Pinpointing *Synechococcus* Rubisco large subunit sections involved in heterologous holoenzyme formation in *Escherichia coli*

Wei Chi Ong[1], Moh Lan Yap[2], Hann Ling Wong[2] and Boon Hoe Lim[1]*

[1]Department of Chemical Science, Faculty of Science, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia.
[2]Department of Biological Science, Faculty of Science, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia.
Email: bhlim@utar.edu.my

ABSTRACT

Aims: Heterologous holoenzyme formation of ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) has been a challenge due to a limited understanding of its biogenesis. Unlike bacterial Rubiscos, eukaryotic Rubiscos are incompatible with the *Escherichia coli* (*E. coli*) chaperone system to fold and assemble into the functional hexadecameric conformation (L_8S_8), which comprises eight large subunits (RbcL) and eight small subunits (RbcS). Our previous study reported three sections (residues 248-297, 348-397 and 398-447) within the RbcL of *Synechococcus elongatus* PCC6301, which may be important for the formation of L_8S_8 in *E. coli*. The present study further examined these three sections separately, dividing them into six sections of 25 residues (i.e., residues 248-272, 273-297, 348-372, 373-397, 398-422 and 423-447).

Methodology and results: Six chimeric Rubiscos with each section within the RbcL from *Synechococcus* replaced by their respective counterpart sequence from *Chlamydomonas reinhardtii* were constructed and checked for their effect on holoenzyme formation in *E. coli*. The present study shows that Section 1 (residues 248-272; section of *Synechococcus* RbcL replaced by corresponding *Chlamydomonas* sequence), Section 2 (residues 273-297), Section 3 (residues 348-372) and Section 6 (residues 423-447) chimeras failed to fold and assemble despite successful expression of both RbcL and RbcS. Only Section 4 (residues 373-397) and 5 (residues 398-422) chimeras could form L_8S_8 in *E. coli*.

Conclusion, significance and impact of study: GroEL chaperonin mediates the folding of bacterial RbcL in *E. coli*. Therefore, residues 248-297, 348-372 and 423-447 of *Synechococcus* RbcL may be important for interacting with the GroEL chaperonin for successful holoenzyme formation in *E. coli*.

Keywords: Ribulose bisphosphate carboxylase/oxygenase (Rubisco), holoenzyme formation, *Escherichia coli*

INTRODUCTION

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco, E.C 4.1.1.39) is the CO_2-fixing enzyme that incorporates atmospheric CO_2 onto organic carbon ribulose-1,5-bisphosphate (RuBP) in the Calvin-Benson-Bassham cycle of photosynthesis to produce two molecules of 3-phosphoglycerate (3PGA), for biomass accumulation in photoautotrophs (Ellis, 1979). Given its pivotal role, Rubisco is the gateway for inorganic CO_2 into the biosphere and its capability to sequester CO_2 dictates photosynthetic CO_2 assimilation efficiency (Jordan and Ogren, 1981). Nevertheless, Rubiscos are slow enzymes with carboxylation turnover rates (k^c_{cat}) from 1 to 13/sec (Tcherkez *et al.*, 2006). Moreover, its inability to distinguish between CO_2 and O_2 for fixation onto RuBP significantly decreases photosynthetic efficiency:

oxygenation of RuBP lead to the formation of one 3PGA and one 2-phosphoglycolate (2PG), which must be recycled into 3PGA through photorespiration, expending cellular energy and losing carbon as CO_2 (Ogren, 1984). Slow catalysis and undesirable oxygenase activity of Rubiscos render CO_2 assimilation as one of the rate-limiting factors in photosynthesis (Long *et al.*, 2006). Unsurprisingly, Rubisco has been a prime target for genetic engineering to improve its catalytic performance in terms of CO_2/O_2 specificity factor and carboxylation rate as a means to increase photosynthetic efficiency to raise crop productivity (Parry *et al.*, 2013; Carmo-Silva *et al.*, 2015).

Plant, algae and cyanobacteria Rubiscos are hexadecamers (L_8S_8, ~560 kDa) consisting of eight large subunits (RbcL, 50-55 kDa) and eight small subunits (RbcS, 12-15 kDa) (Andersson and Backlund, 2008).

---

*Corresponding author

They share conserved structures and key catalytic residues but have varying kinetic properties (Tabita *et al.*, 2008). Extensive studies of Rubisco have been conducted in order to engineer a catalytically improved Rubisco, as well as to identify the key determinants governing the catalytic properties (Spreitzer *et al.*, 2005; Sharwood *et al.*, 2008; Genkov and Spreitzer, 2009; Genkov *et al.*, 2010; Ishikawa *et al.*, 2011; Whitney *et al.*, 2011). Unfortunately, failures in the heterologous holoenzyme formation of Rubiscos from phylogenetically distant hosts limit the study of Rubiscos from diverse sources and the selection of the host for their genetic manipulation. For instance, Rubiscos from the diatom *Phaeodactylum tricornutum* and non-green algae *Galdieria sulphuraria* and *Griffithsia monolis* failed to form functional enzymes in tobacco chloroplast despite successful expression of both RbcL and RbcS subunits (Whitney *et al.*, 2001; Lin and Hanson, 2018). Similarly, eukaryotic Rubiscos from maize, wheat, tobacco and green alga *C. reinhardtii* form insoluble aggregates expressed in *E. coli* (Gatenby *et al.*, 1987; Cloney *et al.*, 1993; Whitney and Sharwood, 2007; Koay *et al.*, 2016). This has prevented efforts in the directed evolution of Rubisco using Rubisco-dependent *E. coli*, which would have been crucial for evolving kinetically improved Rubisco variants and providing structure-function insights (Parikh *et al.*, 2006; Mueller-Cajar *et al.*, 2007; Mueller-Cajar and Whitney, 2008; Wilson *et al.*, 2016; 2018).

