



Article

Directed Evolution of an Improved Rubisco; In Vitro Analyses to Decipher Fact from Fiction

Yu Zhou and Spencer Whitney *

Australian Research Council Center of Excellence for Translational Photosynthesis, Research School of Biology, The Australian National University, 134 Linnaeus Way, Acton, ACT 0200, Australia; Yu.Zhou@anu.edu.au

* Correspondence: spencer.whitney@anu.edu.au; Tel.: +61-2-6125-5073

Received: 30 August 2019; Accepted: 4 October 2019; Published: 10 October 2019



Abstract: Inaccuracies in biochemically characterizing the amount and CO₂-fixing properties of the photosynthetic enzyme Ribulose-1,5-bisphosphate (RuBP) carboxylase/oxygenase continue to hamper an accurate evaluation of Rubisco mutants selected by directed evolution. Here, we outline an analytical pipeline for accurately quantifying Rubisco content and kinetics that averts the misinterpretation of directed evolution outcomes. Our study utilizes a new T7-promoter regulated Rubisco Dependent *Escherichia coli* (RDE3) screen to successfully select for the first *Rhodobacter sphaeroides* Rubisco (*RsRubisco*) mutant with improved CO₂-fixing properties. The *RsRubisco* contains four amino acid substitutions in the large subunit (RbcL) and an improved carboxylation rate (k_{cat}^C , up 27%), carboxylation efficiency (k_{cat}^C/K_m for CO₂, increased 17%), unchanged CO₂/O₂ specificity and a 40% lower holoenzyme biogenesis capacity. Biochemical analysis of *RsRubisco* chimeras coding one to three of the altered amino acids showed Lys-83-Gln and Arg-252-Leu substitutions (plant RbcL numbering) together, but not independently, impaired holoenzyme (L₈S₈) assembly. An N-terminal Val-11-Ile substitution did not affect *RsRubisco* catalysis or assembly, while a Tyr-345-Phe mutation alone conferred the improved kinetics without an effect on *RsRubisco* production. This study confirms the feasibility of improving Rubisco by directed evolution using an analytical pipeline that can identify false positives and reliably discriminate carboxylation enhancing amino acids changes from those influencing Rubisco biogenesis (solubility).

Keywords: photosynthesis; carbon fixation; synthetic biology; metabolic engineering

1. Introduction

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is the initial, often rate-limiting, enzyme in the photosynthetic Calvin–Benson–Bassham (CBB) cycle that converts CO₂ into the pre-cursor organic molecules that underpin life in the biosphere [1]. Rubisco constitutes 20–50% of the protein in leaves and 5–25% of the protein in algae and photosynthetic prokaryotes, making it the most abundant protein on earth [2–4]. This abundance is necessary due to the slowness of the rate at which Rubisco fixes CO₂ to Ribulose-1,5-bisphosphate (RuBP) to produce duplicate 3-phosphoglycerate (3PGA) molecules—a reaction completing only 2–5 cycles per second in vascular plants. The carboxylation reaction rate is also compromised by the erroneous fixation of O₂ instead of CO₂, producing a 3PGA and 2-phosphoglycolate (2PG), a metabolically toxic product whose recycling back into 3PGA (from two 2PG) by photorespiration consumes energy and releases fixed CO₂ [5,6]. In crops such as wheat and rice, 25% or more of the CO₂ assimilated by photosynthesis is lost during photorespiration. Improving the carboxylation properties of plant Rubisco has therefore long been considered a valuable approach for enhancing crop photosynthesis and growth [7]. As more efficient Rubisco isoforms of an equivalent Form I structure exist naturally in some red algae, the feasibility of improving plant Rubisco appears tenable [8].

Form I Rubisco comprises a structurally conserved ~550 kD complex comprising a core of eight large (RbcL) subunits (arranged as a tetrad of RbcL₂ dimers) capped by two tetrads of small (RbcS) subunits [9]. The eight catalytic sites in each L₈S₈ complex are formed by conserved amino acids in each adjoining RbcL₂ pair (Figure A1). While we have a relatively detailed understanding of the conserved, complex catalytic chemistry of Rubisco [10,11] the rationale design of a better Rubisco remains elusive. Compounding the challenge is the finding that natural Rubisco kinetic diversity has evolved from amino acid differences outside of the catalytic pocket, sometimes in the distantly-located RbcS [12–14].

Defining what constitutes a better Rubisco is dependent on the cellular physiology of the target organism. In organisms incorporating a CO₂-concentrating mechanism (CCM) to hinder Rubisco oxygenation (e.g., in most algae, cyanobacteria and plants with crassulacean acid metabolism, CAM, or C₄-physiologies), improving the CO₂ fixation rate (k_{cat}^C) would represent a better Rubisco [1]. In organisms lacking a CCM, such as C₃ plants, the photosynthetic models show that improving Rubisco specificity for CO₂ over O₂ ($S_{C/O}$) and increasing carboxylation efficiency (i.e., k_{cat}^C divided by the K_m for CO₂ under ambient O₂, $K_c^{21\%O_2}$) are key to producing a “better” Rubisco [15]. Rubisco from some red algae possesses these desired kinetic improvements. For example the $S_{C/O}$ and $k_{cat}^C/K_c^{21\%O_2}$ values for red algae *Griffithsia monilis* Rubisco are, respectively, two-fold and 1.3-fold higher than tobacco Rubisco [16]. While this kinetic improvement has the potential to improve plant growth by 30% [17], the assembly requirements of red algae Rubisco are unfortunately not met in tobacco chloroplasts [18,19]. A comparable molecular chaperone incompatibility also impairs, or prevents, the assembly of heterologous plant and cyanobacteria Rubisco in tobacco chloroplasts [20–23].

A successful approach to improving Rubisco kinetics is via the directed evolution of randomly mutated *rbcL* (+/– *rbcS*) libraries using high-throughput Rubisco-dependent *Escherichia coli* (RDE) screens [22,24,25]. As summarized in Figure 1, a core requirement shared among RDE screens is the expression of a *prk* gene coding phosphoribulose kinase (PRK). PRK catalyses the ATP-dependent phosphorylation of the ribulose 5-phosphate produced in the *E. coli* pentose phosphate pathway into RuBP. For unknown reasons, the PRK reaction is toxic to many *E. coli* strains unless Rubisco is introduced to catalyze the RuBP. As RDE selection is undertaken under elevated CO₂, the RuBP is primarily carboxylated by Rubisco to produce the glycolytic intermediate 3PGA. Any 2PG produced via Rubisco oxygenase activity can also be metabolized by *E. coli* [26].

The life–death dependency of RDE screens on Rubisco activity has been exploited using the L-arabinose inducible P_{BAD} promoter to modulate PRK expression (i.e., RuBP production) to select for cells producing higher levels of Rubisco activity. This improvement to RDE fitness can stem from mutations in Rubisco that either improve its biogenesis (i.e., increasing the solubility of Rubisco) or/and increase its catalytic rate. Failure to accurately determine which of these Rubisco biochemical properties improves RDE selection fitness has led to solubility-enhancing RbcL substitutions being erroneously reported as mutations that enhance carboxylase activity (Table 1) [25].

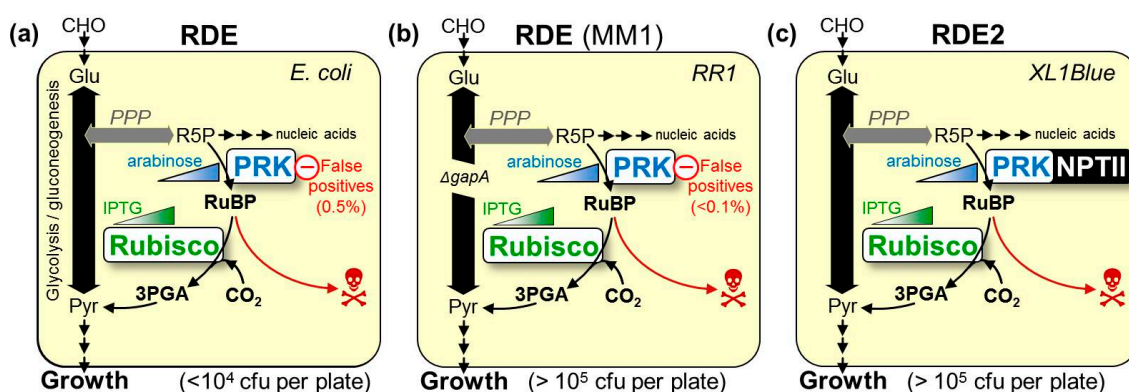


Figure 1. Metabolic rewiring in the different Rubisco dependent *E. coli* (RDE) screens. All RDE screens contain arabinose inducible BAD promoter-regulated *prk* genes whose product, phosphoribulokinase (PRK), phosphorylates ribulose 5-phosphate (R5P, produced in the pentose phosphate pathway, PPP) into RuBP. The toxicity of this reaction can be alleviated by expressing Rubisco to catalyse RuBP carboxylation into metabolically compatible 3-phosphoglycerate (3PGA). Cell growth can also arise from mutations that silence PRK to produce false positives (in which colony growth is independent of Rubisco activity). (a) The frequency of false positives in wild-type *E. coli* RDE screens is high (~0.5% of plated cells, left panel) reducing the number of colony-forming units (cfu) that can be effectively screened per plate. (b) The false positives frequency is >5-fold lower in the MM1 RDE system (middle panel) in which the *gapA* gene is deleted to stop flux through the glycolysis pathway [24]. (c) In RDE2, expressing PRK as a neomycin phosphotransferase II (NPTII) fusion prevents false positive selection, as silencing PRK co-suppresses NPTII expression, resulting in kanamycin sensitivity [22].

Table 1. Influence of commonly selected cyanobacteria Rubisco mutants on Rubisco content and catalysis (relative to wild-type, WT)—deciphering the trait that improves fitness.

Amino Acid Mutation	^a Trait Selected	Rubisco (Solubility)	k_{cat}^C	K_C	K_C^{air}	k_{cat}^C/K_C	$k_{cat}^C/K_C^{air}S_{C/O}$	^b Selection Host	Reference		
SeRubisco (<i>Synechococcus</i> sp. PCC 6301)—Expressed in <i>E. coli</i> at ~1% (w/w) of the Cell Soluble Protein.											
F140I	solubility	^c WT	^c ↑ 46%	↓ 48%	-	↑ 288%	-	↓ 9%	MM1	[27]	
		↑ 4.9-fold	↓ 12%	-	↓ 16%	-	↑ 6%	WT	RDE2	[22]	
V189I	^d solubility	WT	↑ 8%	↓ 35%	-	↑ 71%	-	↓ 12%	MM1	[27]	
		↑ 4.8-fold	↓ 16%	-	↓ 26%	-	↑ 14%	↓ 15%	RDE2	[22]	
M262T	^d solubility	-	↓ 25%	↓ 44%	↓ 41%	↑ 32%	↑ 8%	-	<i>R. capsulatus</i>	[28]	
		↑ 6-fold	↑ 13%	↓ 13%	-	↑ 29%	-	WT	RDE(K12)	[29]	
		↑ 4-fold	-	-	-	-	-	-	MM1	[22]	
F345I	solubility	-	WT	↓ 23%	↓ 15%	↑ 23%	↑ 18%	-	<i>R. capsulatus</i>	[28]	
		↑ 8-fold	↓ 14%	↑ 4%	↑ 10%	↓ 17%	↓ 22%	↓ 13%	MM1	[30]	
		↑ 6-fold	↓ 15%	↑ 10%	-	↓ 21%	-	↓ 7%	MM1	[27]	
F345L	solubility	↑ 11.1-fold	↓ 16%	-	↑ 13%	-	↓ 28%	↓ 8%	RDE2	[22]	
		↑ 7-fold	↓ 17%	↑ 22%	-	↓ 32%	-	↓ 4%	MM1	[30]	
F345V	solubility	-	↓ 25%	↑ 38%	-	↓ 45%	-	↓ 45%	<i>R. capsulatus</i>	[31]	
Te-Rubisco (<i>Thermosynechococcus elongatus</i> BP-1)—Expressed in <i>E. coli</i> at ~6% (w/w) of the Soluble Protein.											
F345I	solubility	↑ 2.2-fold	↓ 27%	-	WT	-	↓ 24%	↓ 10%	RDE2	[22]	
P415A	solubility	↑ 1.6-fold	↑ 5%	-	WT	-	WT	WT	RDE2	[22]	
P415A	F345I	solubility/kinetic	↑ 2.7-fold	↓ 27%	-	WT	-	↓ 24%	↓ 8%	RDE2	[22]
P415A	V98M ⁵	solubility/kinetic	↑ 3.1-fold	↑ 44%	-	↓ 13%	-	↑ 44%	↑ 6%	RDE2	[22]

Table 1. Cont.

