An improved *Escherichia coli* screen for Rubisco identifies a protein–protein interface that can enhance CO_2-fixation kinetics

Received for publication, August 8, 2017, and in revised form, September 28, 2017 Published, Papers in Press, October 6, 2017, DOI 10.1074/jbc.M117.810861

Robert H. Wilson, Elena Martin-Avila, Carly Conlan[1], and Spencer M. Whitney[2]

From the Research School of Biology, Australian National University, Acton, Australian Capital Territory 2601, Australia

Edited by Joseph Jez

An overarching goal of photosynthesis research is to identify how components of the process can be improved to benefit crop productivity, global food security, and renewable energy storage. Improving carbon fixation has mostly focused on enhancing the CO_2 fixing enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). This grand challenge has mostly proved ineffective because of catalytic mechanism constraints and required chaperone complementarity that hinder Rubisco biogenesis in alternative hosts. Here we refashion *Escherichia coli* metabolism by expressing a phosphoribulokinase-neomycin phosphotransferase fusion protein to produce a high-fidelity, high-throughput Rubisco-directed evolution (RDE2) screen that negates false-positive selection. Successive evolution rounds using the plant-like *Te*-Rubisco from the cyanobacterium *Thermosynechococcus elongatus* BP1 identified two large subunit and six small subunit mutations that improved carboxylation rate, efficiency, and specificity. Structural analysis revealed the amino acids clustered in an unexplored subunit interface of the holoenzyme. To study its effect on plant growth, the *Te*-Rubisco was transformed into tobacco by chloroplast transformation. As previously seen for *Synechocccus* PCC6301 Rubisco, the specialized folding and assembly requirements of *Te*-Rubisco hinder its heterologous expression in leaf chloroplasts. Our findings suggest that the ongoing efforts to improve crop photosynthesis by integrating components of a cyanobacteria CO_2-concentrating mechanism will necessitate co-introduction of the ancillary molecular components required for Rubisco biogenesis.

Improving carbon fixation in agriculture has mostly focused on enhancing the activity of the CO_2-fixing enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco)[3] by modifying the enzyme itself or increasing the CO_2 levels around it (1). Improving the kinetics of Rubisco has proved challenging because of its complex catalytic mechanism of fixing CO_2 to ribulose-1,5-bisphosphate (RuBP) and cleaving it into two 3-phosphoglycerate (3-PGA) molecules (2–4). The carboxylation rate is slow (k_{cat}^{c} = ~1 to 5 reactions/s in plants) and often confines flux through the Calvin–Benson–Bassham (CBB) cycle, hence limiting the rate of photosynthesis and plant growth. Because of this limitation, large amounts of Rubisco are needed to support adequate CO_2-assimilation rates (5–7). This results in Rubisco being the Earth's most abundant protein (8). Further encumbering its performance, Rubisco also catalyzes RuBP oxygenation to produce 3-PGA and 2-phosphoglycolate. The recycling of 2-phosphoglycolate back into 3-PGA via photorespiration is considered wasteful because it consumes energy and releases fixed CO_2 (9).

The potential for successfully improving the carboxylation properties of Rubisco is spurred on by findings that crop Rubisco is not the pinnacle of evolution, with kinetically more efficient Rubisco forms found in some red algae (1, 3, 6, 10). Evolutionary adaptation of Rubisco kinetics appears to have been constrained by the complexity of its catalytic chemistry and multiplexed subunit folding and assembly requirements (3, 11–13). The need to retain complementarity with ancillary proteins involved in the biogenesis and metabolic repair of Rubisco appears particularly pertinent for form I Rubisco, which comprises eight RbcL and eight RbcS subunits to form a L_8S_8 complex. Examples of Rubisco-dedicated components include the assembly chaperones BSD2, RbcX, Rubisco accumulation factors 1 and 2 (Raf1, Raf2), and the metabolic repair protein Rubisco activase (14). The protein-folding chaperonins of bacteria (GroEL), chloroplasts (Cpn60), and their protein co-factors GroES (bacteria) and Cpn10/Cpn20 (chloroplasts) are also essential components in Rubisco biogenesis (15). The need for Rubisco to maintain complementarity with these varied chaperone and chaperonin components appears to limit the span of Rubisco isoforms that can be produced in *Escherichia coli*, as well as those that can be bioengineered in the chloroplasts of vascular plants. For example, the chloroplasts of the model C_3 plant tobacco can produce high levels of the RbcS-lacking bacterial form II *Rhodospirillum rubrum* L_2 Rubisco and *Methanococcoides bur-*

This work was supported by Australian Research Council Centre of Excellence for Translational Photosynthesis Grant CE140100015 (to S. M. W.). The authors declare that they have no conflicts of interest with the contents of this article.

This article contains Figs. S1–S5.

[1] Supported by funds from Bill and Melinda Gates Foundation Award RIPE: Realizing Increases in Photosynthetic Efficiency.

[2] To whom correspondence should be addressed: Research School of Biology, Australian National University, Acton, Australian Capital Territory 2601, Australia. Tel.: 61-2-6125-5073; E-mail: spencer.whitney@anu.edu.au.

[3] The abbreviations used are: Rubisco, D-ribulose-1,5-bisphosphate carboxylase/oxygenase; RuBP, D-ribulose-1,5-bisphosphate; CABP, 2′-carboxyarabinitol-1,5-bisphosphate; ANOVA, analysis of variance; RDE, Rubisco-dependent *E. coli*; 3-PGA, 3-phosphoglycerate; CBB, Calvin–Benson–Bassham; PRK, phosphoribulokinase; NPTII, neomycin phosphotransferase; CSP, cell-soluble protein; CCM, carbon-concentration mechanism; IPTG, isopropyl β-D-thiogalactopyranoside.

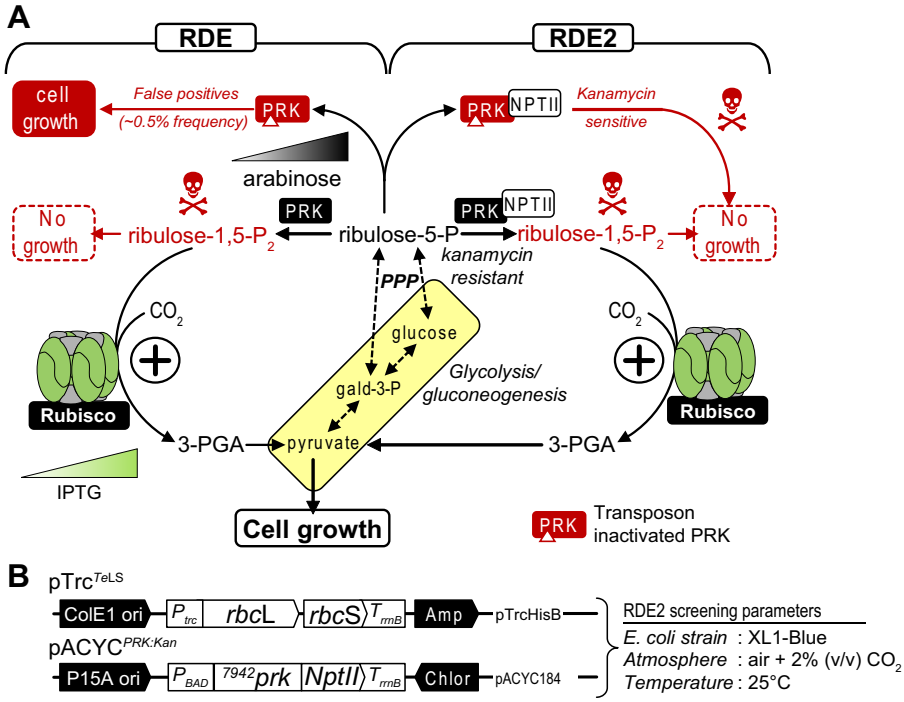

Figure 1. An improved RDE screen. *A*, RDE screens rely on the toxicity of ribulose-1,5-P2 (RuBP), a foreign metabolite produced from the pentose phosphate pathway (*PPP*) intermediate ribulose-5-phosphate (*ribulose-5-P*) via recombinant PRK expression (regulated by arabinose addition). Cell viability is restored either through Rubisco expression (IPTG-induced regulation), where growth rate is proportional to Rubisco activity level, or through transposon silencing of PRK expression that produces false positives at a high frequency (22). The new RDE2 screen utilizes a PRK–NPTII fusion where NPTII is tethered in frame to the C terminus of PRK rendering kanamycin resistance to *E. coli* unless PRK–NPTII expression is transposon silenced. Toxic pathways to *E. coli* are highlighted in *red*. RDE2 screens are performed under high CO_2 (air + 1 to 2% (v/v) CO_2) with the ribulose-1,5-P_2 carboxylation reaction of Rubisco (positive sign) producing 3-PGA, an intermediate of the glycolysis/gluconeogenesis pathways (*shaded yellow*) (reviewed in Refs. 12 and 21). *B*, summary of the plasmids and *E. coli* growth conditions used in the RDE2 screen of an *rbcL* (±*rbcS*) mutant library. P_{trc} and P_{BAD} indicate IPTG- and arabinose-inducible promoters, respectively.

tonii archaeal L_{10} Rubisco (~10–25 μmol active sites·m^2) because of their simple biogenesis requirements (16, 17). Differences in the biogenesis requirements of L_8S_8 Rubisco from red algae and monocot plants, however, preclude their assembly in tobacco and *E. coli* (10, 18). By contrast the assembly requirements of *Synechococcus elongatus* PCC6301 L_8S_8 Rubisco (*Se*-Rubisco) are partially met in chloroplasts (~5 μmol active sites·m^2) (19) and in *E. coli*, with ~98% of the *Se*-RbcL produced forming insoluble, misfolded aggregates in the bacterium (20).

