

# Improving recombinant Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1 chaperone

Spencer M. Whitney<sup>a,1</sup>, Rosemary Birch<sup>a</sup>, Celine Kelso<sup>b</sup>, Jennifer L. Beck<sup>b</sup>, and Maxim V. Kapralov<sup>a</sup>

<sup>a</sup>Research School of Biology, Australian National University, Acton, ACT 2601, Australia; and <sup>b</sup>School of Chemistry, University of Wollongong, Wollongong, NSW 2522, Australia

Edited by Anthony Arthur Gatenby, Coskata, Inc., Warrenton, IL, and accepted by the Editorial Board February 5, 2015 (received for review October 27, 2014)

**Enabling improvements to crop yield and resource use by enhancing the catalysis of the photosynthetic CO<sub>2</sub>-fixing enzyme Rubisco has been a longstanding challenge. Efforts toward realization of this goal have been greatly assisted by advances in understanding the complexities of Rubisco's biogenesis in plastids and the development of tailored chloroplast transformation tools. Here we generate transplastomic tobacco genotypes expressing *Arabidopsis* Rubisco large subunits (AtL), both on their own (producing tob<sup>AtL</sup> plants) and with a cognate Rubisco accumulation factor 1 (AtRAF1) chaperone (producing tob<sup>AtL-R1</sup> plants) that has undergone parallel functional coevolution with AtL. We show AtRAF1 assembles as a dimer and is produced in tob<sup>AtL-R1</sup> and *Arabidopsis* leaves at 10–15 nmol AtRAF1 monomers per square meter. Consistent with a postchaperonin large (L)-subunit assembly role, the AtRAF1 facilitated two to threefold improvements in the amount and biogenesis rate of hybrid L<sub>8</sub>S<sub>8</sub> Rubisco [comprising AtL and tobacco small (S) subunits] in tob<sup>AtL-R1</sup> leaves compared with tob<sup>AtL</sup>, despite >threefold lower steady-state Rubisco mRNA levels in tob<sup>AtL-R1</sup>. Accompanying twofold increases in photosynthetic CO<sub>2</sub>-assimilation rate and plant growth were measured for tob<sup>AtL-R1</sup> lines. These findings highlight the importance of ancillary protein complementarity during Rubisco biogenesis in plastids, the possible constraints this has imposed on Rubisco adaptive evolution, and the likely need for such interaction specificity to be considered when optimizing recombinant Rubisco bioengineering in plants.**

photosynthesis | Rubisco | chloroplast transformation | chaperone | CO<sub>2</sub> assimilation

The increasing global demands for food supply, bioenergy production, and CO<sub>2</sub>-sequestration have placed a high need on improving agriculture yields and resource use (1, 2). It is now widely recognized that yield increases are possible by enhancing the light harvesting and CO<sub>2</sub>-fixation processes of photosynthesis (3–5). A major target for improvement is the enzyme Rubisco [ribulose-1,5-bisphosphate (RuBP) carboxylase/oxygenase] whose deficiencies in CO<sub>2</sub>-fixing speed and efficiency pose a key limitation to photosynthetic CO<sub>2</sub> capture (6, 7). In plants, the complex, multistep catalytic mechanism of Rubisco to bind its 5-carbon substrate RuBP, orient its C-2 for carboxylation, and then process the 6-carbon product into two 3-phosphoglycerate (3PGA) products, limits its throughput to one to four catalytic cycles per second (8). The mechanism also makes Rubisco prone to competitive inhibition by O<sub>2</sub> that produces only one 3PGA and 2-phosphoglycolate (2PG). Metabolic recycling of 2PG by photorespiration requires energy and results in most plants losing 30% of their fixed CO<sub>2</sub> (5). To compensate for these catalytic limitations, plants like rice and wheat invest up to 50% of the leaf protein into Rubisco, which accounts for ~25% of their leaf nitrogen (9).

Natural diversity in Rubisco catalysis demonstrates that plant Rubisco is not the pinnacle of evolution (6, 7). Better-performing versions in some red algae have the potential to raise the yield of crops like rice and wheat by as much as 30% (10). Bioengineering Rubisco in leaves therefore faces two key challenges: identifying

the structural changes that promote performance and identifying ways to efficiently transplant these changes into Rubisco within a target plant. A significant hurdle to both challenges is the complex biogenesis requirements of Rubisco in plant chloroplasts (7, 11). A number of ancillary proteins are required to correctly process and assemble the chloroplast made Rubisco large (L) subunit (coded by the plastome *rbcL* gene) and cytosol made small (S) subunits (coded by multiple *RbcS* genes in the nucleus) into L<sub>8</sub>S<sub>8</sub> complexes in the chloroplast stroma. The complicated assembly requirements of Rubisco in chloroplasts prevent their functional testing in *Escherichia coli* and conversely impedes, sometimes prevents, the biogenesis of Rubisco from other higher plants, cyanobacteria, and algae (12–14). For example, the L-subunits from sunflower and varying *Flaveria* sp. showed fivefold differences in their capacity to form hybrid L<sub>8</sub>S<sub>8</sub> Rubisco (that comprise tobacco S-subunits) in tobacco chloroplasts despite each *rbcL* transgene sharing the same genetic regulatory sequences and showing >92% amino acid identity (13, 14). Evidently, evolution of Rubisco function may have been constrained to maintain compatibility with the molecular chaperones required for its biogenesis (7, 15).

The necessity of chloroplast chaperonin (CPN) complexes for Rubisco biogenesis has been known for some time (16). Upon release from the hetero-oligomeric CPN ring structures in chloroplasts (17) the folded L-subunits are thought to sequentially assemble into dimers (L<sub>2</sub>) then octamers (L<sub>2</sub>)<sub>4</sub> before

## Significance

Using a translational photosynthesis approach, we successfully increased CO<sub>2</sub>-assimilation in leaf chloroplasts of the model plant tobacco. Phylogenetic analysis revealed parallel evolutionary linkages between the large (L-) subunit of the CO<sub>2</sub>-fixing enzyme Rubisco and its molecular chaperone Rubisco accumulation factor 1 (RAF1). We experimentally tested and exploited this correlation using plastome transformation, producing plants that demonstrated the role of RAF1 in L-subunit assembly and resolve the RAF1 quaternary structure as a dimer. We show the increase in Rubisco biogenesis translated to improvements in leaf photosynthesis and growth of the plants. The outcomes have application to the growing interest into identifying and implementing strategies to supercharge photosynthesis to improve crop productivity and stem global food-security concerns.

Author contributions: S.M.W. designed research; S.M.W., R.B., C.K., and M.V.K. performed research; S.M.W., R.B., C.K., J.L.B., and M.V.K. analyzed data; and S.M.W., R.B., C.K., J.L.B., and M.V.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. A.A.G. is a guest editor invited by the Editorial Board.

<sup>1</sup>To whom correspondence should be addressed. Email: spencer.whitney@anu.edu.au.