Lack of functional expression in heterologous host mainly stems from chaperone incompatibility and the requirement of additional auxiliary factors for the complex biogenesis of Rubisco, which is a multi-step process requiring different kinds of chaperones for *de novo* folding and assembly (Bracher *et al.*, 2011; Kolesinski *et al.*, 2014; Hauser *et al.*, 2015; Wilson and Hayer-Hartl, 2018). In *E. coli*, the folding of RbcL monomers is mediated by the endogenous GroEL-GroES chaperonin system prior to their assembly into oligomers (van Duijn *et al.*, 2006; Liu *et al.*, 2010). The importance of the GroEL-GroES chaperonin for Rubisco holoenzyme formation in *E. coli* was demonstrated by the increased yield of soluble Rubiscos upon overexpression of GroEL: conversely, mutations of GroEL or GroES abolished holoenzyme formation (Goloubinoff *et al.*, 1989b; Greene *et al.*, 2007). In addition, studies have shown that Rubiscos are prone to aggregation. They are recognized as stringent substrates by the GroEL chaperonin, whereby *in vitro* reconstitution of dimeric Form II *Rhodospirillum rubrum* Rubisco and hexadecameric Form I cyanobacterial Rubisco require all GroEL, GroES, ATP and Mg^{2+} (Goloubinoff *et al.*, 1989a; Lin and Rye, 2004; Liu *et al.*, 2010). Therefore, the non-formation of eukaryotic Rubiscos in *E. coli* suggests that their RbcLs (and, more specifically, the amino acid sequences) are incompatible with bacterial GroEL/GroES chaperonin. Indeed, functional expression of *A. thaliana* Rubisco in *E. coli* required the co-expression of its chloroplast chaperonin Cpn60αβ/Cpn20, which cannot be replaced by the bacterial GroEL/GroES homolog (Aigner *et al.*, 2017).

Pinpointing RbcL sections that may be key for functional expression in *E. coli* may prove useful for future work to improve eukaryotic Rubisco using *E. coli*-based directed evolution. To that end, our previous study replaced sections of cyanobacterial RbcL from *S. elongatus* PCC6301 with their eukaryotic counterpart sequences from the green alga *C. reinhardtii* sequentially, and a few sections of *Synechococcus* RbcL that might be essential for successful holoenzyme formation in *E. coli* were reported (Koay *et al.*, 2016). The present study aims to narrow the range by further examining each half (25 residues) of three of the sections (i.e. amino acids 248-297, 348-397 and 398-447) that were pinpointed. Therefore, six chimeric Rubiscos with 25-amino acid sections of *Synechococcus* RbcL (residues 248-272, 273-297, 348-372, 373-397, 398-422 and 423-447) were substituted with the corresponding residues in *Chlamydomonas* RbcL were constructed to examine their importance for holoenzyme formation in *E. coli*.

MATERIALS AND METHODS

Chimeric Rubisco plasmid construction

Gene fragments required for constructing chimeric *rbc*L-*rbc*S operons were amplified from plasmid pTrcSynLS harbouring the wild-type *Synechococcus* PCC6301 *rbc*L-*rbc*S operon (Mueller-Cajar and Whitney, 2008) and plasmids carrying different recombinant *rbc*L-*rbc*S (Koay *et al.*, 2016) by polymerase chain reaction (PCR) using *Pfu* DNA Polymerase (Vivantis). Primers used for PCR amplification were designed with different restriction sites (*Pst*I, *Nco*I and *Bsm*BI) to facilitate directional cloning (Table 1). Amplified products were digested with restriction enzymes and ligated into the vector backbone (pTrcHisB) of pTrcSynLS using T4 DNA ligase (Invitrogen) (Table 1). XL-1 Blue *E. coli* cells were transformed with the ligation mixtures and selected on Luria-Bertani (LB) plates containing 100 µg/mL ampicillin. Positive transformants were identified by colony PCR, and plasmids were extracted using a commercial plasmid miniprep kit for subsequent DNA sequencing.

Rubisco expression induction and protein extraction

Constructed chimeric Rubisco plasmids were transformed into XL-1 Blue *E. coli* cells and grown on LB agar (100 µg/mL ampicillin) overnight. Single colonies were inoculated in LB broth (100 µg/mL ampicillin) and cultured at 37 °C, 200 rpm for 16 h. Cultures were then diluted 50× into fresh LB broth (100 µg/mL ampicillin) and grown at 37 °C, 200 rpm. When OD_{600} of 0.5 was reached, 0.5 mM IPTG was added to the cultures to induce expression of Rubisco (during which cultures were grown at 37 °C, 200 rpm for another 16 h). To extract the proteins (crude lysates), cells were harvested by centrifugation at 5000× *g* for 5 min, re-suspended in ice-cold extraction buffer (50 mM Bicine/NaOH, pH 8.0, 10 mM $MgCl_2$, 10 mM $NaHCO_3$ and 2 mM DTT) to 10% w/v cell suspensions, and sonicated.

Table 1: Templates and primer pairs used for constructing chimeric Rubisco plasmids.