To avoid the constraints a photosynthetic environment may pose on the catalytic evolution of Rubisco, modern laboratory evolution applications have made particular use of Rubisco-dependent *E. coli* (RDE) screens (12, 21, 22). An elegant advance has been the extensive rewiring of *E. coli* metabolism to incorporate a non-native CBB cycle where cell survival can be made dependent on CO_2-fixation by Rubisco (23). This contrasts with conventional RDE screens that simply express recombinant phosphoribulokinase (PRK) to produce RuBP, the 5-carbon substrate of Rubisco. As summarized in Fig. 1*A*, RuBP is unexplainably toxic to *E. coli* such that the rate of colony growth in RDE screens is dependent on the level of Rubisco activity. This type of selection has proved useful for evolving *M. burtonii* L_{10} Rubisco mutants with desired improvements in carboxylation rate (k_{cat}^c), CO_2 affinity (lower K_m for CO_2, K_c) and specificity for CO_2 over O_2 ($S_{c/o}$) (17). Other attempts to improve the catalysis of L_8S_8 Rubisco using RDE screens have only identified RbcL mutations that enhance the folding and assembly of the subunit with RbcS into a functional L_8S_8 complex (*i.e.* improving Rubisco solubility) (22).

A feature of RDE screens that limit their throughput and reliability (*i.e.* selection fidelity) is the high proportion of false positives selected (24, 25). These arise from the inactivation of PRK function by transposon integration (Fig. 1*A* and Ref. 22). Here we present the development and implementation of a simple, faster, high-throughput screen called RDE2 that specifically selects for improvements in Rubisco activity, not PRK inactivation mutants. By eliminating false-positive selection, we demonstrate the versatility of the RDE2 screen in selecting for mutants of *Thermosynechococcus elongatus* BP-1 (*Te*-) Rubisco that improve RbcL folding and assembly in *E. coli* (*i.e.* *Te*-Rubisco solubility), the enzyme carboxylation kinetics, or both. We subsequently reveal, using chloroplast genome transformation, how the specialized folding and assembly requirements of *Te*-Rubisco hinder its translational testing in tobacco (*Nicotiana tabacum*). The outcomes demonstrate that Rubisco biogenesis in *E. coli* is not a reliable proxy for expression in chloroplasts. As a consequence, deciphering the crucial molecular partnerships required for Rubisco biogenesis is needed to optimize heterologous Rubisco expression in crop chloroplasts, as well as adapting the RDE2 for future evolution of eukaryotic form I Rubisco.

Results

RDE2 screen development

Rubisco-directed evolution studies that have utilized either RDE or photosynthetic mutant screens have almost exclusively

selected for one or more amino acid substitutions in RbcL that improve Rubisco biogenesis (solubility), not catalysis (12, 22). Impeding success are the poor transformation efficiencies of photosynthetic hosts and the high proportion of false positives selected in RDE screens. As summarized in Fig. 1*A*, false-positive selection in RDE systems arises from terminating RuBP production via transposon silencing of PRK expression (22). To circumvent this impediment, a synthetic gene coding a fusion protein comprising *S. elongatus* PCC6301 PRK (*Se*-PRK) and a C-terminal neomycin phosphotransferase (NPTII) was made (Fig. 1*B*). The PRK–NPTII fusion protein was found to preserve the kanamycin resistance conferring properties of NPTII to *E. coli* in addition to retaining 70% the PRK activity ($k_{cat} = 95 \pm 3\ s^{-1}$) of unmodified *Se*-PRK ($134 \pm 6\ s^{-1}$).

The fidelity of the RDE and RDE2 screens were compared by plating XL-1Blue cells producing *Se*-PRK or the PRK–NPTII fusion, along with wild-type *Te*-Rubisco, on medium containing 0.2% (w/v) arabinose (high PRK inducing) and no kanamycin. Under these conditions, *Te*-Rubisco cannot support cell growth, meaning any colonies formed are false positives caused by transposon silencing of PRK activity (Fig. 1*A* and Ref. 22). More than 500 false-positive colonies/10^6 cells plated grew in the RDE screen, whereas only 28 ± 17 colonies/10^6 cells plated grew using the RDE2 screen. Inclusion of 100 μg/ml kanamycin in the growth medium of replica RDE2 cell platings resulted in no colony growth. This confirmed that the RDE2 screen is immune from generating false positives because transposon silencing of PRK–NPTII activity also confers the cells sensitive to kanamycin (Fig. 1*A*).

Identifying a suitable Rubisco to evolve

The inability of *E. coli* to meet the folding and assembly requirements of plant and algae Rubisco currently prevents their directed evolution using RDE screens (10, 12). By contrast the biogenesis requirements of cyanobacteria L_8S_8 Rubisco can be partially met by *E. coli*. For example, only a small proportion (<2%) of the 52-kDa *Se*-RbcL produced in *E. coli* can assemble with the more soluble 14-kDa *Se*-RbcS into functional *Se*-Rubisco (20). As a consequence *Se*-Rubisco produced in *E. coli* only accounts for 1–2% (w/w) of the cell-soluble protein (CSP). RDE screens have therefore primarily only identified substitutions in RbcL that enhance the biogenesis of *Se*-Rubisco (*i.e.* improved its solubility) (22). Among these are the F140I, V189A, and F345I substitutions in RbcL (numbering relative to spinach and *Te*-Rubisco) that increase *Se*-Rubisco solubility between 3- and 14-fold in *E. coli* (Fig. 2, *A* and *B*, and Table 1). Notably these mutations all impair k_{cat}^{c}, contrary to that proposed previously (11). By comparison the biogenesis requirements of *Te*-Rubisco, that already codes Ile-140, are more readily met in *E. coli* and expressed at ~6% (w/w) CSP (Fig. 2*A*). Moreover, the apparent carboxylation efficiency under ambient O_2 ($k_{cat}^{c}/K_c^{21\%O_2}$) and specificity for CO_2 over O_2 ($S_{c/o}$) of *Te*-Rubisco are ~40 and 18% higher, respectively, than *Se*-Rubisco (Table 1). Notably these favorable *Te*-Rubisco kinetics come at the expense of a slower carboxylation rate (k_{cat}^{c}) that is ~40% less than *Se*-Rubisco (Fig. 2, *A* and *B*). The high solubility of *Te*-Rubisco in *E. coli* and distinctive kinetics made it a superior target for evolution testing using RDE2.

First-generation Te-Rubisco solubility mutant selection in RDE2

The efficacy of RDE2 to select for increased *Te*-Rubisco activity was initially tested using a library of 2.7×10^5 *Te-rbc*L mutants with an average mutation rate of 2.5 nucleotides (1.6 amino acid substitutions) per variant. After 5 days at 25 °C, 18 faster-growing colonies were identified on plates containing 0.1% (w/v) arabinose (defined as moderate PRK–NPTII induction) and found to code either V300A, F345I, F345L, or P415A substitutions in RbcL. The kinetics of the slower growing V300A mutant resembled *Te*-Rubisco, suggesting that it improved RDE2 fitness possibly from the small (<10%) increase in the amount of soluble L_8S_8 made (Table 1). The faster-growing F345I, F345L, and P415A mutants produced significantly more *Te*-Rubisco, especially the F345I mutation that stimulated *Te*-Rubisco biogenesis to ~14% (w/w) of the CSP (Fig. 2, *A* and *B*). When coupled with the P415A mutation, the solubility of the *Te*-F345I/P415A mutant increased to ~17% (w/w) CSP (Fig. 2, *A* and *B*). This improvement came at a cost to the kinetics of each mutant that showed reductions in k_{cat}^{c}, $k_{cat}^{c}/K_c^{21\%O_2}$, and $S_{c/o}$ (Table 1).

Sequence comparisons found that both Phe^{345} and Pro^{415} are highly conserved among cyanobacteria and higher plant form IB RbcL but are located in separate, unconnected α-helical regions (Fig. 2*C*). Interestingly, these loci are modified in the form IA Rubisco lineage that has evolved independent of the assembly chaperones Raf1 and RbcX (11, 14).

