Engineering bacterial microcompartments into chloroplasts: towards carboxysomes in C3 plants

Maureen R. Hanson[1], Myat T. Lin[1], A. Elizabete Carmo-Silva[2], Martin A.J. Parry[2]

[1]Department of Molecular Biology and Genetics, Cornell University, Biotechnology Building, Ithaca, NY 14853 USA
[2] Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, United Kingdom

Corresponding author: Maureen R. Hanson (mrh5@cornell.edu) Department of Molecular Biology and Genetics, Cornell University, Biotechnology Building, Ithaca, NY 14853 USA
Fax: 1-607-255-6249 Phone: 1-607-254-4833

Myat T. Lin: mtl84@cornell.edu
A. Elizabete Carmo-Silva: e.carmosilva@lancaster.ac.uk
Martin A.J. Parry: m.parry@lancaster.ac.uk

Running title: Introducing Microcompartments into Chloroplasts
Keywords: carbon-concentrating mechanism, chloroplast, Rubisco, photosynthesis, carboxysome, chloroplast transformation, Nicotiana, *Synechococcus elongatus*, transgenic, transplastomic

Total word count: Main text: Figure legends: Acknowledgments: References:

Significance Statement:

Photosynthesis in C3 plants is limited by features of the carbon-fixing enzyme Rubisco, which exhibits a low turnover rate and can also react with O_2, leading to photorespiration. In cyanobacteria, bacterial microcompartments known as carboxysomes improve photosynthetic efficiency by concentrating CO_2 near Rubisco. Thus transferring the carbon-concentrating mechanism from cyanobacteria to C3 plants is an attractive approach to improve crop photosynthesis. Here we review recent progress in engineering cyanobacterial β-carboxysomes into plant chloroplasts and in replacing C3 plant Rubisco with cyanobacterial Rubisco.

SUMMARY

Photosynthesis in C3 plants is limited by features of the carbon-fixing enzyme Rubisco, which exhibits a low turnover rate and can react with O_2 instead of CO_2, leading to photorespiration. In cyanobacteria, bacterial microcompartments known as carboxysomes improve the efficiency of photosynthesis by concentrating CO_2 near the enzyme Rubisco. Cyanobacterial Rubisco enzymes are faster than those of C3 plants, though have lower specificity toward CO_2 than the land plant enzyme. Replacement of land plant Rubisco by faster bacterial variants with lower CO_2 specificity will improve photosynthesis only if a microcompartment capable of concentrating CO_2 can also be installed into the chloroplast. We review current information about cyanobacterial microcompartments and carbon-concentrating mechanisms, plant transformation strategies, replacement of Rubisco in a model C3 plant with cyanobacterial Rubisco, and progress toward synthesizing a carboxysome in chloroplasts.

INTRODUCTION

Improving the efficiency of photosynthesis has been progressively more recognized as a strategy for increasing yield of biomass and food crops (Long *et al.* 2015, Ort *et al.* 2015). While plant breeding has been enormously successful in increasing the percentage of a plant's resources that comprise the desired product, many crop species have reached a plateau from which further improvement by traditional means is unlikely (Long *et al.* 2006, Long*, et al.* 2015). Synthetic biology offers the opportunity to increase the conversion of solar energy through such strategies as altering light capture and conversion and carbon uptake and conversion (Ort *et al.* 2015). This review will focus on a possible tactic to improve carbon fixation: installation of a carbon concentrating mechanism (CCM) presently utilized by cyanobacteria but not by vascular plants (Price *et al.* 2008, Price *et al.* 2013, Zarzycki *et al.* 2013). The cyanobacterial CCM requires the presence of a bacterial microcompartment known as the carboxysome, as well as bicarbonate transporters (Rae *et al.* 2013). We will discuss current progress and prospects for engineering functional carboxysomes into chloroplasts.

CARBON-CONCENTRATION MECHANISMS: COMPENSATION FOR THE UNFAVOURABLE TRAITS OF THE CARBON- FIXING ENZYME RUBISCO

Catalysis of the reaction of ribulose-1,5-bisphosphate (RuBP) with atmospheric CO_2 by the enzyme Rubisco results in fixation of carbon into organic molecules. However, as the name indicates, "Ribulose-1,5-bisphosphate carboxylase/oxygenase" can also catalyze a reaction with oxygen (O_2), leading to the process known as photorespiration (Figure 1a). Reaction with O_2 rather than CO_2 results in loss of some fixed carbon as CO_2, loss of NH_3, as well as consumption of ATP, thus decreasing the efficiency of photosynthesis (reviewed by (Parry *et al.* 2013).

The specificity factor ($S_{C/O}$) of Rubisco defines the relative reactivity of the enzyme towards the two gaseous substrates, CO_2 and O_2. $S_{C/O} = V_c \times K_o / V_o \times K_c$, where V_c and V_o represent the maximum velocities of carboxylation and oxygenation, and K_c and K_o represent the Michaelis-Menten constants for CO_2 and O_2. Environmental changes through geological time have shaped evolution and diversification of Rubisco by providing selective pressures that favour changes in Rubisco structure resulting in improved performance (Tcherkez *et al.* 2006, Christin *et al.* 2008b). Rubisco in higher plants is a rather complex enzyme, composed of nuclear-encoded small subunits and chloroplast-encoded large subunits (Spreitzer and Salvucci 2002), and dependent on interaction with diverse molecular chaperones essential for biogenesis and catalysis (Hauser *et al.* 2015b). Even a single point mutation in the Rubisco sequence can disrupt the assembly of Rubisco in tobacco (Foyer *et al.* 1993). As a consequence, Rubisco evolution and diversification is relatively slow. Yet, significant variation is known to exist in nature and Rubisco enzymes with higher turnover or superior catalytic efficiency and/or specificity towards CO_2 have been identified. Even among the same Form 1B type Rubisco enzymes shared by higher plants including both C3 and C4 species and some cyanobacteria, remarkably diverse kinetic properties can be found (Figure 1b). Models suggest that some of these have the potential to improve photosynthetic CO_2 assimilation in major C3 crops such as wheat (Delgado *et al.* 1995, Galmes *et al.* 2005, Carmo-Silva *et al.* 2015, Prins *et al.* 2016).

The CCM in C4 plants

Decreased CO_2 concentration in the atmosphere millions of years ago led to the evolution of photosynthetic CCMs in plants and algal species (Christin *et al.* 2008a, Sage and Stata 2015). In C4 plants, such as maize and sugarcane, carbon is initially fixed by phosphoenolpyruvate carboxylase (PEPC), with production of an organic acid containing 4 carbons (hence the C4

name). Subsequent decarboxylation of the C4 acids results in accumulation of high concentrations of CO_2 around Rubisco, thus favouring carboxylation over oxygenation of RuBP. In addition to the specialized photosynthetic biochemistry, the leaves of many C4 plants show anatomical features associated with the CCM. These characteristics are known as Kranz anatomy (comprehensively reviewed by (Dengler and Nelson 1999). The initial fixation of inorganic carbon occurs in the mesophyll cells, where PEPC is exclusively located. C4 acids are translocated to the bundle sheath cells, where they undergo decarboxylation. The released CO_2 is subsequently assimilated by Rubisco, which is exclusively located in the chloroplasts of the bundle sheath cells (Edwards *et al.* 1985). Chemical modification and increased thickness of the bundle sheath cell walls, and reduction of the surface area of bundle sheath exposure to intercellular spaces decrease leakage of CO_2 back to the mesophyll cells so that CO_2 accumulates (Furbank *et al.* 1989, Brown and Byrd 1993, Evans and Von Caemmerer 1996, Jenkins 1997, Kiirats *et al.* 2002).

The increased CO_2 concentration in the bundle sheath cells results in low RuBP oxygenation and, consequently, low rates of photorespiration and increased rates and efficiency of photosynthesis in C4 plants (Kanai and Edwards 1999). The presence of a CCM makes C4 photosynthesis especially competitive in conditions that promote carbon loss through photorespiration, such as high temperatures, high light intensities and decreased water availability causing low intercellular CO_2 concentrations as a consequence of stomatal closure. As a result, C4 plants tend to have lower stomatal conductance, lower transpiration rate and higher water use efficiency than their C3 counterparts.