Amino Acid Mutation		^a Trait Selected	Rubisco (Solubility)	k_{cat}^C	K_C	K_C^{air}	k_{cat}^C/K_C	$k_{cat}^C/K_C^{air}S_{C/O}$	^b Selection Host	Reference
P415A	A48V ^S	kinetic	↑ 1.5-fold	↑ 33%	-	WT	-	↑ 35% ↓ 10%	RDE2	[22]
P415A	H37L ^S	kinetic	↑ 1.7-fold	↑ 27%	-	WT	-	↑ 28% ↓ 7%	RDE2	[22]
P415A	Y36N ^S /G112D ^S	kinetic	↑ 1.5-fold	↑ 41%	-	↓ 12%	-	↑ 60% ↓ 10%	RDE2	[22]
P415A	L74M/D397N	solubility/ kinetic	↑ 1.9-fold	↑ 24%	-	↓ 14%	-	↑ 45% = WT	RDE2	[22]

Amino acid numbering relative to plant RbcL (Figure A1). Data summarized from the Rubisco content and kinetic listings in Table A1. ^a The governing biochemical feature that improves RDE fitness (i.e., cell growth rate). ^b A directed evolution screen is used to isolate the mutant (see Figure 1). The *R. capsulatus* screen is a low transformation frequency system that uses a Rubisco null mutant [31]. ^c Underestimates of Rubisco content likely explain the erroneous carboxylation rate and efficiency improvements. ^d Improvements in carboxylase properties may contribute to the improved fitness that is primarily imparted by the >4-fold increases in Rubisco biogenesis (solubility). Enhanced carboxylase activities are shown in bold-type. Arrows indicate whether the kinetic value is higher (↑) or lower (↓).

The common Rubisco substrate used in directed evolution studies is that from the cyanobacterium *Synchococcus elongatus* PCC6301 (*Se*Rubisco, Table 1). When expressed in *E. coli*, the *Se*Rubisco RbcL and RbcS are produced in high abundance, with >90% of the RbcL produced forming misfolded insoluble aggregates, thereby limiting *Se*Rubisco expression to ~1% (*w/w*) of the soluble cellular protein [27,29,30,32]. Unassembled RbcS are less prone to insoluble aggregate formation, allowing them to rapidly bind (within <1 sec) to rbcL₈ cores to form stable RbcL₈RbcS₈ (L₈S₈) complexes [33]. The limitations to GroELS chaperonin facilitating *Se*-RbcL folding and assembly in *E. coli* have led to the repeated selection of particular amino acid changes (e.g., RbcL residues 140, 189, 262, 345; Figure A1) that improve *Se*Rubisco biogenesis by >4-fold (Table 1). The impact of these mutations on catalysis is that they mostly impair—or at best provide only modest changes in—the carboxylation properties *Se*Rubisco.

A recent, two-tiered directed evolution study used *Te*-Rubisco from *Thermosynechococcus elongatus* whose assembly requirements are better met in *E. coli* (produced at ~6% (*w/w*) of the soluble cellular protein [22]). Using the “false-positive free” RDE2 screen (Figure 1), *Te*-Rubisco was evolved in sequential stages, with the first identifying the catalytically neutral P415A mutation that improved *Te*-Rubisco assembly 1.6-fold. The second cycle of evolution then successfully selected additional mutations to P415A that were all co-located at the RbcL–RbcS interface of the holoenzyme that could enhance k_{cat}^C , $k_{cat}^C/K_C^{21\%O_2}$ and $S_{C/O}$ (Table 1).

In this study, we trial a re-designed version of the RDE2 screen (here, named RDE3) to evolve Form I *Rs*Rubisco from *Rhodobacter sphaeroides*. Unlike *Se*Rubisco and plant Rubisco (that are both in the “green-type” Form IB lineage), *Rs*Rubisco is in the “red-type” Form IC lineage and more related to the catalytically superior Form ID algae Rubisco from *G. monilis* [34,35]. Unlike the RDE2 screen, *Rs*Rubisco production in the RDE3 system is regulated by bacteriophage T7 transcription and undertaken in JM109(DE3), a *lacUV5*-regulated T7 RNA polymerase-expressing *E. coli* strain sensitive to PRK expression. Described is a model-directed evolution experimental pipeline using RDE3 that distinguished amino acid substitutions that impair *Rs*Rubisco biogenesis from one that significantly enhanced k_{cat}^C and k_{cat}^C/K_C .

2. Results

2.1. Ambiguity in How Cyanobacteria Rubisco Mutations Improve RDE Fitness

Prior RDE-directed evolution studies have mistakenly reported *Se*RbcL amino acid substitutions which improve the folding and assembly (solubility) of *Se*Rubisco as mutations that enhance catalysis. As summarized in Table 1, the commonly selected M262T *Se*RbcL substitution initially reported to enhance catalysis [36] was subsequently shown to enhance *Se*Rubisco solubility in *E. coli* [29] and may also account for the same mutation enhancing fitness in *R. capsulatus* selection [28]. Similarly the

SeRbcL F140L and V189I mutations were initially characterized as catalytic enhancers of RDE fitness [27] but subsequently shown to improve SeRubisco solubility by nearly 5-fold in *E. coli* [22].

The erroneous assessment of L₈S₈ recombinant holoenzyme production in *E. coli* often stems from the inherent complexity of the enzyme assembly requirements that impair its synthesis in heterologous hosts [24,25,37]. This is particularly true for the cyanobacteria Rubisco, whose assembly requirements are either not met, or are poorly met, in *E. coli* [29,32]. For example, the SeRbcL and SeRbcS subunits are produced in abundance in *E. coli*, with SeRbcL comprising >10% (*w/w*) of the total cellular protein (Figure 2a). Compared to SeRbcS that is mostly soluble, more than 90% of the SeRbcL produced in *E. coli* forms misassembled insoluble protein aggregates (compare the lysate and soluble SeRbcL and SeRbcS immunoblot signals in Figure 2b). The remaining soluble SeRbcL are primarily found assembled in SeL₈S₈ complexes, with the unassembled soluble SeRbcS monomers able to bind instantaneously with any assembled RbcL₈ cores via a chaperone-independent process [33]. By comparison, the assembly requirements of RsRubisco are more readily met in *E. coli*, with a greater proportion of soluble RsRbcL produced and assembled into RsL₈S₈ complexes (Figure 2c).

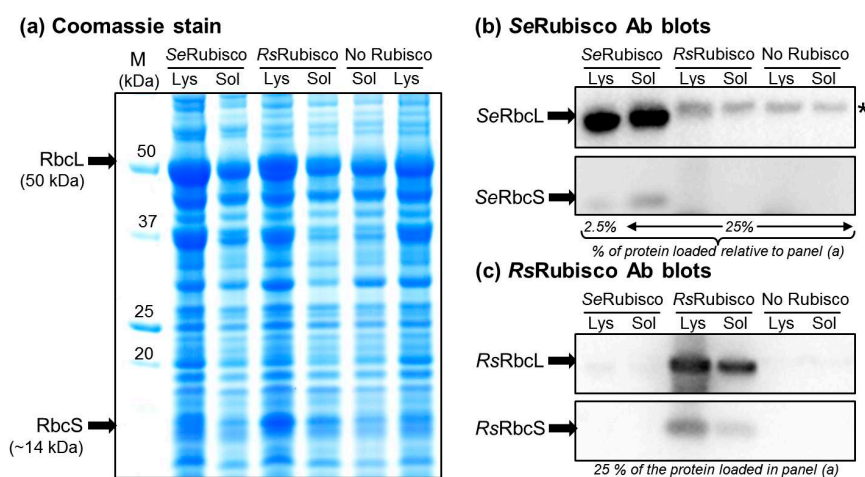


Figure 2. SDS PAGE analysis of SeRubisco and RsRubisco production in BL21(DE3). Rubisco production was induced at 25 °C with 0.5 mM isopropyl- β -D-thiogalactoside (IPTG) for 8 h, the cells were lysed and (a) 4 μ g of soluble protein (Sol) and an equivalent volume of the total cellular protein (Lys, comprising soluble, membrane and insoluble inclusion body proteins) were separated by SDS PAGE and Coomassie stained. Duplicate PAGE gels loaded with lower amounts of each protein preparation (as indicated) were blotted onto nitrocellulose membranes and probed with antibodies (Ab) to (b) SeRubisco and (c) RsRubisco. Empty vector BL21(DE3)-pET28a(+) cells were used as a “no Rubisco” control. *, *E. coli* protein recognized by the SeRubisco antibody. The SeRbcL is produced in high abundance and mostly insoluble (lysate signal \gg soluble signal), while SeRbcS, RsRbcL and RsRbcS are mostly soluble. SeRubisco and RsRubisco L₈S₈ contents comprised, respectively, 1.1% and 4.8% (*w/w*) of the soluble cellular protein (as quantified by ¹⁴C- carboxyarabinitol 1,5-bisphosphate (CABP)-binding). M, protein markers (sizes shown).

The low level of SeRubisco produced in *E. coli* (comprising ~1% (*w/w*) of the *E. coli*-soluble protein, Table 1) hinders the viability of SDS PAGE, Western blotting and native PAGE methods to quantify Rubisco content. Even when analyzed alongside known amounts of SeRubisco, these methods can lack sufficient quantifiable resolution [38]. By contrast, the tight and stoichiometric binding properties of the Rubisco specific inhibitor carboxyarabinitol 1,5-bisphosphate (CABP) provide an accurate method to quantify Rubisco in any soluble cellular protein sample [39]. The method, termed ¹⁴C-CABP binding, takes advantage of CABP’s capacity to tightly bind to each Rubisco catalytic site to form a highly stable complex that, upon separation from unbound ¹⁴C-CABP (e.g., by gel filtration), provides an accurate quantitative measure of Rubisco content [38].

2.2. An Analytical Pathway for the Directed Evolution of Rubisco Using an RDE Screen

Our routine Rubisco-directed evolution experimental protocol spanning RDE selection to in vitro biochemistry analytics is summarized in Figure 3 [22,40]. The experimental pipeline begins with the generation and transformation of an *rbcl* (+/*-rbcs*) mutant library into an RDE screen alongside RDE cells expressing the non-mutated *rbcl-rbcs* gene control (Figure 3a). Colonies showing improved RDE fitness (i.e., those which can grow on PRK-inducing arabinose and IPTG-inducing Rubisco concentrations that have been pre-determined to be non-permissive to the growth of cells expressing the parental Rubisco substrate) are individually selected, and the *rbcl-rbcs* plasmid is isolated, retransformed into RDE and the colony growth on increasing arabinose compared with the control cells to gauge the relative improvement in cell “fitness” (Figure 3b). The *rbcl-rbcs* plasmid is isolated and sequenced from the RDE cells with improved fitness and transformed into a suitable *E. coli* strain to quantify changes in Rubisco expression and catalysis (Figure 3c). SDS PAGE samples of the total and soluble cell protein are taken to qualitatively assess the proportion of soluble and insoluble RbcL and RbcS produced (Figure 2). The soluble protein concentration is measured, the Rubisco content is quantified by ¹⁴C-CABP binding, and 4 to 8 μg of protein is separated by native PAGE to verify the L₈S₈ Rubisco content and aliquots used in ¹⁴CO₂-fixation assays to quantify the CO₂-fixation rate (k_{cat}^C) and the K_m for CO₂ under anaerobic conditions (K_C), ambient O₂ levels ($K_C^{21\%O_2}$), or over a range of O₂ levels to quantify the K_m for O₂ (K_O) [22]. To accurately measure the specificity for CO₂ over O₂ ($S_{C/O}$), the method of Kane et al. (1994) [41] is used, which requires Rubisco that has been purified from the *E. coli* using a rapid ion exchange and gel filtration process [42].

