active Rubisco in tobacco to enhance photosynthesis. Plant Cell **35**:795–807.

Conlan, B., Birch, R., Kelso, C., Holland, S., De Souza, A.P., Long, S.P., Beck, J.L., and Whitney, S.M. (2019). BSD2 is a Rubisco-specific assembly chaperone, forms intermediary hetero-oligomeric complexes, and is nonlimiting to growth in tobacco. Plant Cell Environ. **42**:1287–1301.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. **66**:486–501.

Eshenour, K., Hotto, A., Michel, E.J.S., Oh, Z.G., and Stern, D.B. (2024). Transgenic expression of Rubisco accumulation factor2 and Rubisco subunits increases photosynthesis and growth in maize. J. Exp. Bot. **75**:4024–4037.

Fei, C., Wilson, A.T., Mangan, N.M., Wingreen, N.S., and Jonikas, M.C. (2022). Modelling the pyrenoid-based CO_2-concentrating mechanism provides insights into its operating principles and a roadmap for its engineering into crops. Nat. Plants **8**:583–595.

Feiz, L., Williams-Carrier, R., Wostrikoff, K., Belcher, S., Barkan, A., and Stern, D.B. (2012). Ribulose-1,5-bis-phosphate carboxylase/oxygenase accumulation factor1 is required for holoenzyme assembly in maize. Plant Cell **24**:3435–3446.

Feiz, L., Williams-Carrier, R., Belcher, S., Montano, M., Barkan, A., and Stern, D.B. (2014). A protein with an inactive pterin-4a-carbinolamine dehydratase domain is required for Rubisco biogenesis in plants. Plant J. **80**:862–869.

Flamholz, A.I., Prywes, N., Moran, U., Davidi, D., Bar-On, Y.M., Oltrogge, L.M., Alves, R., Savage, D., and Milo, R. (2019). Revisiting Trade-offs between Rubisco Kinetic Parameters. Biochemistry **58**:3365–3376.

Freeman Rosenzweig, E.S., Xu, B., Kuhn Cuellar, L., Martinez-Sanchez, A., Schaffer, M., Strauss, M., Cartwright, H.N., Ronceray, P., Plitzko, J.M., Förster, F., et al. (2017). The Eukaryotic CO_2-Concentrating Organelle Is Liquid-like and Exhibits Dynamic Reorganization. Cell **171**:148–162.e19.

Fristedt, R., Hu, C., Wheatley, N., Roy, L.M., Wachter, R.M., Savage, L., Harbinson, J., Kramer, D.M., Merchant, S.S., Yeates, T., et al. (2018). RAF2 is a RuBisCO assembly factor in *Arabidopsis thaliana*. Plant J. **94**:146–156.

Galmés, J., Kapralov, M.V., Andralojc, P.J., Conesa, M.À., Keys, A.J., Parry, M.A.J., and Flexas, J. (2014). Expanding knowledge of the Rubisco kinetics variability in plant species: environmental and evolutionary trends. Plant Cell Environ. **37**:1989–2001.

Gunadi, A., Li, F.-W., and Van Eck, J. (2022). Accelerating gametophytic growth in the model hornwort. Appl Plant Sci **10**:e11460.

Gunn, L.H., Martin Avila, E., Birch, R., and Whitney, S.M. (2020). The dependency of red Rubisco on its cognate activase for enhancing plant photosynthesis and growth. Proc. Natl. Acad. Sci. USA **117**:25890–25896.

Hanson, D., Andrews, T.J., and Badger, M.R. (2002). Variability of the pyrenoid-based CO_2 concentrating mechanism in hornworts (Anthocerotophyta). Funct. Plant Biol. **29**:407–416.

Hauser, T., Bhat, J.Y., Miličić, G., Wendler, P., Hartl, F.U., Bracher, A., and Hayer-Hartl, M. (2015). Structure and mechanism of the Rubisco-assembly chaperone Raf1. Nat. Struct. Mol. Biol. **22**:720–728.

Hennacy, J.H., and Jonikas, M.C. (2020). Prospects for Engineering Biophysical CO_2 Concentrating Mechanisms into Land Plants to Enhance Yields. Annu. Rev. Plant Biol. **71**:461–485.

Itakura, A.K., Chan, K.X., Atkinson, N., Pallesen, L., Wang, L., Reeves, G., Patena, W., Caspari, O., Roth, R., Goodenough, U., et al. (2019). A Rubisco-binding protein is required for normal pyrenoid number and starch sheath morphology in *Chlamydomonas reinhardtii*. Proc. Natl. Acad. Sci. USA **116**:18445–18454.

Kane, H.J., Viil, J., Entsch, B., Paul, K., Morell, M.K., and Andrews, T.J. (1994). An Improved Method for Measuring the CO_2/O_2 Specificity of Ribulosebisphosphate Carboxylase-Oxygenase. Funct. Plant Biol. **21**:449.

Karkehabadi, S., Taylor, T.C., Spreitzer, R.J., and Andersson, I. (2005). Altered intersubunit interactions in crystal structures of catalytically compromised ribulose-1,5-bisphosphate carboxylase/oxygenase. Biochemistry **44**:113–120.