The CCM in cyanobacteria

Cyanobacteria have evolved a mechanism that allows them to utilize faster Rubisco enzymes that are more sensitive to oxygen. Inorganic carbon mainly in the form of bicarbonate ion is accumulated within cytosol through optimally regulated bicarbonate transporters and CO_2 uptake systems (Price*, et al.* 2008). As a result, the intracellular concentration of bicarbonate ion in cyanobacteria is typically well over 10 mM or two to three orders of magnitude higher than that observed outside the cells (Woodger *et al.* 2005). By encapsulating Rubisco and carbonic anhydrase in a protein microcompartment known as a carboxysome and surrounding the enzyme with higher concentrations of CO_2, Rubisco is less likely to react with oxygen (Badger and Price 2003). This strategy is so powerful that it has independently evolved in two lineages of bacteria that contain two different types of Rubisco. Form 1A Rubisco is found in α-carboxysomes, while Form 1B Rubisco is located in β-carboxysomes (Badger *et al.* 2002). Genes present in cyanobacteria that make α- or β-carboxysomes have been well studied through analysis of mutants and considerable information is available about the structural arrangement of proteins present in the two types of carboxysomes (Figure 2).

Incorporation of either α- or β-carboxysomes into crop plants is an attractive strategy for improving photosynthesis (Price*, et al.* 2013, Zarzycki*, et al.* 2013, Whitehead *et al.* 2014). Each type of microcompartment has potential advantages and disadvantages as a candidate for engineering into plants. One advantage of utilizing α-carboxysomes is the fact that their components have already been assembled in a heterologous system along with functional Rubisco, following expression of 10 genes from *Halothiobacillus neapolitanus* in *E. coli* (Bonacci *et al.* 2012), while β-carboxysomes have not yet been synthesized in *E. coli* or microbes other than cyanobacteria. However, phylogenetically, Form 1B Rubisco is related to Rubisco present in land plants (Badger*, et al.* 2002), suggesting that land plant chaperones are more likely to be effective on enzymes normally present in β-carboxysomes. Use of either α- or β-carboxysomes in chloroplasts will require replacement of the endogenous land plant Rubisco with a cyanobacterial Rubisco in order to take advantage of better kinetic properties in a high CO_2 environment. Aside from the advantage of improved turnover rates with the cyanobacterial enzyme, land plant Rubisco might likely be more difficult to engineer to pack into either type of

carboxysomes than replacing the endogenous Rubisco with a cyanobacterial enzyme that normally assembles into a microcompartment. This review will focus on β-carboxysomes because at this writing, there are no reports of autotrophic plants that fix carbon with Form 1A Rubisco, while plants containing cyanobacterial Form 1B Rubisco have been reported (Lin *et al.* 2014b, Occhialini *et al.* 2016).

Components Required for Introducing Microcompartments into Chloroplasts

Recently, much progress has been made in characterization and manipulation of bacterial microcompartments for applications in synthetic biology and metabolic engineering. A targeting or encapsulation peptide was first identified at the N terminus of the propionaldehyde dehydrogenase enzyme from the propanediol utilization microcompartment (Fan *et al.* 2010). Similar encapsulation peptides are now known to be present as N- or C-terminal extensions in core enzymes of bacterial microcompartments including CcmN protein from β-carboxysomes (Kinney *et al.* 2012, Aussignargues *et al.* 2015). These encapsulation peptides are capable of targeting foreign proteins and introducing new functions to heterologously produced bacterial microcompartments (Choudhary *et al.* 2012, Lassila *et al.* 2014, Lawrence *et al.* 2014, Lin *et al.* 2014a). A recent study demonstrated that a mixture of shell proteins from an α-carboxysome and a β-carboxysome can assemble into chimeric microcompartments (Cai *et al.* 2015). In addition, shell proteins with modified "pores" have been shown to be incorporated into carboxysomes (Cai*, et al.* 2015). Thus, it is now a real possibility to apply synthetic biology approaches to customize microcompartments for biotechnological purposes.

We chose to investigate the assembly of microcompartments by β-carboxysome shell proteins from *Synechococcus elongatus* PCC7942 in plant chloroplasts because they represent one of the best characterized bacterial microcompartments with regard to structural organization. Genes encoding the protein shell and internal proteins of both α- or β-carboxysomes have been identified through genetic analysis. Currently the proteins CcmK2, CcmO, and CcmL are known to be the most important for synthesis of the shell layer of the β-carboxysome. Based on its crystal structure, it is accepted that CcmK2 represents the most abundant protein of the icosahedral shell and forms its facets (Kerfeld *et al.* 2005). CcmL forms pentamers that are believed to cap the vertices of the icosahedral shell (Tanaka *et al.* 2008). Although the exact role of CcmO remains unclear, it has been proposed to occupy the edges along two adjacent facets (Rae *et al.* 2012). In addition, either CcmK3 or CcmK4 may be required for the fully functional β-carboxysome (Rae*, et al.* 2012). The hexameric shell proteins form central pores that are believed to be selectively permeable.

In β-carboxysomes from *S. elongatus* PCC7942, CcmM58 is an internal protein that is known to interact with Rubisco. CcmM58 carries three so-called small-subunit-like domains (SSLD) in its C-terminal region that are also present in CcmM35, a 35kD isoform that is translated from an internal start site (Figure 2) (Long *et al.* 2010). CcmM58 has also been shown to interact with the carboxysomal carbonic anhydrase (CcaA) and another internal protein, CcmN, which possesses a short encapsulation peptide that binds to the major shell protein CcmK2 (Long *et al.* 2007, Kinney*, et al.* 2012).

One rapid method to investigate interaction of proteins in chloroplasts is to express them transiently by Agroinfiltration (Fischer *et al.* 1999, Yang *et al.* 2000). The species *Nicotiana benthamiana* is particularly amenable to this method, which also works reasonably well for *Nicotiana tabaccum*, in our experience (Goodin *et al.* 2002). T-DNA vectors that will express one or more genes are introduced into a binary vector and incorporated into an *Agrobacterium tumefaciens* strain, which is then infiltrated into a leaf. More than one Agrobacterium strain, containing T-DNA vectors carrying a different set of genes, can be infiltrated simultaneously to express a variety of combinations of genes (Hanson and Sattarzadeh 2014). Following a brief period for expression (2-5 days), protein expression and interactions can be monitored.

We chose to examine tissue for the localization of Yellow Fluorescent Protein (YFP) fused to CcmO. When this protein is expressed by itself, the protein remains diffuse in the stroma. However, when CcmO-YFP is co-expressed with different combinations of (1) CcmK2 and CcmL (2) CcmK2 and CcmM58, and (3) CcmK2, CcmL and CcmM58, fluorescent punctate loci are visualized in the chloroplast stroma (Figure 3). Characterization at a higher resolution by transmission electron microscopy (TEM) indicates that these punctate loci are highly organized spherical bodies about 80-110 nm in size (Lin*, et al.* 2014a). These structures appear to have a double-layered shell with a less organized inner layer. In comparison, β-carboxysomes in native *S. elongatus* PCC7942 are icosahedral bodies about 175 nm in size with a single-layered shell. The smaller size and lack of angular nature of the microcompartments formed by the shell protein expressed in the chloroplasts indicate the importance of internal components in the proper assembly of β-carboxysomes. This is in agreement with recent studies illuminating the central role played by Rubisco-CcmM35 interaction during the assembly process of β-carboxysomes (Cameron *et al.* 2013). In our previous study, we also demonstrated that the encapsulation peptide from CcmN that is 17 amino-acid long (N17) is able to target YFP to the shell assembly formed by CcmK2 and CcmO-YFP (Lin*, et al.* 2014a). The observation that a foreign protein can be targeted into carboxysomes by the N17 peptide opens up the possibility of engineering microcompartments with new functions in chloroplasts.