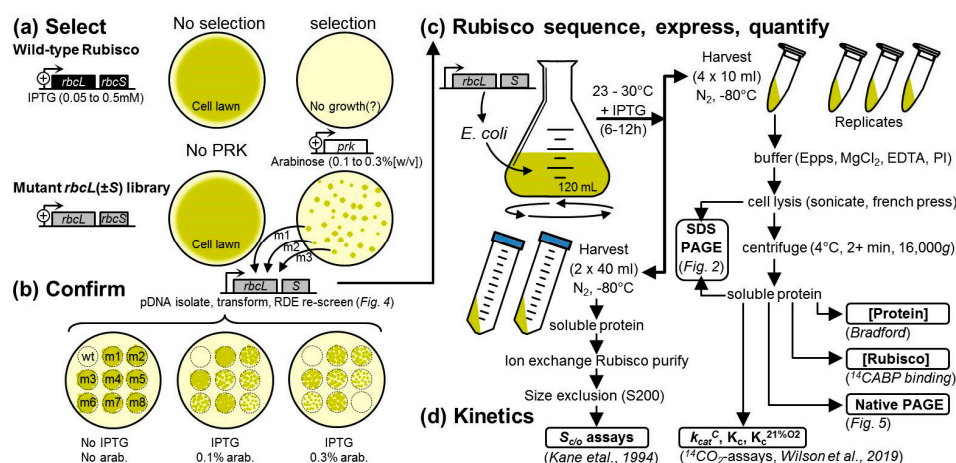


Figure 3. The analytical pipeline of an RDE experiment. An RDE-directed evolution project begins with (a) Rubisco gene library generation, transformation and RDE selection. The Rubisco gene plasmid (*rbcl-rbcs*) from cells with improved fitness is (b) transformed into fresh RDE cells and re-screened on PRK-inducing arabinose to confirm the improved growth phenotype before (c) sequencing to ascertain the amino acid substitutions. Each Rubisco mutant is then expressed in *E. coli* and the Rubisco subunit folding/assembly capacity evaluated by SDS PAGE (e.g., Figure 2); the L₈S₈ holoenzyme content is then quantified (¹⁴C-CABP binding), confirmed (native PAGE) and (d) the kinetics ($S_{C/O}$, k_{cat}^C , K_C and $K_C^{21\%O_2}$) measured as described in [22].

2.3. Determination of a Suitable *E. coli* Strain for RDE3

As in any directed protein evolution experiment, the level of success is dependent on the fidelity and throughput of the selection system employed. For evolving Rubisco, the use of the photosynthetic *R. capsulitis* bacterial screen is hampered by low throughput (screening is limited to a few thousand mutants) and hindered in its ability to characterize mutants with alterations in solubility [28,31,43,44]. As summarized in Figure 1, while RDE screens have higher throughput, their fidelity can be compromised by false positives. These occur at a higher frequency (~0.5% of

the colony forming units plated) in the original RDE screens than in the metabolically altered MM1 RDE strain [24]. The false positives produced in MM1 selection have primarily been attributed to transposon insertions into a pACYC^{RPK} plasmid that silenced PRK production [25]. This has been avoided in the RDE2 screen by fusing NPTII onto the PRK C-terminus and including kanamycin in the selection media, meaning that false positives originating from transposon-induced PRK silencing cannot grow (Figure 1) [22].

In this study, we sought to adapt the RDE2 screen so that Rubisco production was regulated by bacteriophage T7 transcription rather than the previously used *lac* promoter in pTrcHisB. This necessitated that RDE selection be undertaken in a *lacUV5*-regulated T7 RNA polymerase-expressing *E. coli* strain. The plating of differing *E. coli* strains transformed with pACYC^{prk:kan} on media containing differing arabinose concentrations found differences in their susceptibility to PRK-NPTII production (Figure 4). The JM109(DE3), XL1-Blue and DH5 α strains were sensitive to PRK-NPTII synthesis, while DB3.1, BL21(DE3) and BW25113 were not, making the latter three strains unsuitable for use in an RDE screen. The pACYC^{prk:kan} plasmids from DB3.1, BL21(DE3) and BW25113 colonies growing on 0.25% (*w/v*) arabinose were purified, restriction digested (Figure A2) and sequenced. There was no evidence that PRK-NPTII production was compromised, indicating that these three strains show an insensitivity to PRK production under the growth conditions trialed.

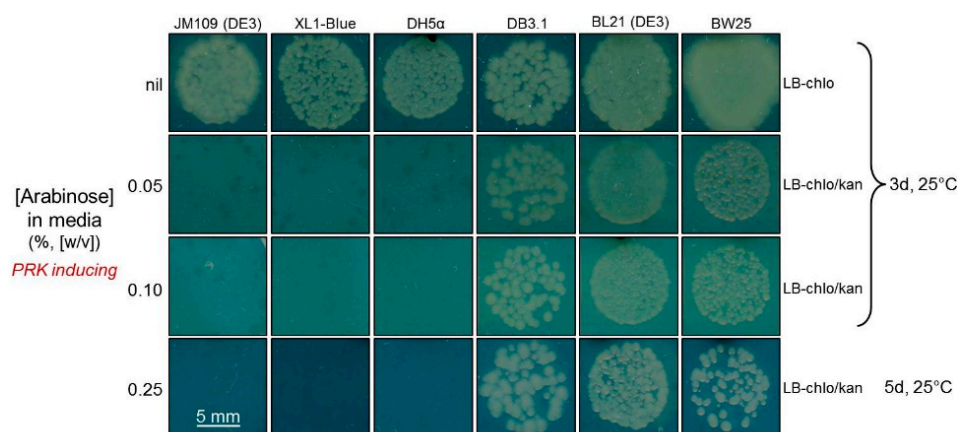


Figure 4. Comparative growth of differing *E. coli* strains under increasing PRK-NPTII expression. Six *E. coli* strains were transformed with pACYC^{prk:kan} and grown on LB-medium containing four different arabinose concentrations. Chloramphenicol (chlo, 30 $\mu\text{g}/\text{mL}^{-1}$) and kanamycin (kan, 125 $\mu\text{g}/\text{mL}^{-1}$) were included in the LB plates, as indicated.

As JM109(DE3) is suited to T7-regulated recombinant expression, it was the strain used in our new RDE3 screen. Comparative expression analyses found that *RsRubisco* was expressed equally well in JM109(DE3) (~5% (*w/w*) of the soluble protein, see Table 2) as in BL21(DE3) (Figure 2).

2.4. Directed Evolution of *RsRubisco* in the RDE3 Screen

In this pilot RDE3-directed evolution experiment, the Form IC *RsRubisco* was the chosen substrate as (1) its assembly requirements are well met in *E. coli* (Figure 2) including in JM109(DE3) (Table 2), (2) it is of the “red-type” Rubisco phylogenetic lineage that encompasses the superior Form ID Rubisco isoforms from red algae [35], and (3) it has not been used in a directed evolution study. Electro-competent JM109(DE3) containing pACYC^{prk:kan} [22] (here on termed RDE3 cells) were transformed with an *Rs-rbcl-rbcs* mutant library cloned into plasmid pYZRsLS (Amp^R, Figure 5a). Approximately 10^5 colony forming units (cfu) were plated on a 150 mm plate of selective LB media containing 75 μM IPTG, 0.25% (*w/v*) arabinose and antibiotics (Figure 5b). RDE3-pYZRsLS cells expressing non-mutated *RsRubisco* (2×10^4 cfu; wild-type (WT) control) were grown on a 90 mm plate of replica selection media. After five days at 25 °C in air supplemented with 2% (*v/v*) CO₂, 37 colonies had grown on the mutant *RsRubisco* plate and eight on the WT *RsRubisco* control (Figure 5b).

The pYZRsLS from 33 mutant and all eight WT colonies were purified and restriction-digested to confirm they contained the *Rs-rbcLS* operon before re-transforming into fresh RDE3 cells and plating again onto LB containing 0.25% (*w/v*) arabinose (Figure 5c). Only one mutant, called YZ1, enhanced the growth fitness of RDE3 under this selection pressure, with the remainder showing comparable arabinose sensitivities to the pYZRsLS(WT) control (Figure 5c).

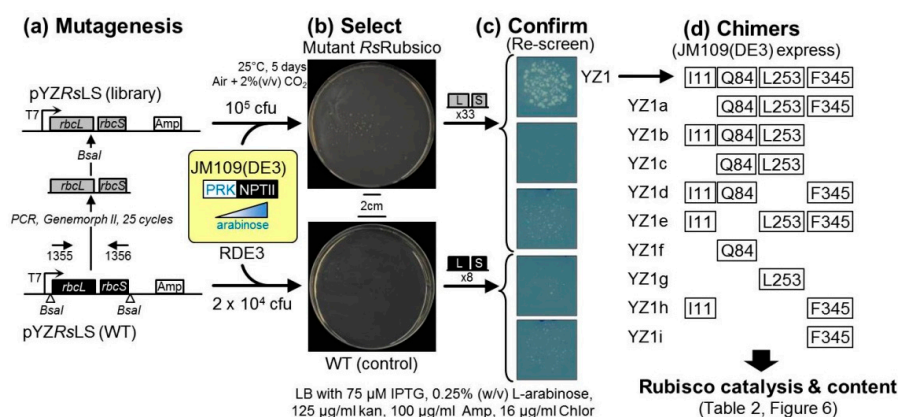


Figure 5. RDE3 evolution of *RsRubisco*. (a) Wild-type (WT) and a mutated *Rs-rbcLS* operon library cloned into pYZRsLS were transformed into RDE3 and (b) screened under strong arabinose selection as shown. The surviving colonies were grown (under non-selective conditions), the pYZRsLS purified and those containing the *Rs-rbcLS* operon re-transformed into RDE3 and (c) re-grown on selection media (examples for five of the 41 RDE3 re-screens are shown). Only the YZ1 mutant improved RDE3 growth, and (d) the sequencing of its *Rs-rbcLS* operon showed that it coded RbcL mutations V11L, K84Q, R253L and Y345F (plant RbcL numbering, Table A1). Nine YZ1 chimers (named YZ1a to YZ1i) coding differing combinations of these mutations were created to study how they influence *RsRubisco* biogenesis (Figure 6) and catalysis (Table 2).

Table 2. Comparative *E. coli* biogenesis (solubility) and catalysis (at 25 °C) of wild-type, YZ1 and their *RsRubisco* chimers.

<i>RsRubisco</i> Isoform	RbcL Residue Number				Rubisco ^a (% CSP)	<i>K_C</i> (μM)	<i>k_{cat}^C</i> (s ⁻¹)	<i>K_{cat}^C/K_C</i> (mM.s ⁻¹)	<i>S_{C/O}</i> mol.mol ⁻¹	<i>K_O</i> (μM)	<i>k_{cat}^O</i> (s ⁻¹)	<i>k_{cat}^O/K_O</i> (mM.s ⁻¹)
	11	83	252	345								
WT	Val	Lys	Arg	Tyr	4.9 ± 0.4 ^a	60 ± 1 ^a	3.7 ± 0.2 ^a	60 ± 1	58 ± 1 ^a	1724 ± 280	1.8	1.0
YZ1	Ile	Gln	Leu	Phe	2.9 ± 0.1 ^b	66 ± 1 ^a	4.7 ± 0.1^c	70 ± 1	57 ± 1 ^a	1296 ± 207	1.6	1.3
YZ1a	Val	Gln	Leu	Phe	3.3 ± 0.4 ^b	66 ± 1 ^a	4.5 ± 0.1^c	68 ± 2	-	-	-	-
YZ1b	Ile	Gln	Leu	Tyr	3.0 ± 0.1 ^b	61 ± 1 ^a	3.8 ± 0.1 ^a	63 ± 1	-	-	-	-
YZ1c	Val	Gln	Leu	Tyr	3.6 ± 0.4 ^b	63 ± 1 ^a	3.9 ± 0.1 ^{ab}	63 ± 1	-	-	-	-
YZ1d	Ile	Gln	Arg	Phe	4.6 ± 0.1 ^a	62 ± 1 ^a	4.6 ± 0.1^c	76 ± 1	-	-	-	-
YZ1e	Ile	Lys	Leu	Phe	4.5 ± 0.3 ^a	63 ± 1 ^a	4.3 ± 0.1^{bc}	70 ± 3	-	-	-	-
YZ1f	Val	Gln	Arg	Tyr	5.1 ± 0.4 ^a	61 ± 2 ^a	3.9 ± 0.2 ^{ab}	65 ± 1	57 ± 1 ^a	1220 ± 199	1.4	1.1
YZ1g	Val	Lys	Leu	Tyr	5.3 ± 0.2 ^a	62 ± 4 ^a	3.5 ± 0.1 ^a	57 ± 3	-	-	-	-
YZ1h	Ile	Lys	Arg	Phe	4.6 ± 0.1 ^a	61 ± 1 ^a	4.4 ± 0.1^a	73 ± 1	-	-	-	-
YZ1i	Val	Lys	Arg	Phe	4.7 ± 0.2 ^a	64 ± 1 ^a	4.6 ± 0.1^a	71 ± 2	59 ± 1 ^a	1551 ± 293	1.9	1.2

Values in bold represent significant improvements in carboxylase activity. CSP (percentage (*w/w*) of the cell soluble protein), *k_{cat}^C* and *K_C* data are the mean ± SE of *N* = 5 (and for *S_{C/O}* *N* = 2) biological samples, each assayed in duplicate. Letters show the ranking of the means using a post hoc Tukey test (different letters indicate statistical differences at the 5% level, *p* < 0.05) following a one-way ANOVA comparison relative to WT. *k_{cat}^O*, maximal oxygenation rate, calculated from *S_{C/O}* = (*k_{cat}^C/K_C*)/(*k_{cat}^C/K_O*). Mutated residues are shaded black.