Karlsson, J., Clarke, A.K., Chen, Z.Y., Hugghins, S.Y., Park, Y.I., Husic, H.D., Moroney, J.V., and Samuelsson, G. (1998). A novel alpha-type carbonic anhydrase associated with the thylakoid membrane in Chlamydomonas reinhardtii is required for growth at ambient CO_2. EMBO J. **17**:1208–1216.

Kolesiński, P., Piechota, J., and Szczepaniak, A. (2011). Initial characteristics of RbcX proteins from *Arabidopsis thaliana*. Plant Mol. Biol. **77**:447–459.

Kubien, D.S., Brown, C.M., and Kane, H.J. (2011). Quantifying the amount and activity of Rubisco in leaves. Methods Mol. Biol. **684**:349–362.

Lafferty, D.J., Robison, T.A., Gunadi, A., Schafran, P.W., Gunn, L.H., Van Eck, J., and Li, F.-W. (2024). Biolistics-mediated transformation of hornworts and its application to study pyrenoid protein localization. J. Exp. Bot. **75**:4760–4771. https://doi.org/10.1093/jxb/erae243.

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics **30**:3276–3278.

Laskowski, R.A., Hutchinson, E.G., Michie, A.D., Wallace, A.C., Jones, M.L., and Thornton, J.M. (1997). PDBsum: a Web-based database of summaries and analyses of all PDB structures. Trends Biochem. Sci. **22**:488–490.

Li, F.-W., Villarreal, J.C., and Szövényi, P. (2017). Hornworts: An Overlooked Window into Carbon-Concentrating Mechanisms. Trends Plant Sci. **22**:275–277.

Li, F.-W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z., Fernandez-Pozo, N., Barker, M.S., Bennett, T., Blázquez, M.A., et al. (2020). Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. Nat. Plants **6**:259–272.

Li, Q., Jiang, Y.-L., Xia, L.-Y., Chen, Y., and Zhou, C.-Z. (2022). Structural insights into cyanobacterial RuBisCO assembly coordinated by two chaperones Raf1 and RbcX. Cell Discov. **8**:93.

Lin, M.T., and Hanson, M.R. (2018). Red algal Rubisco fails to accumulate in transplastomic tobacco expressing *Griffithsia monilis RbcL* and *RbcS* genes. Plant Direct **2**:e00045.

Lin, M.T., Stone, W.D., Chaudhari, V., and Hanson, M.R. (2020). Small subunits can determine enzyme kinetics of tobacco Rubisco expressed in Escherichia coli. Nat. Plants **6**:1289–1299.

Lin, M.T., Salihovic, H., Clark, F.K., and Hanson, M.R. (2022). Improving the efficiency of Rubisco by resurrecting its ancestors in the family Solanaceae. Sci. Adv. **8**:eabm6871.

Liu, C., Young, A.L., Starling-Windhof, A., Bracher, A., Saschenbrecker, S., Rao, B.V., Rao, K.V., Berninghausen, O., Mielke, T., Hartl, F.U., et al. (2010). Coupled chaperone action in folding and assembly of hexadecameric Rubisco. Nature **463**:197–202.

Mackinder, L.C.M., Meyer, M.T., Mettler-Altmann, T., Chen, V.K., Mitchell, M.C., Caspari, O., Freeman Rosenzweig, E.S., Pallesen, L., Reeves, G., Itakura, A., et al. (2016). A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. Proc. Natl. Acad. Sci. USA **113**:5958–5963.

Mao, Y., Catherall, E., Díaz-Ramos, A., Greiff, G.R.L., Azinas, S., Gunn, L., and McCormick, A.J. (2023). The small subunit of Rubisco and its potential as an engineering target. J. Exp. Bot. **74**:543–561.

Ng, J., Guo, Z., and Mueller-Cajar, O. (2020). Rubisco activase requires residues in the large subunit N terminus to remodel inhibited plant Rubisco. J. Biol. Chem. **295**:16427–16435.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. **32**:268–274.

Occhialini, A., Lin, M.T., Andralojc, P.J., Hanson, M.R., and Parry, M.A.J. (2016). Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO_2. Plant J. **85**:148–160.

Oh, Z.G., Ang, W.S.L., Poh, C.W., Lai, S.-K., Sze, S.K., Li, H.-Y., Bhushan, S., Wunder, T., and Mueller-Cajar, O. (2023). A linker protein from a red-type pyrenoid phase separates with Rubisco via oligomerizing sticker motifs. Proc. Natl. Acad. Sci. USA **120**:e2304833120.

Pearce, F.G. (2006). Catalytic by-product formation and ligand binding by ribulose bisphosphate carboxylases from different phylogenies. Biochem. J. **399**:525–534.

Pearce, F.G., and Andrews, T.J. (2003). The Relationship between Side Reactions and Slow Inhibition of Ribulose-bisphosphate Carboxylase Revealed by a Loop 6 Mutant of the Tobacco Enzyme. J. Biol. Chem. **278**:32526–32536.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. **25**:1605–1612.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci. **30**:70–82.