STRATEGY CONSIDERATIONS FOR REPLACEMENT OF LAND PLANT RUBISCO WITH CYANOBACTERIAL RUBISCO

The location of the gene (*rbcL*) encoding the large subunit (LS) of Rubisco in the chloroplast genome and genes for the Rubisco small subunit (SS) in the nuclear genome poses a special challenge for engineering cyanobacterial Rubisco into plants. Transplastomic plants (plants containing transgenes in the plastid) have been reported in which the tobacco *rbcL* gene has been replaced with a large subunit gene from another species, resulting in a hybrid enzyme. Unfortunately, the resultant transplastomic plants exhibited impaired photosynthesis, some requiring sucrose or high CO_2 for growth (reviewed in (Whitney *et al.* 2011a, Hanson *et al.* 2013). Reduced function of hybrid enzymes is not unexpected, given that 8 large and 8 small subunits must assemble into a functional complex.

Expression of both the nuclear and chloroplast-encoding subunits must be prevented for complete replacement of a native Rubisco with a heterologous version. Most plants contain multiple nuclear genes (*rbcS*) encoding the small subunit of Rubisco (Spreitzer 2003), so in order to prevent their expression, either an RNA silencing or mutagenesis strategy must be carried out. Antisense tobacco lines with reduced Rubisco have been previously produced (Rodermel *et al.* 1988, Quick *et al.* 1991, Makino and Sage 2007), but any RNA silencing strategy is susceptible to possible epigenetic loss of silencing. Another strategy for removing a tobacco multigene family is to use recently developed CRISPR-Cas9 technology (Belhaj *et al.* 2013, Jiang *et al.* 2013), though at this writing the method has not yet been used to mutagenize a large plant gene family.

For those plant species in which plastid transformation is feasible, the *rbcL* gene can easily be targeted for deletion or replacement by the natural homologous recombination system that exists in chloroplasts (Maliga 2004)(Figure 4). Often the knockout of either a nuclear or plastid component of a complex that contains subunits encoded by both genomes results in loss of all proteins in the complex due to protein instability when one partner is missing (Stern *et al.* 2004). Therefore, loss of the LS in a species is likely to lead to loss of the corresponding SS (Rodermel*, et al.* 1988). Replacement of the LS with a different species' LS might rescue the endogenous SS, but only if the hybrid enzyme can assemble and stabilize the endogenous SS.

Different strategies can be envisioned for replacement of an endogenous C3 plant Rubisco with a different species' Rubisco. A knockout of the *rbcL* gene is readily accomplished in those species in which chloroplast transformation is possible. By simultaneously deleting the

endogenous *rbcL* and replacing it with genes encoding a foreign Rubisco, it should be possible to obtain autotrophic plants, though perhaps ones requiring high CO_2 or growth on sucrose unless the new Rubisco is highly functional in air and expressed at adequate levels. For those species in which no chloroplast transformation method is available, either a targeted or untargeted mutagenesis of *rbcL* will be required, and the mutant plants will require sucrose unless functional foreign *rbcL* and *rbcS* genes have been incorporated into the nucleus before undertaking the mutagenesis.

A barrier to the use of nuclear transgenes for introducing foreign Rubisco has been the lower expression level of nuclear-encoded transgenic proteins that is typically achieved in stably transformed plants, less than 1% total soluble protein (TSP) when a single gene including a standard 35S promoter cassette is used (Conley *et al.* 2011). However, new methods are promising for increasing the levels of protein from nuclear transgenes. For example, a system based on a DNA virus amplification system resulted in accumulation of foreign protein as 10% of total soluble protein (Dugdale *et al.* 2014). Another possibility is to use the same strategy that appears to have evolved in C3 plants for high-level nuclear expression of the SS of Rubisco: express the protein from multiple nuclear loci. The amount of Rubisco needed to support adequate rates of photosynthesis and plant growth will depend on the catalytic properties of the new Rubisco to be incorporated. Most C3 plants have leaves in which Rubisco represents 25-50% of total soluble protein. If the new Rubisco is 5 times faster, and able to carboxylate at such speed, then only 5-10% Rubisco soluble protein should be needed for growth at the same rate, though more will be needed if the goal is to enhance growth.

A FIRST STEP FOR STABLE INCORPORATION OF CARBOXYSOMES INTO CHLOROPLASTS

We decided to replace the tobacco *rbcL* gene with both the *rbcL* and *rbcS* genes from *S. elongatus* PCC7942, without altering the tobacco nuclear *rbcS* genes (Figure 4) (Lin*, et al.* 2014b, Occhialini*, et al.* 2016). For expression of a foreign gene from a chloroplast genome, a transgene construct must carry a 5' untranslated region (UTR) that provides translation signals and a terminator region for stabilization of the RNA transcript. One of the most effective such 5' UTRs is derived from the bacteriophage T7 gene 10 (Maliga 2004, Yang *et al.* 2013). Multiple tandem copies of the Shine-Dalgarno sequences from gene 10 have been shown to improve translation over single copies (Drechsel and Bock 2011). An issue regarding multigene constructs is the possibility of unwanted homologous recombination between components of the same transgene operon, or between transgenes and homologous regions of the chloroplast genome from which a gene regulatory sequence was derived (Gray *et al.* 2009). Thus, it can be advantageous to use a gene regulatory sequence such as a promoter or terminator from a different plant species, provided that the sequence is effective in the heterologous system. Alternatively, if the endogenous sequence that is used is not large, the probability of undesirable recombination is reduced (Mudd *et al.* 2014).

Two or more chloroplast genes can be transcribed in a polycistronic transcript from a single promoter. However, in chloroplasts, downstream coding regions in transcripts are sometimes not efficiently translated from polycistronic transcripts, unlike in bacteria. To enhance the possibility of effective translation, an Intercistronic Expression Element (IEE), which carries a sequence recognized by a nuclease, can be placed between different transgenes on a polycistronic transcript (Zhou *et al.* 2007). Presence of the IEE has been reported to result in processing of polycistronic transcripts into monocistronic transcripts, enhancing protein accumulation (Bock 2013, Bock 2015)(Figure 4b).

Another enhancement of expression of foreign proteins in chloroplasts can sometimes result from changing the codon usage of the gene of interest to avoid codons that are rarely used in chloroplasts (Reed *et al.* 2001, Franklin *et al.* 2002). Nevertheless, the optimal codon usage in a transgene is not always straightforward to identify (Weiss *et al.* 2012, Sugiura 2014)

and may need empirical optimization. A further aspect of the codon region that can be important is the N-terminal region, either because of the presence of a downstream box sequence that affects protein accumulation (Gray *et al.* 2011) or N-terminal features that affect protein stability (Apel *et al.* 2010, Elghabi *et al.* 2011, Gray*, et al.* 2011).

Chloroplast transformants (transplastomic plants) are usually made by particle bombardment with a biolistic device (Maliga and Tungsuchat-Huang 2014). Tungsten or gold microparticles can be coated with plasmid DNA and then propelled into plant cells, now usually accelerated by a blast of helium. If regenerated transgenic plants are desired, the tissue that is bombarded must be capable of regeneration from tissue cultures. While transplastomic plants can be made from a variety of crop plants in addition to tobacco (Hanson*, et al.* 2013, Bock 2015), Nicotiana species remain the superior model system for chloroplast transformation, due to high regenerability and rapid growth in culture. The spectinomycin/streptomycin resistance gene *aadA*, used in the initial report of stable transplastomic plants (Svab *et al.* 1990), continues to be the most effective selectable marker for chloroplast transformation. Following bombardment of tobacco leaf tissue with plasmids carrying the transgene operons and the selectable marker and incubation without selection for a few days, tissue is removed to regeneration media containing the antibiotic and induced to regenerate shoots. Usually such shoots contain a mixture of transformed and untransformed chloroplasts, a state termed "heteroplasmy". Each cell that was bombarded contained multiple chloroplasts and each chloroplast within that cell contained multiple chloroplast genomes. In order to obtain shoots with a uniform chloroplast genome (a state termed "homoplasmy"), repeated rounds of regeneration are often needed. During dedifferentiation of leaf cells in culture cell and redifferentiation into shoot meristems, the number of plastids/cell is reduced, facilitating random segregation of transgenic and non-transgenic plastids. Selection ensures that only shoots resistant to the selectable marker will regenerate (Lutz *et al.* 2007, Maliga and Tungsuchat-Huang 2014). In our experience, homoplasmic plants can usually be obtained after a second round of regeneration and selection.