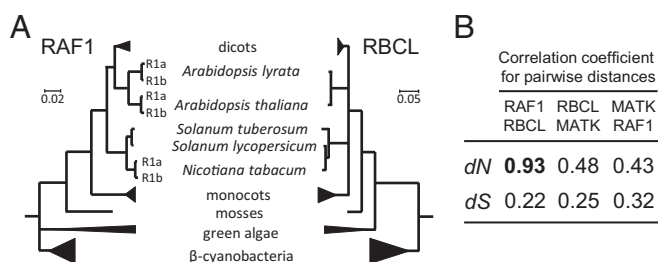
This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1420536112/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1420536112/-DCSupplemental).

S-subunit binding (18). The molecular details of this process remain unclear. The maize Photosynthetic Mutant Library has provided useful insight by identifying three chaperones with roles associated with Rubisco synthesis, assembly, and stability: Rubisco accumulation factors-1 (RAF1) (19) and-2 (RAF2; a Pterin-4a-Carbinolamine Dehydratase-like protein) (20) and Bundle Sheath Defective-2 (BSDII; a DnaJ-like protein) (21). Results of chemical cross-linking experiments in maize leaves suggest all three proteins might associate with the S-subunit during Rubisco biogenesis (20). Other studies, however, suggest RAF1 interacts with post-CPN folded L-subunits to assist in  $(L_2)_4$  then  $(L_2)_4$  formation (19, 22). This function mirrors that shown for RbcX, a Rubisco chaperone that acts as a “molecular staple” to assemble folded L-subunits into  $L_2$  units for  $(L_2)_4$  assembly before S-subunit binding to displace the RbcX and trigger catalytic potential (18). Although the function of RbcX in  $L_8S_8$  Rubisco biogenesis has been resolved in exquisite molecular detail in vitro and in *E. coli*, its functional role in cyanobacteria and in leaf chloroplasts remain unresolved. Comparable molecular details on RAF1, RAF2, and BSDII structure and function remain incomplete, making it difficult to reliably assign their roles and interactions with Rubisco in chloroplasts.

Targeted transformation of the chloroplast genome (plastome) provides a reliable but time-consuming tool for engineering Rubisco (23). This technology is best developed in tobacco with the  $^{cm}trL$  genotype specifically made for bioengineering Rubisco and testing its effects on leaf photosynthesis and growth (6, 7, 13, 14). Here we use chloroplast transformation in  $^{cm}trL$  to examine the function of RAF1 from *Arabidopsis* (AtRAF1) in Rubisco biogenesis. We show that AtRAF1 forms a stable dimer that, when coexpressed with its cognate *Arabidopsis* Rubisco L-subunits (AtL), enhances hybrid  $L_8^AS_8^t$  Rubisco (containing *Arabidopsis* L- and tobacco S-subunits) assembly in tobacco chloroplasts and concomitantly improves leaf photosynthesis and plant growth by more than twofold.

## Results

**Coevolution of RAF1 and the Rubisco L-Subunit.** Analysis of full-length *raf1* and *rbcL* sequences from plant, algae, and cyanobacteria showed that Rubisco L-subunit and RAF1 phylogenies are topologically similar (Fig. 1A). Mirror-tree analysis revealed that the correlation coefficient of these trees was 0.75 ( $P < 10^{-6}$ ) suggesting coevolution of both proteins across cyanobacteria and plants (Fig. S1). Exceptionally high correlations between RAF1 and Rubisco L-subunit pairwise nonsynonymous distances (i.e., those leading to amino acid substitutions) across all of the taxa confirmed coevolution of the two proteins (Fig. 1B). We therefore sought to test the functional significance of this complementarity



**Fig. 1.** RAF1 and Rubisco L-subunits phylogenies of plants, green algae, and  $\beta$ -cyanobacteria. (A) Condensed RAF1 and L-subunit (RBCL) maximum-likelihood trees assembled using RAXML v.8. Full maximum-likelihood trees are shown in Fig. S1 and sequence accessions listed in Table S2. (B) Correlations of pairwise nonsynonymous  $d_N$  (leading to amino acid substitutions) and synonymous  $d_S$  (selectively neutral) distances for RAF1, L-subunit, and maturase K (*matK*, an unassociated chloroplast made protein; negative control) across green plants and algae (all significant at  $P < 0.0001$ ).

by transforming the *Arabidopsis* Rubisco L-subunit (AtL) and one of its two cognate RAF1 isoforms (called AtRAF1) (Fig. S1) into tobacco chloroplasts via plastome transformation. Based on our previous heterologous Rubisco expression studies in tobacco (13, 14), we hypothesized that the phylogenetic divergence of AtL and the tobacco L-subunits (tobL) (Fig. 1A) would be accompanied by differences in ancillary protein requirements that would impede the biogenesis of hybrid  $L_8^AS_8^t$  Rubisco (i.e., comprising AtL and tobacco S-subunits) in tobacco chloroplasts.

## Plastome Transformation of *Arabidopsis* Rubisco AtL-Subunits and AtRAF1 into Tobacco Chloroplasts.

The L-subunit of *Arabidopsis* shares 94% identity with tobL, differing by only 29 amino acids (Fig. S2A). Transplanting the *Arabidopsis rbcL* gene (*AtrbcL*) into the tobacco plastome in place of the native *rbcL* gene was achieved by cloning it into the plastome-transforming plasmid pLEV4 to give plasmid pLEVAtL and transforming it into the plastome of the  $^{cm}trL$  tobacco genotype to produce  $tob^{AtL}$  lines (Fig. 2A). To test the influence of coexpressing AtRAF1 on hybrid  $L_8^AS_8^t$  Rubisco a synthetic *Atraf1* gene coding the full-length 50.2-kDa *Arabidopsis* RAF1 homolog AY063107 (coding its putative 62-aa N-terminal transit peptide sequence) (Fig. S2B) and a C-terminal 6x histidine tag was cloned 39-bp downstream of *AtrbcL* in pLEVAtL. The resulting plasmid, pLEVAtL-R1, was transformed into  $^{cm}trL$  to produce  $tob^{AtL-R1}$  lines (Fig. 2A). As shown in Fig. 1, although most plants only code for one RAF1, tobacco and *Arabidopsis* code two isoforms with the two homologs produced in *Arabidopsis* (~70% identical) only show ~50% identity to the two RAF1 isoforms produced in tobacco (that are 95% identical) (Fig. S2C).