Construct	Primer[a]	Template[b]
Section 1 chimera pTrcSynL (Chl251-275)S	SynN-BsmBI[c] & Chl275-rev-BsmBI 5'-GAAGTGT<u>CGTCTC</u>TGAAACCACCTGTTAAGTAGTCGTGC-3'	pTrcSynL(Chl251-300)S
	SynSS-C-PstI[c] & Syn275-fwd-BsmBI 5'- ACTTCTTGA<u>CGTCTC</u>GTTTCACAGCCAACACCACCT-3'	pTrcSynLS
Section 2 chimera pTrcSynL (Chl276-300)S	SynN-BsmBI[c] & Syn275-rev-BsmBI 5'-TTGCCAAGGT<u>CGTCTC</u>GGCGGTGAAACCAG-3'	pTrcSynLS
	SynSS-C-PstI[c] & Chl275-fwd-BsmBI 5'-CTTAACAGGT<u>CGTCTC</u>ACCGCTAACACTTCATTAGC-3'	pTrcSynL(Chl251-300)S
Section 3 chimera pTrcSynL (Ch351-375)S	SynN-BsmBI[c] & Chl375-rev-BsmBI 5'-CGTGAATA<u>CGTCTC</u>GAAGCAACCGGCATAACACCTG-3'	pTrcSynL(Chl351-400)S
	SynSS-C-PstI[c] & Syn375-fwd-BsmBI 5'-GGCGTGCT<u>CGTCTC</u>TGCTTCCGGTGGTATC-3'	pTrcSynLS
Section 4 chimera pTrcSynL (Chl376-400)S	SynN-BsmBI[c] & Syn375-rev-BsmBI 5'-GGATAC<u>CGTCTC</u>AAGCAACTGGCAGCACGC-3'	pTrcSynLS
	SynSS-C-PstI[c] & Chl375-fwd-BsmBI 5'-CAGGTGTTA<u>CGTCTC</u>TTGCTTCAGGTGGTATTCACG-3'	pTrcSynL(Chl351-400)S
Section 5 chimera pTrcSynL (Chl401-425)S	SynN-BsmBI[c] & Chl425-rev-BsmBI 5'-GAGTACA<u>CGTCTC</u>AAGAGCTACACGGTTAG-3'	pTrcSynL(Chl401-450)S
	SynSS-C-PstI[c] & Syn425-fwd-BsmBI 5'-CGAAC<u>CGTCTC</u>GCTCTTGAAGCTTGCG-3'	pTrcSynLS
Section 6 chimera pTrcSynL (Chl426-450)S	SynN-BsmBI[c] & Syn425-rev-BsmBI 5'-CGTTCCGAG<u>CGTCTC</u>CACAAGCTTCCAAGG-3'	pTrcSynLS
	SynSS-C-PstI[c] & Chl425-fwd-BsmBI 5'-CCGTGTAGCTC<u>CGTCTC</u>CTTGTACTCAAGC-3'	pTrcSynL(Chl401-450)S

[a]Restriction sites incorporated into primer sequences are underlined; [b]pTrcSynLS is from Mueller-Cajar and Whitney (2008) while other templates are from Koay *et al.* (2016); [c]Primers from Koay *et al.* (2016).

Rubisco expression and assembly analysis

Expression of RbcL and RbcS in *E. coli* was checked by denaturing sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Crude lysates were mixed with sample loading buffer (30% w/v sucrose, 5% w/v SDS, 0.05% w/v bromophenol blue and 100 mM DTT) at 3:2 ratio and boiled for 5 min. Then, 4 µL of boiled sample mixtures were resolved on 12% w/v polyacrylamide gel. On the other hand, the assembly or formation of Rubisco holoenzyme was examined by non-denaturing native PAGE. For native PAGE, crude lysates were mixed with sample loading buffer (50% v/v glycerol, 150 mM Tris-HCl and 0.25% w/v bromophenol blue) at a 4:1 ratio and 25 µL of sample mixtures were resolved on 7.5% w/v polyacrylamide gel. For Western blot analysis, resolved proteins from SDS-PAGE and non-denaturing PAGE were transferred to nitrocellulose membranes (0.45 µm)

by electro-blotting at 30 V for 2 h (Towbin *et al.*, 1979). The membrane was then probed with rabbit anti-*Synechococcus* PCC6301 Rubisco IgG (Parikh *et al.*, 2006).