Transplastomic tobacco plants fixing carbon with cyanobacterial Rubisco

In our initial constructs designed to replace endogenous Rubisco with *S. elongatus* PCC7942 Rubisco, we included either genes for the assembly factor RbcX, required for proper Rubisco folding in some cyanobacteria (Saschenbrecker *et al.* 2007), or for CcmM35, a protein involved in organization of Rubisco within the carboxysome (Cameron*, et al.* 2013) (Figure 2). We thought that one of these proteins might facilitate proper folding of the cyanobacterial Rubisco, which normally requires a GroEL chaperonin and its cofactor GroEs. Land plants contain Cpn60, a homolog of GroEL/GroEs comprised of Cpn10 and Cpn20 subunits. Both cyanobacteria and land plants require an additional Rubisco assembly factor known as Raf1 (Feiz *et al.* 2012, Hauser *et al.* 2015a). While some cyanobacterial Rubisco enzymes require RbcX for assembly, *S. elongatus* PCC7942 does not in its native cell (Emlyn-Jones *et al.* 2006); however, we considered that it might be required in a heterologous system. Recently, the inclusion of Arabidopsis Raf1 in a transplastomic plant in which the tobacco *rbcL* was replaced with Arabidopsis *rbcL* was shown to improve hybrid Rubisco assembly and photosynthesis (Whitney *et al.* 2015).

Our two initial lines, named SeLSX (for *S. elongatus* Rubisco Large and Small subunit with RbcX) and SeLSM35 (CcmM35 instead of RbcX) both were able to grow autotrophically provided that they were given elevated CO_2 in their environment. Though their growth rate was much slower than wild-type, the plants had normal morphology and could flower and set seed. By DNA blot, we demonstrated that the tobacco *rbcL* coding region had been completely replaced by the cyanobacterial transgene operons. Immunoblots of total cellular protein revealed no detectable tobacco SS or LS; evidently the absence of the tobacco LS destabilized the tobacco SS even though the *rbcS* genes were intact (Lin*, et al.* 2014b).

Both SeLSX and SeLSM35 were designed with *rbcL* and *rbcS* coding regions altered to be more similar to endogenous chloroplast transcripts. These same coding regions were used in two subsequent transgene operons, one in which neither RbcX nor CcmM35 was expressed (plant line designated SeLS) and a second line in which a *yfp* gene was fused N-terminally to *ccmM35* and expressed along with the cyanobacterial *rbcL* and *rbcS* genes (plant line designated SeLSYM35). Elimination of *rbcX* resulted in attachment of the tobacco *rbcL* terminator to the *rbcS* gene. Each of the 4 transgene operons differed in terminator configuration; furthermore, the *yfp::ccmM35* gene was synthesized with an altered codon usage, unlike the *ccmM35* gene present in SeLSM35 (Figure 4).

The two new plant lines grew much faster than the original lines in 3% CO_2; in fact the growth of the SeLS line was only slightly slower than wild-type tobacco (Figure 5). In order to determine whether transgene expression levels could explain the different growth rates, we examined RNA and protein-level expression. By performing RNA blots, we discovered that transcripts carrying the Arabidopsis *rps16* terminators did not accumulate, thus indicating this terminator is not suitable for tobacco transformation vectors. A second important finding was that the presence of the IEE sequence did not always result in efficient cleavage between the open reading frames, resulting in the continued presence of polycistronic transcripts (Occhialini*, et al.* 2016). Despite the absence of many monocistronic *rbcS* transcripts, evidently there was sufficient expression of the downstream *rbcS* genes for adequate protein synthesis for assembly of Rubisco. When we compared Rubisco content in the four lines in 3% CO_2, we observed that the upper leaves of SeLS contained significantly more Rubisco than SeLSX, perhaps explaining the enhanced growth (Figure 5a-b). One possibility we cannot rule out is that the original transgene constructs had an unknown inhibitory effect on other aspects of chloroplast gene expression, completely unrelated to Rubisco expression. Nevertheless, the rapid growth of SeLS in 3% CO_2 demonstrates that concept that nitrogen use efficiency may improve when less enzyme is needed for carbon fixation (Carmo-Silva*, et al.* 2015).

Despite the fact that plants SeLSM35 and SeLSYM35 had greater Rubisco content than SeLS or SeLSX, their growth was slower than SeLS in 3% CO_2 (Figure 5) (Occhialini *et al.* 2016). The unusual organization of Rubisco in the M35-containing plants may explain the slower growth compared to SeLS. Chloroplasts from the SeLSX, SeLS, SeLSM35, and SeLSM35Y plants were imaged by electron microcopy and the location of Rubisco was determined with specific antibodies. These experiments revealed that the cyanobacterial Rubisco was dispersed within the stroma in SeLSX and SeLS, but aggregated in large bodies in the SeLSM35 and SeLSYM35 plants, suggesting that M35 in the chloroplast binds to the cyanobacterial LS and results in aggregation of Rubisco (Figure 5c-d). The Rubisco/M35 aggregates in tobacco chloroplasts appear analogous to aggregates observed when an M35::YFP fusion was expressed in *S. elongatus* PCC7942 mutant carrying a *ccmK2-ccmO* deletion (Cameron*, et al.* 2013).

Aggregation of Rubisco in the absence of a functional carboxysome could potentially prevent supply of CO_2 to the enzyme, thereby reducing plant growth rate. Purified Rubisco isolated from the transgenic lines expressing the cyanobacterial Rubisco (SeLS and SeLSX) had essentially identical kinetic constants (Figure 6) to wild type Rubisco isolated from the cyanobacteria (Figure 1b). However, where the enzyme was expressed in association with CcmM35, the Rubisco was slower and had a slightly greater affinity for CO_2. This suggests that the incorporation of CcmM35 into the Rubisco complex disrupts the enzyme structure possibly by displacing the Rubisco small subunits and changes the catalytic properties; 3D structural determination of the complexes will be required to clarify the mechanism for these changes.

While SeLS grows substantially faster than SeLSX 3% CO_2, we have observed that growth rates of the two lines are quite similar in 0.9% CO_2. This observation is in line with the observation that CO_2 assimilation in SeLS and SeLSX is quite similar when measured on a leaf area basis. Also, Rubisco μmole sites/m^2 is not significantly different into the two lines when

grown in 0.9% CO_2 (Occhialini*, et al.* 2016). Currently we do not understand why there is relatively more Rubisco in SeLS in 3% CO_2 than in 0.9% CO_2.

ADDITIONAL ENGINEERING NEEDED TO INSTALL THE CYANOBACTERIAL CCM IN CHLOROPLASTS

As the next step, shell proteins such as CcmK2, CcmO and CcmL and other internal proteins of the β-carboxysome including CcmM58, CcaA and CcmN will need to be co-expressed with the core components, made up of Rubisco and CcmM35. These additional proteins may be expressed from similar synthetic operons from chloroplast genome as described above. Alternatively, proteins that are known to be less abundant such as CcmL, CcaA, and either CcmK3 or CcmK4 may be expressed through nuclear transformation and imported into chloroplast stroma with stromal transit peptides. A recent study described the surprising finding that a chimeric protein derived from fusing the three SSU domains from CcmM58, CcaA and the encapsulation peptide from CcmN is able to functionally replace four internal proteins, CcmM58, CcmM35, CcaA and CcmN, thus greatly simplifying the engineering of β-carboxysomes into chloroplasts (Gonzalez-Esquer *et al.* 2015).

It may be necessary to use synthetic operons with different gene arrangements and regulatory elements such as the ribosome binding sites to attain the optimal expression ratios for successful assembly of these proteins into functional β-carboxysomes in chloroplasts. In addition, the number of carboxysomes per chloroplast will also need to be optimized so that the metabolic flux through these carboxysomes is efficiently integrated into the Calvin Cycle. In considering only the relative volume of a cyanobacterium vs. a chloroplast, to maintain the same number of microcompartments per unit volume, a typical tobacco chloroplast would need about 70 carboxysomes (Figure 7). However, a recent theoretical analysis indicates that approximately 2700 carboxysomes per chloroplast would be needed to achieve the same amount of Rubisco per chloroplast as in wild-type plants (McGrath and Long 2014). On the other hand, we expect the optimal number of carboxysomes per chloroplast would be much lower than 2700 due to the significantly higher catalytic rate of the cyanobacterial Rubisco and altered metabolic flux through photorespiration upon installation of the cyanobacterial CCM in the chloroplast.