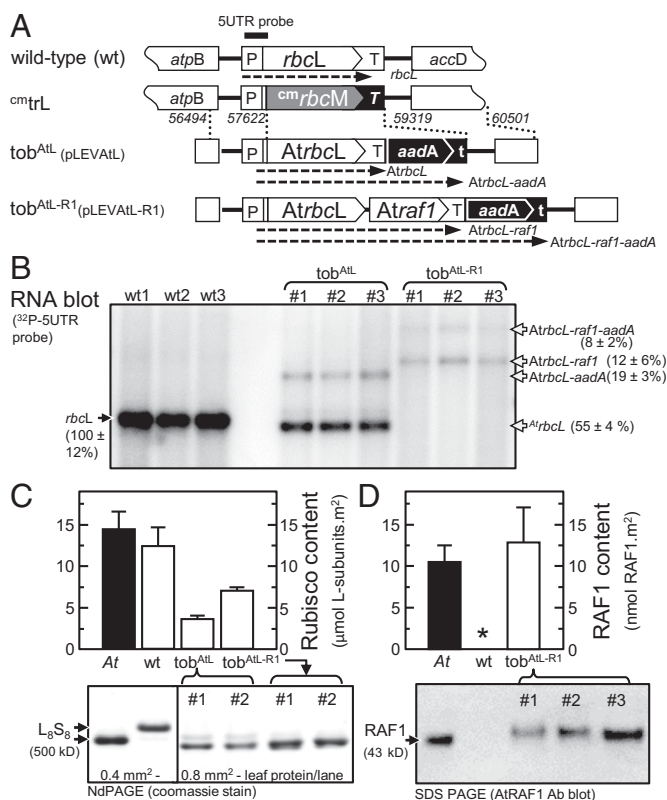
In both the  $tob^{AtL}$  and  $tob^{AtL-R1}$  genotypes, the *AtrbcL* transgene is regulated by the tobacco *rbcL* promoter, 5'- and 3'-untranslated sequences, and incorporates a downstream promoter-less *aadA* transgene that codes for the spectinomycin resistance used to screen for plastome transformed plantlets (Fig. 2A). In  $tob^{AtL-R1}$ , the *Atraf1* gene is located between both transgenes using an intergenic sequence similar to that used in pLEVEL<sup>UBS</sup> that produced a bicistronic tobacco *rbcL-rbcS* mRNA (23).

Three independent transplastomic  $tob^{AtL}$  and  $tob^{AtL-R1}$  lines were grown in soil to maturity in air supplemented with 0.5% (vol/vol) CO<sub>2</sub> and fertilized with wild-type pollen. The increased CO<sub>2</sub> levels were necessary for the survival of the  $tob^{AtL}$  lines in soil early during their development as their leaves contained little Rubisco (<3  $\mu$ mol L-subunits per m<sup>2</sup>/s), significantly impeding viability and drastically slowing growth in air. In contrast the  $tob^{AtL-R1}$  lines grew with greater vigor in air, but still at slow rates. Comprehensive analyses on the T<sub>1</sub> progeny of the  $tob^{AtL}$  and  $tob^{AtL-R1}$  lines were therefore undertaken on plants grown under 0.5% (vol/vol) CO<sub>2</sub> to ensure their viability.

## Variation in the Content and Catalysis of Hybrid $L_8^AS_8^t$ Rubisco in the $tob^{AtL}$ and $tob^{AtL-R1}$ Genotypes.

RNA blot analyses showed there were large differences in steady-state levels of the *AtrbcL* mRNAs produced in  $tob^{AtL}$  and  $tob^{AtL-R1}$  lines. As observed previously, a less-abundant *AtrbcL-aadA* di-cistronic mRNA (~10% that of the *AtrbcL* mRNA) was produced in the young  $tob^{AtL}$  leaves as a result of inefficient transcription termination by the tobacco *rbcL* 3' UTR (13, 14, 23) (Fig. 2B). In contrast, only di-cistronic *AtrbcL-Atraf1* or tricistronic *AtrbcL-Atraf1-aadA* mRNAs were detected in  $tob^{AtL-R1}$  leaves. Relative to the *rbcL* mRNA levels in the wild-type tobacco controls, the total pool of *AtrbcL* mRNAs were 25% and 80% lower in the developmentally comparable leaves from  $tob^{AtL}$  and  $tob^{AtL-R1}$ , respectively (Fig. 2B).

In contrast to the scarcity of *AtrbcL* transcripts in  $tob^{AtL-R1}$ , the levels of hybrid  $L_8^AS_8^t$  Rubisco (comprising *Arabidopsis* L-subunits and tobacco S-subunits) in the same leaves were >twofold higher than the  $L_8^AS_8^t$  content in  $tob^{AtL}$  (Fig. 2C). This variation in  $L_8^AS_8^t$



**Fig. 2.** Transplastomic tobacco generation and analysis of Rubisco and AtRAF1 expression. (A) The transforming plasmids pLEVAtL (GenBank accession no. KP635965) and pLEVAtL-R1 (GenBank accession no. KP635964) contain homologous plastome flanking sequence [indicated by dashed lines; numbering indicates region of sequence integration relative to *Nicotiana tabacum* (wild-type) plastome sequence; GenBank accession no. Z00444] that directed integration of the *Atrbcl* or *Atrbcl-raf1* transgenes and a promoterless *aadA* selectable marker gene into the cm<sup>tr</sup>L tobacco genotype plastome (23) to produce lines tob<sup>AtL</sup> and tob<sup>AtL-R1</sup>. The tobacco *rbcl* promoter/5' UTR (P) and first 42 nucleotides of wild-type *rbcl* sequence are conserved in each tobacco genotype. This sequence corresponds to the 5' UTR probe (14) with the expected mRNA species identified by the probe shown (dashed arrows). t, *rps16* 3' UTR, T, *psbA* 3' UTR, T, *rbcl* 3' UTR. (B) Detection of the various *rbcl* coding mRNA transcripts by the 5' UTR probe in total RNA from 6 mm<sup>2</sup> of young, nearly fully expanded leaves (14–16 cm in diameter) from comparable positions in the canopy of 32 ± 4-cm-tall plants of independent T<sub>1</sub>-transformed lines and three wild-type controls. (C) Variation in the mean (±SD) Rubisco content in tobacco leaves analyzed in B and those from three *Arabidopsis* (At) leaves as quantified by <sup>14</sup>C-CABP binding. Shown is an example ndPAGE analysis of the leaf protein used to confirm the varied levels of L<sub>8</sub>S<sub>8</sub> Rubisco. (D) AtRAF1 production in the At, wild-type, and tob<sup>AtL-R1</sup> leaf protein analyzed in C was quantified by SDS PAGE immunoblot analysis (example shown) against known amounts of purified AtRAF1 (Fig. S3). The asterisk (\*) represents the AtRAF1 antibody does not recognize tobacco RAF1.

content between each genotype was confirmed by nondenaturing PAGE (ndPAGE). Relative to the level of wild-type L<sub>8</sub>S<sub>8</sub> produced in the control, the L<sub>8</sub>S<sub>8</sub><sup>t</sup> content in tob<sup>AtL</sup> and tob<sup>AtL-R1</sup> were reduced by ~75% and ~55%, respectively.

Quantifying AtRAF1 production in leaf protein samples was undertaken by immunoblot analysis against varying amounts of purified recombinant AtRAF1 (Fig. S3). The AtRAF1 antibody recognized the ~43 kDa AtRAF1 in *Arabidopsis* leaf protein (Fig. 2D), the size expected for mature AtRAF1 after processing of the putative 62-aa transit peptide (Fig. S1B). The antibody detected nothing in wild-type tobacco consistent with the <50% sequence identity between AtRAF1 and the two homologs in

tobacco (Fig. S2C). Compared with *Arabidopsis*, the AtRAF1 produced in tob<sup>AtL-R1</sup> leaves was of equivalent size (noting it codes an additional 6x histidines) and produced at similar cellular concentrations (Fig. 2D). This finding indicated the transit peptide processing requirements of AtRAF1 were met by tobacco chloroplast stroma proteases and that the levels produced were physiologically comparable to those naturally made in *Arabidopsis*.