RESULTS AND DISCUSSION

Chimeric Rubisco creation

To examine sections of *Synechococcus* RbcL essential for holoenzyme formation in *E. coli*, six chimeric Rubiscos with sections of 25 residues (i.e., residues 248-272, 273-297, 348-372, 373-397, 398-422 and 423-447) separately changed to corresponding residues in *Chlamydomonas* RbcL were created (Figure 1). In essence, each chimeric Rubisco operon (*rbc*L-*rbc*S) was constructed, cloned into a plasmid and then transformed into electrocompetent XL-1 Blue *E. coli*. Based on the aligned translated
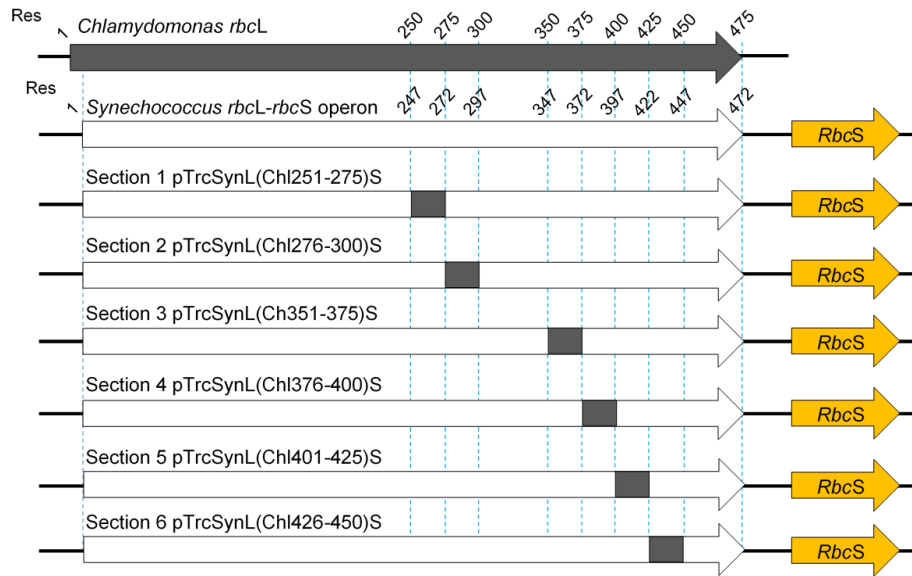
Figure 1: Schematic diagram of the chimeric *rbc*L-*rbc*S operon. Six chimeric Rubisco plasmids with respective sections of *Synechococcus rbc*L substituted by corresponding *Chlamydomonas rbc*L sequences were constructed. In lieu of bp number, residue (codon) number, Res, in *rbc*L genes are indicated.
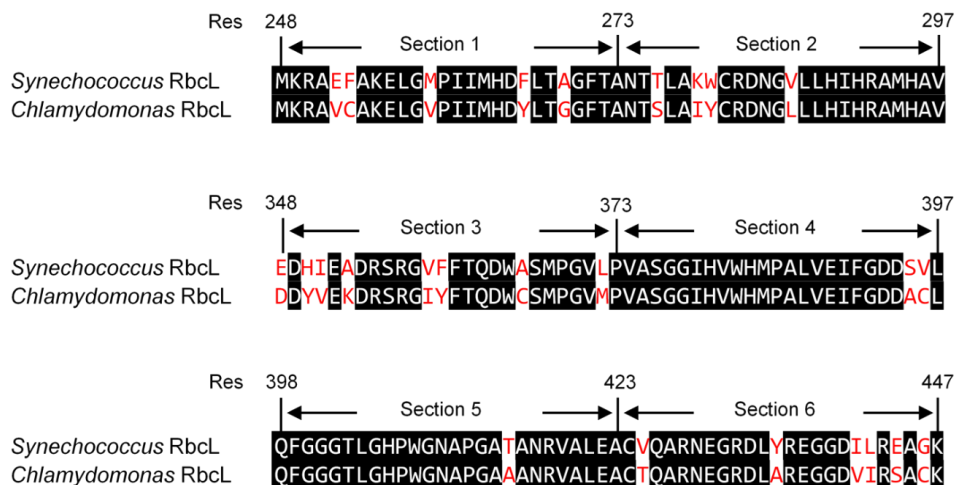
Figure 2: Sequence alignment of examined regions of *Synechococcus* and *Chlamydomonas* RbcL. Residue numbers (Res) are based on *Synechococcus* RbcL.

sequence of *Synechococcus* and *Chlamyomonas rbc*L, the number of actual amino acid changes in the sections ranges from one to eight amino acids only (Figure 2). Indeed, RbcLs are quite conserved across different species (Tabita *et al*., 2008).
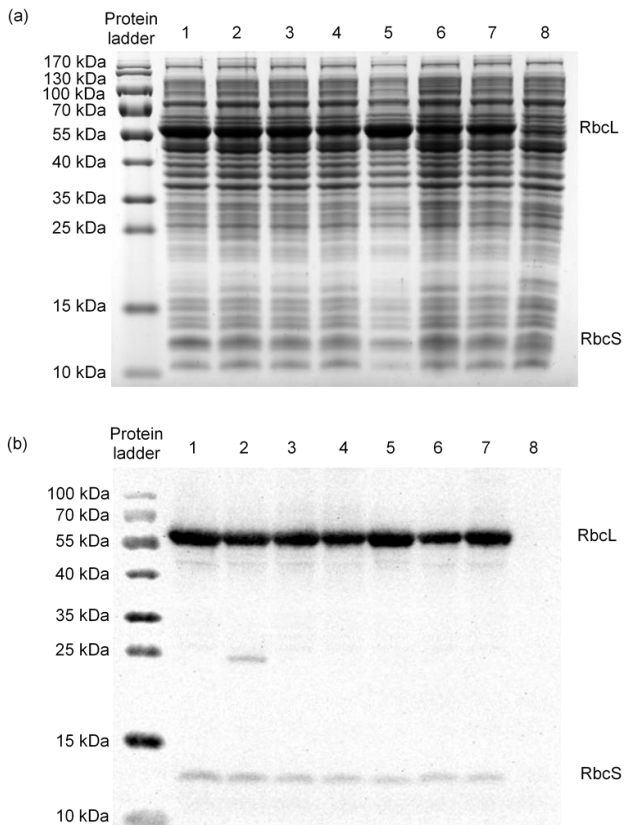
Chimeric Rubisco expression