Incorporating a functional microcompartment containing Rubisco and other internal proteins is not sufficient for creating a complete CCM in plants.

Chloroplasts naturally contain carbonic anhydrase within the stroma (the soluble portion of the chloroplast). Expression of stromal carbonic anhydrase must be prevented through RNA silencing or mutagenesis in order for the carboxysome to function properly. In previous studies with transgenic plants in which stromal carbonic anhydrase was knocked down, little effect on photosynthetic efficiency in mature leaves was detected (Majeau *et al.* 1994, Ferreira *et al.* 2008). In cyanobacteria, when carbonic anhydrase was expressed in the cytoplasm (the bacterial equivalent of the chloroplast stroma), the cells required high CO_2 because the carbon-concentrating mechanism did not function (Price and Badger 1989).

In addition to the carboxysome, transporters will be needed to supply the chloroplast with bicarbonate, which can then be converted to CO_2 by carbonic anhydrase within the carboxysome. Heterologous expression of such transporters has already begun (Pengelly *et al.* 2014, Atkinson *et al.* 2015). Transporters alone, without a microcompartment, are predicted to be able to enhance photosynthesis (McGrath and Long 2014), but additional transporter types and further engineering will be needed to realize potential gains.

ACKNOWLEDGMENTS

Research on engineering carboxysomes into chloroplasts was supported by Biotechnology and Biological Sciences Research Council under grant number BB/I024488/1 to M.A.J.P., the

National Science Foundation under grant number EF-1105584 to M.R.H. and the National Institute of General Medical Sciences of the National Institutes of Health under award number F32GM103019 to M.T.L. M.R.H. and M.T.L. also acknowledge support from the Cornell University Biotechnology Resource Center (NIH S10RR025502) for the shared Zeiss LSM 710 Confocal Microscope. E.C.S. and M.A.J.P. also acknowledge support from the 20:20 Wheat Institute Strategic Program (BBSRC BB/J/00426X/1). We thank Kevin Hines for the carboxysome model and Alessandro Occhialini for transmission electron microscopy and immunolocalization.

FIGURE LEGENDS

Figure 1. (a) Diagram of Rubisco-catalyzed reactions in C3 plants. Carboxylation of RuBP initiates the Calvin Cycle and leads to production of carbohydrates. Oxygenation of RuBP initiates photorespiration and results in the net loss of fixed CO_2 and NH_3, and consumption of energy. The carboxylation of RuBP by Rubisco and subsequent reactions in the Calvin Cycle are shown in green arrows. The oxygenation of RuBP by Rubisco and the subsequent photorespiration process are shown in red arrows. RuBP = Ribulose-1,5-bisphosphate; 3PGA = 3-phosphoglycerate; 2PG = 2-phosphoglycolate; G3P = glyceraldehyde 3-phosphate. (b) Rubisco specificity ($S_{C/O}$) and maximum carboxylation rate (V_C) in tobacco (C3 model species), wheat (C3 crop), limonium (*L. gibertii*, C3 with high specificity), maize and sorghum (C4 crops), *S. elongatus* (cyanobacteria, β-carboxysome species) and *Rhodospirilum rubrum* (purple bacteria, α-carboxysome species). Data are averages of values reported by: (Jordan and Ogren 1981, Jordan and Ogren 1983, Jordan and Ogren 1984, Makino *et al.* 1985, Parry *et al.* 1987, Parry *et al.* 1989, Sage and Seemann 1993, Kane *et al.* 1994, Delgado*, et al.* 1995, Uemura *et al.* 1996, Whitney *et al.* 1999, Whitney *et al.* 2001, Pearce 2006, Mueller-Cajar *et al.* 2007, Parry *et al.* 2007, Kubien *et al.* 2008, Mueller-Cajar and Whitney 2008, Sharwood *et al.* 2008, Carmo-Silva *et al.* 2010, Genkov *et al.* 2010, Whitney *et al.* 2011b, Occhialini*, et al.* 2016, Prins*, et al.* 2016).

Figure 2. (a) Structural model of β-carboxysome in *S. elongatus* PCC7942. Data on protein location and assembly as described by *Cameron et al*., (2013). Please note that the role of CcmO remains unclear although it is represented as pseudohexamers here. Carboxysome model courtesy of Kevin Hines. (b) Operons present in *S. elongatus* PCC7942 (Heinhorst *et al.* 2014).

Figure 3. Transient expression of cyanobacterial shell proteins in *Nicotiana benthamiana*. (a) Diffuse localization of CcmO-YFP within chloroplasts when CcmK2 and CcmL are absent. (b) Formation of YFP fluorescent punctate loci (green) within the chloroplasts (red) of *N. benthamiana* transiently expressing CcmK2, CcmL and CcmO-YFP. (c) and (d) Spherical structures formed in plants transiently expression CcmK2-YFP, CcmL, and CcmO-YFP. Microscopic images from Lin *et al*., (2014), with permission.

Figure 4. Schematics of gene arrangements in synthetic operons to express cyanobacterial Rubisco from chloroplast. (a) Replacement of the tobacco *rbcL* gene with cyanobacterial transgenes and a selectable marker by homologous recombination. (b) A typical gene arrangement in a synthetic operon with two generic cyanobacterial genes, *ccm1* and *ccm2*. Each gene is followed by a different terminator sequence denoted as T1 or T2. In the intergenic region between *ccm1* and *ccm2*, an intercistronic expression element (IEE) and a ribosome binding site (RBS) are inserted immediately upstream of *ccm2* for processing of the dicistronic transcript into monocristronic ones for more efficient translation of the downstream gene, *ccm2*. (c) Schematics of synthetic operons in four different constructs to express cyanobacterial transgenes from the tobacco *rbcL* locus (Occhialini*, et al.* 2016). *Se LS*, *Se SS*, *RbcX*, and *M35* represent *rbcL*, *rbcS*, *rbcX*, and *ccmM35* genes from *S. elongatus* PCC7942 respectively. RBS: ribosome binding site. Single (SD) or triple (SD18) Shine-Dalgarno sequences from T7 gene 10 (Drechsel and Bock 2011).

Figure 5. (a) Relative size of wild-type (wt) and SeLS, SeLSX and SeLSM35 tobacco plants when grown for 42 days in 3% CO_2. (b) Rubisco content per leaf area in wt tobacco growing in air vs. wt and transgenic lines growing in 3% CO_2. (c-d) Localization of cyanobacterial Rubisco in the chloroplast stroma of the SeLS (c) and SeLSYM35 (d) tobacco transplastomic lines probed with anti-Se Rubisco antibody and a secondary antibody conjugated with 10 nm gold particles (black circles or dots). Scale bars = 500 nm. Modified from Occhialini *et al.* 2016, with permission.

Figure 6. Rubisco specificity ($S_{C/O}$) and maximum carboxylation rate (V_C) in wild-type and SeLS, SeLSX, SeLSM35 and SeLSYM35 transplastomic tobacco plants. Data from Occhialini *et al.* (2016).

Figure 7. Diagram of carboxysomes in cyanobacteria and hypothetically in a chloroplast. Diagram visualizes approximate relative size of *S. elongatus* cell and typical tobacco chloroplast. A typical *S. elongatus* cell contains an average of 4 carboxysomes (Savage *et al.* 2010). However, modeling indicates more will be needed for adequate Rubisco to be present (McGrath and Long 2014).

REFERENCES

Apel, W., Schulze, W.X. and Bock, R. (2010) Identification of protein stability determinants in chloroplasts. *Plant J.*, **63**, 636-650.

Atkinson, N., Feike, D., Mackinder, L.C., Meyer, M.T., Griffiths, H., Jonikas, M.C., Smith, A.M. and McCormick, A.J. (2015) Introducing an algal carbon-concentrating mechanism into higher plants: location and incorporation of key components. *Plant Biotechnol. J.*, **in press**.

Aussignargues, C., Paasch, B.C., Gonzalez-Esquer, R., Erbilgin, O. and Kerfeld, C.A. (2015) Bacterial microcompartment assembly: The key role of encapsulation peptides. *Commun. Integr. Biol.*, **8**, e1039755.