Second-generation Te-Rubisco catalytic mutant selection

Improving the kinetics of *Te*-Rubisco required a second round of RDE2 screening that targeted mutagenesis of both RbcL and RbcS. A 3×10^5 member mutagenic library of the full-length *Te-rbc*LS operon coding the first-generation P415A RbcL mutant (abbreviated *Te*-P415A) was used. Unlike the F345I and F345I/P415A mutants, the catalytic properties of *Te*-P415A more closely matched wild-type *Te*-Rubisco (Fig. 2*A*). Under strong PRK–NPTII selection (*i.e.* on medium containing 0.2% (w/v) arabinose), the growth of cells producing wild-type (abbreviated *Te*-LS) or *Te*-P415A Rubisco were impeded (Fig. 3*A*). After 6 days growth at 25 °C in air supplemented with 2% (v/v) CO_2, 15 colonies were isolated. The sequence and biochemistry of these second-generation mutants (abbreviated *Te*-2G Rubisco variants) were found to enhance RDE2 growth either through further improving *Te*-P415A Rubisco solubility and/or significantly enhancing its carboxylation properties (Fig. 3*A* and Table 1). The fastest growing *Te*-2Ga mutant was independently selected four times and coded a V98M RbcS point mutation that improved Rubisco biogenesis, k_{cat}^{c}, $k_{cat}^{c}/K_c^{21\%O_2}$, and $S_{c/o}$ by ~310, 28, 43, and 6% relative to *Te*-Rubisco (Fig. 3, *B–E*). Four additional *Te*-P415A Rubisco mutants with improvements in k_{cat}^{c} and $k_{cat}^{c}/K_c^{21\%O_2}$ and unchanged solubility were also identified. They were found to code for point mutations in either RbcS (A48V, *Te-2*Gb; H37L, *Te-2*Gc; or Y36N/G112D, *Te-2*Gd) or RbcL (L74M/D397N, *Te-2*Gg). The *Te-2*Gg mutant also showed a small but significant increase of ~4% in its $S_{c/o}$ relative to native *Te*-Rubisco (Table 1 and Fig. 3*E*).

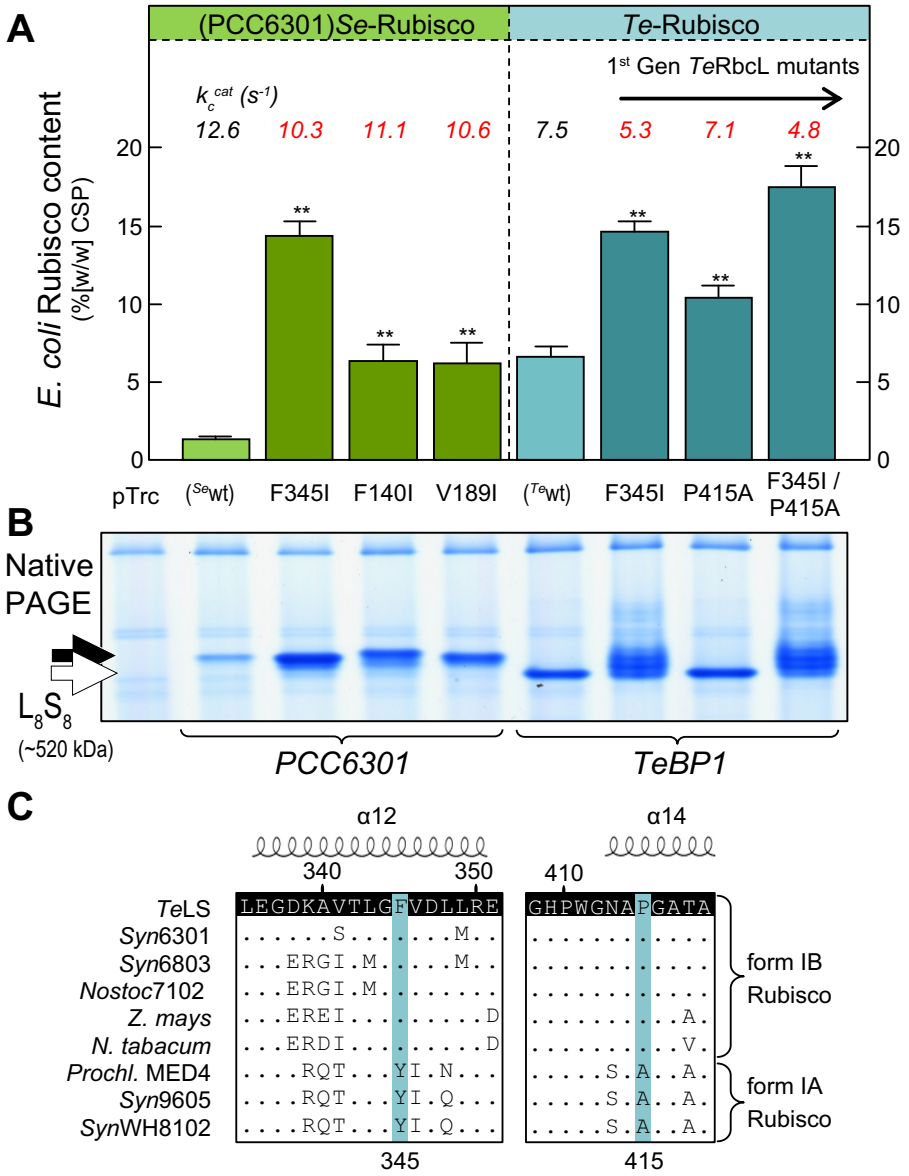

Figure 2. Directed evolution of *Te*-Rubisco using the RDE2 screen. *A* and *B*, comparative L_8S_8 Rubisco biogenesis capacity quantified by [^{14}C]CABP binding (*A*) and confirmed by native PAGE (*B*). The CO_2-fixation rate (k_{cat}^c) of WT and prior RDE selected *S. elongatus* PCC 6301 (*Se*)-Rubisco solubility RbcL mutants (in *green*, F345I, F140I, and V189I) (11, 20) relative to *T. elongatus* BP1 (*Te*)-Rubisco (in *blue*) and the RbcL solubility F345I and P415A mutants identified using the RDE2 screen. The data are the means ± S.E. of more than three biological samples with significance relative to each WT Rubisco analyzed by one-way ANOVA with Turkey's post hoc analysis. **, $p < 0.01$. See Table 1 for additional catalysis measurements. *C*, partial RbcL alignment showing the conserved Phe^{345} and Pro^{415} loci in form IB Rubisco compared with Tyr^{345} and Ala^{415} in the form IA Rubisco examples shown.

Clustering of the catalytic mutations at a RbcL-RbcS interface in Te-Rubisco

Structural analysis revealed that the H37L, Y36N, V98M, L74M, and D397N substitutions all occur at residues that cluster on a surface exposed region of an RbcS–RbcL interface (Fig. 4, *A* and *B*). The A48V mutation occurs on the same interface but is oriented toward the inner enzyme surface (Fig. 4, *A* and *C*). The common locality of these mutations suggest this region to be a hot spot for *Te*-Rubisco catalytic modification, despite being located quite distant from the catalytic pockets in adjoining regions of paired RbcL subunits (Fig. 4*A*). In support of this assertion, the RbcS residue Tyr^{36} is located directly adjacent to Val^{98} in the *Te*-Rubisco holoenzyme (Fig. 4*B*) and can convey comparable improvements in k_{cat}^c and $k_{cat}^c/K_c^{21\%O_2}$ when mutated to Asn (Y36N mutant *Te*-2Gd; Fig. 3, *C* and *D*). Curiously the selected RbcS mutants comprise residue changes adjacent to highly conserved regions of cyanobacterial and higher plant form IB Rubisco (Fig. S2). This suggests that this RbcL-RbcS interface may pose a hot spot for manipulating the kinetics in related plant Rubisco lineages.

Second-generation Te-Rubisco solubility mutant selection

Of the 10 different *Te*-P415A Rubisco-derived mutants selected, five arose from improvements in the biogenesis of soluble *Te*-Rubisco, not catalysis (Fig. 3*A*). The improved solubility derived either from a R51H substitution in RbcS (or A48V, see above) or from I393M, A398T, or A414T mutations in RbcL (Table 1). Notably none of these RbcL substitutions were

Table 1

***E. coli* L_8S_8 Rubisco content and catalytic parameters at 25 °C**

The values are the means ± S.E. of typically $N \geq 3$ biological samples assayed in duplicate or triplicate (details shown in top row). One-way ANOVA was undertaken with reference to either WT *Synechococcus* PCC 6301 Rubisco (*Se*-LS) or WT *Te*-Rubisco (*Te*-LS). Symbols show the statistical significance levels relative to WT enzymes: *, $p < 0.05$; **, $p < 0.01$. 1G, first generation RDE2 *Te*-LS Rubisco mutants selected on media containing 0.1% (w/v) arabinose; 2G, second generation mutants derived from *Te*-P415A under higher PRK selection on 0.25% (w/v) arabinose media (Fig. 3); NA, not applicable; NM, not measured.