2.5. The RbcL Substitution Y345F Enhances *RsRubisco* Carboxylation without Impairing Enzyme Biogenesis

The sequencing of the YZ1 pYZRsLS plasmid found it coded five *Rs-rbcL* nucleotide mutations, with four coding the *RsRbcL* mutations V11L (GTT-ATT), K85Q (AAA-CAA), R254L (CGT-CTT) and Y346F (TAT-TTT) and one a synonymous CTG-CTC change to L391. Relative to tobacco Rubisco, numbering these mutations corresponded to K83Q, R252L and Y345F (Figure A1a). A kinetic analysis of YZ1 *RsRubisco* found that the RbcL mutations had no effect on the *S_{C/O}* but significantly stimulated *k_{cat}^C* and carboxylation efficiency (*k_{cat}^C/K_C*) by 27% and 17%, respectively (Table 2). The mutations also increased the affinity of YZ1-Rubisco for O₂ (i.e., *K_O* was 25% lower than *RsRubisco*), thereby

reducing its CO₂ affinity under ambient O₂ ($K_C^{21\%O_2} = 78.8 \mu\text{M}$) relative to wild-type *RsRubisco* ($K_C^{21\%O_2} = 68.8 \mu\text{M}$). Nevertheless, the carboxylation efficiency of YZ1-*Rubisco* in air ($k_{cat}^C/K_C^{21\%O_2} = 59.6 \text{ mM}^{-1}\cdot\text{s}^{-1}$) still exceeded *RsRubisco* ($53.8 \text{ mM}^{-1}\cdot\text{s}^{-1}$) by 11%.

The analysis of YZ1-*RsRubisco* solubility by ¹⁴C-CABP binding and native PAGE showed its biogenesis in JM109(DE3) was ~40% lower than wild-type *RsRubisco* (Table 2, Figure 6). A series of nine YZ1 *RsRubisco* chimeras coding one to three RbcL amino acid substitutions was made (called YZ1a to YZi, Figure 5d) to examine whether the changes in biogenesis and kinetics could be uncoupled. As summarized in Table 2, all YZ1-chimera incorporating the Y345F mutation sustained the kinetics of YZ1 *RsRubisco*, demonstrating that this mutation imparted the improved k_{cat}^C and k_{cat}^C/K_C kinetic phenotype. The assembly of *RsRubisco* was only impaired in the YZ1-chimera coding both K83Q and R252L together (YZ1, YZ1a, YZ1b, YZ1c, Table 2, Figure 6). On their own (YZ1f, Yz1g) or in combination (YZ1c), the K83Q and R252L substitutions had no significant influence on *RsRubisco* kinetics (Table 2). The separation of all the YZ1 *RsRubisco* chimeras incorporating K83Q through native PAGE was marginally faster, suggesting that the mutation altered the net charge of the enzyme or induced a structural conformational change that increased its migration rate (Figure 6).

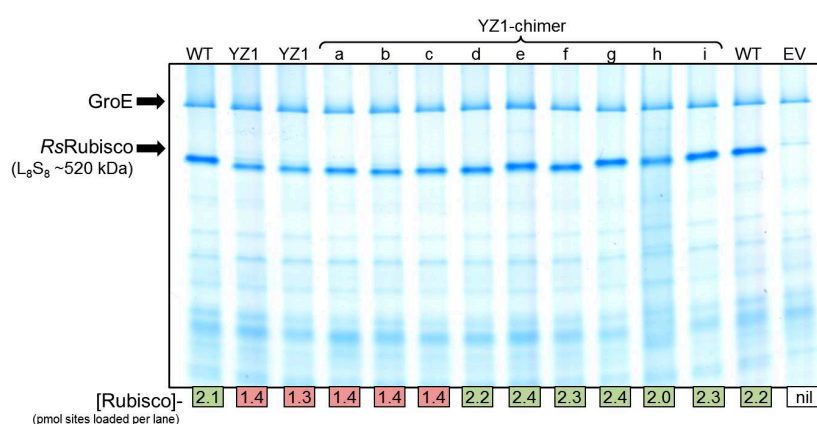


Figure 6. Native PAGE analysis of *RsRubisco* and the YZ1 chimeras' relative expression in JM109(DE3) after 0.5 mM IPTG induction for 6 h at 28 °C. The separated soluble cellular protein (4 μg, including an empty vector (EV) JM109(DE3)-pET28a(+)) was Coomassie stained, with the *RsRubisco* band intensity correlating with that quantified by ¹⁴C-CABP binding (indicated at base of gel). *RsRubisco* biogenesis for the YZ1 chimeras similar to wild-type (WT) are shaded green, while those produced in amounts equivalent to YZ1-*RsRubisco* are shaded red.

3. Discussion

A common challenge encountered with *Rubisco*-directed evolution studies using RDE selection is accurately measuring the biochemical properties of the mutant enzymes to correctly interpret the basis of how they improved RDE fitness [25]. Inaccuracies primarily arise from the erroneous evaluation of how amino acid substitutions affect *Rubisco* assembly (solubility) and kinetics (Table 1). Here, we demonstrate an experimental pipeline that spans the “selection” to the “analysis” phases of RDE-directed evolution (Figure 3). Using the pipeline, we successfully identified amino acids in Form IC “red-type” *RsRubisco* that influence either solubility or catalysis (Table 2). Our success in improving the carboxylation properties of *RsRubisco* extends to the comparable achievements made using MM1 and RDE2 in identifying amino acid substitutions that improve the carboxylase activities of the archaeon *Methanococcus burtonii* Form III *Rubisco* decamer (*rbcL*₁₀, [40]) and the Form IB “green-type” *L₈S₈* *TrRubisco* [22]. The chimeric dissection of the selected YZ1 *RsRubisco* mutant identified Y345F as a mutation that increases carboxylase activity and K84Q and R253L as substitutions that, in combination, impair *L₈S₈* biogenesis (Table 2). Our findings also show that sensitivity to PRK production varies between *E. coli* strains (Figure 4) and that RDE false positive selection can arise independently from *prk* silencing (Figure 5b,c and Figure A2). This raises questions

regarding the validity of the underpinning basis for RDE selection; i.e., that cell fitness is dependent on Rubisco expression to alleviate the inexplicable cellular toxicity of PRK-dependent RuBP production (Figure 1, [24]), a metabolic impediment that does not appear universal among all *E. coli* strains.

3.1. Mutating RbcL Amino Acid 345 Has Alternate Effects on Form I Rubisco Biogenesis and Catalysis

A common outcome of RDE screens is the selection of Rubisco mutants coding amino acid substitutions at residue 345 (numbered according to plant Rubisco, Table 1). Amino acid 345 is positioned in α -helix 6 of the six-stranded α/β -barrel C-domain of RbcL (as structurally defined by [45]). In the L_8S_8 structure residue, 345 is embedded beneath the enzymes' outer surface, away from subunit interfaces and ~ 5.6 Å from Arg-295, a conserved catalytic site residue involved in binding the P5 phosphate group of RuBP (Figure 7). Adjoining α -helix 6 is the catalytic loop 6 (residues $_{331}AVGKLE_{336}$, Figure A1). This conserved, flexible, "lid-like" sequence interacts with the C-terminal tail to close over the catalytic pocket upon RuBP binding, then opening upon catalysis to allow product release [10]. Among Form IB and ID Rubisco isoforms, position 345 is phenylalanine (Phe, F), and in cyanobacteria Form IA Rubisco and *RsRubisco* Form IC, it is a tyrosine (Tyr, Y) (Figure A1). A mutagenic analysis of Form IB *SeRubisco* found changes to α -helix 6 amino acids located closer to loop 6 dramatically impaired carboxylase activity [46]. Not included in this analysis were changes to the other α -helix 6 $G^{345}FVDL$ residues that are conserved among Form IB Rubisco. From this analysis, it was hypothesized that structural changes in α -helix 6 can influence loop 6 positioning and movement [46]. Such movements are thought to alter how K334 (the catalytically critical loop 6 apex residue) ligands with the RuBP enediolate transition state intermediate and influences its catalysis with CO_2 or O_2 .

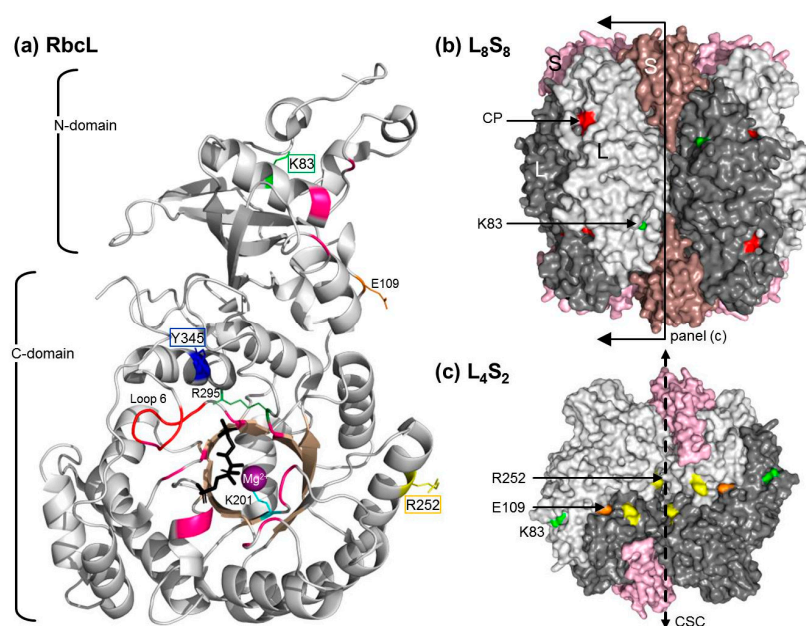


Figure 7. Mapping the location of the RDE3-selected YZ1 mutations in *R. sphaeroides* Rubisco. Pymol-assembled *RsRubisco* (PDB 5NV3) structures of (a) a single *RsRbcL* subunit, (b) L_8S_8 holoenzyme (side view) and (c) L_4S_2 sub-complex (cross-section view). RbcL pairs are colored with differing shades of grey, and RbcS is alternately colored pink/salmon. Shown are the locations of the YZ1-selected RbcL amino acid substitutions K83Q (green, strand C), R252L (yellow, α -helix 3) and Y345F (blue, α -helix 6) relative to the conserved catalytic pocket (CP) residues (dark pink), including R295 (green), K201 (light blue) loop 6 (red) and CABP (black, RuBP intermediary analog). E109 (orange) forms a salt bridge with a conserved R253 in the adjacent RbcL. The R252L substitution is located on the surface of the central solvent channel (CSC) that traverses the inner core of the L_8S_8 holoenzyme (dashed arrow, panel c). Residues are numbered relative to plant Rubisco using the alignment in Figure A1.

The existing data suggest that changes to residue 345 have a pervasive influence on Rubisco catalysis. For example, replacing Phe 345 with Ile, Leu or Val impairs the carboxylation properties of *Se*Rubisco [27,30] and *Te*Rubisco [22] (Table 1). It is difficult to structurally ascertain whether these kinetic changes arise from the mutations altering the positioning and movement of loop 6 or by altering RuBP ligand binding with Arg-295 in the catalytic pocket. Inexplicably, mutations to F345 also somehow impart two to 11-fold improvements in *Te*Rubisco and *Se*Rubisco production (Table 1). It is presumed that these changes enhance *Se*RbcL and *Te*RbcL folding and/or assembly in *E. coli* [22,24,27]. By contrast, the Y345F substitution in *Rs*RbcL had no effect on *Rs*Rubisco biogenesis in *E. coli* and improved its carboxylation rate and efficiency by 27 and 17%, respectively (Table 2). This incongruence in how changes to amino acid 345 differentially influence the catalysis and assembly of Form IB and IC Rubiscos underpins the challenge of using rationale design approaches to identify “all-purpose” amino acid substitutions that can benefit catalysis between, and within, differing Rubisco lineages [1,13,22,47].