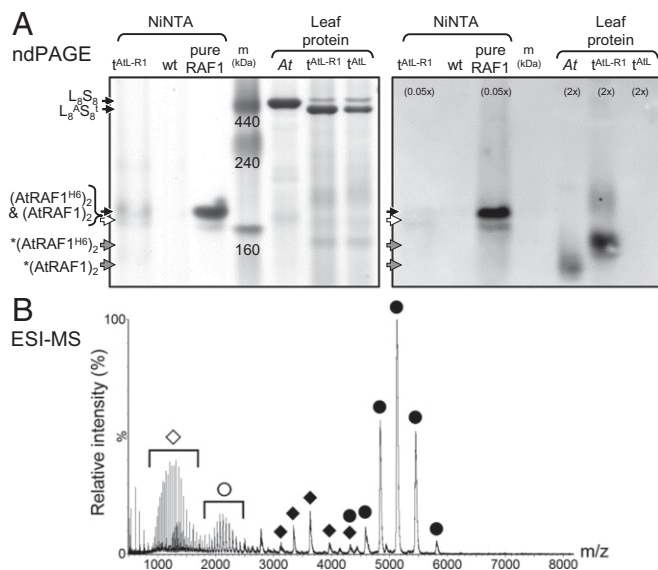
The catalytic properties of the hybrid L<sub>8</sub>S<sub>8</sub><sup>t</sup> were compared with *Arabidopsis* and tobacco Rubisco (Table S1). Significant reductions (24%) in carboxylation rate ( $k_C^{\text{cat}}$ ) coupled with an improved affinity for CO<sub>2</sub> (i.e., a 12% lower  $K_m$  for CO<sub>2</sub>,  $K_C$ ) were measured for L<sub>8</sub>S<sub>8</sub><sup>t</sup> albeit without significant change to its  $K_m$  for O<sub>2</sub> ( $K_O$ ), specificity for CO<sub>2</sub> over O<sub>2</sub> ( $S_{C/O}$ ) or carboxylation efficiency under atmospheric [O<sub>2</sub>] ( $k_C^{\text{cat}}/K_C^{21\%O_2}$ ).

**AtRAF1 Forms a Stable Dimer Complex.** The AtRAF1 made and purified from *E. coli* could be stably stored at -80 °C in buffer containing 20% (vol/vol) glycerol. Multiple freeze-thaw cycles had no discernible influence on AtRAF1 separation as two bands above the 160-kDa aldolase standard by ndPAGE, a prominent upper band, and >90% less abundant lower band (Fig. 3A). Immunoblot analysis showed this AtRAF1 oligomer separated at a slower rate than the immuno-reactive product detected in *Arabidopsis* leaf protein and the slightly larger His<sub>6</sub>-tagged AtRAF1 product (H<sub>6</sub>-AtRAF1) produced in tob<sup>AtL-R1</sup>. The mobility through ndPAGE of H<sub>6</sub>-AtRAF1 from tob<sup>AtL-R1</sup> after Ni-NTA affinity purification, however, matched that of the AtRAF1 purified from *E. coli* (Fig. 3A). This finding suggests the faster migrating, more diffusely separated, AtRAF1 products detected in the *Arabidopsis* and tob<sup>AtL-R1</sup> leaf samples might involve complexes with other proteins, the identity of which remain unclarified. In the leaf protein samples, the Rubisco antibody only recognized the L<sub>8</sub>S<sub>8</sub> holoenzyme and did not react with any of the products recognized by the RAF1 or CPN antibodies (Fig. S4). Similarly, no Rubisco was detected in the protein purified by Ni-NTA from tob<sup>AtL-R1</sup> leaves. These findings suggest the AtL-subunits do not form stable interactions with either AtRAF1 or CPN complexes in *Arabidopsis* or tob<sup>AtL-R1</sup> leaves.

The migration of proteins through ndPAGE is significantly influenced by their folded quaternary structure, which can mislead estimates of molecular size and subunit stoichiometry. For example, the 500-kDa bands for tobacco and *Arabidopsis* Rubisco resolve at different positions following ndPAGE (with the latter resolving at a smaller size to the 440-kDa ferritin protein standard) (Fig. 3A). We therefore undertook nano-electrospray ionization (ESI)-MS analysis of the pure AtRAF1 to accurately determine its subunit stoichiometry. Under non-denaturing conditions, the most abundant ions in the mass spectrum corresponded to a dimer with a molecular mass of ~86,871 Da (Fig. 3B) consistent with the predicted 43,434 Da for AtRAF1 subunits forming a stable dimer of (AtRAF1)<sub>2</sub>. This stoichiometry matches that determined for affinity purified RAF1 from *Thermosynechococcus elongatus* cells (22) but contrasts with the trimer structure predicted for RAF1 from maize (19).

#### Leaf Photosynthesis and Plant Growth Are Enhanced in tob<sup>AtL-R1</sup>.

Consistent with higher amounts of hybrid L<sub>8</sub>S<sub>8</sub><sup>t</sup> made in each tob<sup>AtL-R1</sup> line, the leaf photosynthetic CO<sub>2</sub> assimilation rates at varying CO<sub>2</sub> partial pressures ( $p_{CO_2}$ ) were ~twofold faster relative to tob<sup>AtL</sup>, albeit still slower than in wild-type tobacco (Fig. 4A). Accordingly, the tob<sup>AtL-R1</sup> genotypes grew faster than the tob<sup>AtL</sup> plants, although again less quickly than the tobacco controls (Fig. 4B). Consistent with this faster growth and higher Rubisco contents, the tob<sup>AtL-R1</sup> phenotype more closely resembled wild-type with little evidence of the pale green, marginal curling, and dimpling leaf phenotype seen for the tob<sup>AtL</sup> plants. This impaired growth phenotype matches that seen in other tobacco genotypes producing low levels of hybrid Rubisco (i.e., <3 μmol



**Fig. 3.** AtRAF1 stably assembles as a dimer. (A) ndPAGE analyses reproducibly showed recombinant AtRAF1 oligomers purified from *E. coli* (pure, Fig. S3A) was highly stable and separated at the same position above aldolase (160 kDa) in the marker protein standards (m) as Ni<sup>2+</sup>-nitrilotriacetic acid agarose (Ni-NTA) agarose purified His<sub>6</sub>-tagged AtRAF1 complexes (AtRAF1<sup>H6</sup>) from *tob*<sup>AtL-R1</sup> (*t*<sup>AtL-R1</sup>) leaves (see Fig. S4 for details). In *Arabidopsis* (*At*) and *t*<sup>AtL-R1</sup> leaf soluble protein the AtRAF1 and larger AtRAF1<sup>H6</sup> separated as smaller, more diffuse protein complexes of unknown content (indicated by an asterisk). Variations in the amount of sample loaded per lane relative to the Coomassie-stained gel are shown in parentheses. (B) Nano-ESI mass spectrum of pure AtRAF1 (3.2 μM; buffer exchanged into 0.1 M ammonium acetate, pH 7.2; cone voltage, 80 V) shows that the most abundant isoform was the dimer [i.e., (AtRAF1)<sub>2</sub>], with ions of low abundance from the monomer, and small amounts of unfolded monomer and dimer. The folded dimer was the most abundant isoform under cone voltages of 30–150 V. ●, folded dimer (AtRAF1)<sub>2</sub>; ◆, folded monomer AtRAF1; ○, unfolded dimer (AtRAF1)<sub>2</sub>; ◇, unfolded monomer AtRAF1.

sites per m<sup>2</sup>/s) comprising tobacco S-subunits and L-subunits from either sunflower (13) or *Flaveria pringlei* (14).