Phillips, R., and Milo, R. (2009). A feeling for the numbers in biology. Proc. Natl. Acad. Sci. USA **106**:21465–21471.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods **14**:290–296.

Rambaut, A. (2018). FigTree. http://tree.bio.ed.ac.uk/software/figtree/.

Salesse-Smith, C.E., Sharwood, R.E., Busch, F.A., Kromdijk, J., Bardal, V., and Stern, D.B. (2018). Overexpression of Rubisco subunits with RAF1 increases Rubisco content in maize. Nat. Plants **4**:802–810.

Saschenbrecker, S., Bracher, A., Rao, K.V., Rao, B.V., Hartl, F.U., and Hayer-Hartl, M. (2007). Structure and function of RbcX, an assembly chaperone for hexadecameric Rubisco. Cell **129**:1189–1200.

Sharwood, R.E. (2017). Engineering chloroplasts to improve Rubisco catalysis: prospects for translating improvements into food and fiber crops. New Phytol. **213**:494–510.

Smith, E., and Griffiths, H. (1996). A pyrenoid-based carbon-concentrating mechanism is present in terrestrial bryophytes of the class Anthocerotae. Planta **200**.

Stein, N. (2008). CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. J. Appl. Crystallogr. **41**:641–643.

Studer, R.A., Christin, P.-A., Williams, M.A., and Orengo, C.A. (2014). Stability-activity tradeoffs constrain the adaptive evolution of RubisCO. Proc. Natl. Acad. Sci. USA **111**:2223–2228.

Taylor, T.C., Backlund, A., Bjorhall, K., Spreitzer, R.J., and Andersson, I. (2001). First crystal structure of Rubisco from a green alga, Chlamydomonas reinhardtii. J. Biol. Chem. **276**:48159–48164.

Tcherkez, G., and Farquhar, G.D. (2021). Rubisco catalytic adaptation is mostly driven by photosynthetic conditions - Not by phylogenetic constraints. J. Plant Physiol. **267**:153554.

Valegård, K., Hasse, D., Andersson, I., and Gunn, L.H. (2018). Structure of Rubisco from *Arabidopsis thaliana* in complex with 2-carboxyarabinitol-1,5-bisphosphate. Acta Crystallogr. D Struct. Biol. **74**:1–9.

Vaughn, K.C., Campbell, E.O., Hasegawa, J., Owen, H.A., and Renzaglia, K.S. (1990). The pyrenoid is the site of ribulose 1, 5-bisphosphate carboxylase/oxygenase accumulation in the hornwort (Bryophyta: Anthocerotae) chloroplast. Protoplasma **156**:117–129.

Wang, R., Song, H., Zhang, W., Wang, N., Zhang, S., Shao, R., and Liu, C. (2023). Structural insights into the functions of Raf1 and Bsd2 in hexadecameric Rubisco assembly. Mol. Plant **16**:1927–1936. https://doi.org/10.1016/j.molp.2023.10.011.

Whitney, S.M., Baldet, P., Hudson, G.S., and Andrews, T.J. (2001). Form I Rubiscos from non-green algae are expressed abundantly but not assembled in tobacco chloroplasts. Plant J. **26**:535–547.

Wilson, R.H., Martin-Avila, E., Conlan, C., and Whitney, S.M. (2018). An improved *Escherichia coli* screen for Rubisco identifies a protein–protein interface that can enhance CO_2-fixation kinetics. J. Biol. Chem. **293**:18–27.

Wilson, R.H., Thieulin-Pardo, G., Hartl, F.-U., and Hayer-Hartl, M. (2019). Improved recombinant expression and purification of functional plant Rubisco. FEBS Lett. **593**:611–621.

Wunder, T., Cheng, S.L.H., Lai, S.-K., Li, H.-Y., and Mueller-Cajar, O. (2018). The phase separation underlying the pyrenoid-based microalgal Rubisco supercharger. Nat. Commun. **9**:5076.

Wunder, T., Oh, Z.G., and Mueller-Cajar, O. (2019). CO_2 -fixing liquid droplets: Towards a dissection of the microalgal pyrenoid. Traffic **20**:380–389.

Yamano, T., Tsujikawa, T., Hatano, K., Ozawa, S.-I., Takahashi, Y., and Fukuzawa, H. (2010). Light and low-CO_2-dependent LCIB-LCIC complex localization in the chloroplast supports the carbon-concentrating mechanism in Chlamydomonas reinhardtii. Plant Cell Physiol. **51**:1453–1468.

Young, J.N., Heureux, A.M.C., Sharwood, R.E., Rickaby, R.E.M., Morel, F.M.M., and Whitney, S.M. (2016). Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. J. Exp. Bot. **67**:3445–3456.

Zhang, B., Yu, Y., Fox, B.W., Liu, Y., Thirumalaikumar, V.P., Skirycz, A., Lin, H., and Schroeder, F.C. (2024). Amino acid and protein specificity of protein fatty acylation in *C. elegans*. Proc. Natl. Acad. Sci. USA **121**:e2307515121.