3.2. How Do the K83Q and R252L Substitutions Function Cooperatively to Reduce *Rs*Rubisco Biogenesis?

K83 and R252 in *Rs*Rubisco are located, respectively, in strand C of the RbcL N-domain and α -helix 3 of the α/β -barrel (Figure 7a). Their distant locations from the catalytic sites are in accordance with our findings that the mutations K83Q and R252L, both singly (*Rs*Rubisco chimeras YZ1f and YZ1g) and in combination (YZ1c), have no significant effect on catalysis (Table 2). *Rs*Rubisco biogenesis was not affected in YZ1 chimeras containing only one of these mutations (chimeras YZ1d to g), but was reduced by ~40% when both were present (YZ1 and YZ1a to c, Table 2). How the K83Q and R252L substitutions work in combination, but not in isolation, to reduce the level of *Rs*Rubisco is unclear. It is possible that the mutations reduce *Rs*RbcL compatibility with one or more components of the *E. coli* protein folding machinery that impairs holoenzyme assembly or alters the stability of the L_8S_8 complex. As summarized in Table 1, a number of single amino acid mutations are known to enhance cyanobacteria Rubisco biogenesis in *E. coli*. For *Te*Rubisco, the solubility-enhancing F345I and P415A mutations were able to function in a cumulative fashion to further augment solubility (Table 1). It is proposed that these improvements arose by improving the compatibility of RbcL with components of the *E. coli* protein folding machinery [22]. It is known that chaperonin (GroELS) availability, but not DnaK, DnaJ, GrpE or Trigger Factor, has a pervasive influence on *Se*Rubisco biogenesis in *E. coli* [29,32]. Sequence complementarity with *E. coli* GroELS has been used to explain the vast difference in the assembly capabilities of differing cyanobacteria Rubisco (e.g., PCC7002 Rubisco is produced at <0.5% CSP [32], *Se*Rubisco at ~1.3% CSP and *Te*Rubisco at ~7% CSP [22,48]) and with *Rs*Rubisco (5% CSP, Table 2). It is therefore possible that the K84Q and R253L mutations in combination reduce *Rs*RbcL compatibility with GroELS to impair *Rs*Rubisco biogenesis.

The evaluation of the *Rs*Rubisco L_8S_8 crystal structure (Figure 7b) showed that R252 is located near the surface of the inner solvent channel that traverses the L_8S_8 structure and is situated at the edge of the interface of adjoining large subunits that are arranged head to tail in each L_2 unit (Figure 7c). The adjacent R253 residue is conserved among Form I Rubisco and forms a conserved salt bridge with E109 in the adjoined RbcL of each L_2 (Figure 7c). While this intradimer interaction is hypothesized to maintain L_2 's structural integrity [49], our findings suggest the R252L substitution does not disturb this Arg–Glu salt bridge as *Rs*Rubisco synthesis and stability seems unaffected (Table 2). In contrast to R252, K83 is a solvent-exposed outer surface residue and independent of any inter-subunit interactions (Figure 7b). Curiously, the YZ1 chimeras incorporating K83Q showed a faster mobility through native PAGE, independent of other amino acid substitutions (Figure 6), suggesting the mutation alters the net surface charge of the protein or influences *Rs*Rubisco quaternary structure, but without an effect on catalysis (Table 2). The disparate locations of the K83R and R252L mutations in the *Rs*Rubisco do not support the hypothesis that they jointly function to reduce *Rs*Rubisco content in *E. coli* through destabilization of the L_8S_8 complex.

3.3. Understanding the Basis of RDE Selection

The adage that “you get what you screen for” highlights how crucial it is to ensure any directed evolution screen can select for a desired fitness trait with high fidelity [50,51]. The production of false positives commonly compromises the fidelity of RDE screens, especially those using PRK expressing wild-type *E. coli* (Figure 1a). In the MM1 *gapA*⁻ mutant RDE screen, the impaired cell viability associated with blocking glycolysis/gluconeogenesis by deleting glyceraldehyde-3-phosphate dehydrogenase (Figure 1b) reduces the frequency of false positives [24,26] which primarily occur from the transposon-interrupted expression of PRK [25]. The selection of these false positives in the XL1-Blue-based RDE2 screen was prevented using PRK:NPTII fusion so that PRK production and kanamycin resistance are linked, thereby preventing the selection of PRK-silenced mutants (Figure 1c). While initial tests indicated the suitability of the JM109(DE3) strain for RDE3 selection (Figure 4), a false positive production rate of approximately 0.04% was found when more cells were grown on non-permissive, high arabinose LB media (Figure 5b). Curiously, no sequence changes were evident in the purified pACYC^{prk:kan} or the pYZRsLS from the eight WT control colonies that grew. A comparable frequency of false positives was also identified in the RDE3-mutant *RsRubisco* screen (Figure 5b), in which again the sequence integrity of pACYC^{prk:kan} was maintained, and only one (YZ1) of the 33 mutant pYZRsLS plasmids re-transformed into RDE3 survived re-selection on high arabinose. The growth of the false positives in the initial RDE3 screens therefore arose independently of changes to either PRK:NPTII or *RsRubisco* expression. This suggests the false positives may result from chromosomal mutations that alter JM109(DE3) cellular metabolism to dampen its sensitivity to PRK production. This hypothesis is consistent with our finding that other *E. coli* strains (e.g., DB3.1, BL21(DE3) and BW25113) show resistance to PRK:NPTII production (Figure 4). Current research is underway to identify the genetic and/or metabolic basis for the contrasting sensitivities of *E. coli* strains to PRK production as a means to improve the fidelity of RDE screens.

3.4. Pathways for Evolving *RsRubisco* Further—Why and How?

Our interest in continuing to explore the evolutionary potential of *RsRubisco* stems from its relatedness to the Form ID “red-Rubisco” lineage that comprises red algae Rubisco isoforms whose high S_{CO} and carboxylation efficiencies dramatically exceed those of crop plants [16,18]. While optimizing the fidelity of the RDE3 to avoid false positive selection is desirable, we found it is not critical if suitable care is taken to re-screen selected *RsRubisco* mutants (i.e., confirming that Rubisco is improving RDE3 fitness, Figure 5c) before pursuing Rubisco biochemical analyses. As the carboxylase improved YZ1i (Y345F) and biogenesis-impaired YZ1c (R83Q, R252L) mutants offer alternative starting points in the *RsRubisco* evolutionary landscape, separately evolving them using RDE3 and comparing the biochemical fitness traits selected (solubility and/or kinetic) would provide valuable strategical guidance to successive rounds of evolution.

Incorporating the *R. sphaeroides* chaperonin folding machinery components into RDE3 would alter—and possibly benefit—the *RsRubisco* evolutionary fitness landscape that can be surveyed. A similar approach taken with *SeRubisco* found that the inclusion of the Rubisco-specific post-chaperonin assembly chaperone RbcX altered the fitness landscape of the MM1 screen, with different mutants selected compared with a screen lacking RbcX [27]. This supported an earlier hypothesis that Rubisco evolution is influenced—and possibly even constrained—by the need to maintain complementarity with the protein folding/assembly machinery of a cell [24]. This appears particularly apparent for Form IB plant Rubiscos, whose assembly in *E. coli* requires five compatible molecular partners (i.e., the Rubisco-specific chaperones Raf1, RbcX, Raf2, BSD2 and Cpn60 $\alpha\beta$ 20 chaperonin cages [37,52,53]). Thus far, maintaining sequence compatibility between Raf1 and RbcL appears to be the most important factor influencing plant Rubisco biogenesis [54]. Our findings that the R83Q and R252L substitutions inexplicably impair *RsRubisco* biogenesis in *E. coli* suggest sequence compatibility between RbcL and that the bacterial GroELS protein folding machinery also impacts Rubisco biogenesis potential.

4. Materials and Methods

4.1. Comparing the PRK Sensitivity of Different *E. coli* Strains

The pACYC^{prk:kan} plasmid [22] was transformed into CaCl₂-competent JM109(DE3), XL1-B, DH5 α , DB3.1, BL21(DE3), and BW25113 *E. coli* cells, and a colony of each was grown in LB-Chlor (30 $\mu\text{g}\cdot\text{mL}^{-1}$) at 25 °C. At an absorbance (A_{600}) of 0.6, the cultures were diluted by 10⁻⁴ and 20 μL was plated onto LB-Chlor plates containing 0, 0.05%, 0.1%, 0.25% (*w/v*) L-arabinose with/without 100 $\mu\text{g}\cdot\text{mL}^{-1}$ kanamycin. Colony growth at 25 °C in air containing 2% (*v/v*) CO₂ was monitored up to 5 days. Electro-competent RDE3 cells (i.e., pACYC^{prk:kan} transformed JM109(DE3)) in 10% (*v/v*) glycerol were N₂ frozen and stored at -80 °C.

4.2. Library Construction and Rubisco Selection Using the RDE Screening

Figure 3 provides a general summary of an RDE-directed evolution experiment. The pilot RDE3 *RsRubisco* evolution experiment undertaken in this study is summarized in Figure 5. The 1965-bp synthetic *Rs-rbcLS* operon in which the codon use was modified to match the tobacco *rbcL* gene (Genbank Accession KM464722) was PCR-amplified using primer 1355 (5'-ATCGAGGTCTCGCCATATGGCACCACAAACAGAGAC-3', *Bsal* site underlined) and primer 1356 (5'-GCATGGTCTCCAAGCATCTCGAGCTCAGATCTGTC-3', *Bsal* site underlined) and cloned into a modified pTriEx-2 Amp plasmid (Novagen, Addgene, Cambridge, Massachusetts, USA) to give plasmid pYZRsLS (Genbank Accession MN541175, Figure 5a). The 13 N-terminal amino acids of *RsRbcL* coded in pYZRsLS have been replaced with the 12 N-terminal amino acids from the tobacco *RbcL* (Figure A1). In pYZRsLS the *Bsal* sites flanking the *Rs-rbcLS* operon are positioned downstream of the T7 promoter and adjacent to the T7 terminator. A mutant *Rs-rbcLS* library was amplified from pYZRsLS with primers 1355 and 1356 using the Genemorph II Random Mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA) over 25 cycles as per the manufacturer's recommendations. The amplified *Rs-rbcLS* mutant library was digested with *Bsal*, ligated into *Bsal* cut pYZRsLS and electroporated (1.8 kV pulse) into 25 μL RDE3 cells. RDE3 was also transformed with non-mutated pYZRsLS as the wild-type (WT) control. The cells were added to 1 mL LB, grown for 1 h at 37 °C, 5 μL diluted and plated on non-selective media (LB-chlor-amp; 16 $\mu\text{g}/\text{mL}$ chloramphenicol and 100 $\mu\text{g}/\text{mL}$ ampicillin) to determine the colony forming unit (cfu) size of the WT and mutant *RsRubisco* library. The mutant pYZRsLS plasmids from 10 colonies grown on permissive media were transformed into XL1-Blue (LB-amp selected), purified and fully sequenced. The *Rs-rbcLS* mutational frequency was 2.4 mutations per kb.

The remaining transformed RDE cells were pelleted (6000 \times g, 1 min), suspended in 0.2 mL LB and plated onto LB-arabinose selection (LB-chlor-amp plates containing 125 $\mu\text{g}/\text{mL}$ kanamycin, 75 μM IPTG and 0.25% (*w/v*) arabinose, Figure 5b). The RDE3-WT and RDE3-mutant *RsRubisco* library plates were grown at 25 °C for 3–7 days in air containing 2% (*v/v*) CO₂. All colonies from both plates were colony picked and grown in 1 mL LB-chlor-amp, their plasmids were purified and the pYZRsLS re-transformed into XL1-Blue (LB-Amp selection), and then purified for *NdeI* mapping (expect 3822, 2177 and 1020-bp products) before transforming into RDE3 for the re-screening of selection fitness.

Colonies of the RDE3-rescreened cells were grown to an OD₆₀₀ of 0.6 in 1 mL LB-chlor-amp, diluted 10⁻⁴ with LB and 20 μL of cells plated on LB-arabinose selection (Figure 5c). Of the 41 plasmids re-screened (including eight from the WT screen, Figure 5b) only one, the pYZ1 mutant, improved RDE3 growth on the LB-arabinose (0.25% *w/v*, Figure 5c). The *Rs-rbcLS* operon in pYZ1 contained five nucleotide substitutions that coded four *RbcL* amino acid changes, from which nine YZ1 *RsRubisco* derivatives were generated by restriction cloning so that each chimera only coded one to three of each mutation (Figure 5d).

4.3. Rubisco Content and Catalysis

JM109(DE3) transformed with pYZRsLS and each pYZ1 derivative were grown at 28 °C and *RsRubisco* expression induced for 6 h with 0.5 mM IPTG before harvesting by centrifugation (10 min,

4 °C, 6200× g) and replica cell pellet aliquots N₂ frozen and stored at −80 °C (Figure 3c). The soluble and total (lysate) proteins in cell pellets from 10 mL of cultured cells were analyzed by SDS-PAGE, native PAGE and immunoblot analysis as described [55]. The Rubisco content in the *E. coli* soluble protein was quantified by ¹⁴C-CABP binding as described [38]. The same soluble protein sample was used to measure k_{cat}^C and K_C using ¹⁴CO₂-fixation assays as described [22] using 0–170 μM ¹⁴CO₂ in assays equilibrated with 0, 20%, 35% or 50% O₂ (v/v) in N₂. The increase in K_C with increasing [O₂] was plotted and the slope of the linear fit used to derive K_O . S_{CO} was quantified using the [1-³H]-RuBP consumption assay as described [41] using *RsRubisco* purified from ~40 mL of cultured cells (Figure 3c) by rapid ion exchange and size exclusion (Superdex 200, GE Life Science, Chicago, IL, USA) as described [42]. Soluble protein concentration was measured using the dye-binding Coomassie Plus Kit (Thermo Fisher Scientific, Waltham, MA, USA) against bovine serum albumin.