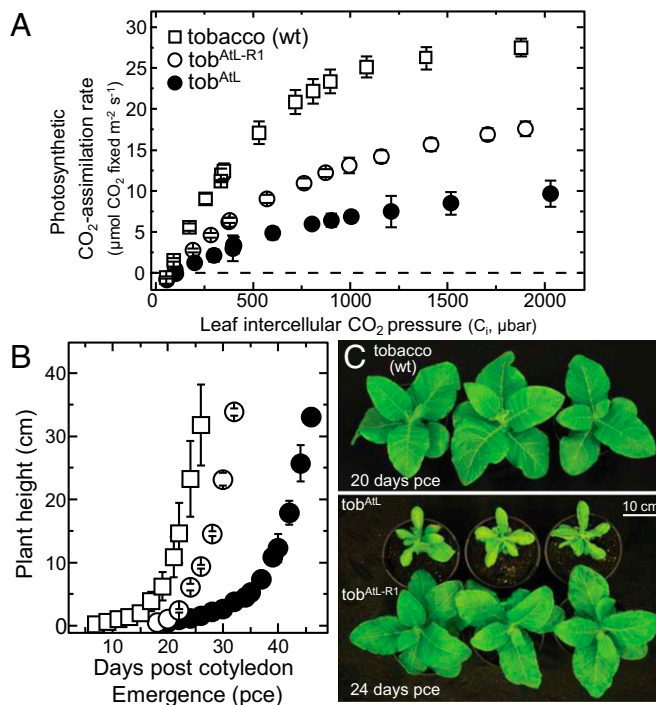
**Coexpressing AtRAF1 Enhances the Postchaperonin Assembly of AtL-Subunits into Stable L<sub>8</sub>S<sub>8</sub><sup>t</sup> Complexes.** Labeling of intact leaves with [<sup>35</sup>S]methionine showed varying rates of incorporation into <sup>35</sup>S-Rubisco complexes among the different tobacco genotypes (Fig. S4). Compared with *tob*<sup>AtL</sup>, the rates of L<sub>8</sub>S<sub>8</sub><sup>t</sup> biogenesis were threefold faster in the *tob*<sup>AtL-R1</sup>, although still threefold slower than the rate of L<sub>8</sub>S<sub>8</sub> synthesis in the wild-type tobacco controls. Unlabeled methionine “chase” analyses showed no change in the <sup>35</sup>S-Rubisco signal in any tobacco genotype indicating both tobacco L<sub>8</sub>S<sub>8</sub> and hybrid L<sub>8</sub>S<sub>8</sub><sup>t</sup> complexes were equally stable over the 7-h analysis period in young upper canopy leaves (Fig. S5B).

## Discussion

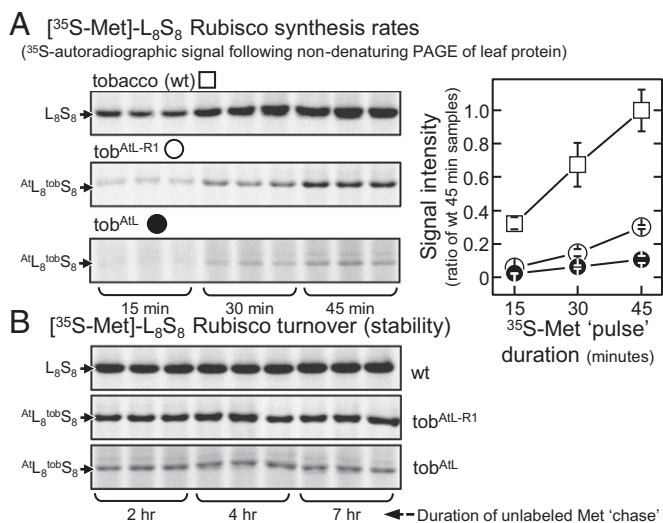
Here we highlight a pivotal role for the chloroplast RAF1 chaperone in Rubisco L-subunit assembly and the underpinning requirement for sequence complementarity between both proteins for optimal rates of L<sub>8</sub>S<sub>8</sub> biogenesis. The higher levels and quicker production of L<sub>8</sub>S<sub>8</sub><sup>t</sup> Rubisco in *tob*<sup>AtL-R1</sup> leaves (Figs. 2C and S4) and their corresponding faster rates of photosynthesis and growth (Fig. 4) relative to the *tob*<sup>AtL</sup> genotype underscore the pervasive role that RAF1 plays in the assembly of post-CPN folded L-subunits. This finding advances our understanding of Rubisco biogenesis in leaf chloroplasts and also highlights how chaperone compatibility demands on L-subunit folding and assembly might have constrained Rubisco’s catalytic evolution (7, 15).

Our phylogenetic pre-evaluation of parallel evolutionary linkages between the L-subunit and RAF1 and subsequent translational testing of this knowledge by plastome transformation proved highly successful in increasing recombinant Rubisco biogenesis. The specificity shown by Rubisco toward its regulatory protein Rubisco activase (RCA) provides a longstanding example of sequence compatibility requirements between both enzymes (24). Complementarity between residues in the L-subunit N-domain (residues 89–94) and those in the specificity H9 helix (residues 317–320) of RCA determine the capacity of RCA to stimulate release of inhibitory sugar phosphate molecules from the catalytic sites of Rubisco (25). Similar sequence compliance requirements between L-subunits and other ancillary proteins likely contribute to the low levels of Rubisco from cyanobacteria (12) and other plants (13, 14, 26) that can be produced in tobacco chloroplasts. To what extent expressing the cognate RAF1 proteins for each Rubisco isoform might augment their biogenesis in tobacco leaves remains untested. Determining the extent of parallel evolutionary linkages between the L-subunit and other molecular partners considered influential to Rubisco biogenesis (e.g., CPN, BSDII, RBCX, RAF2) may help identify those whose coexpression might augment recombinant Rubisco assembly in chloroplasts and other expression systems. This approach is particularly pertinent to the ongoing efforts to design and express more efficient Rubisco variants in crop plants (6).

Our analysis of AtRAF1 produced in *E. coli* indicates that it forms a stable dimer that differs in its migration size through ndPAGE to the RAF1 in soluble leaf cellular protein extract (Fig. S4). This finding suggests RAF1 in chloroplasts might interact with other proteins or cofactors that alter quaternary



**Fig. 4.** AtRAF1 improved leaf photosynthesis and growth in *tob*<sup>AtL-R1</sup>. (A) Leaf gas-exchange measurements of CO<sub>2</sub>-assimilation rates at 25 °C under varying intercellular CO<sub>2</sub> pressures (C<sub>i</sub>) made at 1,000 μmol quanta m<sup>-2</sup>/s illumination. Shown are the average of three measurements (±SD) made on the leaves analyzed in Fig. 2. (B) Comparison of the faster growth (as a function of plant height ± SD) of the *tob*<sup>AtL-R1</sup> lines (n = 3) relative to *tob*<sup>AtL</sup> (n = 3) at 25 °C in a growth cabinet in air with 0.5% (vol/vol) CO<sub>2</sub> under ~400 ± 100 μmol quanta m<sup>-2</sup>/s illumination. Both transplastomic genotypes grew slower than wild-type tobacco (wt, n = 3). (C) Phenotype of the plants at the respective age postcotyledon emergence (pce).