Badger, M.R., Hanson, D. and Price, G.D. (2002) Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. *Funct. Plant Biol.*, **29**, 161-173.

Badger, M.R. and Price, G.D. (2003) CO2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J. Exp. Bot.*, **54**, 609-622.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S. and Nekrasov, V. (2013) Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods*, **9**, 39.

Bock, R. (2013) Strategies for metabolic pathway engineering with multiple transgenes. *Plant Mol. Biol.*, **83**, 21-31.

Bock, R. (2015) Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. *Annu. Rev. Plant Biol.*, **66**, 211-241.

Bonacci, W., Teng, P.K., Afonso, B., Niederholtmeyer, H., Grob, P., Silver, P.A. and Savage, D.F. (2012) Modularity of a carbon-fixing protein organelle. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 478-483.

Brown, R.H. and Byrd, G.T. (1993) Estimation of bundle sheath cell conductance in C4 species and O_2 insensitivity of photosynthesis. *Plant Physiol.*, **103**, 1183-1188.

Cai, F., Sutter, M., Bernstein, S.L., Kinney, J.N. and Kerfeld, C.A. (2015) Engineering bacterial microcompartment shells: chimeric shell proteins and chimeric carboxysome shells. *ACS Synth. Biol.*, **4**, 444-453.

Cameron, J.C., Wilson, S.C., Bernstein, S.L. and Kerfeld, C.A. (2013) Biogenesis of a bacterial organelle: the carboxysome assembly pathway. *Cell*, **155**, 1131-1140.

Carmo-Silva, A.E., Keys, A.J., Andralojc, P.J., Powers, S.J., Arrabaca, M.C. and Parry, M.A. (2010) Rubisco activities, properties, and regulation in three different C4 grasses under drought. *J. Exp. Bot.*, **61**, 2355-2366.

Carmo-Silva, E., Scales, J.C., Madgwick, P.J. and Parry, M.A. (2015) Optimizing Rubisco and its regulation for greater resource use efficiency. *Plant Cell Environ.*, **38**, 1817-1832.

Choudhary, S., Quin, M.B., Sanders, M.A., Johnson, E.T. and Schmidt-Dannert, C. (2012) Engineered protein nano-compartments for targeted enzyme localization. *PLoS One*, **7**, e33342.

Christin, P.A., Besnard, G., Samaritani, E., Duvall, M.R., Hodkinson, T.R., Savolainen, V. and Salamin, N. (2008a) Oligocene CO_2 decline promoted C4 photosynthesis in grasses. *Curr. Biol.*, **18**, 37-43.

Christin, P.A., Salamin, N., Muasya, A.M., Roalson, E.H., Russier, F. and Besnard, G. (2008b) Evolutionary switch and genetic convergence on *rbcL* following the evolution of C4 photosynthesis. *Mol. Biol. Evol.*, **25**, 2361-2368.

Conley, A.J., Zhu, H., Le, L.C., Jevnikar, A.M., Lee, B.H., Brandle, J.E. and Menassa, R. (2011) Recombinant protein production in a variety of Nicotiana hosts: a comparative analysis. *Plant Biotechnol. J.*, **9**, 434-444.

Delgado, E., Medrano, H., Keys, A.J. and Parry, M.A.J. (1995) Species variation in Rubisco specificity factor. *J. Exp. Bot.*, **46**, 1775-1777.

Dengler, N.G. and Nelson, T. (1999) Leaf structure and development in C4 plants. In *C4 Plant Biology* (Sage, R.F. and Monson, R.K. eds). New York: Academic Press, pp. 133-172.

Drechsel, O. and Bock, R. (2011) Selection of Shine-Dalgarno sequences in plastids. *Nucleic Acids Res.*, **39**, 1427-1438.

Dugdale, B., Mortimer, C.L., Kato, M., James, T.A., Harding, R.M. and Dale, J.L. (2014) Design and construction of an in-plant activation cassette for transgene expression and recombinant protein production in plants. *Nat. Protoc.*, **9**, 1010-1027.

Edwards, G.E., Ku, M.S.B. and Monson, R.K. (1985) C4 photosynthesis and its regulation. In *Photosynthetic Mechanisms and the Environment* (Barber, J. and Baker, N.R. eds). Amsterdam: Elsevier Science Publishers B.V. (Biomedical Division), pp. 287-327.

Elghabi, Z., Karcher, D., Zhou, F., Ruf, S. and Bock, R. (2011) Optimization of the expression of the HIV fusion inhibitor cyanovirin-N from the tobacco plastid genome. *Plant Biotechnol. J.*, **9**, 599-608.

Emlyn-Jones, D., Woodger, F.J., Price, G.D. and Whitney, S.M. (2006) RbcX can function as a rubisco chaperonin, but is non-essential in Synechococcus PCC7942. *Plant Cell Physiol.*, **47**, 1630-1640.

Evans, J.R. and Von Caemmerer, S. (1996) Carbon dioxide diffusion inside leaves. *Plant Physiol.*, **110**, 339-346.

Fan, C.G., Cheng, S.Q., Liu, Y., Escobar, C.M., Crowley, C.S., Jefferson, R.E., Yeates, T.O. and Bobik, T.A. (2010) Short N-terminal sequences package proteins into bacterial microcompartments. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 7509-7514.

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

Ferreira, F.J., Guo, C. and Coleman, J.R. (2008) Reduction of plastid-localized carbonic anhydrase activity results in reduced Arabidopsis seedling survivorship. *Plant Physiol.*, **147**, 585-594.

Fischer, R., Vaquero-Martin, C., Sack, M., Drossard, J., Emans, N. and Commandeur, U. (1999) Towards molecular farming in the future: transient protein expression in plants. *Biotechnol. Appl. Biochem.*, **30 (Pt 2)**, 113-116.

Foyer, C.H., Nurmi, A., Dulieu, H. and Parry, M.A.J. (1993) Analysis of 2 Rubisco-deficient tobacco mutants, H7 and Sp25 - Evidence for the production of Rubisco large subunits in the Sp25 mutant that form clusters and are inactive. *J. Exp. Bot.*, **44**, 1445-1452.

Franklin, S., Ngo, B., Efuet, E. and Mayfield, S.P. (2002) Development of a GFP reporter gene for Chlamydomonas reinhardtii chloroplast. *Plant J.*, **30**, 733-744.

Furbank, R.T., Jenkins, C.L.D. and Hatch, M.D. (1989) CO_2 concentrating mechanism of C4 photosynthesis - Permeability of isolated bundle sheath-cells to inorganic carbon. *Plant Physiol.*, **91**, 1364-1371.

Galmes, J., Flexas, J., Keys, A.J., Cifre, J., Mitchell, R.A.C., Madgwick, P.J., Haslam, R.P., Medrano, H. and Parry, M.A.J. (2005) Rubisco specificity factor tends to be larger in plant species from drier habitats and in species with persistent leaves. *Plant Cell Environ.*, **28**, 571-579.

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. *J. Biol. Chem.*, **285**, 19833-19841.

Gonzalez-Esquer, C.R., Shubitowski, T.B. and Kerfeld, C.A. (2015) Streamlined construction of the cyanobacterial CO_2-fixing organelle via protein domain fusions for use in plant synthetic biology. *Plant Cell*, **27**, 2637-2644.

Goodin, M.M., Dietzgen, R.G., Schichnes, D., Ruzin, S. and Jackson, A.O. (2002) pGD vectors: versatile tools for the expression of green and red fluorescent protein fusions in agroinfiltrated plant leaves. *Plant J.*, **31**, 375-383.

Gray, B.N., Ahner, B.A. and Hanson, M.R. (2009) Extensive homologous recombination between introduced and native regulatory plastid DNA elements in transplastomic plants. *Transgenic Res.*, **18**, 559-572.

Gray, B.N., Yang, H., Ahner, B.A. and Hanson, M.R. (2011) An efficient downstream box fusion allows high-level accumulation of active bacterial beta-glucosidase in tobacco chloroplasts. *Plant Mol. Biol.*, **76**, 345-355.

Hanson, M.R., Gray, B.N. and Ahner, B.A. (2013) Chloroplast transformation for engineering of photosynthesis. *J. Exp. Bot.*, **64**, 731-742.