Rubisco	Mutant library	Amino acid mutation(s)		Expression (% (w/w) sol *E. coli* protein)	k_{cat}^c	$K_c^{21\%\ O_2}$	$k_{cat}^c/K_c^{21\%O_2}$	$S_{C/O}$	T_m[a]
		RbcL	RbcS		s^{-1}	μM	$mM^{-1}\ s^{-1}$	*mol mol*$^{-1}$	°C
n = biological samples, technical replicates				3–5, 2	3–9, 2	3–5, 2	3–5	3, 3	1,3
Se-LS(WT)	NA	None	None	1.3 ± 0.2	12.6 ± 0.1	251 ± 4	50 ± 1	45.9 ± 0.7	60.2 ± 0.3
Se-F345I	Refs. 11 and 20	F345I	None	14.4 ± 0.9**	10.3 ± 0.4**	285 ± 9**	36 ± 2**	41.8 ± 1.1**	NM
Se-F140I	Ref. 11	F137I	None	6.3 ± 1.1**	11.1 ± 0.5**	211 ± 5**	53 ± 3	46.0 ± 0.7	NM
Se-V186I	Ref. 11	V189I	None	6.2 ± 1.3**	10.6 ± 0.2**	186 ± 7**	57 ± 2**	39.3 ± 0.7**	NM
Te-LS (WT)	NA	None	None	6.6 ± 0.7	7.4 ± 0.1	107 ± 2	70 ± 2	53.3 ± 0.5	72.8 ± 0.2
First generation *Te*-rbcL–RDE2 mutants									
Te-F345I	1G	F345I	None	14.6 ± 0.7**	4.8 ± 0.3**	101 ± 1	47 ± 2**	48.0 ± 0.5**	72.3 ± 0.3
Te-P415A	1G	P415A	None	10.4 ± 0.8**	6.9 ± 0.2*	111 ± 2	62 ± 2*	52.6 ± 0.5	73.7 ± 0.1**
Te-V300A	1G	V300A	None	7.1 ± 0.2	6.6 ± 0.2*	135 ± 9**	49 ± 6**	51.9 ± 0.6	71.4 ± 0.6
Te-FIPA	1G chimer	F345I, P415A	None	17.5 ± 1.2*	4.8 ± 0.4**	102 ± 5	47 ± 6**	49.2 ± 0.4**	75.9 ± 0.1**
Second generation *Te*-rbcLS–RDE2 mutants									
Te-2Ga (1)	2G	P415A	V98M	20.4 ± 3.3**	9.5 ± 0.3**	93 ± 3**	100 ± 3**	56.3 ± 0.3**	NM
Te-2Gb (2)	2G	P415A	A48V	9.7 ± 1.2	8.8 ± 0.2**	105 ± 3	84 ± 3**	48.0 ± 0.6**	NM
Te-2Gc (3)	2G	P415A	H37L	11.1 ± 0.2	8.4 ± 0.3**	100 ± 2*	84 ± 4**	49.7 ± 0.3**	NM
Te-2Gd (5)	2G	P415A	Y36N, G112D	9.7 ± 0.8	9.3 ± 0.2**	94 ± 0.4**	99 ± 2**	47.7 ± 0.1**	NM
Te-2Ge (8)	2G	P415A	Y36N, R51H	13.6 ± 1.6**	8.0 ± 0.4	120 ± 8	68 ± 2**	51.8 ± 0.2**	NM
Te-2Gf (10)	2G	A398T, P415A	A48V	13.3 ± 3.2	8.1 ± 0.1*	105 ± 3	75 ± 5	45.5 ± 0.1**	NM
Te-2Gg (11)	2G	L74M, D397N, P415A	None	12.5 ± 0.8**	8.2 ± 0.5	92 ± 3**	90 ± 6**	55.4 ± 0.3**	NM
Te-2Gh (12)	2G	A414T, P415A	None	13.9 ± 3.0*	8.0 ± 0.1	102 ± 2	78 ± 1	55.8 ± 0.5**	NM
Te-2Gi (13)	2G	I393M, P415A	None	16.6 ± 2.4**	7.8 ± 0.4	90 ± 2**	87 ± 1**	48.1 ± 0.4**	NM
Te-2Gj (15)	2G	A398T, P415A	None	16.6 ± 3.6**	6.4 ± 0.2**	110 ± 4	59 ± 3*	50.5 ± 0.5**	NM

[a] Measurement of thermal stability (T_m) were made by circular dichroism spectroscopy as described (38). The K_m for RuBP of *Se*-LS and *Te*-LS Rubisco are 44 ± 3 μM (20) and 36 ± 2 μM (this study), respectively, ~10-fold less than the 0.4 mM RuBP concentration used in the $^{14}CO_2$-fixation assays. Rubisco expression levels were confirmed by native PAGE (Fig. S1).

selected in the first-generation RDE2 screen, suggesting that they complement the P415A mutation. Structural analyses revealed, however, that the residues associated with increasing *Te*-Rubisco solubility show no apparent relation to one another and are located quite distant from Pro^{415} within the quaternary Rubisco structure (Fig. S3). Consistent with that postulated previously (11), this suggests these residues pose important determinants of RbcL subunit folding by GroEL or/and provide structural stability to avoid their misfolding and aggregation in the absence of ancillary molecular assembly components.

Comparative analysis of Te-Rubisco biogenesis in chloroplasts

The varied biogenesis requirements within the Rubisco superfamily limit not only which isoforms that can be produced in *E. coli* but also those that can be assembled within leaf chloroplasts (10, 14, 18). The availability of efficient plastome transformation capabilities in tobacco has made it the model plant for Rubisco bioengineering (13, 16). Consistent with leaf photosynthesis modeling (Fig. 5*A*), the growth of tobacco producing *Se*-Rubisco necessitate high levels of CO_2 for growth in soil to compensate for the enzyme's poor kinetics and limited biogenesis potential in chloroplasts (19). Comparable modeling using *Te*-Rubisco kinetics show it would support higher rates of CO_2 assimilation in a C_3 plant over all intracellular CO_2 concentrations, more so in plants producing the *Te-2*Ga Rubisco mutant selected in this study using the RDE2 screen (Fig. 5*A*).

To test the correlative potential between *Te*-Rubisco biogenesis in *E. coli* and leaf chloroplasts, a synthetic *Te-rbc*LS operon coding for either *Te*-Rubisco or the high-solubility F345I/P415A mutant was transformed in the tobacco chloroplast genome (plastome) in place of the native tobacco *rbc*L gene (Fig. 5*B* and Fig. S4). To optimize translation, the codon use of the transformed genes matched the native tobacco *rbc*L, and 12 amino acids of the native tobacco RbcL N terminus were retained (Fig. S4*a*). Independently transformed lines for each tobacco genotype (called tob^{TeLS} and $tob^{TeLS\text{-}FIPA}$; Fig. 5*B*) were obtained and grown in tissue culture until homoplasmic (Fig. 5*C*). Plastome sequencing confirmed correct integration of the transgenes, and RNA blots confirmed high *Te-rbc*L-*rbc*S and *Te-rbc*L-*rbc*S-*aad*A mRNA levels (Fig. S4*b*). Nevertheless the tob^{TeLS} and $tob^{TeLS\text{-}FIPA}$ plants shared a pale green phenotype and could only survive in tissue culture (Fig. 5*C*). CABP binding (Fig. 5*D*) and native PAGE analysis (Fig. 5*E*) confirmed the assembly of functional L_8S_8 *Te*-Rubisco in tob^{TeLS} and slightly more of the F345I/P415A mutated *Te*-Rubisco in $tob^{TeLS\text{-}FIPA}$. However, the amount of Te-Rubisco produced were more than 1000-fold lower than Rubisco production in the wild-type tobacco controls (Fig. 5*D*). Reciprocal mutagenic tests showed introducing the F345I, P415A, or dual F345I/P415A mutations into the *N. tabacum* RbcL did not permit the assembly of tobacco Rubisco in *E. coli* (Fig. S5).

Discussion

Following successive evolution rounds, this study reveals a novel region in the subunit interface of the plant-like Rubisco from the *T. elongatus* BP1 that offers multiple solutions for improving carboxylation rate, efficiency, and specificity. Key to this success is the unwavering fidelity, faster screening, and higher throughput of the new RDE2 screening platform. To date, the RDE screen designed around a Δ*gapA E. coli* mutant (a