SDS-PAGE and Western blot analysis of denatured total cellular proteins from *E. coli* transformed with the chimeric plasmids showed that both large and small subunits of all the chimeric Rubiscos were expressible (Figure 3). In each chimera lane (Figure 3, lanes 2 to 7), the large subunit (~55 kDa) and small subunit bands (~10-15 kDa) were visible and were similar to that of heterologously expressed wild-type *Synechococcus* Rubisco in terms of molecular weight and intensity (Figure 3, lane 1). Thus, there was no significant decrease in expression nor any unexpected proteolysis of the chimeric Rubiscos.
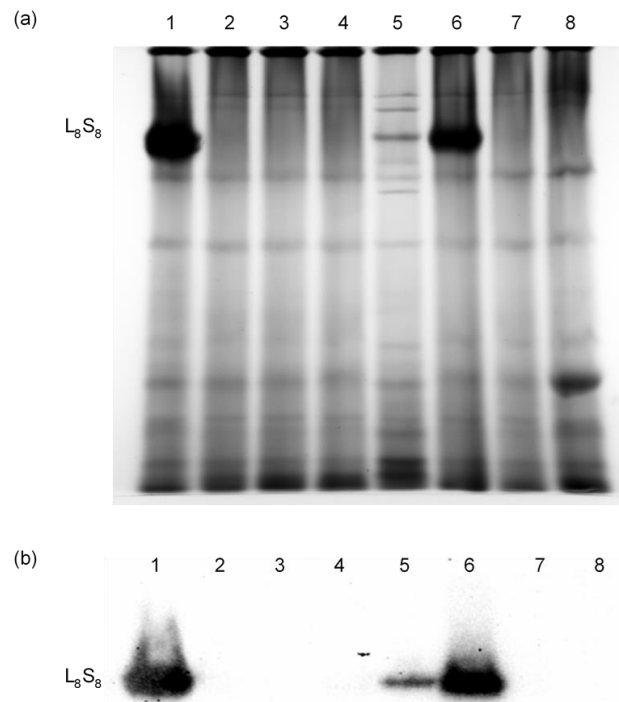
Chimeric Rubisco assembly

Native-PAGE and Western blot analysis showed sections 4 and 5 chimeric Rubiscos assembled into L_8S_8 (Figure 4, lanes 5 and 6). Therefore, the substituted large subunit residues within these sections, encompassing residues

Figure 3: Expression analysis of chimeric Rubiscos in XL-1 Blue *E. coli*. (a) SDS-PAGE and (b) Western blot analysis of total cellular protein of *E. coli* transformed with (1) pTrcSynLS; (2) pTrcSynL(Chl251-275)S; (3) pTrcSynL(Chl276-300)S; (4) pTrcSynL(Ch351-375)S; (5) pTrcSynL(Chl376-400)S; (6) pTrcSynL(Chl401-425)S; (7) pTrcSynL(Chl426-450)S; and (8) no plasmid.

373-397 and residues 398-422, may not be essential for GroEL interactions as the substitutions with corresponding *Chlamydomonas* sequences still allowed folding and assembly into their final hexadecameric complexes (L_8S_8) in *E. coli*. However, it remains to be ascertained whether their kinetic properties are altered.

It is noteworthy that the section 4 chimera also showed a reduced amount of L_8S_8 as compared to wild-type Rubisco (Figure 4, lane 5 cf. lane 1). Section 4 chimera has two amino acid substitutions, S395A and V396C. One possibility is that these substitutions affect the expression level of RbcL and RbcS. But looking at the SDS-PAGE and Western blot analysis of total cellular extracts, there were no discernible differences in the expressed levels of RbcL and RbcS between wild-type and chimeric Rubiscos (Figure 3). This is consistent with previous findings whereby most mutations in RbcL generally do not affect the steady-state mRNA level (Greene *et al.*, 2007; Mueller-Cajar and Whitney, 2008). Another possibility is that S395A and/or V396C decreased the stability of the holoenzyme, as has been


Figure 4: Assembly analysis of chimeric Rubiscos expressed in XL-1 Blue *E. coli*. (a) Native PAGE and (b) Western blot analysis of total cellular protein of in *E. coli* transformed with (1) pTrcSynLS; (2) pTrcSynL(Chl251-275)S; (3) pTrcSynL(Chl276-300)S; (4) pTrcSynL(Ch351-375)S; (5) pTrcSynL(Chl376-400)S; (6) pTrcSynL(Chl401-425)S; (7) pTrcSynL(Chl426-450)S; and (8) no plasmid.

reported before for other substitutions in RbcL (Du *et al.*, 2000; Genkov *et al.*, 2006; Genkov and Spreitzer, 2009). A comparison of *Synechococcus* and *Chlamydomonas* Rubiscos crystal structures (PDB IDs 1RBL and 1GK8) suggests that there might be a loss of van der Waals interaction to Leu-372 when Val-396 is changed to cysteine (Figure 5). This interaction could be necessary for the stability of the beta-sheet that constitutes the core of the alpha/beta barrel C-terminal domain of RbcL (Figure 5). Indeed, it would be interesting to create a V396C substitution alone to investigate this hypothesis in the future.