Hanson, M.R. and Sattarzadeh, A. (2014) Fluorescent labeling and confocal microscopic imaging of chloroplasts and non-green plastids. *Methods Mol. Biol.*, **1132**, 125-143.

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

Hauser, T., Popilka, L., Hartl, F.U. and Hayer-Hartl, M. (2015b) Role of auxiliary proteins in Rubisco biogenesis and function. *Nature Plants*, **1**, 1-11.

Heinhorst, S., Cannon, G.C. and Shively, J.M. (2014) Carboxysomes and Their Structural Organization in Prokaryotes. In *Nanomicrobiology* (Barton, L.L., Bazylinski, D.A. and Xu, H. eds): Springer New York, pp. 75-101.

Jenkins, C.L.D. (1997) The CO_2 concentrating mechanism of C-4 photosynthesis: Bundle sheath cell CO_2 concentration and leakage. *Aust. J. Plant Physiol.*, **24**, 543-547.

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B. and Weeks, D.P. (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.*, **41**, e188.

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

Jordan, D.B. and Ogren, W.L. (1983) Species variation in kinetic properties of ribulose 1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.*, **227**, 425-433.

Jordan, D.B. and Ogren, W.L. (1984) The CO_2/O_2 specificity of ribulose 1,5-bisphosphate carboxylase oxygenase : Dependence on ribulosebisphosphate concentration, pH and temperature. *Planta*, **161**, 308-313.

Kanai, R. and Edwards, G.E. (1999) The biochemistry of C4 photosynthesis. In *C4 Plant Biology* (Sage, R.F. and Monson, R.K. eds). New York: Academic Press, pp. 49-87.

Kane, H.J., Viil, J., Entsch, B., Paul, K., Morell, M.K. and Andrews, T.J. (1994) An improved method for measuring the CO_2/O_2 specificity of ribulosebisphosphate carboxylase-oxygenase. *Aust. J. Plant Physiol.*, **21**, 449-461.

Kerfeld, C.A., Sawaya, M.R., Tanaka, S., Nguyen, C.V., Phillips, M., Beeby, M. and Yeates, T.O. (2005) Protein structures forming the shell of primitive bacterial organelles. *Science*, **309**, 936-938.

Kiirats, O., Lea, P.J., Franceschi, V.R. and Edwards, G.E. (2002) Bundle sheath diffusive resistance to CO_2 and effectiveness of C_4 photosynthesis and refixation of photorespired CO_2 in a C_4 cycle mutant and wild-type *Amaranthus edulis*. *Plant Physiol.*, **130**, 964-976.

Kinney, J.N., Salmeen, A., Cai, F. and Kerfeld, C.A. (2012) Elucidating essential role of conserved carboxysomal protein CcmN reveals common feature of bacterial microcompartment assembly. *J. Biol. Chem.*, **287**, 17729-17736.

Kubien, D.S., Whitney, S.M., Moore, P.V. and Jesson, L.K. (2008) The biochemistry of Rubisco in Flaveria. *J. Exp. Bot.*, **59**, 1767-1777.

Lassila, J.K., Bernstein, S.L., Kinney, J.N., Axen, S.D. and Kerfeld, C.A. (2014) Assembly of robust bacterial microcompartment shells using building blocks from an organelle of unknown function. *J. Mol. Biol.*, **426**, 2217-2228.

Lawrence, A.D., Frank, S., Newnham, S., Lee, M.J., Brown, I.R., Xue, W.F., Rowe, M.L., Mulvihill, D.P., Prentice, M.B., Howard, M.J. and Warren, M.J. (2014) Solution structure of a bacterial microcompartment targeting peptide and its application in the construction of an ethanol bioreactor. *ACS Synth. Biol.*, **3**, 454-465.

Lin, M.T., Occhialini, A., Andralojc, P.J., Devonshire, J., Hines, K.M., Parry, M.A. and Hanson, M.R. (2014a) b-Carboxysomal proteins assemble into highly organized structures in *Nicotiana* chloroplasts. *Plant J.*, **79**, 1-12.

Lin, M.T., Occhialini, A., Andralojc, P.J., Parry, M.A. and Hanson, M.R. (2014b) A faster Rubisco with potential to increase photosynthesis in crops. *Nature*, **513**, 547-550.

Long, B.M., Badger, M.R., Whitney, S.M. and Price, G.D. (2007) Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. *J. Biol. Chem.*, **282**, 29323-29335.

Long, B.M., Tucker, L., Badger, M.R. and Price, G.D. (2010) Functional cyanobacterial b-carboxysomes have an absolute requirement for both long and short forms of the CcmM protein. *Plant Physiol.*, **153**, 285-293.

Long, S.P., Marshall-Colon, A. and Zhu, X.G. (2015) Meeting the global food demand of the future by engineering crop photosynthesis and yield potential. *Cell*, **161**, 56-66.

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

Lutz, K.A., Azhagiri, A.K., Tungsuchat-Huang, T. and Maliga, P. (2007) A guide to choosing vectors for transformation of the plastid genome of higher plants. *Plant Physiol.*, **145**, 1201-1210.

Majeau, N., Arnoldo, M.A. and Coleman, J.R. (1994) Modification of carbonic anhydrase activity by antisense and over-expression constructs in transgenic tobacco. *Plant Mol. Biol.*, **25**, 377-385.

Makino, A., Mae, T. and Ohira, K. (1985) Enzymic properties of ribulose-1,5-bisphosphate carboxylase/oxygenase purified from rice leaves. *Plant Physiol.*, **79**, 57-61.

Makino, A. and Sage, R.F. (2007) Temperature response of photosynthesis in transgenic rice transformed with 'sense' or 'antisense' rbcS. *Plant Cell Physiol.*, **48**, 1472-1483.

Maliga, P. (2004) Plastid transformation in higher plants. *Annu. Rev. Plant Biol.*, **55**, 289-313.

Maliga, P. and Tungsuchat-Huang, T. (2014) Plastid transformation in Nicotiana tabacum and Nicotiana sylvestris by biolistic DNA delivery to leaves. *Methods Mol. Biol.*, **1132**, 147-163.

McGrath, J.M. and Long, S.P. (2014) Can the cyanobacterial carbon-concentrating mechanism increase photosynthesis in crop species? A theoretical analysis. *Plant Physiol.*, **164**, 2247-2261.

Mudd, E.A., Madesis, P., Avila, E.M. and Day, A. (2014) Excision of plastid marker genes using directly repeated DNA sequences. *Methods Mol. Biol.*, **1132**, 107-123.

Mueller-Cajar, O., Morell, M. and Whitney, S.M. (2007) Directed evolution of rubisco in *Escherichia coli* reveals a specificity-determining hydrogen bond in the form II enzyme. *Biochemistry*, **46**, 14067-14074.

Mueller-Cajar, O. and Whitney, S.M. (2008) Directing the evolution of Rubisco and Rubisco activase: first impressions of a new tool for photosynthesis research. *Photosynth. Res.*, **98**, 667-675.

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

Ort, D.R., Merchant, S.S., Alric, J., Barkan, A., Blankenship, R.E., Bock, R., Croce, R., Hanson, M.R., Hibberd, J.M., Long, S.P., Moore, T.A., Moroney, J., Niyogi, K.K., Parry, M.A., Peralta-Yahya, P.P., Prince, R.C., Redding, K.E., Spalding, M.H., van Wijk, K.J., Vermaas, W.F., von Caemmerer, S., Weber, A.P., Yeates, T.O., Yuan, J.S. and Zhu, X.G. (2015) Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, 8529-8536.

Parry, M.A., Andralojc, P.J., Scales, J.C., Salvucci, M.E., Carmo-Silva, A.E., Alonso, H. and Whitney, S.M. (2013) Rubisco activity and regulation as targets for crop improvement. *J. Exp. Bot.*, **64**, 717-730.

Parry, M.A.J., Keys, A.J. and Gutteridge, S. (1989) Variation in the specificity factor of C3 higher plant Rubiscos determined by the total consumption of ribulose-P_2. *J. Exp. Bot.*, **40**, 317-320.

Parry, M.A.J., Madgwick, P.J., Carvalho, J.F.C. and Andralojc, P.J. (2007) Prospects for increasing photosynthesis by overcoming the limitations of Rubisco. *J. Agr. Sci.*, **145**, 31-43.