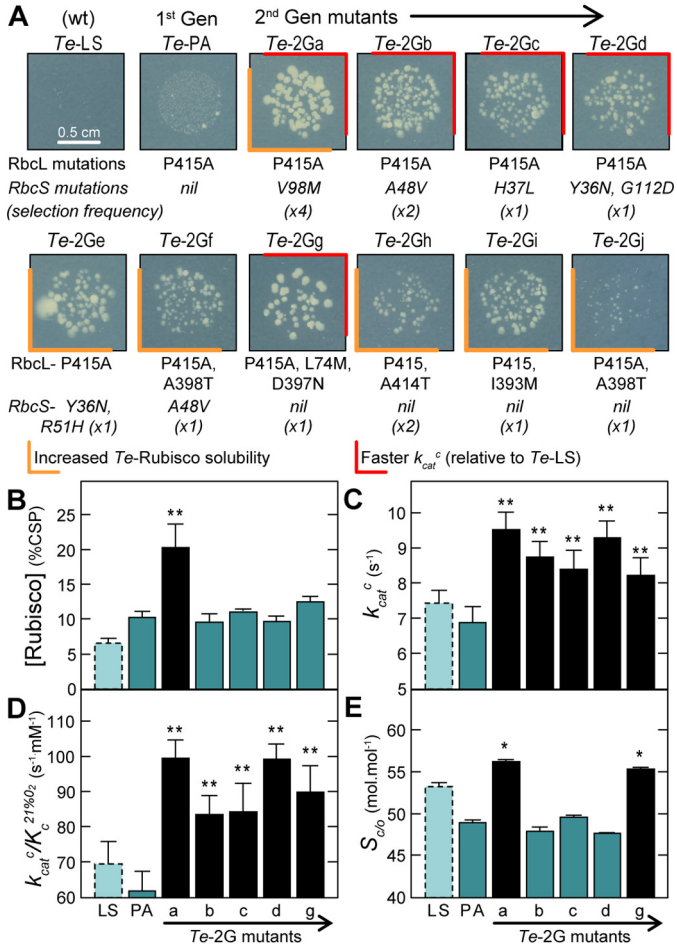

Figure 3. The catalysis and assembly of the second-generation *Te*-Rubisco mutants. *A*, comparative growth after 7 days at 25 °C of XL1-Blue *E. coli* producing wild-type, P415A, and second-generation *Te*-Rubisco (Te-2G) mutants on selective medium (0.25% (w/v) arabinose 0.5 mM IPTG, 0.2 mg·ml^{-1} kanamycin, 0.1 mg·ml^{-1} ampicillin). Shown are the frequencies of selection for each mutant (from a mutagenic library of 3×10^5 colony-forming units) and the amino acid substitutions in RbcL or/and RbcS identified. *B–E*, the effects of the mutations on L_8S_8 Rubisco biogenesis (quantified by [^{14}C]-2-CABP binding) (*B*), CO_2-fixation rate (k_{cat}^{c}) (*C*), apparent carboxylation efficiency (k_{cat}^{c} divided by the apparent K_m for CO_2 under ambient O_2; $K_c^{21\%O_2}$) (*D*), and specificity for CO_2 over O_2 ($S_{c/o}$) at 25 °C, pH 8.0 (*E*), were compared relative to the parental wild-type (*Te*-LS) and *Te*-P415A Rubiscos. Shown are the average data (± S.E. of $N > 3$ samples, see Table 1). Significance of variation to *Te*-Rubisco determined by one-way ANOVA with Turkey's post hoc analysis. *, $p < 0.05$; **, $p < 0.01$.

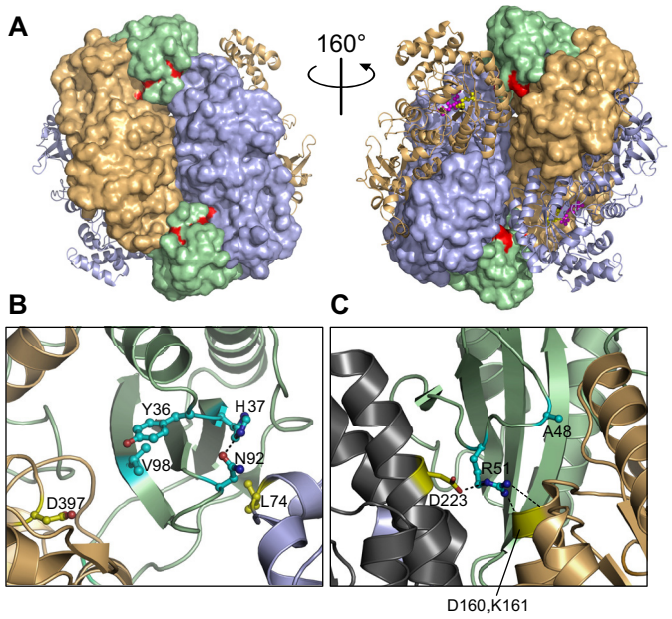

Figure 4. Location of the catalysis enhancing residues in *Te*-Rubisco. *A*, surface view from two orientations showing two *Te*-RbcL (*purple* and *orange*) and their antiparallel *Te*-RbcL partners (*ribbon structures*) bound by upper and lower RbcS monomers (*green*) with the loci of the second-generation mutations shown in *red*. Bound CABP (an inhibitory structural analog of RuBP) in the RbcL catalytic sites is shown in *magenta*, and the conserved Lys^{201} lysine is shown in *yellow*. *B* and *C*, structural detail of the clustered location and interactions of the catalysis influencing Val^{98}, Ala^{48}, His^{37}, and Tyr^{36} residues in *Te*-RbcS and the *Te*-RbcL amino acids Leu^{74} and Asp^{397} (with non-selected H-bond–interacting residues Asn^{92} and Asp^{223} shown; Protein Data Bank code 3ZXW).

strain called MM1) (26) has proven the most effective in Rubisco-directed applications because of the low number of false positives selected (12). Offsetting this improved fidelity is the slow growth rate of the MM1 strain and its need for specialized growth medium that prolongs the screening time frame to 9–12 days at the optimal growth temperature of 25 °C (11, 17, 20, 26). Comparatively, the RDE2 screen is immune to false-positive production and uses LB growth medium, allowing the screens to be completed in 6–7 days at 25 °C. The feasibility of adapting the new CO_2-dependent sugar synthesizing CBB–*E. coli* strains (23) as an alternative RDE screen for selecting Rubisco catalytic mutants remains to be demonstrated (21).

There has been long-standing interest in modifying Rubisco catalysis via mutagenesis of RbcL (3, 5–7, 10). More recent RbcS mutagenic studies have demonstrated how it can pervasively influence plant Rubisco kinetics (27), a finding supported by structure–function surveys that indicate a role by the RbcS in the kinetic adaptation of Rubisco to environmental cues (18). X-ray crystallography comparisons show that the quaternary RbcS structure remains highly conserved among the differing L_8S_8 lineages, especially among the plant and cyanobacteria Form 1B isoforms (Fig. 6*A*). This suggests that RbcS-mediated changes to catalysis can be strongly influenced by subtle changes in electron density within key areas. One particular area, as demonstrated in this study, is located between the antiparallel β-sheets around Val^{98}, Tyr^{36}, and His^{37} in *Te*-RbcS that share structural similarity to Cys^{122}, Trp^{38}, and Val^{39} in tobacco RbcS (Fig. 6*B*). Generally, the Met^{98}, Asn^{36}, or Leu^{37} substitutions differentially improved *Te*-Rubisco catalysis by modification toward more plant-like structures (Fig. 6*C*). Notably these substitutions did not arise from limitations in codon redundancy at these residues in *Te*-RbcS. This suggests that further improvements may be possible by additional rounds of laboratory evolution, or possibly even by rational design guided by the vast array of crystal structure information already available for plant and algae Rubisco (4).

An obvious extension of our study is to test how amino acid changes around this region in RbcS influence plant Rubisco catalysis. Such efforts are hindered by the inadequate transgenic methods for replacing or modifying all the multi-*RbcS* gene copies in plants. The low throughput of modern nucleus gene editing approaches limit their usefulness in mutagenic screening, despite their capacity to introduce multiplexed nucleotide changes with relatively high accuracy (28). Although