Chimeric Rubisco non-assembly

On the other hand, sections 1, 2, 3 and 6 chimeras had no detectable complexes in *E. coli* when resolved with native-PAGE (Figure 4, lanes 2 to 4 and 7). As these chimeras showed no L_8S_8 holoenzyme assembly, these sections of *Synechococcus* RbcL, comprising residues 248-272 (Section 1), 273-297 (Section 2), 348-372 (Section 3) and 423-447 (Section 6), might be important for interacting with GroEL.
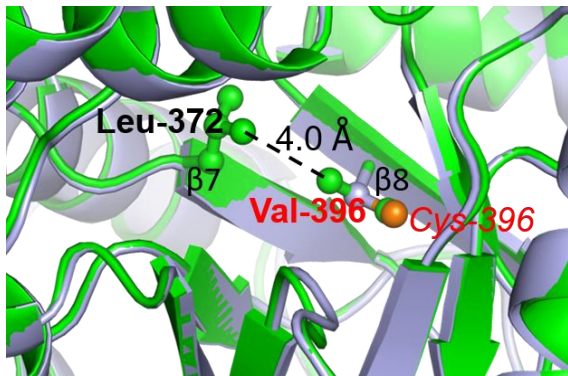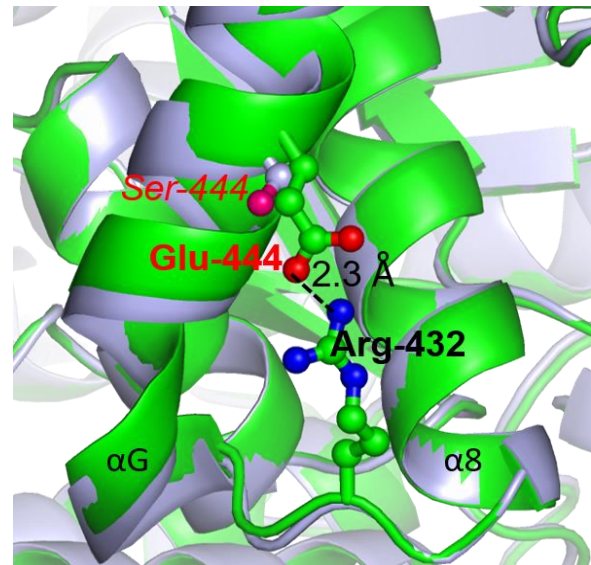
Figure 5: Potential loss of structural interaction due to V396C substitution (labeled in *red*). *Synechococcus* and *Chlamydomonas* Rubiscos large subunits (PDB IDs 1RBL and 1GK8) were superimposed. *Synechococcus* structure is mainly *green* and *Chlamydomonas* is *gray.* Important residues that are potentially affected are shown as ball and sticks (carbon atoms are *green* for *Synechococcus*, *grey* for *Chlamydomonas,* and sulfur atom is *orange*). The beta-strands that these residues are found on are labeled (β7 and β8). van der Waals contact that is lost upon mutation is shown, with distance measured in Å.

It has been reported that RbcL forms multivalent interactions to a minimum of three consecutive subunits of the heptameric GroEL ring for efficient binding (Farr *et al.*, 2000; Chen *et al.*, 2013). A recent cryo-electron microscopy study of the GroEL-RbcL binary complex showed the RbcL folding intermediate C-terminal domain in contact with three consecutive GroEL apical domains and the N-terminal domain in contact with one apical domain (Natesh *et al.*, 2018). Perhaps the three separate RbcL C-terminal regions (residues 248-297, 348-372 and 423-447) examined herein are involved in GroEL interactions and consequently, disruptions to these interactions due to amino acid changes abolished or substantially reduced binding of RbcL by GroEL subunits.

Other than the loss of GroEL interactions, Rubisco non-assembly could be due to the structural destabilizing effect of certain amino acid substitutions. Despite being captured by the GroEL chaperonin, the chimeric RbcL monomer might not be folded into its native structure stably because of disruptions to intra-subunit interactions within the RbcL and thus was unable to assemble into the final hexadecamer. Sections 1, 2, 3 and 6 chimeras have four to eight substitutions (Figure 2). For the section 1 chimera, which has five substitutions, A269G has been reported to improve the fitness of Rubisco-dependent *E. coli* (Mueller-Cajar and Whitney, 2008). Therefore, non-assembly is more likely caused by the remaining four substitutions. As for the section 6 chimera, one of the six mutations, E444S, potentially abolishes an intra-subunit salt bridge between Glu-444 and Arg-432 in the RbcL monomer (Figure 6).

As there are three states of assembly: dimeric L_2 pair, $(L_2)_4$ core complex and final hexadecameric L_8S_8, substitutions that affect inter-subunit interactions of any of

Figure 6: Potential loss of structural interaction due to E444S substitution (labeled in *red*). *Synechococcus* and *Chlamydomonas* Rubiscos were superimposed, colored and labeled as per Figure 5. Additional colors are *red* for *Synechococcus* oxygen atoms, *pink* for *Chlamydomonas* oxygen, and *blue* for nitrogen. The ionic interaction that is lost upon mutation is shown, with distance measured in Å.
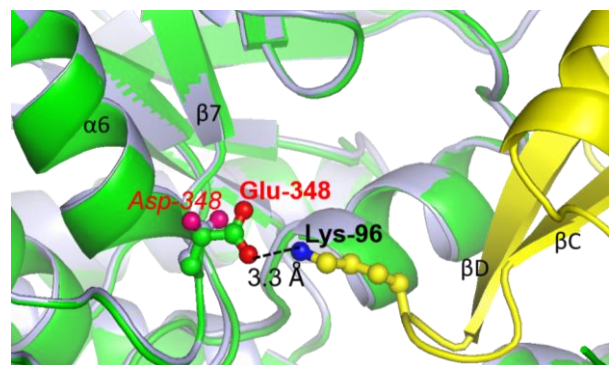
Figure 7: Potential loss of structural interaction due to E348D substitution (labeled in *red*). *Synechococcus* and *Chlamydomonas* Rubiscos were superimposed, colored, and labeled as per Figures 5 and 6. Additionally, the *Synechococcus* small subunit (mainly *yellow*) is also shown. The Rubisco large-small subunit interaction that is lost upon mutation is shown, with distance measured in Å.