5. Conclusions

This study provides another demonstration of how directed evolution can improve the carboxylase properties of Rubisco. As shown here, key challenges with RDE screens include differentiating the Rubisco mutants that enhance cell fitness from false positives and then distinguishing carboxylation enhancing mutations from those that improve Rubisco biogenesis. Success is also dependent on tailoring selection screen conditions to the substrate Rubisco being evolved; that is, determining the optimal arabinose and IPTG concentrations for PRK and Rubisco production, keeping in mind that Rubisco production itself impairs *E. coli* viability [25]. An ongoing consideration is the extent to which the red-type *RsRubisco* might serve as a suitable surrogate to sequentially evolve its properties towards the high S_{CO} , high carboxylation efficiency properties of the related “better” red algae Rubisco. One possible objective might be to replace plant Rubisco with a kinetically improved *RsRubisco*, although this will depend on whether its folding and assembly requirements are met by plant chloroplasts. Current evidence suggests the extent of prokaryotic Rubisco expression in *E. coli* is a poor indicator of its biogenesis potential in tobacco [20–22]. An alternative to evolving *RsRubisco* might be to evolve plant Rubisco itself using an RDE screen incorporating the chloroplast protein folding machinery [53,56]. As the components of the plant Rubisco expression systems are primarily T7-promoter regulated [37,57], our JM109(DE3)-based RDE3 screen appears to be a suitable selection host for such a directed evolution endeavor.

Author Contributions: Y.Z. and S.W. contributed equally to the conceptualization, formal analysis, and writing of the article. Funding was awarded to S.W.

Funding: This research was supported by the Australian Government through the Australian Research Council Centre of Excellence for Translational Photosynthesis CE140100015.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

Raf1	Rubisco accumulation factor 1
Raf2	Rubisco accumulation factor 2
BSD2	Bundle Sheath Defective II Rubisco chaperone
RbcX	Rubisco specific chaperone
cfu	Colony forming units
RuBP	Ribulose-1,5-bisphosphate
S_{CO}	Specificity for CO ₂ over O ₂
k_{cat}^C	Substrate saturated carboxylation rate
k_{cat}^O	Substrate saturated oxygenation rate
K_C	Apparent K_m for CO ₂
$K_C^{21\%O_2}$	Apparent K_m for CO ₂ under ambient (20.6% (v/v)) O ₂
K_O	Apparent K_m for O ₂
PRK	Phosphoribulokinase

Appendix A

(a) RbcL alignment

11

1 - M S P Q T E T K A S V G F K A G V K E Y K L T - Y Y T P E Y Q T K D T D I L A tob-L
1 - . P K T Q S - - A A . Y D - D . T P L . . Se-RbcL
1 - . A Y T Q S K S Q K Y Q D . R - D . T P Te-RbcL
1 M D T K T . . I . G K E R Y L K . A Q M G . W D G D . V P V . . Rs-RbcL
1 - M S P Q T E T K A S V G F L K . A Q M G . W D G D . V P V . . cmRs-RbcL

60 65

39 A F R V T P Q P G V P P E E A G A A V A A E S S T G T W T T V W T D G L T S L D tob-L
36 . . . F S A D I L D M . Se-RbcL
39 F A L D . Te-RbcL
41 L . . I . . . E . . D . V . . A G A . . . V R . . A C . Rs-RbcL

83 109

79 R Y K G R C Y R I E R V V G E K D Q Y I A Y V A Y P L D L F E E G S V T N M F T tob-L
76 K . . H . . P . Q . . E N S . F . F I I L . Se-RbcL
79 C . . D . . P L P . . D N . F I L . Te-RbcL
81 S . R A K A . . V . P . P . T P G . . F C D . I I A . L T A Rs-RbcL

123 140

119 S I V G N V F G F K A L R A L R L E D L R T P P A Y V K T F Q G P P H G I Q V E tob-L
116 I . S I . F . V . L Se-RbcL
119 K V . L Te-RbcL
121 . . I . . S . . P . K . A M . F . V Y K T . . V G . Rs-RbcL

175 177 189

159 R D K L N K Y G R P L L G C T I K P K L G L S A K N Y G R A V Y E C L R G G L D tob-L
156 . . L M . Se-RbcL
159 . Te-RbcL
161 . E R . D . F . K A . T G V G . K Rs-RbcL

201 203204

199 F T K D D E N V N S Q P F M R W R D R F L F C A E A L Y K A Q A E T G E I K G H tob-L
196 I Q V . D . I H . S Se-RbcL
199 I Q V . D . I H Te-RbcL
201 . M I H Y V M . . V N L . S . Q . . . V . . Rs-RbcL

252 262

239 Y L N A T A G T C E E M I K R A V F A R E L G V P I V M H D Y L T G G F T A N T tob-L
236 . . . V . . P M E . K M I F . . A . . Se-RbcL
239 . . . V . . P L E . K M I F . . A . . Te-RbcL
241 . . . I M Y R E . K S . . . S V . . . V . L I I . - Y . . I Q Rs-RbcL

294 295 300

279 S L A H Y C R D N G L L L H I H R A M H A V I D R Q K N H G I H F R V L A K A L tob-L
276 T . . K W V . R C . Se-RbcL
279 T . S K W M . M C . Te-RbcL
280 . I S E W . . Q . D M I . . M G . G T Y T S . . . I . . W . Rs-RbcL

327 334 345

319 R M S G G D H I H S G T V V G K L E G E R D I T L G F V D L L R D D F V E Q D R tob-L
316 L . D K A S M . E . H I . A . Se-RbcL
319 T D K A V ENY I . . . Te-RbcL
320 . L A . V . . L . C . . . A D P L T V Q . Y Y N V C . E P . N T V . L Rs-RbcL

379 381

359 S R G I Y F T Q D W V S L P G V L P V A S G G I H V W H M P A L T E I F G D D S tob-L
356 . V F A . M . V Se-RbcL
359 A . M M A V D A Te-RbcL
360 P . . . F . E . . . A D . R K . M A G Q . H Q . L S L V Rs-RbcL

403 404 415

399 V L Q F G G G T L G H P W G N A P G A V A N R V A L E A C V K A R N E G R D L A tob-L
396 . T Q Y Se-RbcL
399 . T I Q M Te-RbcL
400 I . . . M . I Q A . . T M . L N I D Rs-RbcL

439 Q E G N E I I R E A C K W S P E L A A A C E V W K E I V F N F A A V D V L D K . tob-L
436 R . . G D . L . . . G L D L K . E . E T M . K . Se-RbcL
439 R . . G D A R L K . E . E . Q . T I Te-RbcL
440 V . . P . . L . A . A . . C K P . E . . L D T . G N . T . . Y T S T . T S . F V Rs-RbcL

480 P T A S V A M Rs-RbcL

Figure A1. Cont.

(b) RbcS alignment

1	- - M Q V W P P I N K K K Y E T L S Y L P D L S Q E Q L L S E V E Y L L K N G W	tob-S
1	M S . K T L . - - K E R R F . . F . . . P . . D R . I A A Q I . . M I E Q . F	SeRbcS
1	- - . K T L . - - K E R R . . F . . . P . . D A . I A R Q I Q . A I D Q . Y	TeRbcS
1	- - . R I T Q - - - - - G C F . F . . . T D A . I S A Q . D . C . G R . .	RsRbcS
39	V P C L E F E T E H G F V Y R E N N K S P G Y Y D G R Y W T M W K L P M F G C T	tob-S
39	H . L I . . N - - - - - . H S N P - - - - E E F L . D . K	SeRbcS
37	H . . V . . N - - - - - . T S N A - - - - E I L . N . .	TeRbcS
32	A V S . . H T - - - - - D D P H P - - - - R N T . . E . . G M . . . D L R	RsRbcS
79	D A T Q V L A E V E E A K K A Y P Q A W I R I I G F D N V R - - Q V Q C I S F I	tob-S
67	S P Q . . . D . . R . C R S E . G D C Y . . V A I K - - . C . T V . . .	SeRbcS
65	N . Q D . . N . . Q Q C R S E . . N C F . . V V A . . . I K - - . C . V M . . .	TeRbcS
60	. P K G . M I . L D . C R . . W . G R Y . . . N A . . S T . G F E T V T M . . .	RsRbcS
117	A Y K P - E G Y	tob-S
105	V H R . - - . R Y	SeRbcS
103	V . . N Q A N S G Y S G Y R Y	TeRbcS
100	V N R . E V E P S L R M E R T E V D G R S I R Y T H S I V R	RsRbcS

- RDE & MM1 selected *Se*Rubisco mutations (see Table A1).
- RDE2 selected *Te*Rubisco mutations (see Table A1).
- RDE3 selected *Rs*Rubisco mutations (see Table 2).
- Residues selected in multiple RDE screens
- Conserved catalytic pocket residues (including R295 ■)
- Critical catalytic K201
- E109 that forms a salt bridge with R253 in adjoining RbcL

Figure A1. Alignments of (a) RbcL and (b) RbcS for Rubisco from *Nicotiana tabacum* (tob-L, Z00044; tob-S, XP_016440853), *Synechococcus elongatus* PCC6301 (*Se*RbcL, P00880.2; *Se*RbcS P04716), *Thermosynechococcus elongatus* BP1 (*Te*RbcL, NP_682296.1; *Te*RbcS NP_682294.1) and *Rhodobacter sphaeroides* (*Rs*RbcL, WP_002721829.1; *Rs*RbcS WP_002721826.1). The first 39 amino acids of the altered codon modified (cm) *Rs*RbcL N-terminus coded in pYZRsLS are shown in blue text. Sequences were Clustal W-aligned, with only residues differing from the tob sequences shown. Dashes were included to improve the alignment of any missing residues. Amino acid numbering is relative to tob-L and tob-S. Coloring is as indicated in the legend under panel (b).

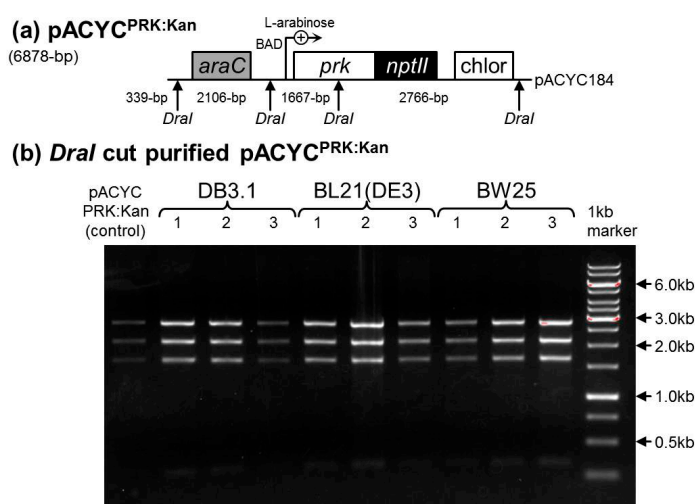


Figure A2. Restriction mapping of pACYC^{PRK:Kan}. (a) The plasmid map and (b) *DraI* restriction mapping of the source pACYC^{PRK:Kan} plasmid (control) and that purified from three separate colonies of the *E. coli* strains indicate that they survived growth on selective media containing 0.25% (*w/v*) L-Arabinose (Figure 4).

Table A1. Summary of directed evolution Rubisco mutants and the content and catalytic properties selected using *E. coli* and *Ralstonia eutrophus* screens.