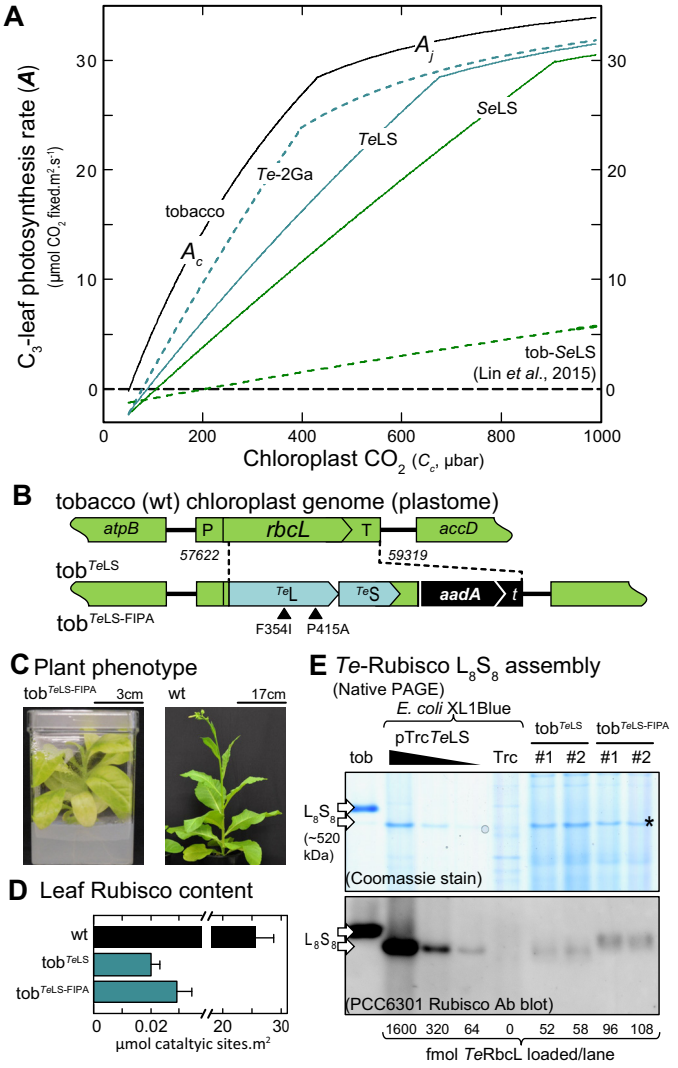

Figure 5. *Te*-Rubisco expression in tobacco. *A*, simulated effect of *Se*-Rubisco (*Se*LS), *Te*-Rubisco (*Te*LS), and the catalytically improved *Te*2Ga-Rubisco (*Te*-2Ga) on CO_2 assimilation rate (*A*) in tobacco (a model C_3 plant) at 25 °C in response to varying chloroplast CO_2 partial pressures (C_c). *A* was calculated as the minimum of the Rubisco carboxylase-limited (A_c) and light-limited (A_j) rates modeled as described under "Experimental methods" and assuming a leaf Rubisco content of 30 μmol active sites m^{-2} (or 5 μmol active sites m^{-2} as observed in tob-SeLS (19)). *B*, the tobacco plastome was transformed with plasmids pLEV-TeLS and pLEV-TeLS-FIPA containing homologous plastome flanking sequences (indicated by *dashed lines*; numbering relative to *N. tabacum* (WT) plastome; GenBank™ accession no. Z00044) that directed integration of codon optimized *rbcLS* operon (GenBank™ accession no. MG000149) coding WT or FIPA *Te*-Rubisco to produce the tobacco genotypes tob^{TeLS} and tob$^{TeLS\text{-}FIPA}$ (see Fig. S4 for further detail). *C–E*, the tob^{TeLS} and tob$^{TeLS\text{-}FIPA}$ plants could be grown in tissue culture (*C*) but not in soil because both produced finite levels of L_8S_8 *Te*-Rubisco as quantified by [^{14}C]-CABP binding (*D*) and confirmed by native PAGE (*E*). Loading controls included tobacco Rubisco (*tob*), known amounts of *Te*-Rubisco (pTrc^{TeLS}) expressed in *E. coli*, and an empty vector (Trc) negative control. No L_8S_8 *Te*-Rubisco band was evident by Coomassie staining in the tob^{TeLS} and tob$^{TeLS\text{-}FIPA}$ leaf protein samples but detected by immunoblotting using a *Se*-Rubisco antibody. *, non-Rubisco tobacco protein that separates at the same location as *Te*-Rubisco by native PAGE.

our RDE2 screen poses a potential remedy in terms of screening throughput, our data highlight its versatility for evolving vascular plant Rubisco will require incorporation of requisite Rubisco assembly components from leaf chloroplasts. This observation accords with the underpinning need for ancillary protein complementarity for folding and assembly of form I Rubisco in leaf chloroplasts (13). Understanding these requirements is critical to enabling heterologous Rubisco expression potential in cyanobacteria, *E. coli*, and chloroplasts. Despite the common prokaryotic connections of these hosts, our work demonstrates that *E. coli* cannot be used as a proxy for chloroplast Rubisco assembly. This underpins the importance of understanding the specific chaperonin (Cpn60α, 60β), co-chaperonin (Cpn10, 20, and 21), and assembly chaperone requirements of plant and algal Rubisco or strategies to circumvent their necessity for successful production in a non-native host. The last decade has seen significant advances into understanding the mechanisms for many of these components (14), but they have not yet been incorporated into a screen such as RDE2. Conceivably such an approach would provide a high-throughput screen of plant (or algae) Rubisco activity that would be envisaged to function equivalently chloroplasts (29).

Efforts to introduce a cyanobacterial CO_2-concentration mechanism (CCM) into tobacco plastids are also likely to require the inclusion of Rubisco biogenesis components. In this study we show that *Te*-Rubisco has the highest CO_2 affinity and CO_2/O_2 specificity of known cyanobacteria Rubisco. These kinetic features make *Te*-Rubisco better able to support C_3 photosynthesis, relative to the more commonly studied *Se*-Rubisco. The reverse would be the case if the extensive requirements for building a functional CCM could be met in plastids to support the higher k_{cat}^{c} of *Se*-Rubisco (30, 31). The limited biogenesis capacity of *Te*-Rubisco (Fig. 5) and *Se*-Rubisco (19, 30, 32) in tobacco suggest that building a CCM in chloroplasts will require not only introducing the multiple carboxysome components, inorganic carbon membrane transporters, and modulating CA levels, but also incorporating the necessary chaperonin/chaperone components for assembling cyanobacteria Rubisco.

Conclusion

Much of the effort to identify superior Rubiscos to date has focused on measuring the natural variation in Rubisco kinetics. This approach is slowed by the complexity of the catalytic assay methods and the vast spectrum of natural diversity available. Here we show how directed evolution using RDE2 poses a potentially faster alternative. The challenge is to extend past the prior success of RDE screens in investigating prokaryotic and Archaea Rubisco sources and adapt RDE2 for the directed evolution of plant and algae Rubisco. As with Rubisco engineering in plant chloroplasts, our data advocate that this will necessitate the co-expression of the complimentary ancillary proteins needed for the biogenesis and metabolic repair of the target Rubisco.

Experimental procedures

Expression, purification, and assay of PRK and PRK:NPTII

The *prkA* gene from *Synechococcus* PCC6301 was synthesized (Genscript) and cloned downstream of the arabinose inducible BAD promoter in pAC^{BAD} (26) to generate pACYC^{PRK}. Sequence coding a synthetic *nptII* gene was cloned in frame to the 3′ of *prkA* to produce the *prk-npt*II gene coding a PRK–NPTII fusion protein (GenBank™ accession no. MG000147). The *prk*A and *prk-npt*II genes were cloned in frame to a N-terminal His_6–ubiquitin fusion protein using the expression

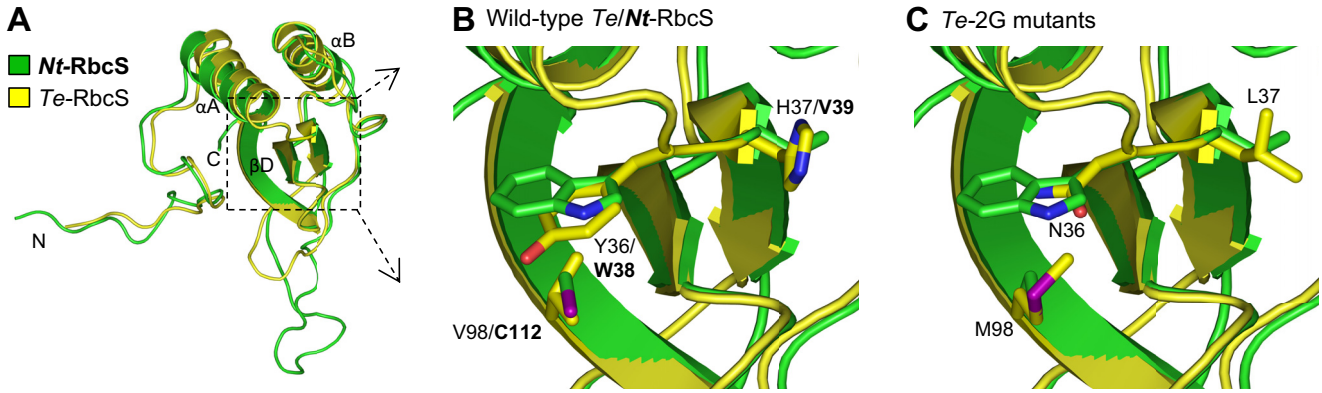

Figure 6. Comparative structures of tobacco and *Te*-RbcS. *A*, structural overlay of *N. tabacum* (*Nt*RbcS, *green*, PDB code 4RUB) and *T. elongatus* (*Te*RbcS, *yellow*, PDB code 3ZXW) RbcS structures *B*, comparative positioning of the amino acid side chains along the anti-parallel β-sheet in wild-type *Nt*RbcS and *Te*RbcS at the loci selected in the second-generation *Te*-Rubisco mutants. *C*, extension of *B* comparing the *Te*RbcS-2G amino acid substitutions with *Nt*RbcS.

plasmid pHue, expressed in BL21(DE3) *E. coli*, purified by immobilized metal affinity chromatography and the 10-kDa His_6–ubiquitin sequence removed as previously described (33). PRK activity (k_{cat}) was measured using a NADH enzyme-linked assay as described (34).

Comparing the selection stringency of RDE and RDE2

The *Te*-Rubisco expressing plasmid $pTrc^{TeLS}$ (Fig. 1*B*) was produced by separately amplifying the *rbc*L and *rbc*S genes from genomic DNA isolated from *T. elongatus* BP1 (gifted by Dr. Warwick Hillier, Australian National University) and cloning into pTrcHisB as a NcoI-SalI fragment (GenBank™ accession no. MG000148). *E. coli* XL1-Blue cells containing $pTrc^{TeLS}$ were transformed with either $pACYC^{PRK}$ (RDE) or $pACYC^{PRK:Kan}$ (RDE2; Fig. 1*B*) and grown at 25 °C for 6 days in air with 2% (v/v) CO_2 on LB-agar containing 0.5 mM IPTG (*Te*-Rubisco induction), 0.2% (w/v) arabinose (high PRK/PRK–NPTII induction) and with/without 0.2 mg·ml^{-1} kanamycin.

Library construction and Rubisco selection using the RDE2 screen

The *rbc*LS operon in $pTrc^{TeLS}$ (first-generation library) or $pTrc^{TePA}$ (second-generation library) were amplified from 10 ng of plasmid DNA by error-prone PCR using the primers Trc55 (5′-GAGGTATATATTAATGTATCG-3′) and Trc33 (5′-ATCTTCTCTCATCCGCCA-3′) and the Genemorph II random mutagenesis kit (Agilent), as per the manufacturer's recommendations. Libraries were transformed into $CaCl_2$-competent RDE2 cells and plated onto LB medium containing 32 μg/ml chloramphenicol, 200 μg/ml ampicillin, 100 μg/ml kanamycin, 0.5 mM IPTG, and 0.05 to 0.25% (w/v) arabinose. After growing at 25 °C for 3–7 days in air containing 2% (v/v) CO_2, the faster-growing colonies were replated, and then their $pTrc^{TeLS}$ plasmids were purified, sequenced, and retransformed into RDE2, and the screen was repeated to confirm their selective advantage.