these assembled complexes could eventually lead to non-assembly as well. Mutation E348D in the section 3 chimera is located at the interface between RbcL and RbcS. Comparing the Rubisco crystal structures of *Synechococcus* and *Chlamydomonas*, this mutation would replace the carboxyethyl side-chain of Glu-348 by a shorter carboxymethyl side-chain of Asp, which may disrupt the salt-bridge with Lys-96 in the small subunit (Koay *et al.*, 2016) (Figure 7).

CONCLUSION

When six sections of *Synechococcus* RbcL were substituted by counterpart sequences from *Chlamydomonas* RbcL, separately, only sections 4 and 5 chimeric Rubiscos form L_8S_8 in *E. coli*. Non-assembly of sections 1, 2, 3 and 6 chimeras suggests that these substituted sections of *Synechococcus* RbcL might be necessary for interacting with the endogenous GroEL chaperonin to be folded correctly in *E. coli*. Substitutions within these sections might have disrupted interaction with GroEL and led to misfolding or aggregation of chimeric RbcL, which precluded subsequent assembly. Non-assembly could also be due to structural destabilization by *Chlamydomonas* residue substitutions in RbcL. Although the actual cause of non-assembly remains to be elucidated, sections of *Synechococcus* RbcL that might be essential for successful holoenzyme formation in *E. coli* were narrowed down to residues 248-297, 348-372 and 423-447 for future dissection and engineering of Rubisco.

ACKNOWLEDGEMENTS

REFERENCES

Aigner, H., Wilson, R. H., Bracher, A., Calisse, L., Bhat, J. Y., Hartl, F. U. *et al.* (2017). Plant RuBisCo assembly in *E. coli* with five chloroplast chaperones including BSD2. *Science* **358(6368), 1272-1278.**

Andersson, I. and Backlund, A. (2008). Structure and function of Rubisco. *Plant Physiology and Biochemistry* **46(3), 275-291.**

Bracher, A., Starling-Windhof, A., Hartl, F. U. and Hayer-Hartl, M. (2011). Crystal structure of a chaperone-bound assembly intermediate of form I Rubisco. *Nature Structural and Molecular Biology* **18(8), 875-880.**

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 and Environment* **38(9), 1817-1832.**

Chen, D. H., Madan, D., Weaver, J., Lin, Z., Schröder, G. F., Chiu, W. *et al.* (2013). Visualizing GroEL/ES in the act of encapsulating a folding protein. *Cell* **153(6), 1354-1365.**

Cloney, L. P., Bekkaoui, D. R. and Hemmingsen, S. M. (1993). Co-expression of plastid chaperonin genes and a synthetic plant Rubisco operon in *Escherichia coli*. *Plant Molecular Biology* **23(6), 1285-1290.**

Du, Y. C., Hong, S. and Spreitzer, R. J. (2000). RbcS suppressor mutations improve the thermal stability and CO_2/O_2 specificity of rbcL-mutant ribulose-1,5-bisphosphate carboxylase/oxygenase. *Proceedings of the National Academy of Sciences of the United States of America* **97(26), 14206-14211.**

Ellis, R. J. (1979). The most abundant protein in the world. *Trends in Biochemical Sciences* **4, 241-244.**

Farr, G. W., Furtak, K., Rowland, M. B., Ranson, N. A., Saibil, H. R., Kirchhausen, T. *et al.* (2000). Multivalent binding of nonnative substrate proteins by the chaperonin GroEL. *Cell* **100, 561-573.**

Gatenby, A. A., van der Vies, S. M. and Rothstein, S. J. (1987). Co-expression of both the maize large and wheat small subunit genes of ribulose-bisphosphate carboxylase in *Escherichia coli*. *European Journal of Biochemistry* **168(1), 227-231.**

Genkov, T. and Spreitzer, R. J. (2009). Highly conserved small subunit residues influence Rubisco large subunit catalysis. *Journal of Biological Chemistry* **284(44), 30105-30112.**

Genkov, T., Du, Y. C. and Spreitzer, R. J. (2006). Small-subunit cysteine-65 substitutions can suppress or induce alterations in the large-subunit catalytic efficiency and holoenzyme thermal stability of ribulose-1,5-bisphosphate carboxylase/oxygenase. *Archives of Biochemistry and Biophysics* **451(2), 167-174.**

Genkov, T., Meyer, M., Griffiths, H. and Spreitzer, R. J. (2010). Functional hybrid rubisco enzymes with plant small subunits and algal large subunits: Engineered rbcS cDNA for expression in chlamydomonas. *Journal of Biological Chemistry* **285(26), 19833-19841.**

Goloubinoff, P., Christeller, J. T., Gatenby, A. A. and Lorimer, G. H. (1989a). Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. *Nature* **342(6252), 884-889.**

Goloubinoff, P., Gatenby, A. A. and Lorimer, G. H. (1989b). GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. *Nature* **337(6202), 44-47.**

Greene, D. N., Whitney, S. M. and Matsumura, I. (2007). Artificially evolved *Synechococcus* PCC6301 Rubisco variants exhibit improvements in folding and catalytic efficiency. *Biochemical Journal* **404, 517-524.**