**Fig. 5.** AtRAF1 stimulated assembly of Rubisco.  $^{35}\text{S}\text{-Met}$  "pulse" unlabeled-Met "chase" analysis of hybrid  $\text{L}_8^{\text{A}}\text{S}_8^{\text{t}}$  Rubisco synthesis and turnover relative to tobacco  $\text{L}_8\text{S}_8$  Rubisco performed on young attached leaves under constant illumination ( $\sim 500 \mu\text{mol quanta m}^{-2}\text{s}^{-1}$ ) (Fig. 55). (A) Autoradiography signals of ndPAGE separated soluble protein from  $6 \text{ mm}^2$  of leaf taken 15, 30, and 45 min after infiltration with  $[^{35}\text{S}]\text{methionine}$  showing increasing  $^{35}\text{S}$  incorporation into  $\text{L}_8\text{S}_8$  Rubisco. Plotted are the average densitometry signals for  $\text{L}_8\text{S}_8$  Rubisco at each time point ( $n = 3 \pm \text{SD}$ ) relative to the average of the 45-min wild-type sample signals. Rates of  $\text{L}_8\text{S}_8$  synthesis extrapolated from linear fits to the normalized data were  $27 \times 10^{-4}$  ( $r^2 = 0.999$ ),  $78 \times 10^{-4}$  ( $r^2 = 0.997$ ), and  $229 \times 10^{-4}$  ( $r^2 = 1.000$ ) for the  $\text{tob}^{\text{AtL}}$  (●),  $\text{tob}^{\text{AtL-R1}}$  (○), and wild-type (□) leaves, respectively. (B) ndPAGE analyses made on soluble protein from the same leaves taken 2, 4, and 7 h after a "chase" infiltration with 10-mM unlabeled-methionine. No discernible changes in the densitometry of either hybrid  $\text{L}_8^{\text{A}}\text{S}_8^{\text{t}}$  or wild-type  $\text{L}_8\text{S}_8$  Rubisco autoradiography signals were detected indicative of little or no Rubisco turnover during this period.

structure and prevent dimer formation because of assembly with other proteins that are sufficiently stable to ndPAGE separation, but not to Ni-NTA purification where  $(\text{RAF1})_2$  oligomers matching those purified from *E. coli* are formed. Recent analysis of formaldehyde-treated maize leaf protein indicated RAF1 may interact with RAF2 and BSDII (20). Whether such interactions are responsible for the different migration rates through ndPAGE is a possibility that remains to be tested. Resolving the crystal structure for the  $(\text{RAF1})_2$  complex should help reveal its potential for forming alternative quaternary structures that might explain its alternative ndPAGE separation patterns and propensity to separate as an apparently larger sized complex that has previously been interpreted as a trimer (19, 20). For example, are the variations in  $(\text{RAF1})_2$  separation by ndPAGE because of its capacity to form "closed" and "open" conformations or from interactions with ancillary proteins or cofactors?

Constraints on the steady-state *AtrbcL* mRNA levels in  $\text{tob}^{\text{AtL-R1}}$  leaves appear a leading cause to limiting  $\text{L}_8^{\text{A}}\text{S}_8^{\text{t}}$  biogenesis. The steady-state pool of *AtrbcL* mRNA in  $\text{tob}^{\text{AtL-R1}}$  leaves was reduced fivefold relative to the tobacco *rbcl* mRNA levels (Fig. 2B), but still managed to produce  $\text{L}_8^{\text{A}}\text{S}_8^{\text{t}}$  at half the levels of  $\text{L}_8\text{S}_8$  made in wild-type (Fig. 2C). This result would suggest producing more hybrid  $\text{L}_8^{\text{A}}\text{S}_8^{\text{t}}$ , possibly matching wild-type Rubisco levels, would be feasible by enhancing *AtrbcL* mRNA levels. The operon structure in  $\text{tob}^{\text{AtL-R1}}$  matches that used previously in the transplastomic LEVUbS tobacco genotype. As seen in  $\text{tob}^{\text{AtL-R1}}$  leaves (Fig. 2B), the LEVUbS leaves also produced a di-cistronic *rbcl-UbrbcS* mRNA and a five- to sixfold less-abundant tricistronic *rbcl-UbrbcS-aadA* transcript; however, they were produced at levels that matched the *rbcl* mRNA content in wild-type (23). This finding suggests the *Atrafl* transgene likely destabilizes the di- and

tricistronic *AtrbcL* transcripts produced in  $\text{tob}^{\text{AtL-R1}}$ . Future RAF1 transplastomic studies should therefore consider equipping the *raf1* transgene with separate promoter/terminator regulatory elements to those controlling *rbcl* expression. Alternatively a small RNA intercistronic expression element between the *rbcl* and *raf1* transgenes that has been shown to trigger processing of polycistronic transcripts into more stable and translatable smaller transcripts could be included (27).

Previous studies of hybrid Rubiscos comprising plant L-subunits have shown the pervasive role of the L-subunit on shaping catalysis (13, 14, 28). Here, a modest yet significant reduction in  $k_{\text{C}}^{\text{cat}}$  and improvement in  $K_{\text{C}}$  was found for the  $\text{L}_8^{\text{A}}\text{S}_8^{\text{t}}$  Rubisco relative to the native *Arabidopsis* and tobacco enzymes, which have comparable catalytic constants at  $25^\circ\text{C}$  (Table S1). This catalytic variability of  $\text{L}_8^{\text{A}}\text{S}_8^{\text{t}}$  Rubisco likely arises from complementarity differences between *Arabidopsis* and tobacco S-subunits, consistent with a growing appreciation of the influential role the S-subunits can have on catalysis (6, 29).

Here we demonstrate the importance of a chaperone compatibility to enhancing recombinant Rubisco production in tobacco plastids. The finding enhances the potential for bioengineering Rubisco in chloroplasts and provides mechanistic evidence for the role of RAF1 in L-subunit assembly. Future applications of this coengineering approach will focus on identifying ways to more efficiently coexpress Rubisco L-subunits and their complementary RAF1s without compromising leaf *rbcl* mRNA pools. Extending this transplastomic coexpression method to other Rubisco chaperones—BSDII, RBCX, and RAF2—may prove a useful approach for determining their biochemical function in chloroplasts.