Rubisco content and catalysis

Cyanobacteria Rubisco expression was induced in XL-1Blue *E. coli* transformed with $pTrc^{TeLS}$, pTrcSynLS (20), or their mutant plasmid derivatives with 1 mM IPTG at 28 °C. After 6 h the cells were harvested by centrifugation (5 min at 4 °C, 6200 × *g*), and the cell pellets were N_2-frozen and stored at −80 °C. Their soluble protein was isolated following lysis using a French pressure cell (140 MPa) in ice-cold extraction buffer (100 mM EPPES-NaOH, pH 8.05, 15 mM $MgCl_2$, 0.5 mM EDTA, 1 mM PMSF, 2.5 mM DTT). After centrifugation (3 min, 2 °C, 15,000 × *g*), the supernatant was mixed with an equal volume of reaction buffer (100 mM EPPES-NaOH, pH 8.05, 15 mM $MgCl_2$, 0.5 mM EDTA) containing 40 mM $NaH^{14}CO_3$ and incubated at 25 °C for 8 or 12 min (technical repeats) to activate Rubisco. The $^{14}CO_2$ fixation assays (0.5 ml of total volume) were performed in 7.7 ml of septum-capped scintillation vials in reaction buffer containing 10 μg ml^{-1} carbonic anhydrase and saturating (0.4 mM) RuBP synthesized and purified according to Kane *et al.* (36). All assay components were equilibrated in CO_2-free air (*i.e.* 20.9% (v/v) O_2 in N_2) prior to adding a series of five differing amounts of $NaH^{14}CO_3$. The final $^{14}CO_2$ concentrations were 0–240 μM or 0–350 μM when assaying the wild-type and mutant variants of *Te*-Rubisco or *Se*-Rubisco, respectively. The assays were initiated by addition of 20 μl of $^{14}CO_2$-activated *E. coli* soluble protein. RuBP-independent $^{14}CO_2$ fixation control assays were run for each protein sample and contained H_2O in place of RuBP. The assays were terminated after 1 min by rapid mixing with 0.2 ml of 20% (v/v) formic acid and then dried at 80 °C before adding 0.5 ml of H_2O and mixing with 1 ml of Ultima-Gold scintillant (PerkinElmer). The fixed ^{14}C was measured in a Tri-carb 4910TR scintillation counter with the carboxylase activity between the technical repeats varying by <2%, confirming the extracted Rubisco was fully activated and stable.

The specific activity of each $NaH^{14}CO_3$ stock was determined in assays ($n = 4$) containing the highest $^{14}CO_2$ concentration and 5 nmol of pure RuBP. These assays were allowed to react for 30 or 90 min to ensure full RuBP fixation. After acid treatment and drying, the scintillation counter ^{14}C values were divided by 5 to derive the specific activity value, which varied between 1500 and 1800 cpm/nmol CO_2 fixed.

The CO_2 levels in the assays were calculated using the Henderson–Hasselbalch–derived equation,

$$[CO_2] = \frac{(C_t)}{1 + \frac{V}{vqRT} + 10^{(pH-pK_1)} + 10^{(2pH-pK_1-pK_2)}} \quad \text{(Eq. 1)}$$

where V/v is the ratio of reaction vial headspace (V) to assay volume (v); C_t is the total $NaHCO_3$ concentration (including the 0.4 μmol of $NaH^{14}CO_3$ in the *E. coli* soluble protein); q is the CO_2 solubility at 1 atm (0.0329 mol/liter/atm at 25 °C); R is the universal gas constant (0.082057 liter/atm/K/mol); T is the assay temperature (298 K); and dissociation constants are 6.251 (pK_1) and 10.329 (pK_2). The ^{14}C data were fitted to the Michaelis–Menten equation to derive the apparent Michaelis–Menten constant (K_m) for CO_2 under air levels of O_2 ($K_c^{21\%O_2}$) and the maximal rate of carboxylation (V_c^{max}) from at least three independent *E. coli* soluble protein preparations. The caboxylation rate for each Rubisco (k_{cat}^c) was standardized by dividing V_c^{max} by the number of Rubisco catalytic sites in the *E. coli* protein quantified by [^{14}C]-2-CABP binding and native PAGE(20, 35). The soluble protein concentration in each *E. coli* extract was quantified using a Coomassie dye binding assay against BSA. Assays to measure the K_m for RuBP of *Te*-Rubisco (see legend to Table 1) were undertaken in reactions containing 20 mM $NaH^{14}CO_3$ and a series of six RuBP concentrations (0–150 μM).

Rubisco $S_{C/O}$ measurements were made as described by Kane *et al.* (36) using recombinant Rubisco purified from *E. coli* by anion exchange chromatography and then Superdex 200 (GE Life Sciences) size exclusion column chromatography (35). Each purified Rubisco was equilibrated with a gas mixture of 0.05% (v/v) CO_2 and 99.95% (v/v) O_2 (accurately mixed using Wostoff gas mixing pumps) at 25 °C in a replica septum seal 20-ml glass vial assays comprising 1 ml of specificity buffer (30 mM triethanolamine, pH 8.1, 10 mM $MgSO_4$, 10 μg/ml carbonic anhydrase). After 1 h the reactions were initiated by the addition of 1 nmol [1-^{3}H]-RuBP (10 MBq/nmol), and then 10 units of alkaline phosphatase (Sigma) was added 30–60 min later. The resulting [^{3}H]glycerate and [^{3}H]glycolate products were then separated by HPLC and measured by scintillation counting as described (36). The $S_{C/O}$ factor was calculated using the following equation,

$$S_{c/o} = \left(R_{\text{glycerate/glycolate}} \times \frac{M_{O_2}}{M_{CO_2}}\right) \times 0.037 \qquad \text{(Eq. 2)}$$

where from R is the ratio of [^{3}H]glycerate to [^{3}H]glycolate; M_{O_2} and M_{CO_2} are the mole fractions of O_2 and CO_2, respectively, in the assay; and 0.037 is the ratio between the aqueous solubility of O_2 and CO_2 at 25 °C.

PAGE analyses

The leaf and *E. coli* soluble protein extracts were prepared and analyzed by SDS-PAGE, native PAGE, and immunoblot analysis as described (35).

Tobacco chloroplast transformation

The plasmids pLEV*Te*LS (GenBank™ accession no. MG000149) and pLEV*Te*LS^{FIPA} were biolistically transformed into the plastome of the ^{cm}trL tobacco genotype as previously described (16). The resulting transplastomic genotypes tob^{TeLS} and tob$^{TeLS\text{-}FIPA}$ coded a synthetic *rbc*LS operon and the *aad*A marker gene (coding spectinomycin resistance) in place of the tobacco *rbc*L gene (Fig. 5*A*; see also Fig. S4 for additional transformation and RNA blotting detail). Correct transgene insertion was confirmed by fully sequencing the PCR product amplified from total leaf genomic DNA isolated using the DNeasy® plant mini kit and primers LSD (5′-CACGGAATTCGTGTCGAGTAG-3′) and LSZ (5′-ATCCTTCTTTATTTCCTGC-3′).

Leaf photosynthesis simulations

Photosynthetic CO_2 assimilation rates (A) in tobacco at 25 °C under varying chloroplast CO_2 partial pressures (C_C) were simulated as the minimum of the Rubisco carboxylase-limited (A_c) and light-limited (A_j) CO_2-assimilation rates modeled according to Ref. 37,

$$A_c = \frac{m \cdot k_{cat}^c (C_c \cdot S_c - 0.5\, O_c/S_{c/o})}{(C_c \cdot S_c + K_c^{21\%O_2})} - R_d \qquad \text{(Eq. 3)}$$

$$A_j = \frac{(C_c \cdot S_c - 0.5\, O_c/S_{c/o}) J_{max}}{4(C_c \cdot S_c + O_c/S_{c/o})} - R_d \qquad \text{(Eq. 4)}$$

using the Rubisco kinetic values listed in Table S1 and a leaf Rubisco content (m) of 30 μmol active sites m^{-2} (or 5 μmol active sites m^{-2} for the assembly impaired *Se*-Rubisco producing tobacco line tob-SeLS, (19)); a maximal RuBP regeneration rate (J_{max}) of 160 μmol m^{-2} s^{-1}; a mitochondrial respiration rate (R_d) of 1 μmol m^{-2} s^{-1} and using the solubility constants 0.0334 M bar^{-1} (s_c) and 0.00126 M bar^{-1} (s_o) to calculate the CO_2 (C_c) and O_2 (O_c) concentrations in the chloroplast.

Author contributions—R. H. W. and S. M. W. conceived the study. R. H. W. designed RDE2, performed the directed evolution research, and generated the tobacco transgenics that were maintained and analyzed by C. C., E. M.-A., and S. M. W. R. H. W. and S. M. W. performed the Rubisco biochemical analyses and wrote the manuscript.