Hauser, T., Popilka, L., Hartl, F. U. and Hayer-Hartl, M. (2015). Role of auxiliary proteins in Rubisco biogenesis and function. *Nature Plants* **1, 15065.**

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 Physiology* **156(3), 1603-1611.**

Jordan, D. B. and Ogren, W. L. (1981). Species variation in the specificity of ribulose biphosphate carboxylase/oxygenase. *Nature* **291(5815), 513-515.**

Koay, T. W., Wong, H. L. and Lim, B. H. (2016). Engineering of chimeric eukaryotic/bacterial Rubisco large subunits in *Escherichia coli. Genes and Genetic Systems* **91(3), 139-150.**

Kolesinski, P., Belusiak, I., Czarnocki-Cieciura, M. and Szczepaniak, A. (2014). Rubisco Accumulation Factor 1 from *Thermosynechococcus elongatus* participates in the final stages of ribulose-1,5-bisphosphate carboxylase/oxygenase assembly in *Escherichia coli* cells and *in vitro. FEBS Journal* **281(17), 3920-3932.**

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

Lin, Z. and Rye, H. S. (2004). Expansion and compression of a protein folding intermediate by GroEL. *Molecular Cell* **16(1), 23-34.**

Liu, C., Young, A. L., Starling-Windhof, A., Bracher, A., Saschenbrecker, S., Rao, B. V. *et al.* (2010). Coupled chaperone action in folding and assembly of hexadecameric Rubisco. *Nature* **463(7278), 197-202.**

Long, S. P., Zhu, X. G., Naidu, S. L. and Ort, D. R. (2006). Can improvement in photosynthesis increase crop yields? *Plant, Cell and Environment* **29(3), 315-330.**

Mueller-Cajar, O. and Whitney, S. M. (2008). Evolving improved *Synechococcus* Rubisco functional expression in *Escherichia coli. Biochemical Journal* **414(2), 205-214.**

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(49), 14067-14074.**

Natesh, R., Clare, D. K., Farr, G. W., Horwich, A. L. and Saibil, H. R. (2018). A two-domain folding intermediate of RuBisCO in complex with the GroEL chaperonin. *International Journal of Biological Macromolecules* **118, 671-675.**

Ogren, W. L. (1984). Photorespiration: Pathways, regulation, and modification. *Annual Review of Plant Physiology* **35(1), 415-442.**

Parikh, M. R., Greene, D. N., Woods, K. K. and Matsumura, I. (2006). Directed evolution of RuBisCO hypermorphs through genetic selection in engineered *E. coli. Protein Engineering, Design and Selection* **19(3), 113-119.**

Parry, M. A. J., Andralojc, P. J., Scales, J. C., Salvucci, M. E., Carmo-Silva, A. E., Alonso, H. *et al.* (2013). Rubisco activity and regulation as targets for crop improvement. *Journal of Experimental Botany* **64(3), 717-730.**

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 Physiology* **146(1), 83-96.**

Spreitzer, R. J., Peddi, S. R. and Satagopan, S. (2005). Phylogenetic engineering at an interface between large and small subunits imparts land-plant kinetic properties to algal Rubisco. *Proceedings of the National Academy of Sciences of the United States of America* **102(47), 17225-17230.**

Tabita, F. R., Satagopan, S., Hanson, T. E., Kreel, N. E. and Scott, S. S. (2008). Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *Journal of Experimental Botany* **59(7), 1515-1524.**

Tcherkez, G. G. B., Farquhar, G. D. and Andrews, T. J. (2006). Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proceedings of the National Academy of Sciences of the United States of America* **103(19), 7246-7251.**

Towbin, H., Staehelint, T. and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proceedings of the National Academy of Sciences* **76(9), 4350-4354.**

van Duijn, E., Simmons, D. A., van den Heuvel, R. H. H., Bakkes, P. J., van Heerikhuizen, H., Heeren, R. M. A. *et al.* (2006). Tandem mass spectrometry of intact GroEL - Substrate complexes reveals substrate-specific conformational changes in the trans ring. *Journal of the American Chemical Society* **128(14), 4694-4702.**

Whitney, S. M. and Sharwood, R. E. (2007). Linked rubisco subunits can assemble into functional oligomers without impeding catalytic performance. *Journal of Biological Chemistry* **282(6), 3809-3818.**

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 Journal* **26(5), 535-547.**

Whitney, S. M., Sharwood, R. E., Orr, D., White, S. J., Alonso, H. and Galmés, J. (2011). Isoleucine 309 acts as a C_4 catalytic switch that increases ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco) carboxylation rate in *Flaveria. Proceedings of the National Academy of Sciences of the United States of America* **108(35), 14688-14693.**

Wilson, R. H. and Hayer-Hartl, M. (2018). Complex chaperone dependence of Rubisco biogenesis. *Biochemistry* **57(23), 3210-3216.**

Wilson, R. H., Alonso, H. and Whitney, S. M. (2016). Evolving *Methanococcoides burtonii* archaeal Rubisco for improved photosynthesis and plant growth. *Scientific Reports* **6(1), 22284.**

Wilson, R. H., Martin-Avila, E., Conlan, C. and Whitney, S. M. (2018). An improved *Escherichia coli* screen for Rubisco identifies a protein-protein interface that can enhance CO_2-fixation kinetics. *Journal of Biological Chemistry* **293(1), 18-27.**