Parry, M.A.J., Schmidt, C.N.G., Cornelius, M.J., Millard, B.N., Burton, S., Gutteridge, S., Dyer, T.A. and Keys, A.J. (1987) Variations in properties of ribulose-1,5-bisphosphate carboxylase from various species related to differences in amino-acid-sequences. *J. Exp. Bot.*, **38**, 1260-1271.

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

Pengelly, J.J., Forster, B., von Caemmerer, S., Badger, M.R., Price, G.D. and Whitney, S.M. (2014) Transplastomic integration of a cyanobacterial bicarbonate transporter into tobacco chloroplasts. *J. Exp. Bot.*, **65**, 3071-3080.

Price, G.D. and Badger, M.R. (1989) Expression of human carbonic anhydrase in the cyanobacterium Synechococcus PCC7942 creates a high CO(2)-requiring phenotype : evidence for a central role for carboxysomes in the CO(2) concentrating mechanism. *Plant Physiol.*, **91**, 505-513.

Price, G.D., Badger, M.R., Woodger, F.J. and Long, B.M. (2008) Advances in understanding the cyanobacterial CO2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants. *J. Exp. Bot.*, **59**, 1441-1461.

Price, G.D., Pengelly, J.J., Forster, B., Du, J., Whitney, S.M., von Caemmerer, S., Badger, M.R., Howitt, S.M. and Evans, J.R. (2013) The cyanobacterial CCM as a source of

genes for improving photosynthetic CO2 fixation in crop species. *J. Exp. Bot.*, **64**, 753-768.

Prins, A., Orr, D.J., Andralojc, P.J., Reynolds, M.P., Carmo-Silva, E. and Parry, M.A.J. (2016) Rubisco catalytic properties of wheat relatives provide scope for improving photosynthesis. *J. Exp. Bot.*

Quick, W.P., Schurr, U., Scheibe, R., Schulze, E.D., Rodermel, S.R., Bogorad, L. and Stitt, M. (1991) Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with "antisense" rbcS : I. Impact on photosynthesis in ambient growth conditions. *Planta*, **183**, 542-554.

Rae, B.D., Long, B.M., Badger, M.R. and Price, G.D. (2012) Structural determinants of the outer shell of b-carboxysomes in *Synechococcus elongatus* PCC 7942: roles for CcmK2, K3-K4, CcmO, and CcmL. *PLoS One*, **7**, e43871.

Rae, B.D., Long, B.M., Whitehead, L.F., Forster, B., Badger, M.R. and Price, G.D. (2013) Cyanobacterial carboxysomes: microcompartments that facilitate CO2 fixation. *J. Mol. Microbiol. Biotechnol.*, **23**, 300-307.

Reed, M.L., Wilson, S.K., Sutton, C.A. and Hanson, M.R. (2001) High-level expression of a synthetic red-shifted GFP coding region incorporated into transgenic chloroplasts. *Plant J.*, **27**, 257-265.

Rodermel, S.R., Abbott, M.S. and Bogorad, L. (1988) Nuclear-organelle interactions: nuclear antisense gene inhibits ribulose bisphosphate carboxylase enzyme levels in transformed tobacco plants. *Cell*, **55**, 673-681.

Sage, R.F. and Seemann, J.R. (1993) Regulation of ribulose-1,5-bisphosphate carboxylase/oxygenase activity in response to reduced light intensity in C4 plants. *Plant Physiol.*, **102**, 21-28.

Sage, R.F. and Stata, M. (2015) Photosynthetic diversity meets biodiversity: the C4 plant example. *J. Plant Physiol.*, **172**, 104-119.

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

Savage, D.F., Afonso, B., Chen, A.H. and Silver, P.A. (2010) Spatially ordered dynamics of the bacterial carbon fixation machinery. *Science*, **327**, 1258-1261.

Sharwood, R.E., von Caemmerer, S., Maliga, P. and Whitney, S.M. (2008) The catalytic properties of hybrid Rubisco comprising tobacco small and sunflower large Subunits mirror the kinetically equivalent source Rubiscos and can support tobacco growth. *Plant Physiol.*, **146**, 83-96.

Spreitzer, R.J. (2003) Role of the small subunit in ribulose-1,5-bisphosphate carboxylase/oxygenase. *Arch. Biochem. Biophys.*, **414**, 141-149.

Spreitzer, R.J. and Salvucci, M.E. (2002) Rubisco: structure, regulatory interactions, and possibilities for a better enzyme. *Annu. Rev. Plant Biol.*, **53**, 449-475.

Stern, D.B., Hanson, M.R. and Barkan, A. (2004) Genetics and genomics of chloroplast biogenesis: maize as a model system. *Trends Plant Sci*, **9**, 293-301.

Sugiura, M. (2014) Plastid mRNA translation. *Methods Mol. Biol.*, **1132**, 73-91.

Svab, Z., Hajdukiewicz, P. and Maliga, P. (1990) Stable transformation of plastids in higher plants. *Proc. Natl. Acad. Sci. U. S. A.*, **87**, 8526-8530.

Tanaka, S., Kerfeld, C.A., Sawaya, M.R., Cai, F., Heinhorst, S., Cannon, G.C. and Yeates, T.O. (2008) Atomic-level models of the bacterial carboxysome shell. *Science*, **319**, 1083-1086.

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. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 7246-7251.

Uemura, K., Suzuki, Y., Shikanai, T., Wadano, A., Jensen, R.G., Chmara, W. and Yokota, A. (1996) A rapid and sensitive method for determination of relative specificity of RuBisCO from various species by anion-exchange chromatography. *Plant Cell Physiol.*, **37**, 325-331.

Weiss, C., Bertalan, I. and Johanningmeier, U. (2012) Effects of rare codon clusters on the expression of a high-turnover chloroplast protein in Chlamydomonas reinhardtii. *J. Biotechnol.*, **160**, 105-111.

Whitehead, L., Long, B.M., Price, G.D. and Badger, M.R. (2014) Comparing the in vivo function of alpha-carboxysomes and beta-carboxysomes in two model cyanobacteria. *Plant Physiol.*, **165**, 398-411.

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

Whitney, S.M., Birch, R., Kelso, C., Beck, J.L. and Kapralov, M.V. (2015) Improving recombinant Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1 chaperone. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, 3564-3569.

Whitney, S.M., Houtz, R.L. and Alonso, H. (2011a) Advancing our understanding and capacity to engineer nature's CO2-sequestering enzyme, Rubisco. *Plant Physiol.*, **155**, 27-35.

Whitney, S.M., Sharwood, R.E., Orr, D., White, S.J., Alonso, H. and Galmes, J. (2011b) Isoleucine 309 acts as a C_4 catalytic switch that increases ribulose-1,5-bisphosphate carboxylase/oxygenase (rubisco) carboxylation rate in *Flaveria*. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 14688-14693.

Whitney, S.M., von Caemmerer, S., Hudson, G.S. and Andrews, T.J. (1999) Directed mutation of the Rubisco large subunit of tobacco influences photorespiration and growth. *Plant Physiol.*, **121**, 579-588.

Woodger, F.J., Badger, M.R. and Price, G.D. (2005) Sensing of inorganic carbon limitation in *Synechococcus* PCC7942 is correlated with the size of the internal inorganic carbon pool and involves oxygen. *Plant Physiol.*, **139**, 1959-1969.

Yang, H., Gray, B.N., Ahner, B.A. and Hanson, M.R. (2013) Bacteriophage 5' untranslated regions for control of plastid transgene expression. *Planta*, **237**, 517-527.

Yang, Y., Li, R. and Qi, M. (2000) In vivo analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves. *Plant J.*, **22**, 543-551.

Zarzycki, J., Axen, S.D., Kinney, J.N. and Kerfeld, C.A. (2013) Cyanobacterial-based approaches to improving photosynthesis in plants. *J. Exp. Bot.*, **64**, 787-798.

Zhou, F., Karcher, D. and Bock, R. (2007) Identification of a plastid intercistronic expression element (IEE) facilitating the expression of stable translatable monocistronic mRNAs from operons. *Plant J.*, **52**, 961-972.