Rubisco Origin	Mutation (s)	Plant RbcL Numbering	K_C (μM)	$K_C^{21\%O_2}$ (μM)	k_{cat}^C (s^{-1})	k_{cat}^C/K_C ($\text{s}^{-1}\cdot\text{mM}^{-1}$)	$k_{cat}^C/K_C^{21\%O_2}$ ($\text{s}^{-1}\cdot\text{mM}^{-1}$)	K_O (μM)	k_{cat}^O (s^{-1})	k_{cat}^O/K_O	$S_{C/O}$	K_{RuBP} (μM)	Rubisco (%CSP)	RDE Strain
<i>Synchococcus</i> PCC6301 (L ₈ S ₈)	WT	-	273	-	11.4	42	-	-	-	-	44.0	64	1.0	<i>E. coli</i> K12 Ref [29]
	M259T	262	237	-	12.8	54	-	-	-	-	43.0	67	6.0	
	A8S/M259T	10/262	216	-	11.4	53	-	-	-	-	41.0	48	5.5	
	F342S/M259T	345/262	207	-	10.0	48	-	-	-	-	38.0	34	3.5	
	WT	-	200	242	11.8	59	48.7	1191	1.7	1.42	41.5	44	0.2	
<i>Synchococcus</i> PCC6301 (L ₈ S ₈)	F345I	345	208	266	10.1	49	38.0	908	1.2	1.33	36.6	31	3.4	MM1 Ref [30]
	I174V	174	233	283	10.2	44	36.0	1174	1.2	1.02	42.9	33	1.1	
	F345L	345	245	-	9.8	40	-	-	-	-	40.0	93	1.9	
	Q212L	212	370	444	9.9	27	22.3	1252	0.9	0.72	37.4	31	1.7	
	L348M	348	287	-	11.3	39	-	-	-	-	35.1	84	-	
	N115S	115	191	-	9.0	47	-	-	-	-	43.5	44	-	
	A272G	272	244	-	9.7	40	-	-	-	-	40.2	34	-	
	K339R	339	148	-	10.1	68	-	-	-	-	43.0	35	-	
	M262T	262	-	-	-	-	-	-	-	-	-	-	0.9	
	M169L	169	-	-	-	-	-	-	-	-	-	-	0.8	
	L161M	161	-	-	-	-	-	-	-	-	-	-	1.1	
	<i>Synech.</i> PCC7002 (L ₈ S ₈)	WT	-	132	-	6.7	51	-	-	-	-	-	52	
E49V ^S /D82G ^S		E51V ^S /D96G ^S	200	-	14.6	74	-	-	-	-	-	50	-	
<i>Synchococcus</i> PCC6301 (L ₈ S ₈)	WT	-	248	-	12.9	52	-	-	-	0.9	56.1	-	-	MM1 Ref [27]
	F140I	140	129	-	19.4	150	-	-	-	2.9	51.3	-	-	
	C172G	172	493	-	12.8	26	-	-	-	0.4	59.6	-	-	
	V189I	189	160	-	14.3	89	-	-	-	1.8	49.2	-	-	
	V221I	221	181	-	2.5	14	-	-	-	0.3	57.9	-	-	
	V346L	346	256	-	4.3	17	-	-	-	0.3	56.9	-	-	
	V346I	346	274	-	7.2	26	-	-	-	0.5	54.4	-	-	
	S398C	398	322	-	18.6	58	-	-	-	1.0	56.2	-	-	
	I465V	465	298	-	14.3	48	-	-	-	0.9	53.7	-	-	
	F345I	345	274	-	11.2	41	-	-	-	0.8	52.1	-	-	
	F140I/F345I	140/345	218	-	15.6	72	-	-	-	1.2	59.3	-	-	
<i>Thermosynech</i> <i>elongatus</i> (L ₈ S ₈)	WT	-	-	107	6.6	-	70	-	-	-	53.3	36	6.6	RDE2 XLI-B Ref [22]
	F345I	345	-	101	6.9	-	47	-	-	-	48.0	-	14.6	
	P415A	415	-	111	6.9	-	62	-	-	-	52.6	-	10.4	
	V300A	300	-	135	6.6	-	49	-	-	-	51.9	-	7.1	
	F345I/P415A	345I/415	-	102	4.8	-	47	-	-	-	49.2	-	17.5	
	P415A/V98M ^S	415/112 ^S	-	93	9.5	-	100	-	-	-	56.3	-	20.4	
	P415A/A48V ^S	415/58 ^S	-	105	8.8	-	84	-	-	-	48.0	-	9.7	
	P415A/H37L ^S	415/39 ^S	-	100	8.4	-	84	-	-	-	49.7	-	11.1	
	P415A/Y36N ^S /G112D ^S	415/38 ^S /122 ^S /123 ^S	-	94	9.3	-	99	-	-	-	47.7	-	9.7	
	P415A/Y36N ^S /R51H ^S	415/38 ^S /65 ^S	-	120	8.0	-	68	-	-	-	51.8	-	13.6	
	P415A/A398T/A48V ^S	415/398/58 ^S	-	105	8.1	-	75	-	-	-	45.5	-	13.3	
	P415A/L74M/D397N	415/74/397	-	92	8.2	-	90	-	-	-	51.8	-	12.5	
	P415A/A414T	415/414	-	102	8.0	-	78	-	-	-	55.8	-	13.9	
	P415A/I393M	415/393	-	90	7.8	-	87	-	-	-	48.1	-	16.6	
P415A/A398T	415/398	-	110	6.4	-	59	-	-	-	50.5	-	16.6		
<i>Synchococcus</i> PCC6301 (L ₈ S ₈)	WT	-	-	251	12.6	-	50	-	-	-	46.0	44	1.3	
	F345I	345	-	285	10.3	-	36	-	-	-	42.0	-	14.4	
	F140I	140	-	211	11.1	-	53	-	-	-	46.0	-	6.3	
	V189I	189	-	186	10.6	-	57	-	-	-	39.0	NA	6.2	

Table A1. Cont.

Rubisco Origin	Mutation (s)	Plant RbcL Numbering	K_C (μM)	$K_C^{21\%O_2}$ (μM)	k_{cat}^C (s^{-1})	k_{cat}^C/K_C ($\text{s}^{-1}\cdot\text{mM}^{-1}$)	$k_{cat}^C/K_C^{21\%O_2}$ ($\text{s}^{-1}\cdot\text{mM}^{-1}$)	K_O (μM)	k_{cat}^O (s^{-1})	k_{cat}^O/K_O	$S_{C/O}$	K_{RuBP} (μM)	Rubisco (%CSP)	RDE Strain
<i>Rhodospirillum Rubrum</i> (L ₂)	WT	-	149	-	12.3	83	-	159	1.4	9.0	9.0	63	0.2	MM1 Ref [26]
	D117V	129	196	-	7.3	38	-	199	1.5	8.0	5.3	55	0.6	
	D117H	129	192	-	7.5	39	-	153	1.1	7.0	5.3	48	0.5	
	H44Q	56	301	-	9.3	31	-	185	1.2	6.0	5.3	56	0.9	
	H44N	56	204	-	9.8	48	-	116	1.0	9.0	5.5	59	0.7	
<i>Methanococcoides burtonii</i> (L ₂ -L ₁₀)	WT	-	57	1337	0.6	11	0.5	11	0.1	8.0	1.3	4	7.0	MM1 Ref [40]
	E138V	146	66	543	1.0	15	1.8	35	0.4	11.0	1.5	2	8.0	
	K332E	336	78	892	1.2	15	1.4	24	0.3	13.0	1.3	5	6.0	
	T421A	400	-	-	1.3	-	-	-	-	-	1.7	-	9.0	
	M423V	402	-	-	1.1	-	-	-	-	-	-	-	7.0	
	H379R	360	-	-	0.6	-	-	-	-	-	-	-	6.0	
	Y471N	471	-	-	0.9	-	-	-	-	-	-	-	4.0	
G327S	331/insert\332	-	-	0.7	-	-	-	-	-	-	-	6.0		
<i>Synechococcus</i> PCC6301 (L ₈ S ₈)	WT	-	186	-	3.5	22	-	-	-	-	-	38	-	<i>R.capsulatus</i> SBI/II-mutant Ref [31]
	F342V	345	256	-	2.4	12	-	-	-	-	-	21	-	
	D103V	106	459	-	2.4	7	-	-	-	-	-	34	-	
	D103E	106	325	-	1.8	6	-	-	-	-	-	36	-	
<i>Synechococcus</i> PCC6301 (L ₈ S ₈)	WT	-	186	-	4.9	26	-	850	-	-	-	27	-	<i>R.capsulatus</i> SBI/II-mutant Ref [43]
	D103V	106	388	-	4.9	13	-	866	-	-	-	25	-	
	A375V	378	143	-	0.6	4	-	1088	-	-	-	34	-	
	D103V/A375V	106/378	294	-	3.1	10	-	1153	-	-	-	269	-	
<i>Synechococcus</i> PCC6301 (L ₈ S ₈)	WT	-	190	247	4.3	21	-	841	-	-	-	29	-	<i>R.capsulatus</i> SBI/II-mutant Ref [28]
	A375T	378	146	-	0.5	3	-	1076	-	-	-	37	-	
	M259T	262	147	209	4.4	27	-	595	-	-	-	30	-	
	A375V/M259T/M571 ^S	378/262/81 ^S	93	126	2.2	22	-	716	-	-	-	64	-	
	R214H	217	682	-	3.6	6	-	2152	-	-	-	30	-	
	R214H/A375S	217/378	264	-	0.6	2	-	1731	-	-	-	25	-	
	T327A	330	273	-	2.8	11	-	1303	-	-	-	13	-	
	V186I	189	106	145	2.8	28	-	686	-	-	-	25	-	
	V186I/T327A	189/330	110	133	2.1	18	-	1189	-	-	-	25	-	
	S325L	328	111	-	4.0	36	-	931	-	-	-	251	-	
S325L/T327A	328/330	142	-	2.6	21	-	577	-	-	-	77	-		
<i>Ralstonia eutropha</i> (Cupriavidus necator; L ₈ S ₈)	WT	-	37	-	3.8	102	-	1149	-	-	-	34	-	<i>R.eutropha</i> H16ΔLS mutant Ref [44]
	Y347V	345	35	-	4.1	117	-	1139	-	-	-	-	-	
	A380V	378	34	-	0.3	9	-	1435	-	-	-	939	-	
	T330A	328	44	-	2.8	64	-	1087	-	-	-	29	-	
	Y348C	346	24	-	1.6	25	-	674	-	-	-	36	-	
	T330A/A380V	328/378	28	-	0.7	25	-	1069	-	-	-	769	-	
Y348C/A380V	346/378	20	-	0.7	35	-	652	-	-	-	581	-		

%CSP, Percentage (*w/w*) cell soluble protein; Red text—kinetic values corrected from those published in Mueller-cajar et al., 2008; (-) value not measured.