## Materials and Methods

**Bioinformatics Analyses.** Full-length *raf1* and *rbcl* sequences from 26 plant, 3 algal, and 46 cyanobacterial genomes were obtained from the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>) and Phytozome (<http://www.phytozome.net>) using the BLAST algorithm (Table S2). Phylogenetic trees of the translated proteins were constructed by the RAXML program (30) using the maximum-likelihood method with the following parameters: the Dayhoff model with  $\gamma$ -distributed rates, partial deletion, and bootstrap (1,000 replicates; random seed). L-subunit and RAF1 phylogenetic trees were compared using the Mirrortree server (31). Pairwise nonsynonymous (leading to amino acid substitutions) and synonymous (selectively neutral) sequence distances were calculated using the PAML package (32). We used the Mantel test to compute the Pearson correlation coefficient  $R$ . The chloroplast gene, *matK*, encoding maturase K (absent in most cyanobacteria genomes), which doesn't interact with Rubisco, was included as a negative control.

**Tobacco Plastome Transformation and Growth.** The *rbcl* gene from *Arabidopsis* was PCR amplified from leaf genomic DNA with primers 5' Nhel**rbcl** (14) and 3' AtSall**rbcl** (5'-TGTCGACTGTTTTATCTCTTCITATCTTATCT-3') and the 1,439-bp NheI-Sall *AtrbcL* product cloned into pLEV4 (14) to give pLEVAtL (GenBank accession no. KP635965). A synthetic *Atrafl* gene whose codon use matched tobacco *rbcl* was synthesized by GenScript and cloned downstream of *AtrbcL* in pLEVAtL using the intergenic sequence used in pLEVUbS (23) to give pLEVAtL-R1 (GenBank accession no. KP635964). pLEVAtL and pLEVAtL-R1 were each biolistically transformed into five leaves of the tobacco-masterline  $\text{cm}^{\text{trL}}$  as described in ref. 23, with four and seven spectinomycin-resistant plants, respectively, obtained. Three independent plastome transformed lines of each genotype were grown to maturity in soil in a growth atmosphere supplemented with 0.5% (vol/vol)  $\text{CO}_2$ , as described previously (13), and fertilized with wild-type pollen. The resulting  $T_1$  progeny were used for all analyses.

**RNA Blot, PCR, Protein, and PAGE Analyses.** Total leaf genomic DNA was isolated using the DNeasy Plant Mini Kit and used to PCR amplify and sequence the transformed plastome region using primers LSH and LSE (14). Total RNA extracted from  $0.5\text{-cm}^2$  leaf discs was separated on denaturing formaldehyde gels, blotted onto Hybond-N nitrocellulose membrane (GE Healthcare) and probed with the  $^{32}\text{P}$ -labeled 5' UTR probe (Fig. 2A), as described previously (13). The preparation, quantification (against BSA) of

soluble leaf protein, and analysis by SDS/PAGE, ndPAGE, and immunoblot analysis was performed as described previously (33).

**Rubisco Content and Catalysis.** Rates of Rubisco fixation in soluble protein extracts from three different leaves of each tobacco genotype and *Arabidopsis* were measured under varying concentrations of  $\text{NaH}^{14}\text{CO}_3$  (0–43  $\mu\text{M}$ ) and  $\text{O}_2$  [0–25% (vol/vol)] and the Michaelis constants ( $K_m$ ) for  $\text{CO}_2$  ( $K_C$ ), and  $\text{O}_2$  ( $K_O$ ) determined from the fitted data (14). The maximal rate of carboxylation ( $V_C$ ) was extrapolated from the Michaelis–Menten fit and then divided by the amount of Rubisco active sites quantified by [ $^{14}\text{C}$ ]-2-CABP binding (33, 34) to determine the turnover rate ( $k^{C_{cat}}$ ). Rubisco  $\text{CO}_2/\text{O}_2$  specificity ( $S_{C/O}$ ) was measured using ion exchange purified protein, as described previously (13).

**Growth and Photosynthesis Analysis.** All plants were grown in a growth chamber at 25 °C in air containing 0.5% (vol/vol)  $\text{CO}_2$  as described previously (13). Leaf photosynthesis rates were measured using a LI-6400 gas-exchange system (LI-COR) on the fifth upper canopy leaf of each tobacco genotype once they had reached comparable stages of physiological development.

**Recombinant RAF1 and CPN60 $\alpha$  Purification and Antibody Production.** Genes coding *Arabidopsis* RAF1 (AY063107) and Chaperonin 60 $\alpha$ 2 (NM\_121887) were cloned into plasmid pHueAct and expressed as N-terminal 6-Histidine-ubiquitin ( $\text{H}_6\text{Ub}$ ) tagged proteins in BL21(DE3) cells and purified by affinity chromatography (Fig. S3). Antibodies to both purified proteins were raised in rabbits.

**Mass Spectrometry.** Purified AtRAF1 stored at –80 °C in buffer containing 20% (vol/vol) glycerol was dialyzed (14,000 MWCO) against 100 mM ammonium acetate buffer adjusted to pH 7.2. The protein concentration was measured using a Nanodrop2000c (Thermo Fisher Scientific) and adjusted to 3  $\mu\text{M}$  (monomer concentration) before mass spectrometry. Positive ion nano-ESI mass spectra were acquired using a Waters Synapt HDMS fitted with a Z-spray nano-ESI source. Spectra were acquired using a MCP potential of

1,850 V, capillary voltage of 1.5 kV, extraction cone voltage of 4 V, and sampling cone voltages of 30, 80, and 150 V. The source temperature was set to 30 °C, the nanoflow back pressure to 0.1 bar, and the backing pressure to 3.93 mbar. The trap and transfer collision energies were 6.0 V and 4.0 V, respectively. Spectra were acquired over the 500–10,000  $m/z$  range and 40–50 acquisitions. The instrument was calibrated using a Csl solution (10 mg/mL in water).

**Pulse-Chase Labeling with  $^{35}\text{S}$ .** Plants of comparable size (~38 cm in height) stored overnight in a darkened laboratory were equilibrated for 15 min with ~500  $\mu\text{mol photons m}^{-2}\text{s}^{-1}$  illumination (at the surface of the youngest near fully expanded leaf sampled). Upper canopy leaves of equivalent age were infiltrated through the abaxial stomata by syringe (Fig. S5) with 3–4 mL of  $\text{Trans}^{35}\text{S}$ -label (ICN) diluted to 0.25  $\text{mCi/mL}^{-1}$  (9.25  $\text{MBq/mL}^{-1}$ ) with infiltration buffer (10 mM Mes-NaOH pH 5.5, 10 mM  $\text{MgSO}_4$ ). This process took 45–60 s. Leaf discs (0.5  $\text{cm}^2$ ) were collected after 15, 30, and 45 min and frozen in liquid nitrogen. After 60 min the leaves were infiltrated with infiltration buffer containing 10 mM methionine and leaf samples taken after 2, 4, and 7 h. The soluble leaf protein was separated by ndPAGE, the proteins fixed by Coomassie staining before drying the gels and exposing to a Storage Phosphor screen GP (Kodak) for 2 d. The autoradiograph signals were visualized using a PharoFX Molecular Imager and quantified with Quantity One software (Bio-Rad).

**Affinity Purification of 6xHis-tagged AtRAF1 from tob<sup>AtL-R1</sup> Leaves.** Soluble leaf protein from tob<sup>AtL-R1</sup> and wild-type tobacco (negative control) was purified by  $\text{Ni}^{2+}$ -nitrilotriacetic acid (Ni-NTA) agarose (Qiagen) chromatography and analyzed by SDS PAGE, ndPAGE and immunoblotting for evidence of stable interactions between AtRAF, AtL-subunits, and CPN (Fig. S4).