Acknowledgment—We thank Andreas Bracher (Max Planck Institute for Biochemistry, Planegg, Germany) for constructive assistance in Rubisco structural modeling and Yu Zhou for technical assistance.

Note added in proof—Fig. 2*D* was removed during redaction.

References

1. Long, S. P., Marshall-Colon, A., and Zhu, X.-G. (2015) Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* **161,** 56–66 CrossRef Medline
2. Tcherkez, G. (2016) The mechanism of Rubisco-catalysed oxygenation. *Plant Cell Environ.* **39,** 983–997 CrossRef Medline
3. Whitney, S. M., Houtz, R. L., and Alonso, H. (2011) Advancing our understanding and capacity to engineer nature's CO_2-sequestering enzyme, Rubisco. *Plant Physiol.* **155,** 27–35 CrossRef Medline
4. Andersson, I., and Backlund, A. (2008) Structure and function of Rubisco. *Plant Physiol. Biochem.* **46,** 275–291 CrossRef Medline
5. Parry, M. A., Andralojc, P. J., Scales, J. C., Salvucci, M. E., Carmo-Silva, A. E., Alonso, H., and Whitney, S. M. (2013) Rubisco activity and regulation as targets for crop improvement. *J. Exp. Bot.* **64,** 717–730 CrossRef Medline
6. Sharwood, R. E. (2017) Engineering chloroplasts to improve Rubisco catalysis: prospects for translating improvements into food and fiber crops. *New Phytol.* **213,** 494–510 CrossRef Medline
7. Carmo-Silva, E., Scales, J. C., Madgwick, P. J., and Parry, M. A. J. (2015) Optimizing Rubisco and its regulation for greater resource use efficiency. *Plant Cell Environ.* **38,** 1817–1832 CrossRef Medline

8. Ellis, R. J. (1979) The most abundant protein in the world. *Trends Biochem. Sci.* **4,** 241–244 CrossRef
9. Ort, D. R., Merchant, S. S., Alric, J., Barkan, A., Blankenship, R. E., Bock, R., Croce, R., Hanson, M. R., Hibberd, J. M., Long, S. P., Moore, T. A., Moroney, J., Niyogi, K. K., Parry, M. A., Peralta-Yahya, P. P., *et al.* (2015) Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 8529–8536 CrossRef Medline
10. 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 CrossRef Medline
11. Durão, P., Aigner, H., Nagy, P., Mueller-Cajar, O., Hartl, F. U., and Hayer-Hartl, M. (2015) Opposing effects of folding and assembly chaperones on evolvability of Rubisco. *Nat. Chem. Biol.* **11,** 148–155 CrossRef Medline
12. Mueller-Cajar, O., and Whitney, S. M. (2008) Directing the evolution of Rubisco and Rubisco activase: first impressions of a new tool for photosynthesis research. *Photosyn. Res.* **98,** 667–675 CrossRef Medline
13. Whitney, S. M., Birch, R., Kelso, C., Beck, J. L., and Kapralov, M. V. (2015) Improving recombinant Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1 chaperone. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 3564–3569 CrossRef Medline
14. Bracher, A., Whitney, S. M., Hartl, F. U., and Hayer-Hartl, M. (2017) Biogenesis and metabolic maintenance of Rubisco. *Annu. Rev. Plant Biol.* **68,** 29–60 CrossRef Medline
15. Tsai, Y.-C., Mueller-Cajar, O., Saschenbrecker, S., Hartl, F. U., and Hayer-Hartl, M. (2012) Chaperonin cofactors, Cpn10 and Cpn20, of green algae and plants function as hetero-oligomeric ring complexes. *J. Biol. Chem.* **287,** 20471–20481 CrossRef Medline
16. Whitney, S. M., and Sharwood, R. E. (2008) Construction of a tobacco master line to improve Rubisco engineering in chloroplasts. *J. Exp. Bot.* **59,** 1909–1921 CrossRef Medline
17. Wilson, R. H., Alonso, H., and Whitney, S. M. (2016) Evolving *Methanococcoides burtonii* archaeal Rubisco for improved photosynthesis and plant growth. *Sci. Rep.* **6,** 22284 CrossRef Medline
18. Sharwood, R. E., Ghannoum, O., Kapralov, M. V., Gunn, L. H., and Whitney, S. M. (2016) Temperature responses of Rubisco from *Paniceae* grasses provide opportunities for improving C_3 photosynthesis. *Nat. Plants* **2,** 16186 CrossRef Medline
19. Lin, M. T., Occhialini, A., Andralojc, P. J., Parry, M. A., and Hanson, M. R. (2014) A faster Rubisco with potential to increase photosynthesis in crops. *Nature* **513,** 547–550 CrossRef Medline
20. Mueller-Cajar, O., and Whitney, S. M. (2008) Evolving improved *Synechococcus* Rubisco functional expression in *Escherichia coli*. *Biochem. J.* **414,** 205–214 CrossRef Medline
21. Antonovsky, N., Gleizer, S., and Milo, R. (2017) Engineering carbon fixation in *E. coli*: from heterologous Rubisco expression to the Calvin–Benson–Bassham cycle. *Curr. Opin. Biotechnol.* **47,** 83–91 CrossRef Medline
22. Wilson, R. H., and Whitney, S. M. (2017) Improving CO_2 fixation by enhancing Rubisco performance. In *Directed Enzyme Evolution: Advances and Applications* (Alcalde, M., ed) pp. 101–126, Springer International Publishing, Cham, Switzerland
23. Antonovsky, N., Gleizer, S., Noor, E., Zohar, Y., Herz, E., Barenholz, U., Zelcbuch, L., Amram, S., Wides, A., Tepper, N., Davidi, D., Bar-On, Y., Bareia, T., Wernick, D. G., Shani, I., *et al.* (2016) Sugar synthesis from CO_2 in *Escherichia coli*. *Cell* **166,** 115–125 CrossRef Medline
24. Cai, Z., Liu, G., Zhang, J., and Li, Y. (2014) Development of an activity-directed selection system enabled significant improvement of the carboxylation efficiency of Rubisco. *Protein Cell* **5,** 552–562 CrossRef Medline
25. Greene, D. N., Whitney, S. M., and Matsumura, I. (2007) Artificially evolved *Synechococcus* PCC6301 Rubisco variants exhibit improvements in folding and catalytic efficiency. *Biochem. J.* **404,** 517–524 CrossRef Medline
26. Mueller-Cajar, O., Morell, M., and Whitney, S. M. (2007) Directed evolution of Rubisco in *Escherichia coli* reveals a specificity-determining hydrogen bond in the form II enzyme. *Biochemistry* **46,** 14067–14074 CrossRef Medline
27. Ishikawa, C., Hatanaka, T., Misoo, S., Miyake, C., and Fukayama, H. (2011) Functional incorporation of sorghum small subunit increases the catalytic turnover rate of Rubisco in transgenic rice. *Plant Physiol.* **156,** 1603–1611 CrossRef Medline
28. Yin, K., Gao, C., and Qiu, J.-L. (2017) Progress and prospects in plant genome editing. *Nat. Plants* **3,** 17107 CrossRef Medline
29. Whitney, S. M., and Andrews, T. J. (2003) Photosynthesis and growth of tobacco with a substituted bacterial Rubisco mirror the properties of the introduced enzyme. *Plant Physiol.* **133,** 287–294 CrossRef Medline
30. Hanson, M. R., Lin, M. T., Carmo-Silva, A. E., and Parry, M. A. (2016) Towards engineering carboxysomes into C_3 plants. *Plant J.* **87,** 38–50 CrossRef Medline
31. Price, G. D., and Howitt, S. M. (2014) Towards turbocharged photosynthesis. *Nature* **513,** 497–498 CrossRef Medline
32. Occhialini, A., Lin, M. T., Andralojc, P. J., Hanson, M. R., and Parry, M. A. (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 CrossRef Medline
33. Baker, R. T., Catanzariti, A. M., Karunasekara, Y., Soboleva, T. A., Sharwood, R., Whitney, S., and Board, P. G. (2005) Using deubiquitylating enzymes as research tools. *Methods Enzymol.* **398,** 540–554 CrossRef Medline
34. He, Z., von Caemmerer, S., Hudson, G. S., Price, G. D., Badger, M. R., and Andrews, T. J. (1997) Ribulose-1,5-bisphosphate carboxylase/oxygenase activase deficiency delays senescence of ribulose-1,5-bisphosphate carboxylase/oxygenase but progressively impairs its catalysis during tobacco leaf development. *Plant Physiol.* **115,** 1569–1580 CrossRef Medline
35. Sharwood, R. E., von Caemmerer, S., Maliga, P., and Whitney, S. M. (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,** 83–96 Medline
36. 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. *Aust. J. Plant Physiol.* **21,** 449–461 CrossRef
37. Farquhar, G. D., von Caemmerer, S., and Berry, J. A. (1980) A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta* **149,** 78–90 CrossRef Medline
38. Whitfield, J. H., Zhang, W. H., Herde, M. K., Clifton, B. E., Radziejewski, J., Janovjak, H., Henneberger, C., and Jackson, C. J. (2015) Construction of a robust and sensitive arginine biosensor through ancestral protein reconstruction. *Protein Sci.* **24,** 1412–1422 CrossRef Medline