References

1. Sharwood, R.E. Engineering chloroplasts to improve Rubisco catalysis: Prospects for translating improvements into food and fiber crops. *New Phytol.* **2017**, *213*, 494–510. [[CrossRef](#)]
2. Ellis, R.J. The most abundant protein in the world. *Trends Biochem. Sci.* **1979**, *4*, 241–244. [[CrossRef](#)]
3. Young, J.N.; Heureux, A.M.C.; Sharwood, R.E.; Rickaby, R.E.M.; Morel, F.M.M.; Whitney, S.M. Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *J. Exp. Bot.* **2016**, *67*, 3445–3456. [[CrossRef](#)] [[PubMed](#)]
4. Makino, A. Rubisco and nitrogen relationships in rice: Leaf photosynthesis and plant growth. *Soil Sci. Plant Nutr.* **2003**, *49*, 319–327. [[CrossRef](#)]
5. Bauwe, H.; Hagemann, M.; Kern, R.; Timm, S. Photorespiration has a dual origin and manifold links to central metabolism. *Curr. Opin. Plant Biol.* **2012**, *15*, 269–275. [[CrossRef](#)] [[PubMed](#)]
6. Sage, R.F.; Sage, T.L.; Kocacinar, F. Photorespiration and the evolution of C4 photosynthesis. *Annu. Rev. Plant Biol.* **2012**, *63*, 19–47. [[CrossRef](#)] [[PubMed](#)]
7. Carmo-Silva, E.; Scales, J.C.; Madgwick, P.J.; Parry, M.A.J. Optimizing Rubisco and its regulation for greater resource use efficiency. *Plantcell Environ.* **2015**, *38*, 1817–1832. [[CrossRef](#)] [[PubMed](#)]
8. Parry, M.A.J.; Andralojc, P.J.; Scales, J.C.; Salvucci, M.E.; Carmo-Silva, A.E.; Alonso, H.; Whitney, S.M. Rubisco activity and regulation as targets for crop improvement. *J. Exp. Bot.* **2013**, *64*, 717–730. [[CrossRef](#)] [[PubMed](#)]
9. Andersson, I.; Backlund, A. Structure and function of Rubisco. *Plant Physiol. Biochem.* **2008**, *46*, 275–291. [[CrossRef](#)]
10. Andersson, I. Catalysis and regulation in Rubisco. *J. Exp. Bot.* **2008**, *59*, 1555–1568. [[CrossRef](#)]
11. Cummins, P.L.; Kannappan, B.; Gready, J.E. Revised mechanism of carboxylation of ribulose-1,5-bisphosphate by rubisco from large scale quantum chemical calculations. *J. Comput. Chem.* **2018**, *39*, 1656–1665. [[CrossRef](#)] [[PubMed](#)]
12. Ishikawa, C.; Hatanaka, T.; Misoo, S.; Miyake, C.; Fukayama, H. Functional incorporation of sorghum small subunit increases the catalytic turnover rate of Rubisco in transgenic rice. *Plant Physiol.* **2011**, *156*, 1603–1611. [[CrossRef](#)] [[PubMed](#)]
13. Whitney, S.M.; Sharwood, R.E.; Orr, D.; White, S.J.; Alonso, H.; Galmés, J. Isoleucine 309 acts as a C4 catalytic switch that increases ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco) carboxylation rate in *Flaveria*. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 14688–14693. [[CrossRef](#)] [[PubMed](#)]
14. Spreitzer, R.J.; Peddi, S.R.; Satagopan, S. Phylogenetic engineering at an interface between large and small subunits imparts land-plant kinetic properties to algal Rubisco. *Proc. Nat. Acad. Sci.* **2005**, *102*, 17225–17230. [[CrossRef](#)] [[PubMed](#)]
15. Farquhar, G.D.; von Caemmerer, S.; Berry, J.A. A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species. *Planta* **1980**, *149*, 78–90. [[CrossRef](#)] [[PubMed](#)]
16. Andrews, T.J.; Whitney, S.M. Manipulating ribulose bisphosphate carboxylase/oxygenase in the chloroplasts of higher plants. *Arch. Biochem. Biophys.* **2003**, *414*, 159–169. [[CrossRef](#)]
17. Long, S.P.; Marshall-Colon, A.; Zhu, X.-G. Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell* **2015**, *161*, 56–66. [[CrossRef](#)] [[PubMed](#)]
18. Whitney, S.M.; Baldet, P.; Hudson, G.S.; Andrews, T.J. Form I Rubiscos from non-green algae are expressed abundantly but not assembled in tobacco chloroplasts. *Plant J.* **2001**, *26*, 535–547. [[CrossRef](#)]
19. Lin, M.T.; Hanson, M.R. Red algal Rubisco fails to accumulate in transplastomic tobacco expressing *Griffithsia monilis* RbcL and RbcS genes. *Plant Direct.* **2018**, *2*, e00045. [[CrossRef](#)]
20. Long, B.M.; Hee, W.Y.; Sharwood, R.E.; Rae, B.D.; Kaines, S.; Lim, Y.-L.; Nguyen, N.D.; Massey, B.; Bala, S.; von Caemmerer, S.; et al. Carboxysome encapsulation of the CO₂-fixing enzyme Rubisco in tobacco chloroplasts. *Nat. Commun.* **2018**, *9*, 3570. [[CrossRef](#)]
21. Occhialini, A.; Lin, M.T.; Andralojc, P.J.; Hanson, M.R.; Parry, M.A.J. Transgenic tobacco plants with improved cyanobacterial Rubisco expression but no extra assembly factors grow at near wild-type rates if provided with elevated CO₂. *Plant J.* **2016**, *85*, 148–160. [[CrossRef](#)] [[PubMed](#)]
22. Wilson, R.H.; Martin-Avila, E.; Conlan, C.; Whitney, S.M. An improved *Escherichia coli* screen for Rubisco identifies a protein–protein interface that can enhance CO₂-fixation kinetics. *J. Biol. Chem.* **2018**, *293*, 18–27. [[CrossRef](#)] [[PubMed](#)]

23. Sharwood, R.E.; Ghannoum, O.; Kapralov, M.V.; Gunn, L.H.; Whitney, S.M. Temperature responses of Rubisco from Paniceae grasses provide opportunities for improving C₃ photosynthesis. *Nat. Plants* **2016**, *2*, 16186. [[CrossRef](#)] [[PubMed](#)]
24. Mueller-Cajar, O.; Whitney, S.M. Directing the evolution of Rubisco and Rubisco activase: First impressions of a new tool for photosynthesis research. *Photosynth. Res.* **2008**, *98*, 667–675. [[CrossRef](#)] [[PubMed](#)]
25. Wilson, R.H.; Whitney, S.M. Improving CO₂ fixation by enhancing Rubisco performance. In *Directed Enzyme Evolution: Advances and Applications*; Alcalde, M., Ed.; Springer International Publishing: Berlin, Germany, 2017; pp. 101–126.
26. Mueller-Cajar, O.; Morell, M.; Whitney, S.M. Directed evolution of Rubisco in *Escherichia coli* reveals a specificity-determining hydrogen bond in the form II enzyme. *Biochemistry* **2007**, *46*, 14067–14074. [[CrossRef](#)]
27. Duraõ, P.; Aigner, H.; Nagy, P.; Mueller-Cajar, O.; Hartl, F.U.; Hayer-Hartl, M. Opposing effects of folding and assembly chaperones on evolvability of Rubisco. *Nat. Chem. Biol.* **2015**, *11*, 148–155. [[CrossRef](#)]
28. Satagopan, S.; Huening, K.A.; Tabita, F.R. Selection of cyanobacterial *Synechococcus* sp. Strain PCC 6301 Rubisco variants with improved functional properties that confer enhanced CO₂-dependent growth of *Rhodobacter capsulatus*, a photosynthetic bacterium. *mBio* **2019**, *10*, e01537-19. [[CrossRef](#)]
29. Greene, D.N.; Whitney, S.M.; Matsumura, I. Artificially evolved *Synechococcus* PCC6301 Rubisco variants exhibit improvements in folding and catalytic efficiency. *Biochem. J.* **2007**, *404*, 517–524. [[CrossRef](#)]
30. Mueller-Cajar, O.; Whitney, S.M. Evolving improved *Synechococcus* Rubisco functional expression in *Escherichia coli*. *Biochem. J.* **2008**, *414*, 205–214. [[CrossRef](#)]
31. Smith, S.A.; Tabita, F.R. Positive and negative selection of mutant forms of prokaryotic (cyanobacterial) ribulose-1, 5-bisphosphate carboxylase/oxygenase. *J. Mol. Biol.* **2003**, *331*, 557–569. [[CrossRef](#)]
32. Emlyn-Jones, D.; Woodger, F.J.; Price, G.D.; Whitney, S.M. RbcX can function as a Rubisco-chaperonin, but is non essential in *Synechococcus* PCC7942. *Plant Cell Physiol.* **2006**, *47*, 1630–1640. [[CrossRef](#)] [[PubMed](#)]
33. Andrews, T.J.; Ballment, B. The function of the small subunits of ribulose bisphosphate carboxylase-oxygenase. *J. Biol. Chem.* **1983**, *258*, 7514–7518. [[PubMed](#)]
34. Bracher, A.; Whitney, S.M.; Hartl, F.U.; Hayer-Hartl, M. Biogenesis and metabolic maintenance of Rubisco. *Annu. Rev. Plant Biol.* **2017**, *68*, 29–60. [[CrossRef](#)]
35. Liu, D.; Ramya, R.C.S.; Mueller-Cajar, O. Surveying the expanding prokaryotic Rubisco multiverse. *FEMS Microbiol. Lett.* **2017**, *364*, 16. [[CrossRef](#)] [[PubMed](#)]
36. Parikh, M.R.; Greene, D.N.; Woods, K.K.; Matsumura, I. Directed evolution of Rubisco hypermorphs through genetic selection in engineered *E. coli*. *Protein Eng. Des. Sel.* **2006**, *19*, 113–119. [[CrossRef](#)] [[PubMed](#)]
37. Aigner, H.; Wilson, R.H.; Bracher, A.; Calisse, L.; Bhat, J.Y.; Hartl, F.U.; Hayer-Hartl, M. Plant Rubisco assembly in *E. coli* with five chloroplast chaperones including BSD2. *Science* **2017**, *358*, 1272–1278. [[CrossRef](#)]
38. Whitney, S.M.; Sharwood, R.E. Plastid transformation for Rubisco engineering and protocols for assessing expression. *Methods Mol. Biol.* **2014**, *1132*, 245–262.
39. Blayney, M.; Whitney, S.; Beck, J. NanoESI mass spectrometry of Rubisco and Rubisco activase structures and their Interactions with nucleotides and sugar phosphates. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 1588–1601. [[CrossRef](#)]
40. Wilson, R.H.; Alonso, H.; Whitney, S.M. Evolving *Methanococcoides burtonii* archaeal Rubisco for improved photosynthesis and plant growth. *Sci. Rep.* **2016**, *6*, 22284. [[CrossRef](#)]
41. Kane, H.J.; Viil, J.; Entsch, B.; Paul, K.; Morell, M.K.; Andrews, T.J. An improved method for measuring the CO₂/O₂ specificity of ribulosebisphosphate carboxylase-oxygenase. *Aust. J. Plant Physiol.* **1994**, *21*, 449–461. [[CrossRef](#)]
42. Sharwood, R.E.; von Caemmerer, S.; Maliga, P.; Whitney, S.M. 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.* **2008**, *146*, 83–96. [[CrossRef](#)] [[PubMed](#)]
43. Satagopan, S.; Scott, S.S.; Smith, T.G.; Tabita, F.R. A Rubisco mutant that confers growth under a normally “inhibitory” oxygen concentration. *Biochemistry* **2009**, *48*, 9076–9083. [[CrossRef](#)] [[PubMed](#)]
44. Satagopan, S.; Tabita, F.R. Rubisco selection using the vigorously aerobic and metabolically versatile bacterium *Ralstonia eutropha*. *FEBS J.* **2016**, 2869–2880. [[CrossRef](#)] [[PubMed](#)]

45. Knight, S.; Andersson, I.; Brändén, C. Crystallographic analysis of ribulose 1,5-bisphosphate carboxylase from spinach at 2.4 Å resolution. Subunit interactions and active site. *J. Mol. Biol.* **1990**, *215*, 113–160. [[CrossRef](#)]
46. Ramage, R.T.; Read, B.A.; Tabita, F.R. Alteration of the alpha helix region of cyanobacterial ribulose 1,5-bisphosphate carboxylase/oxygenase to reflect sequences found in high substrate specificity enzymes. *Arch. Biochem. Biophys.* **1998**, *349*, 81–88. [[CrossRef](#)] [[PubMed](#)]
47. Furbank, R.T.; Sage, R.F. Editorial overview: Physiology and metabolism: CO₂ concentrating mechanisms in photosynthetic organisms: Evolution, efficiency and significance for crop improvement. *Curr. Opin. Plant Biol.* **2016**, *31*, iv–vii. [[CrossRef](#)]
48. Hasse, D.; Larsson, A.M.; Andersson, I. Structure of Arabidopsis thaliana Rubisco activase. *Acta Crystallogr. D Biol. Crystallogr.* **2015**, *71*, 800–808. [[CrossRef](#)]
49. Van Lun, M.; van der Spoel, D.; Andersson, I. Subunit interface dynamics in hexadecameric rubisco. *J Mol. Biol.* **2011**, *411*, 1083–1098. [[CrossRef](#)]
50. Romero, P.A.; Arnold, F.H. Exploring protein fitness landscapes by directed evolution. *Nat. Reviews. Mol. Cell Biol.* **2009**, *10*, 866–876. [[CrossRef](#)]
51. You, L.; Arnold, F.H. Directed evolution of subtilisin E in *Bacillus subtilis* to enhance total activity in aqueous dimethylformamide. *Protein Eng. Des. Sel.* **1996**, *9*, 77–83. [[CrossRef](#)]
52. Vitlin Gruber, A.; Feiz, L. Rubisco Assembly in the Chloroplast. *Front. Mol. Biosci.* **2018**, *5*, 24. [[CrossRef](#)] [[PubMed](#)]
53. Wilson, R.H.; Hayer-Hartl, M. Complex chaperone dependence of Rubisco biogenesis. *Biochemistry* **2018**, *57*, 3210–3216. [[CrossRef](#)] [[PubMed](#)]
54. Whitney, S.M.; Birch, R.; Kelso, C.; Beck, J.L.; Kapralov, M.V. Improving recombinant Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1 chaperone. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 3564–3569. [[CrossRef](#)] [[PubMed](#)]
55. Whitney, S.M.; Sharwood, R.E. Linked Rubisco subunits can assemble into functional oligomers without impeding catalytic performance. *J Biol. Chem.* **2007**, *282*, 3809–3818. [[CrossRef](#)] [[PubMed](#)]
56. Conlan, B.; Whitney, S. Preparing Rubisco for a tune up. *Nat. Plants* **2018**, *4*, 12–13. [[CrossRef](#)]
57. Wilson, R.H.; Thieulin-Pardo, G.; Hartl, F.-U.; Hayer-Hartl, M. Improved recombinant expression and purification of functional plant Rubisco. *FEBS Lett.* **2019**, *593*, 611–621. [[CrossRef](#)] [[PubMed](#)]
58. Cai, Z.; Liu, G.; Zhang, J.; Li, Y. Development of an activity-directed selection system enabled significant improvement of the carboxylation efficiency of Rubisco. *Prot. Cell* **2014**, *5*, 552–562. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).