**ACKNOWLEDGMENTS.** This research was supported by Australian Research Council Grants FT0991407 (to S.M.W.), CE140100015 (to S.M.W.), and LE0882289 (to J.L.B.); and the Bill and Melinda Gates Foundation-funded project “Realizing Increased Photosynthetic Efficiency” (R.B. and M.V.K.).

- Edgerton MD (2009) Increasing crop productivity to meet global needs for feed, food, and fuel. *Plant Physiol* 149(1):7–13.
- Ziska LH, et al. (2012) Food security and climate change: On the potential to adapt global crop production by active selection to rising atmospheric carbon dioxide. *Proc Biol Sci* 279(1745):4097–4105.
- Evans JR (2013) Improving photosynthesis. *Plant Physiol* 162(4):1780–1793.
- Peterhansel C, Offermann S (2012) Re-engineering of carbon fixation in plants—Challenges for plant biotechnology to improve yields in a high- $\text{CO}_2$  world. *Curr Opin Biotechnol* 23(2):204–208.
- Zhu X-G, Long SP, Ort DR (2010) Improving photosynthetic efficiency for greater yield. *Annu Rev Plant Biol* 61(1):235–261.
- Parry MAJ, et al. (2013) Rubisco activity and regulation as targets for crop improvement. *J Exp Bot* 64(3):717–730.
- Whitney SM, Houtz RL, Alonso H (2011a) Advancing our understanding and capacity to engineer nature’s  $\text{CO}_2$ -sequestering enzyme, Rubisco. *Plant Physiol* 155(1):27–35.
- Andersson I, Backlund A (2008) Structure and function of Rubisco. *Plant Physiol Biochem* 46(3):275–291.
- Evans J, Seemann J (1989) The allocation of nitrogen in the photosynthetic apparatus: Costs, consequences and control. *Photosynthesis*, ed Briggs WR (Alan R. Liss, New York), pp 183–205.
- Zhu XG, Portis AR, Long SP (2004) Would transformation of  $\text{C}_3$  crop plants with foreign Rubisco increase productivity? A computational analysis extrapolating from kinetic properties to canopy photosynthesis. *Plant Cell Environ* 27(2):155–165.
- Nishimura K, Ogawa T, Ashida H, Yokota A (2008) Molecular mechanisms of Rubisco biosynthesis in higher plants. *Plant Biotechnol* 25(5):285–290.
- Lin MT, Occhialini A, Andralojc PJ, Parry MAJ, Hanson MR (2014) A faster Rubisco with potential to increase photosynthesis in crops. *Nature* 513(7519):547–550.
- Sharwood RE, von Caemmerer S, Maliga P, Whitney SM (2008) The catalytic properties of hybrid Rubisco comprising tobacco small and sunflower large subunits mirror the kinetically equivalent source Rubiscos and can support tobacco growth. *Plant Physiol* 146(1):83–96.
- Whitney SM, et al. (2011b) Isoleucine 309 acts as a  $\text{C}_4$  catalytic switch that increases ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco) carboxylation rate in *Flaveria*. *Proc Natl Acad Sci USA* 108(35):14688–14693.
- Mueller-Cajar O, Whitney SM (2008) Directing the evolution of Rubisco and Rubisco activase: First impressions of a new tool for photosynthesis research. *Photosynth Res* 98(1–3):667–675.
- Hemmingsen SM, et al. (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* 333(6171):330–334.
- Tsai Y-CC, Mueller-Cajar O, Saschenbrecker S, Hartl FU, Hayer-Hartl M (2012) Chaperonin cofactors, Cpn10 and Cpn20, of green algae and plants function as hetero-oligomeric ring complexes. *J Biol Chem* 287(24):20471–20481.
- Bracher A, Starling-Windhof A, Hartl FU, Hayer-Hartl M (2011) Crystal structure of a chaperone-bound assembly intermediate of form I Rubisco. *Nat Struct Mol Biol* 18(8):875–880.
- Feiz L, et al. (2012) Ribulose-1,5-bis-phosphate carboxylase/oxygenase accumulation factor1 is required for holoenzyme assembly in maize. *Plant Cell* 24(8):3435–3446.
- Feiz L, et al. (2014) A protein with an inactive pterin-4a-carbinolamine dehydratase domain is required for Rubisco biogenesis in plants. *Plant J* 80(5):862–869.
- Brutnell TP, Sawers RJH, Mant A, Langdale JA (1999) BUNDLE SHEATH DEFECTIVE2, a novel protein required for post-translational regulation of the *rbcl* gene of maize. *Plant Cell* 11(5):849–864.
- Kolesinski P, Belusiak I, Czarnocki-Cieciura M, Szczepaniak A (2014) Rubisco Accumulation Factor 1 from *Thermosynechococcus elongatus* participates in the final stages of ribulose-1,5-bisphosphate carboxylase/oxygenase assembly in *Escherichia coli* cells and in vitro. *FEBS J* 281(17):3920–3932.
- Whitney SM, Sharwood RE (2008) Construction of a tobacco master line to improve Rubisco engineering in chloroplasts. *J Exp Bot* 59(7):1909–1921.
- Mueller-Cajar O, Stotz M, Bracher A (2014) Maintaining photosynthetic  $\text{CO}_2$  fixation via protein remodelling: The Rubisco activases. *Photosynth Res* 119(1–2):191–201.
- Stotz M, et al. (2011) Structure of green-type Rubisco activase from tobacco. *Nat Struct Mol Biol* 18(12):1366–1370.
- Whitney SM, Baldet P, Hudson GS, Andrews TJ (2001) Form I Rubiscos from non-green algae are expressed abundantly but not assembled in tobacco chloroplasts. *Plant J* 26(5):535–547.
- Lu Y, Rijzaani H, Karcher R, Ruf S, Bock R (2013) Efficient metabolic pathway engineering in transgenic tobacco and tomato plastids with synthetic multigene operons. *Proc Natl Acad Sci USA* 110(8):E623–E632.
- Zhang XH, et al. (2011) Hybrid Rubisco of tomato large subunits and tobacco small subunits is functional in tobacco plants. *Plant Sci* 180(3):480–488.
- Ishikawa C, Hatanaka T, Misoo S, Miyake C, Fukayama H (2011) Functional incorporation of sorghum small subunit increases the catalytic turnover rate of Rubisco in transgenic rice. *Plant Physiol* 156(3):1603–1611.
- Stamatakis A (2014) RAXML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.
- Ochoa D, Pazos F (2010) Studying the co-evolution of protein families with the Mirrortree web server. *Bioinformatics* 26(10):1370–1371.
- Xu B, Yang Z (2013) PAMLX: A graphical user interface for PAML. *Mol Biol Evol* 30(12):2723–2724.
- Whitney SM, Sharwood RE (2007) Linked Rubisco subunits can assemble into functional oligomers without impeding catalytic performance. *J Biol Chem* 282(6):3809–3818.
- Whitney SM, Andrews TJ (2001) Plastome-encoded bacterial ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) supports photosynthesis and growth in tobacco. *Proc Natl Acad Sci USA* 98(25):14738–14743.